A Longitudinal Systems Biology Approach to Stroke Recovery

Submitted by

Vinh Nguyen

BPsychSc (Hons)

In total fulfillment of the requirement of the degree of

Doctor of Philosophy

Department of Occupational Therapy, Social Work and Social Policy

School of Allied Health, Human Services and Sport

College of Science, Health and Engineering

La Trobe University

Victoria, Australia

August 2020

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2D-DIGE: 2-dimensional–difference in gel electrophoresis
AF: Atrial fibrillation
AGC: Automatic gain control
AMPA: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANOVA: Analysis of variance
APOA1: apolipoprotein A1
APOB1: apolipoprotein B1
ATP: Adenosine triphosphate
BBB: Blood brain barrier
BDI: Beck Depression Inventory
BH: Benjamini-Hochberg
BI: Barthel Index
C1…9: Complement component 1… 9
CAM: Cell adhesion molecule
CASZ1: castor zinc finger 1
CBC: Complete blood count
CDK6: cyclin dependent kinase 6
CFD: Complement factor D
CFI: Complement factor I
CID: Collision-induced dissociation
CLU: Clusterin
CNS: Central nervous system
CRP: C-reactive protein
CT: Computerized tomography
DAF: Decay accelerating factor
DAMP: Damage associated molecular pattern
DAVID: Database for Annotation, Visualization and Integrated Discovery
$DMSO_4$: Dimethyl sulfoxide
DNA: Deoxyribonucleic acid
DTI: Diffusion Tensor Imaging
EDTA: Ethylenediaminetetraacetic acid
ESI: Electrospray ionisation
FBE: Full Blood Examination
FDR: False discovery rate
FN1: Fibronectin 1
GGEA: Gene graph enrichment analysis
GO: Gene Ontology
GP1BB: Glycoprotein Ib platelet beta subunit
GP9: Glycoprotein 9
GSEA: Gene set enrichment analysis
GSN: Gelsolin
GWAS: Genome wide association study
HADS: Hospital Anxiety and Depression Scale
HMGB1: high mobility group box 1
HSA: Human serum albumin
ICH: Intracerebral hemorrhage

IAA: Iodoacetamide
IS: Ischemic stroke
IL-6: Interleukin-6
iTRAQ: Isobaric tagging for relative and absolute quantification
IVIG: Intravenous immunoglobulin
KEGG: Kyoto Encyclopedia of Genes and Genomes
LACI: Lacunar cerebral infarct
LBP: Lipopolysaccharide binding protein
LC-MS: Liquid chromatography – mass spectrometry
LFQ: Label free quantitation
MAC: Membrane attack complex
MADRS: Montgomery–Åsberg Depression Rating Scale
MASP1…2: Mannan-binding lectin serine protease 1…2
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
MBL: mannose-binding lectin
mRNA: Messenger ribonucleic acid
MRI: Magnetic resonance imagining
mRS: Modified Rankin Scale
MW: Molecular weight
MS: Mass spectrometry
msigdb: Molecular signatures database
MRM: Multiple reaction monitoring
NC: Neutrophil count
NCBI: National Centre for Biotechnology Information
NDMA: *N*-methyl-D-aspartic acid
NES: Normalized enrichment score
NIHSS: National Institute of Health Stroke Scale
NMR: nuclear magnetic resonance
NTA: Neuroscience Trials Australia
OXPHOS: Oxidative phosphorylation
PACI: Partial anterior circulation infarct
PANTHER: Protein Analysis THrough Evolutionary Relationships
PF4: Platelet factor 4
PHQ-2: 2 item Patient Health Questionnaire
POCI: Posterior circulation infarct
PRDX3: Peroxiredoxin 3
PRDX5: Peroxiredoxin 5
PROZ: Protein Z
PITX: Paired like homeodomain 2
PSD: Post-stroke depression
QT-PCR: Quantitative polymerase chain reaction
RAC1: Ras-related C3 botulinum toxin substrate 1
RNS: Reactive nitrogen species
ROS: Reactive oxygen species
S100A12: Calcium binding protein A12
SAA: serum amyloid A
SBML: Systems Biology Markup Language

sC5b-9: Soluble membrane attack complex
SCIG: subcutaneous immunoglobulin
SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIGMA: Structured Interview Guide for the MADRS
SLE: Systemic lupus erythematosus
SNP: Single nucleotide polymorphism
SRM: Single reaction monitoring
SST: Serum separating tube
START: STroke imAging pRevention and Treatment
START_EXTEND: EXtending the time for Thrombolysis in Emergency Neurological Deficits
START_PrePARE: PREdiction and Prevention to Achieve Optimal Recovery Endpoints after stroke
TACI: Total anterior circulation infarct
TIA: Transient ischemic attack
TCA: Tricarboxylic acid
TCEP: Tris(2-carboxyethyl)phosphine
TF: Transferrin
TLR4: Toll-like receptor 4
TOAST: Trial of ORG 10172
TOF: Time of flight
TNF-a: Tumor necrosis factor alpha
tPA: Tissue plasminogen activator
VEGAS: Versatile gene based association study
VEGF: Vascular endothelial growth factor
VTN: Vitronectin
WBC: White blood cell

Thesis Abstract

Stroke is one of the highest causes of mortality, second only to heart attack. Globally, there are over 15 million stroke cases reported every year. One third of stroke cases results in death, another third face permanent disability and the final third recover to a functional degree. Past stroke research has focused on acute interventions as the primary method of reducing mortality and providing the best circumstances for recovery. However, there is a lack of understanding of the biological systems underlying the heterogenous and longitudinal timeline of stroke recovery. Accordingly, this thesis has employed systems biology approaches such as transcriptomics and proteomics of human blood samples to examine the molecules and the biological systems involved in the timeline of stroke recovery. From a review conducted on systems biology studies in the acute period, this thesis has identified that platelet function, response to oxidative stress and signalling for leukocyte chemotaxis are biological systems that have evidence for expression at both the transcriptomics and proteomics level. Empirical investigation of blood samples and clinical data from a longitudinal stroke cohort (n=219) permitted investigation of the relationship between biological systems and recovery. Analysis of routine bloods in the acute phase (n=156) indicated a relationship between leukocyte and neutrophil counts to cognitive function. Proteomic studies using blood samples from a subset of the stroke cohort (n=44) demonstrated that mechanisms of complement signalling and immune function are factors likely to influence post-stroke recovery. Further elucidation of these mechanisms via a longitudinal proteomics study (n = 60) suggested that changes in expression of molecular signals that regulate immune system functioning may contribute to important clinical outcomes such as immunosuppression that extend from the weeks to months post-stroke. Overall, the studies in this thesis provide evidence for the role of the peripheral immune system and especially the complement system in association with stroke recovery into 3 months and 12 months

post-stroke. Further, these findings provide a basis for systems biology techniques to better understand the molecular and biological processes involved in longitudinal stroke recovery. In future, such studies will be useful to consider towards the development of new therapeutics and personalised medicine for stroke survivors.

Statement of Authorship

This thesis includes work by the author that has been published or accepted for publication as described in the text. Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

This thesis includes 3 co-authored manuscripts published or under review in peer-reviewed journals. The theoretical framework, experimental data, data analyses, and written material presented in these manuscripts were produced solely by the candidate under the supervision of Prof. Leeanne Carey and Prof. Sheila Crewther except where reference is made below.

The START team collected patient data and stored samples within the infrastructure at the Florey Institute of Neuroscience and Mental Health. Dr. Nina Riddell assisted with creation of figures presented in Chapters 3 and 6. For the samples analysed using proteomics in Chapters 6 and 7, Ms. Rachael Downs and Dr. Pierre Faou ran the liquid chromatography mass spectrometry based proteomics platform and conducted post-processing of proteomics data. Dr Ira Cooke assisted with the post-processing and analysis of proteomics expression in relation to stroke outcomes in Chapter 5.

We acknowledge financial support for conduct of the research from the Commonwealth Scientific and Industrial Research Organization (CSIRO) of Australia, Flagship Collaboration Fund through the Preventative Health Flagship; data collection and project planning by the Stroke Imaging, Prevention and Treatment research team (www.START.csiro.au); support for analysis, write up and/or researchers from the James S. McDonnell Foundation 21st Century Science Initiative in Cognitive Rehabilitation - Collaborative Award (#220020413); Victorian Government’s Operational Infrastructure Support Program; an Australian Research Council Future Fellowship awarded to LMC [#FT0992299]; and a La Trobe University Post Graduate Research Scholarship awarded to VN and supported by the Understanding Diseases Research Focus Group at La Trobe University.

Ethics was approved by the Human Research Ethics Committee of Austin Hospital, Heidelberg (HREC code: H2010/03588) and La Trobe University Human Ethics Committee (HEC10-071).

This work was supported by an Australian Government Research Training Program Scholarship.

Name: Vinh Nguyen

Date: 17/08/2020

Acknowledgements and Dedications

This may be the only time and place that I ever get to stray from the strict confines of scientific writing so here I go. There are too many people to thank after all.

The first and foremost thank you is dedicated to Sheila G. Crewther. I have worked with Sheila since my third year project which puts us on the record for 7-8 years straight of teaching and learning. The fact that we are not yet bored of each other after all this time says something to this student-teacher relationship. After talking to many students around the world, I am totally convinced that there can be no other group of students that are as fortunate as those whom have studied their PhDs under Crewther. I hope that our scientific journey does not end with my completion of this candidature.

Many thanks to Nina Riddell and Loretta Giumarra, both of whom are my seniors in this field. Much of their knowledge and expertise were foundational for me to develop a formative understanding of systems biology and laboratory techniques. Hopefully we'll all get to work together on an exciting study again very soon.

The second thank you is dedicated to Leeanne M. Carey. As a supervisor that is in a different field, I have to thank Leeanne for putting up with my tendency to delve deeper and deeper into the biology. Without Leeanne, this thesis would not have the important and relevant focus on the clinical implications of the results to current and hopefully, future practice.

The next set of thanks goes to Dr. Graeme Hankey and Dr. Tissa Wijeratne. As practicing neurologists at the top of their fields, both Dr. Graeme and Dr. Tissa have aided immensely with the financial and intellectual development of the present thesis. Without them, certain studies and lines of thinking would have never come to fruition. It is thanks to their continuing support that I can continue to work on publish on studies that could not make it into this thesis due to time constraints.

A large thank you to Christopher Armstrong and Ashish Sethi for their expertise in metabolomics. I'll get onto the metabolomics data as soon as I am able to put this thesis to rest, I promise.

Thank you to my family. Parents, siblings and the enormous extended Vietnamese family. The three words that I would use to describe how helpful my family has been throughout this candidature would be: stability, rest and food. Definitely in that order.

A set of thanks goes to friends. If I have shared a meal and/or drink with you in the last 5 years, you are all the reason why I don’t have generalized anxiety disorder. Unfortunately, I cannot specifically name the hundreds of people that I would call friends here as I am running out of space and I don’t want to endorse favouritism.

The last set goes out to the Melbourne music community. Thank you so much.

Chapter 1

Introduction and Thesis Overview

Introduction

Stroke is the second leading cause of global mortality, accounting for 11.8% of deaths worldwide (Feigin, Norrving, & Mensah, 2017) and creating an annual cost burden of $65.5 billion in the United States yearly (Di Carlo, 2009), with an estimated $4850 per month cost for post-stroke patients accessing combined inpatient and outpatient care (Rajsic et al., 2019). Disability and impairment persist in at least 50% of post-stroke survivors and may include multiple functional domains including cognition, motor and speech (Adamson, Beswick, & Ebrahim, 2004). Currently, intravenous thrombolysis and endovascular thrombectomy are the main frontline treatments for acute ischemic stroke (IS), with the goal being rapid recanalization of the vascular supply to the affected areas of the brain. Although successful administration of thrombolytic agent alteplase within 4.5 has been shown to lead to improved motor outcomes (Campbell et al., 2019), the 90-day mortality rates are surprisingly increased for alteplase treated patients compared to standard care (Emberson et al., 2014). Even in patients achieving good acute reperfusion > 90%, 90-day outcomes show that approximately 17.8% of patients die and 9.7% are permanently disabled (Liebeskind et al., 2019). Current and future research in this area must address the post-stroke timeline of recovery and in particular the great need for development of prognostic models of recovery around which to design effective and individualized treatment programs (Lin, Finklestein, & Cramer, 2018), with the ultimate goals of improving clinical outcomes and long-term wellbeing of stroke patients. In response to this need, the aim of this thesis was to comprehensively explore the underlying biological processes affecting stroke recovery from the days to years post-acute first episode of stroke.

Acute interventions have been the major focus of stroke research within the last decade (Catanese, Tarsia, & Fisher, 2017). Currently the administration of the tissue plasminogen activator (tPA) class of thrombolytic therapies within the ‘golden’ 4.5 hours

is of the utmost priority for stroke clinicians (Saver et al., 2010). While the greatest efficacy for recovery can be seen for therapeutics administered as early as possible post IS onset, administration of tPA to intracranial hemorrhage (ICH) patients is a major safety consideration that can lead to risk of vascular compromise and bleeding (Hemphill et al., 2015). Thus, one of the foremost goals in acute clinical investigation of stroke is to ascertain whether an acute intracranial event is hemorrhagic or ischemic in origin and to confirm diagnoses with neuroimaging. This is critical as the commonly used thrombolytic tPA agents of alteplase (Fugate & Rabinstein, 2015) and tenecteplase (Chester et al., 2019) are contraindicated in cases of hemorrhagic stroke (HS) as can exacerbate further bleeding. Thus, one of the current dilemmas in acute stroke management is the acquisition of robust biomarker information to make effective clinical decisions, including the option of use of the relatively simple intravenous delivery of thrombolytic administration of tissue plasminogen tPA, which can be effectively delivered in almost all clinical settings.

Human stroke biology research has been focussed on developing rapid and specific biomarkers for the point-of-care differentiation between HS and IS (Vibha & Misra, 2020). Reliance is placed on brain imaging techniques such as computerized tomography (CT) and/or magnetic resonance imaging (MRI) modalities as the gold standard for stroke diagnosis (Latchaw et al., 2009). Unfortunately however, brain imaging is often not available especially in rural or low-income areas (Misra et al., 2017). However, even though past systematic analyses and studies examining biomarkers to differentiate stroke type have identified serum apolipoprotein C-1 (APOC-1), C-3 (APOC-3) and glial fibrillary acidic protein (GFAP) as distinguishing biomarkers between IS and HS (Misra et al., 2017), none of these biomarkers are routinely incorporated into acute stroke management guidelines nor indicated as a recommendation in the American Stroke Association Early Acute Stroke Management clinical guidelines (Powers et al., 2018).

A further major theme in stroke biology research concerns prediction of post-stroke recovery. The term ‘recovery’ as a concept is often attributed to positive clinical gains in terms of cognitive systems, motor and functional recovery (Lin et al., 2018). However, stroke recovery also involves a variety of biological processes such as resolution of initial central nervous system (CNS) damage, axonal and neuronal repair, angiogenesis and ongoing mediatory inflammatory processes (Cramer, 2018). Current prognostic models of stroke recovery that use basic demographic and clinical factors such as age, independent living and the Glasgow Coma Scale have, to date, not been able to be adequately translated into active clinical practice and have difficulty in addressing heterogenous patterns of individualized patient recovery (Sim, Teece, Dennis, Roffe, & Team, 2016). Therefore, there is an ongoing and unmet need to improve the quality of post-stroke patient care by beginning to use objective markers that are more closely linked to the biological processes that drive stroke recovery. Despite frequent calls for additional research (Lindgren & Maguire, 2016), to date there are no accepted molecular biomarkers of stroke recovery and there is currently inadequate evidence for blood-based molecules that can be recommended as prognostic markers of stroke recovery (Boyd et al., 2017).

Ongoing biomarker research has been and continues to be limited by the available technologies in biology and chemistry fields. Thus, biomarker literature has often utilized targeted techniques such as commercial immunoassays that elucidate single molecules at a time (Sahab, Semaan, & Sang, 2007), utilizing a hypothesis-based approach and prior knowledge. However, this approach is limited and unlikely to be able to represent the biological complexity of complex disease states considering that the human genome is annotated at 22,000 protein coding regions (Maglott, Ostell, Pruitt, & Tatusova, 2011). With the addition of downstream biological processes such as post-translational modifications, degradation and protein-protein interactions, a cross section of distinct

peptide compositions in a human may total up to 1,000,000 protein isoforms (Sharma, Cosme, & Gramolini, 2013). Therefore, investigating an increased number of molecules and subsequent pathways and biological systems would be expected to provide unprecedented ability to address biological understanding in stroke recovery.

Recent developments in systems biology ‘omics’ techniques have allowed for simultaneous identification and quantification of a large number of genes (genetics), messenger ribonucleic acid (mRNA) (transcriptomics) and proteins (proteomics) in any given sample. Indeed, discovery approaches such as RNA sequencing of blood products is able to elucidate over 20,000 probe sets (Koczan, Fitzner, Zettl, & Hecker, 2018) and protein expression profiling in blood is able to elucidate between 200-400 proteins (Geyer, Holdt, Teupser, & Mann, 2017). With the addition of bioinformatics or computational biology approaches, these omics approaches can further add a layer of biological richness to the interpretation of biomarker functioning, as the large number of biomolecules elucidated in a study can be further processed into whole biological pathways, for example as shown in Figure 1 for energy metabolism. Therefore, in the case of post-stroke recovery where biological understanding is not well developed and there are currently no clinically available applications for molecular biomarkers, systems biology discovery approaches may represent an important and cost-effective tool for discovery of the biomarkers and biosystems involved in stroke recovery.

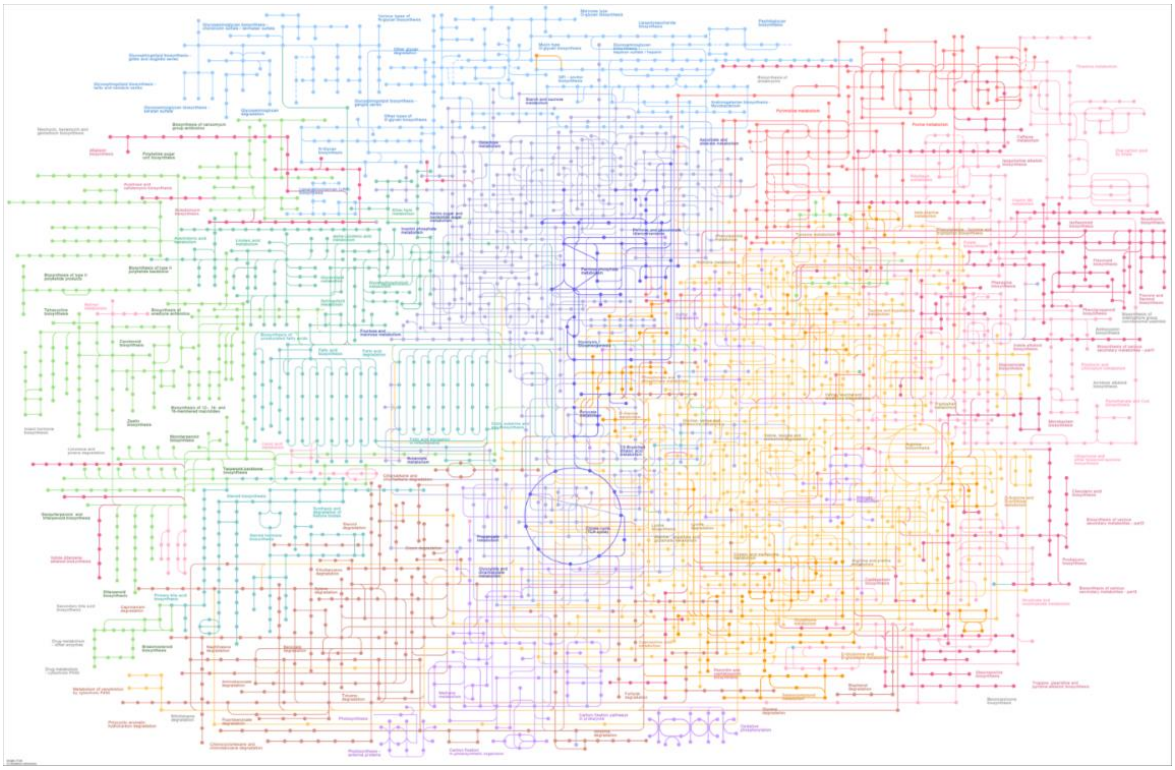

Figure 1. An example of overall metabolic pathways from the Kyoto Encyclopedia of Genes and Genomes database. Used with permission from Kanehisa Laboratories.

Thesis Aims

The overall objective of this thesis is to provide/establish a better understanding and interpretation of the post ischemic stroke longitudinal biological profile associated with recovery by utilizing discovery techniques such as genetics, transcriptomics and proteomics to examine the blood samples of IS patients from acute and longitudinal cohort studies with a focus on identification of predictive markers of clinical outcomes and recovery.

Specifically, the aims of the studies in this thesis are:

(i) to examine the relationship between the IS genome and downstream molecular mechanisms by systematically reviewing genetics, transcriptomics and proteomics findings in stroke studies,

(ii) to examine current clinically used measures of routine blood investigations in acute IS and their role in prognosticating recovery,

(iii) to determine associations of gene/protein systems to clinical recovery outcomes in an IS cohort and

(iv) to provide a better understanding of the longitudinal aspects of protein biomarkers and biosystems in IS recovery.

Although this thesis employs hypothesis-free approaches to data analysis, the current knowledge of the stroke pathophysiology timeline predicts that inflammation, immune, cardiological (Brouns & De Deyn, 2009) and physiological stress (Chamorro et al., 2007) would be the main pathways associated with stroke recovery. Advancing current understanding of the temporal trajectory of these immunological issues is a major expected outcome of the thesis.

Thesis Outline.

Chapter 1 introduces the overall context for this thesis in relation to stroke; the need for blood-based biomarkers of stroke recovery, aims of this thesis and thesis outline.

Chapter 2 aims to provide the biological context for the series of studies by examining published literature of the stroke timeline, from stroke risk to stroke infarct and associated damage and subsequent recovery processes.

Chapter 3 of this thesis sought to systematically review the main biological systems involved in acute stroke by examining previous genetic, transcriptomic and proteomic studies, using untargeted approaches and compared between IS patients and healthy controls. This review has collected the available published datasets from the studies examined while employing a multi-omics systems biology approach to examine

the overlapping molecules and biosystems between the expressed genome, transcriptome and proteome.

Chapter 4 discusses the methodological considerations and issues associated with the use of the present stroke cohort in relation to sample selection and technologies used in the proteomics studies.

Chapter 5 aimed to explore potential associations between the results of common routine hospital haematological blood tests (white blood cells (WBC) and neutrophil counts) employed in acute stroke, with longitudinal post-stroke measures of cognition, depression and functional recovery outcomes at key recovery timepoints including 3 months and 12 months.

Chapter 6 aimed to utilize blood plasma proteomics to explore the relationship between blood samples collected at the time of acute IS and depression outcomes at 3 months post-stroke. The purpose was to validate the untargeted label-free quantitation mass spectrometry methodology on blood plasma in a preliminary analysis. Further, the popular Gene Set Enrichment Analysis (GSEA) bioinformatics approach (Subramanian et al., 2005) was employed to test the associations between database-defined biological pathways and depression outcomes, thereby providing a statistical relationship between biological systems and clinical outcomes.

Chapter 7 aimed to further develop upon the methodologies used in Chapter 6 to address the biosystems involved in the timeline of stroke recovery, from the acute phase extending through the first year post-stroke. This study examined blood proteomics in patients across 3 timepoints: 3-7 days, 3 months and 12 months post-stroke. We also aimed to improve upon existing bioinformatics methods and interpretations by adopting Gene Graph Enrichment Analysis (GGEA) (Geistlinger, Csaba, Küffner, Mulder, & Zimmer, 2011) as an approach to organize expressed molecules into topological

pathways. The GGEA algorithm further utilizes network based statistical approaches to examine the relationship of regulation between proteins across different timepoints. Findings from this study provide insight into the changes in stroke biology while also allowing comment on the regulatory relationships of individual biomarkers over the first year post-stroke.

Chapter 8 concludes with a general discussion of key findings, including clinical implications and directions for future research.

References

Adamson, J., Beswick, A., & Ebrahim, S. (2004). Is stroke the most common cause of disability? *Journal of Stroke and Cerebrovascular Diseases, 13*(4), 171-177. doi:10.1016/j.jstrokecerebrovasdis.2004.06.003

Boyd, L. A., Hayward, K. S., Ward, N. S., Stinear, C. M., Rosso, C., Fisher, R. J., . . . Cramer, S. C. (2017). Biomarkers of stroke recovery: Consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *International Journal of Stroke, 12*(5), 480-493. doi:10.1177/1747493017714176

Brouns, R., & De Deyn, P. P. (2009). The complexity of neurobiological processes in acute ischemic stroke. *Clinical Neurology and Neurosurgery, 111*(6), 483-495. doi:https://doi.org/10.1016/j.clineuro.2009.04.001

Campbell, B. C. V., Ma, H., Ringleb, P. A., Parsons, M. W., Churilov, L., Bendszus, M., . . . Williams, M. (2019). Extending thrombolysis to 4.5-9 h and wake-up stroke using perfusion imaging: A systematic review and meta-analysis of individual patient data. *The Lancet, 394*(10193), 139-147. doi:10.1016/S0140-6736(19)31053-0

Catanese, L., Tarsia, J., & Fisher, M. (2017). Acute ischemic stroke therapy overview. *Circulation Research, 120*(3), 541-558. doi:doi:10.1161/CIRCRESAHA.116.309278

Chamorro, Á., Amaro, S., Vargas, M., Obach, V., Cervera, Á., Gómez-Choco, M., . . . Planas, A. M. (2007). Catecholamines, infection, and death in acute ischemic stroke. *Journal of the Neurological Sciences, 252*(1), 29-35. doi:10.1016/j.jns.2006.10.001

Chester, K. W., Corrigan, M., Schoeffler, J. M., Shah, M., Toy, F., Purdon, B., & Dillon, G. M. (2019). Making a case for the right '-ase' in acute ischemic stroke:

Alteplase, tenecteplase, and reteplase. *Expert Opinion on Drug Safety, 18*(2), 87-96. doi:10.1080/14740338.2019.1573985

Cramer, S. C. (2018). Treatments to promote neural repair after stroke. *Journal of Stroke, 20*(1), 57-70. doi:10.5853/jos.2017.02796

Di Carlo, A. (2009). Human and economic burden of stroke. *Age and Ageing, 38*(1), 4-5. doi:10.1093/ageing/afn282

Emberson, J., Lees, K. R., Lyden, P., Blackwell, L., Albers, G., Bluhmki, E., . . . Donnan, G. (2014). Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: A meta-analysis of individual patient data from randomised trials. *The Lancet, 384*(9958), 1929-1935.

Feigin, V. L., Norrving, B., & Mensah, G. A. (2017). Global burden of stroke. *Circulation Research, 120*(3), 439-448. doi:doi:10.1161/CIRCRESAHA.116.308413

Fugate, J. E., & Rabinstein, A. A. (2015). Absolute and relative contraindications to iv rt-pa for acute ischemic stroke. *The Neurohospitalist, 5*(3), 110-121. doi:10.1177/1941874415578532

Geistlinger, L., Csaba, G., Küffner, R., Mulder, N., & Zimmer, R. (2011). From sets to graphs: Towards a realistic enrichment analysis of transcriptomic systems. *Bioinformatics, 27*(13), 366-373. doi:10.1093/bioinformatics/btr228

Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. (2017). Revisiting biomarker discovery by plasma proteomics. *Molecular Systems Biology, 13*(9), 942-942. doi:10.15252/msb.20156297

Hemphill, J. C., Greenberg, S. M., Anderson, C. S., Becker, K., Bendok, B. R., Cushman, M., . . . Woo, D. (2015). Guidelines for the management of spontaneous

intracerebral hemorrhage. *Stroke, 46*(7), 2032-2060. doi:doi:10.1161/STR.0000000000000069

Koczan, D., Fitzner, B., Zettl, U. K., & Hecker, M. (2018). Microarray data of transcriptome shifts in blood cell subsets during s1p receptor modulator therapy. *Scientific Data, 5*, 180145. doi:10.1038/sdata.2018.145

Latchaw, R. E., Alberts, M. J., Lev, M. H., Connors, J. J., Harbaugh, R. E., Higashida, R. T., . . . Walters, B. (2009). Recommendations for imaging of acute ischemic stroke. *Stroke, 40*(11), 3646-3678. doi:doi:10.1161/STROKEAHA.108.192616

Liebeskind, D. S., Bracard, S., Guillemin, F., Jahan, R., Jovin, T. G., Majoie, … HERMES Collaborators (2019). eTICI reperfusion: defining success in endovascular stroke therapy. *Journal of Neurointerventional Surgery, 11*(5), 433–438. https://doi.org/10.1136/neurintsurg-2018-014127

Lin, D. J., Finklestein, S. P., & Cramer, S. C. (2018). New directions in treatments targeting stroke recovery. *Stroke, 49*(12), 3107-3114. doi:doi:10.1161/STROKEAHA.118.021359

Lindgren, A., & Maguire, J. (2016). Stroke recovery genetics. *Stroke, 47*(9), 2427-2434. doi:doi:10.1161/STROKEAHA.116.010648

Maglott, D., Ostell, J., Pruitt, K. D., & Tatusova, T. (2011). Entrez gene: Gene-centered information at NCBI. *Nucleic Acids Research, 39*, D52-D57. doi:10.1093/nar/gkq1237

Misra, S., Kumar, A., Kumar, P., Yadav, A. K., Mohania, D., Pandit, A. K., . . . Vibha, D. (2017). Blood-based protein biomarkers for stroke differentiation: A systematic review. *Proteomics – Clinical Applications, 11*(9-10), 1700007. doi:10.1002/prca.201700007

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., Becker, K., . . . Tirschwell, D. L. (2018). 2018 guidelines for the early

management of patients with acute ischemic stroke: A guideline for healthcare professionals from the american heart association/american stroke association. *Stroke*.

Rajsic, S., Gothe, H., Borba, H. H., Sroczynski, G., Vujicic, J., Toell, T., & Siebert, U. (2019). Economic burden of stroke: A systematic review on post-stroke care. *The European Journal of Health Economics, 20*(1), 107-134. doi:10.1007/s10198-018-0984-0

Sahab, Z. J., Semaan, S. M., & Sang, Q.-X. A. (2007). Methodology and applications of disease biomarker identification in human serum. *Biomarker Insights, 2*, 21-43.

Saver, J. L., Smith, E. E., Fonarow, G. C., Reeves, M. J., Zhao, X., Olson, D. M., & Schwamm, L. H. (2010). The "golden hour" and acute brain ischemia. *Stroke, 41*(7), 1431-1439. doi:doi:10.1161/STROKEAHA.110.583815

Sharma, P., Cosme, J., & Gramolini, A. O. (2013). Recent proteomic advances in cardiac cells. *Journal of Proteomics, 81*, 3-14. doi:10.1016/j.jprot.2012.10.026

Sim, J., Teece, L., Dennis, M. S., Roffe, C., & Team, S. S. S. (2016). Validation and recalibration of two multivariable prognostic models for survival and independence in acute stroke. *PloS One, 11*(5), e0153527. doi:10.1371/journal.pone.0153527

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), 15545.

Vibha, D., & Misra, S. (2020). Blood biomarkers for stroke differentiation. In P. V. Peplow, B. Martinez, & S. A. Dambinova (Eds.), *Stroke biomarkers* (pp. 79-111). New York, NY: Springer New York.

Chapter 2

The Biological Timeline of Stroke and Stroke Recovery

Preface

Chapter 2 of this thesis aimed to narratively review the biological mechanisms that are involved in the stroke timeline, describing the periods of pre-stroke risk, acute stroke damage and subsequent downstream processes including inflammation and neurological repair. This information is required to give context to the later experimental chapters that explore the ability of blood based changes induced acutely in stroke to predict recovery over later periods of the post-stroke timeline than is commonly reported in relation to biological mechanisms, i.e. such as 3 and 12 months post-stroke.

Chapter Overview

This chapter is aimed at reviewing the currently known biological systems underlying stroke presentation; encompassing the periods of pre-stroke risk, acute stroke damage and stroke recovery. These periods represent distinct biological processes that are important to consider in context of this thesis. This review will also outline current clinical management timelines and further reference the current state of human stroke biomarkers research.

The Biological Basis of Stroke

Stroke is the clinical diagnosis used to describe the persistent symptoms that are associated with interruption of blood flow to the parenchyma of the brain (Andersen, Olsen, Dehlendorff, & Kammersgaard, 2009). Although the motor symptoms of stroke have been described by Hippocrates as ‘apoplexy’ since 460 BC, stroke has been acknowledged as a ‘cerebrovascular accident’ since 1927 (Finger, Boller, Tyler et al., 2010), essentially highlighting the vascular nature of stroke aetiology. There are two major types of presentations, ischemic stroke (IS) and intracerebral hemorrhage (ICH) (Massaro, Sacco, Scaff, & Mohr, 2002). IS occurs when a cerebral artery is occluded internally by a blood clot, whereas ICH occurs when a blood vessel is damaged and the contents are leaked into the surrounding brain tissue (Andersen et al., 2009). While ICH constitutes only 10%-20% of all stroke diagnoses, a fatal outcome is four times more prevalent on acute presentation when compared to IS (Andersen et al., 2009). Due to the relatively low incidence of ICH, there is a relative lack of research and no available primary treatments with the exception of acute medical management of blood and intracranial pressure (Rymer, 2011). Nevertheless, the mortality risk for patients who survive the first 4 weeks of ICH is no longer significantly different to IS at 90-days (Figure 1) (Andersen et al., 2009).

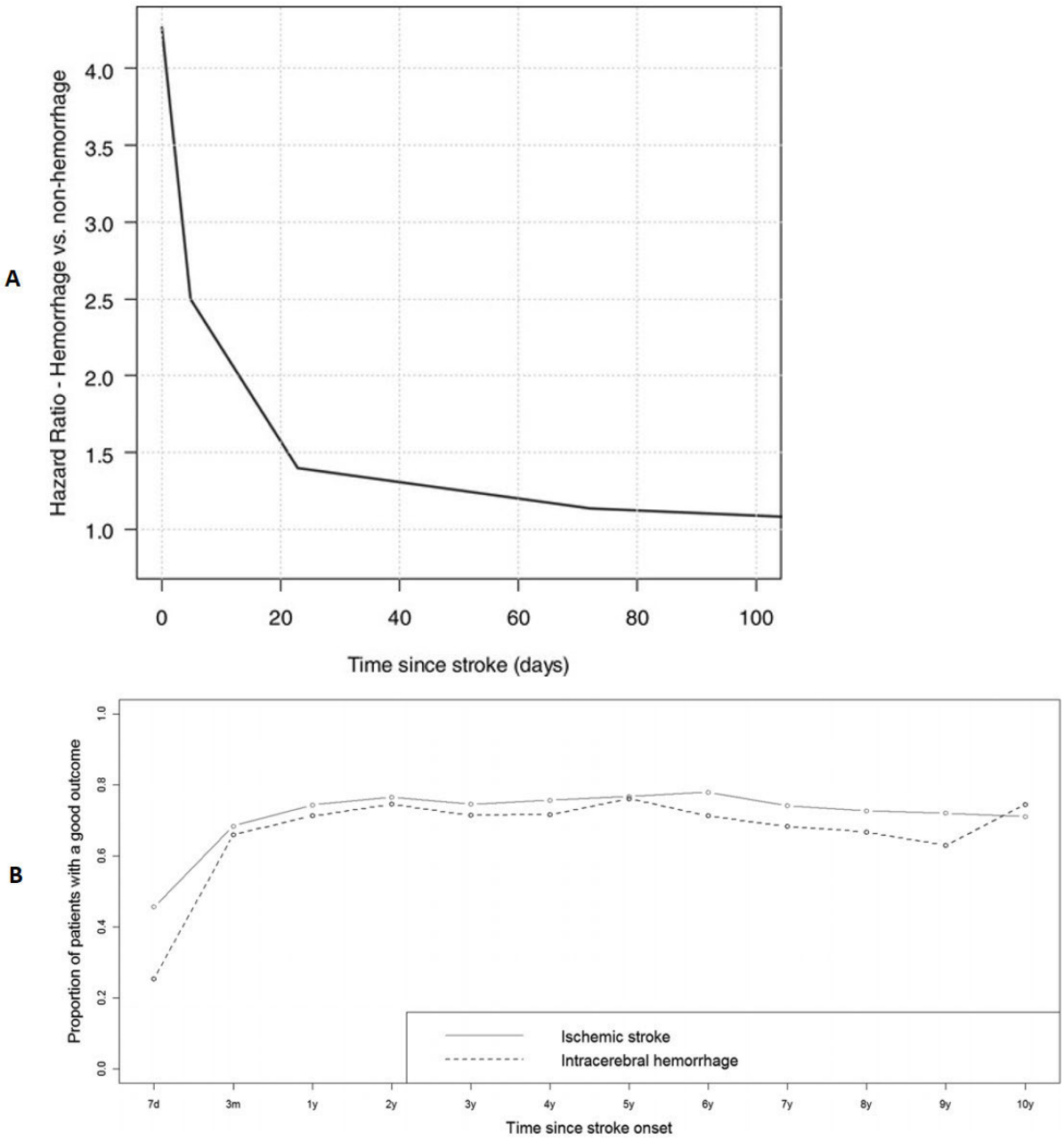

Figure 1.
1a. A comparison of survival risk for patients with ICH compared to IS. Although ICH presents with a much higher mortality risk than IS initially, there is an exponential decrease of mortality risk from the first month onwards. From Anderson, Olsen, Dehlendorff & Kammersgaard (2009). Hemorrhagic and Ischemic Strokes Compared. *Stroke*, 40(6), pp. 2070. Used with permission from Wolters Kluwer.
1b. Time course comparing IS and ICH stroke recovery using the Barthel Index. From Bhalla, Wang, Rudd & Wolfe. (2013). Differences in Outcome and Predictors Between Ischemic and Intracerebral Hemorrhage. *Stroke*, 44(8), pp. 2197. Used with permission from Wolters Kluwer.

Although the focus of this thesis is on the biological basis of IS, the knowledge gained from understanding the biological factors involved in post-IS recovery may also apply, at least in part, to ICH. In many cases both IS and ICH stroke patients present with a complex combination of cardiovascular factors including arterial vessel stiffness, atherosclerosis, hypertension and heart rhythm dysrhythmia, biological functions of an aging cardiovascular system (Paneni, Diaz Cañestro, Libby, Lüscher, & Camici, 2017). Thus, this review section will briefly explore the known temporal trajectories of the biological systems involved in the stroke process, from well-established research in acute stroke risk factors and damage towards the emerging understanding of the longitudinal recovery phase.

Clinical Management and Biomarkers in Ischemic Stroke

Acute clinical management for IS has been optimized around the ability to resolve embolic occlusion and restore cerebral blood flow in the acute phase post-stroke. In particular, frontline clinical management includes investigations for the safe administration of thrombolytic therapies and thrombectomy operations. Computerized tomography (CT) or magnetic resonance imaging (MRI) are used to establish neuroanatomical evidence of instances of intracranial haemorrhage, an immediate contraindication of thrombolytic therapies (Wintermark et al., 2013). Imaging modalities can produce clinically relevant parameters such as lesion location, core infarct size, penumbra size, vessel status, edema, cerebral blood flow and volume (Harston et al., 2015). Compared with non-insular lesions, damage in the insula region has been associated with poorer functional outcomes on a modified Rankin Scale (mRS) at 3 months (Laredo et al., 2018). Additionally, measures of corticospinal tract integrity on diffusion tensor imaging (DTI) MRI predicts motor outcomes at 12 months post-stroke (Bigourdan et al., 2016). However, these measures are not commonly used to predict functional recovery in clinical practice and remain as ad-hoc investigations accompanying

clinical trials (Boyd et al., 2017). While there have been studies showing the value of neuroimaging as a prognostic tool for stroke recovery, there is a lack of evidence and consensus on which blood based biomarkers are useful to adopt for the purpose of assessing stroke recovery within the acute period (Boyd et al., 2017).

Although clinical blood tests such as the complete blood count (CBC) (Lappé et al., 2011) and component leukocyte or white blood cell (WBC) count, platelet count and the international normalized ratio (INR) of thrombin time are prioritized following stroke admissions for the purpose of assessing coagulability and infection states (Fugate & Rabinstein, 2015; Kristensen et al., 2017), there appear to be few studies clinically utilizing blood tests to aid either diagnosis or to specify stroke subtypes upon presentation to emergency departments (Pandey et al., 2018). Recent consensus recommendations for biomarkers of stroke recovery and rehabilitation have highlighted that there is a lack of evidence for clinical stroke management strategies based on blood biomarkers (Boyd et al., 2017) where biomarkers are defined as a group of agents or molecules associated with measurable physiological signs that reflect ongoing biological processes. Stroke biomarker research has traditionally focused on the differentiation of major stroke types in the acute setting, intended as a diagnostic tool to facilitate thrombolysis in lieu of imaging (Misra et al., 2017). However to date, there has been no translatable evidence for a single stroke marker similar to the cardiac ‘troponin’ that has achieved diagnostic capability in cardiovascular diseases (Misra et al., 2017). Increasingly, biomarkers in stroke have been employed to predict motor, sensory, cognition and language outcomes with varying degrees of success (Boyd et al., 2017).

An older systematic review of blood biomarkers and biomarker panels in IS diagnosis concluded that while promising, there is a lack of subsequent validation and consensus for any of the identified molecules in the context of stroke practice (Whiteley, Tseng, & Sandercock, 2008). Rather, the only recommendation of either the Australian

National Stroke Foundation (NSF) guidelines (Boddice, Brauer, Gustafsson, Kenardy, & Hoffmann, 2010) or the recently published American Heart Association guidelines (Powers et al., 2018) is a blood glucose reading before alteplase administration to investigate for possible hyperglycemia (> 11.1mmol/l) or hypoglycemia, i.e homeostatic dysfunction that can present with symptoms similar to stroke. Other evidence also highlights the need to exclude various types of symptomatic acute infections that may have some overlap with stroke presentations (Hatzitolios et al., 2008). Indeed, current clinical guidelines only recommend a very limited range of blood tests to accommodate clinical decision making, prior to the time critical acute delivery of thrombolysis and thrombectomy (Wu et al., 2018).

As briefly discussed, previous blood-based biomarker studies in stroke patients typically examined gross clinical measures including mortality and functional recovery in relation to single biomarkers. Increases in copeptin, a peptide hormone precursor to stress molecule arginine vasopressin, has been shown to increase risk of 90-day mortality by 37% (de Marchis et al., 2013). Similarly, a meta-analysis of 16 studies and 2,258 stroke patients has found a 2.30 odds ratio increase in post-stroke mortality for patients with higher acute levels of brain natriuretic peptide (García-Berrocoso et al., 2013). Brain natriuretic peptide is a hormone that is secreted by the cardiomyocytes in the heart in response to increase of blood volume and acts to reduce vascular resistance and blood pressure while increasing natriuresis, the excretion of sodium through urine (Menon, Ramalingam, Conjeevaram, & Munisusmitha, 2016). Concerning functional recovery, a review of pro-inflammatory interleukin 6 (IL-6) has been shown to correlate negatively with post-stroke measures of functional recovery such as the modified Rankin Scale (mRS) and the Barthel Index (BI) (Marie-Héléne & Cramer, 2008). These studies suggest that although single molecule studies in stroke can provide some clinical utility, biomarker research has yet to be adopted in clinical stroke management.

In summary, stroke clinical management collects a large amount of patient centric data that includes blood measurements, patient physiology and demographics, although these data remain largely unused for the purpose of prognosticating recovery. Furthermore, biomarker studies in stroke are not able to elaborate on the biological functioning of molecules of interest, as this often requires extensive research that can translate from animal and cellular models to the bedside (Whiteley et al., 2009). Classical understanding of biomarkers as a single indicator of outcomes often disregards that biomarkers exist as units in the context of much larger and interconnected biological systems and networks. Therefore, there is a need to adopt a new analytical paradigm in longitudinal stroke studies that can simultaneously assess both the molecular components and constituent the biological systems of stroke recovery. Indeed, assessing biomarkers and biosystems should be complementary to clinical stroke management and provide both patients and healthcare professionals with clear biological information to inform practices to support positive stroke recovery.

The Stroke Timeline

As a preface to understanding the biological factors affecting stroke recovery, this section presents the major molecules and biological systems extending from stroke risk, to acute damage, and towards longitudinal recovery. Briefly, this section considers that biological systems are activated or deactivated because of preceding factors whereby stroke risk informs degree of ischemic damage and subsequently, both stroke risk and degree of ischemic damage informs the physiological environment that determines recovery potential.

Ischemic Stroke Risk

Ischemic stroke has a well-established group of risk factors that precede IS presentation. The risk factors of stroke encompass a complicated list beginning with

modifiable risk factors, modifiable (e.g. genetics, age and gender), and potentially modifiable risk factors (e.g. lifestyle & comorbidities). An extensive case-controlled study with 2337 IS cases and 3000 controls identified that some of the effective modifiable factors regarding incidence of IS are hypertension (>160/90 mm Hg), atrial fibrillation, previous myocardial infarction, smoking status, psychosocial factors such as depression and stress and hyperlipidaemia by measurement of the ratio of apolipoprotein B (ApoB) to apoliproteoin A1 (ApoA1) (O'Donnell et al., 2010). Of these, hypertension is the largest single contributor to incidence of stroke. However, pharmacological treatment of hypertension to normotensive status (systolic blood pressure <140mm Hg) does not completely eliminate risk of stroke (Howard et al., 2015). Atrial fibrillation (AF) is also strongly associated with incidence of stroke and further cardiovascular disorders (Boehme, Esenwa, & Elkind, 2017). AF refers to the irregular rhythms in heart beat that greatly predispose individuals to venous thrombosis, with more recent stroke models describing AF as a result of irregular cardiac substrates that arise from irregular parasympathetic-sympathetic control (Kamel, Okin, Elkind, & Iadecola, 2016). Although many of these factors can be assessed in general practice clinics and are often treated with frontline pharmacology, their presentation is also symptomatic of chronically impaired vascular health and usually aging.

Aging is the foremost non-modifiable risk factor, with a sharp increase in incidence of IS doubling each decade after the age of 55 (Boehme et al., 2017), with over 75% of strokes occurring in patients over the age of 65 (Benjamin et al., 2017). Cardiovascular aging is accompanied by maladaptive changes in structure and function of arteries that can be generalized to endothelial dysfunction and central arterial stiffness (Paneni et al., 2017). Heritable risk contributes differently for IS subtypes such as cardioembolic and large vessel stroke but overall family history of stroke increases risk by 30% (Boehme et al., 2017). Males have an overall 33% higher risk of stroke when

compared to women and this trend is especially prevalent in younger age groups (ages 35-44 years) (Appelros, Stegmayr, & Terént, 2009). However, women present with a much high mortality rates and worsened outcomes that may be linked to experiencing first-ever strokes at a later age than men and the lessening neuroprotective effects of estradiol (Girijala, Sohrabji, & Bush, 2017), as evidenced by studies examining stroke following ovariectomy (Liu, Yuan, Benashski, & McCullough, 2009) and menopause (Lisabeth & Bushnell, 2012). There are several genome wide association studies (GWAS) also examining the contribution of single nucleotide polymorphisms (SNPs), with a large-scale meta-analysis completed recently (Malik et al., 2018). Not surprisingly, these studies conclude that IS cannot be considered as a single-gene driven disorder, with 32 loci associated with stroke risk, albeit with extremely modest odds ratios (Malik et al., 2018). The genetic basis for stroke biology will be explored in-depth later in Chapter 3.

The modifiable risk factors for IS largely overlap with those found in myocardial infarction and suggest that cardiovascular factors are the cause of strokes (Yusuf et al., 2004). Indeed, primary prevention of stroke is managed in the same way as cardiology and endovascular considerations such as statins, diet modifications and exercise, with clinical neurology only being concerned with stroke after the first event (Heuser, 2017). Therefore, the mechanisms of ventricular thrombus formation must be considered for the main IS stroke subtypes that include cardioembolic and atherothrombotic. Studies examining the histological composition of stroke emboli suggest that cardioembolic strokes have a higher amount of fibrin and lower red blood cell (RBC) counts than non-cardioembolic strokes (Sporns et al., 2017). The formation and composition of cardioembolisms may be partially explained by a model incorporating atrial fibrillation which suggests that dysrhythmia induces a thrombogenic environment indicative of a worsened atrial tissue substrate (Kamel et al., 2016). Hyperlipidemia is a common cardiovascular condition and is highly correlated with the presence of atherosclerotic

plaque accumulation (Stoll & Bendszus, 2006). Atherothrombus formation often occurs in peripheral blood vessels, especially when the extracranial artery is stenosed due to atherosclerosis (Stoll, Kleinschnitz, & Nieswandt, 2008). The understanding of atherosclerosis has evolved from a simplistic view of lipid deposition on arterial walls to a more complex model of endothelial dysfunction involving shear stress from hypertension, oxidative damage, inflammatory and immune processes (Stoll & Bendszus, 2006) and more recently understanding of autonomic imbalance that occurs with aging and obesity and diabetes (Anderson, DeCicco, Schwaber, & Vadigepalli, 2017).

While cardiovascular integrity may be explained largely by lifestyle risk, there also exists a small heritable genetic component that can be linked to blood vessel integrity, blood pressure and lipid levels (Malik et al., 2018). Genetic dispositions are compounded by the biological senescence of ageing whereby the cardiovascular system becomes increasingly rigid due to decreased sympathetic efficiency associated with age induced autonomic imbalance. Hence the cardiovascular system is less able to recover from rapid changes in environmental pressure to normal cellular function and address accumulating stressors and damage over the lifetime. To summarize, the combination of genetic, aging and environmental stroke risk factors suggests that a poor vascular environment (also shared by ischemic heart disease) is inducive to the formation of cardioembolic clots or atherothrombotic plaques that ultimately present as stroke with neurological damage.

Stroke damage

An ischemic event occurs when an embolus significantly reduces or blocks the vasculature to an area of the cerebral parenchyma. Owing to the flow pressure and size of emboli, occlusions most often occur in the M1 branches of the main cerebral artery, followed by the downstream branches in M2 (Rai et al., 2018). This leads to a large variety of different neuroanatomical presentations under the umbrella term of IS. Women

are more likely to experience a total anterior circulation infarct (TACI) (most severe) but less likely to experience a posterior circulation infarct (POCI) when compared to men, with evidence showing little difference between genders for partial anterior circulation infarct (PACI) or lacunar infarct incidence rates (Giralt et al., 2012). However, the phenotypes for acute neurological cellular damage are similar for all IS subtypes and extends both spatially and temporally beyond the initial stroke damage (Stokum, Gerzanich, & Simard, 2016). Hence, an understanding of the temporal trajectory of the pathophysiological damage that begins within the first minutes of occlusion will provide a context for the biological processes involved in the recovery timeline (Figure 2) (Brouns & De Deyn, 2009).

Cerebrovascular circulation is central to brain function. Although the brain accounts for only 2% of body mass, it utilizes 15-20% of cardiac output (Xing et al., 2017), oxygen and glucose (Mergenthaler, Lindauer, Dienel, & Meisel, 2013), highlighting the highly disproportionate metabolic demand of brain tissue. Continual artery to venous return is needed for metabolic supply but also to deplete waste products such as carbon dioxide (CO_2) and heat from ongoing metabolic reactions (Kisler, Nelson, Montagne, & Zlokovic, 2017). Cessation or reduction of blood flow immediately following cerebrovascular occlusion sets the pathological context for the entire stroke timeline (Markus, 2004).

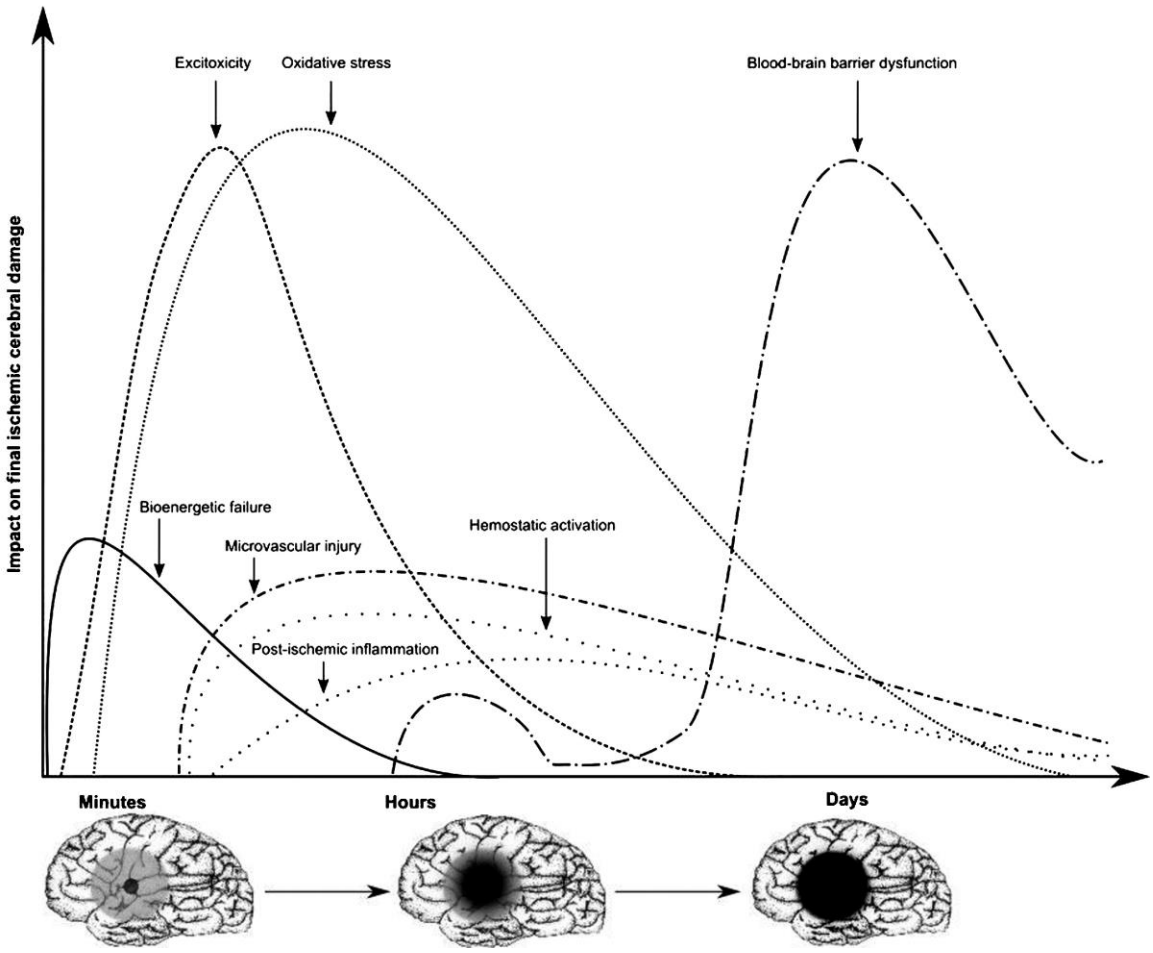

Figure 2. Stroke damage and the biological processes involved in the minutes to days post-stroke. Reprinted from *Clinical Neurology and Neurosurgery*, 111(6), Brouns, R & De Deyn, P. P. The complexity of neurobiological processes in acute ischemic stroke, pp. 485. Copyright (2009), with permission from Elsevier.

Depletion of resources vital for maintaining cell life such as oxygen and glucose quickly follow in the minutes post vascular occlusion (Robbins & Swanson, 2014). Chief amongst these are oxygen and glucose, both needed to maintain cellular function by mitochondrial adenosine triphosphate (ATP) reactions. Ischemic cell death or 'oncosis' is a form of rapid and violent cell death that is characterized by bioenergetic failure and hypoxia (Loh, Wang, & Liao, 2018). Indeed, one of the most fundamental processes in cell survival is the maintenance of the electrical potential across the lipid bilayers of the basement membrane (Engl & Attwell, 2015). This is estimated to require up to 43% of ATP based oxidative phosphorylation (OXPOHS) in resting state neurons in grey matter (Harris & Attwell, 2012). When energy supply is interrupted by vascular occlusion and

osmotic concentration is perturbed, ionic pumps on the membrane surface (especially the Na^-/K^+-ATPase enzyme) fail in a cascade-like fashion, causing a large degree of extracellular fluid influx into the cellular space (Figure 3) (Loh et al., 2018). Indeed, the histological hallmark of oncotic cell death is cellular edema, exemplified by irreversible swelling (Majno & Joris, 1995). The area of irrecoverable tissue damage known as the infarct core is exemplified by rampant acute necrotic cell death, representing the biological endpoint of cells most severely affected by oncosis.

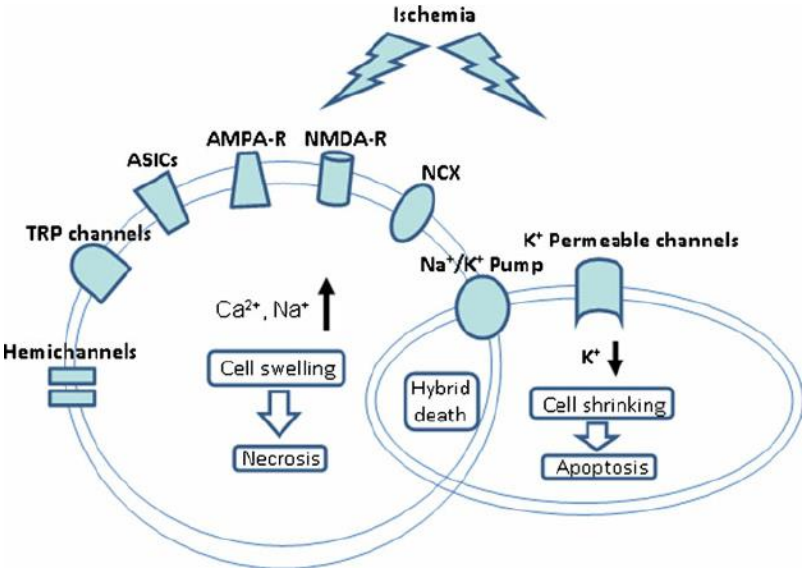

Figure 3. Different types of neuronal cell death are involved in stroke, such as necrosis, apoptosis and oncosis. A variety of ion channels are involved in the rapid movement of water, resulting in the loss or retention of fluid ultimately leading to the presentation of morphological hallmarks of cellular damage, cell swelling or cell shrinkage. Reprinted with permission from the Licensor: Springer Nature, *Translational Stroke Research,* Ionic Regulation of Cell Volume Changes and Cell Death after Ischemic Stroke, Song, M., & Yu, S. P, Copyright, 2013.

Not all cells in the focal area of the occlusion die immediately following stoke. Some cells in the penumbra region survive, substantiated by minimal flow from other arterial anatomising branches or cerebral collaterals (Jung et al., 2017). The infarct core is surrounded by a layer of neuronally depressed tissue that is profoundly impaired by the initial damage but nevertheless retains some metabolic potential by switching from aerobic to anaerobic forms of metabolism, including lactate, glycogen and fatty acids (Rama & García, 2016). This represents the functionality of dormant cells that are

focused on survival by maintaining an impaired membrane potential that is nevertheless sufficient to prevent a complete lysis of the cellular barrier (Dirnagl, Iadecola, & Moskowitz, 1999). Impairment of ionic pumps leads to depolarization of neurons, astrocytes and mitochondria (Ham & Raju, 2017), causing failure of voltage gated Ca^{2+} channels to remove Ca^{2+} from the cytosol, and thus causing an excessive efflux of excitatory neurotransmitters such as glutamate into the synaptic space (Rama & García, 2016). Excitotoxicity occurs when glutamate stimulates the membrane bound receptors α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and Ca^{2+} permeable *N*-methyl-D-aspartic acid (NMDA) to further allow influx of Ca^{2+} and NMDA, further damaging cellular organelles in an auto-stimulation paradigm (Rama & García, 2016). There are also a number of cell death processes that are activated by physiologically stressed cells in the stroke penumbra including: necrosis, programmed cell death by apoptosis or rescue of cellular function by reperfusion (Manning, Campbell, Oxley, & Chapot, 2014).

Although both bioenergetic failure and excitotoxicity present with clear cascades for single-cell death, these processes often release molecules that further propagate damage to neighbouring cells. The act of necrotic cell death releases the contents of the cytosol into the extracellular space, whereby molecules such as extracellular ATP (Cauwels, Rogge, Vandendriessche, Shiva, & Brouckaert, 2014) and heat shock proteins (Sachet, Liang, & Oehler, 2017) are extremely damaging to cells in the immediate area and often can further initiate apoptotic and/or necrotic processes. In surviving cells, Ca^{2+} ion influx leads to intracellular homeostatic failure and mitochondrial dysfunction, leading to increased overturn of reactive oxygen species (ROS) and reactive nitrogen species (RNS) (Prentice, Modi, & Wu, 2015). Furthermore, the post-ischemic extracellular environment promotes an upregulation of ROS superoxides by glutamate mediated excitotoxic activation of neuronal NMDA receptors (Brennan et al., 2009). The

combination of NMDA receptor activation and mitochondrial dysfunction is responsible for initiating neighbouring cell death and the development of cytotoxic edema in the ischemic penumbra (Heo, Han, & Lee, 2005; Villanueva, Kross, & Pérez-Astudillo, 2016).

Cellular and systemic immune activation and inflammation are common biological processes associated with tissue damage and representing the first line of biological defence to tissue injury. Danger-associated molecular patterns (DAMPs) or alarmins are molecular signals that are released endogenously from physiologically stressed or injured cellular bodies that activate mechanisms to address the outcomes of cellular damage (Figure 4). These include signals the intracellular contents of disrupted cells that have spilled into the extracellular fluid such as the presence of extracellular ATP (Ceruti et al., 2009), mitochondrial deoxyribonucleic acid (DNA) (Walko et al., 2014), and high mobility group box 1 (HMGB1) protein (Shichita et al., 2012). Immune cells and neurons possess the corresponding receptors to DAMPs and once activated, execute both a cellular function and induce an inflammatory response. As an example, extracellular ATP is released primarily from leaky necrotic cells and interacts with purinergic receptors P2X and P2Y on neurons and immune cells (Shichita et al., 2012). While ATP can directly activate T-cells and macrophages via P2X (Sachet et al., 2017), it also stimulates caspase-1 mediated inflammasome activity on the cellular surface via potassium (K^+) efflux to release large concentrations of proinflammatory cytokines (Gülke, Gelderblom, & Magnus, 2018). Although evidence in human stroke models is lacking, Caspase-1 can further induce the activation of local pyroptosis cascades (Barrington, Lemarchand, & Allan, 2017), a form of inflammation driven programmed cell death that is possibly also linked to caspase-8 induced apoptosis (Denes, Lopez-Castejon, & Brough, 2012).

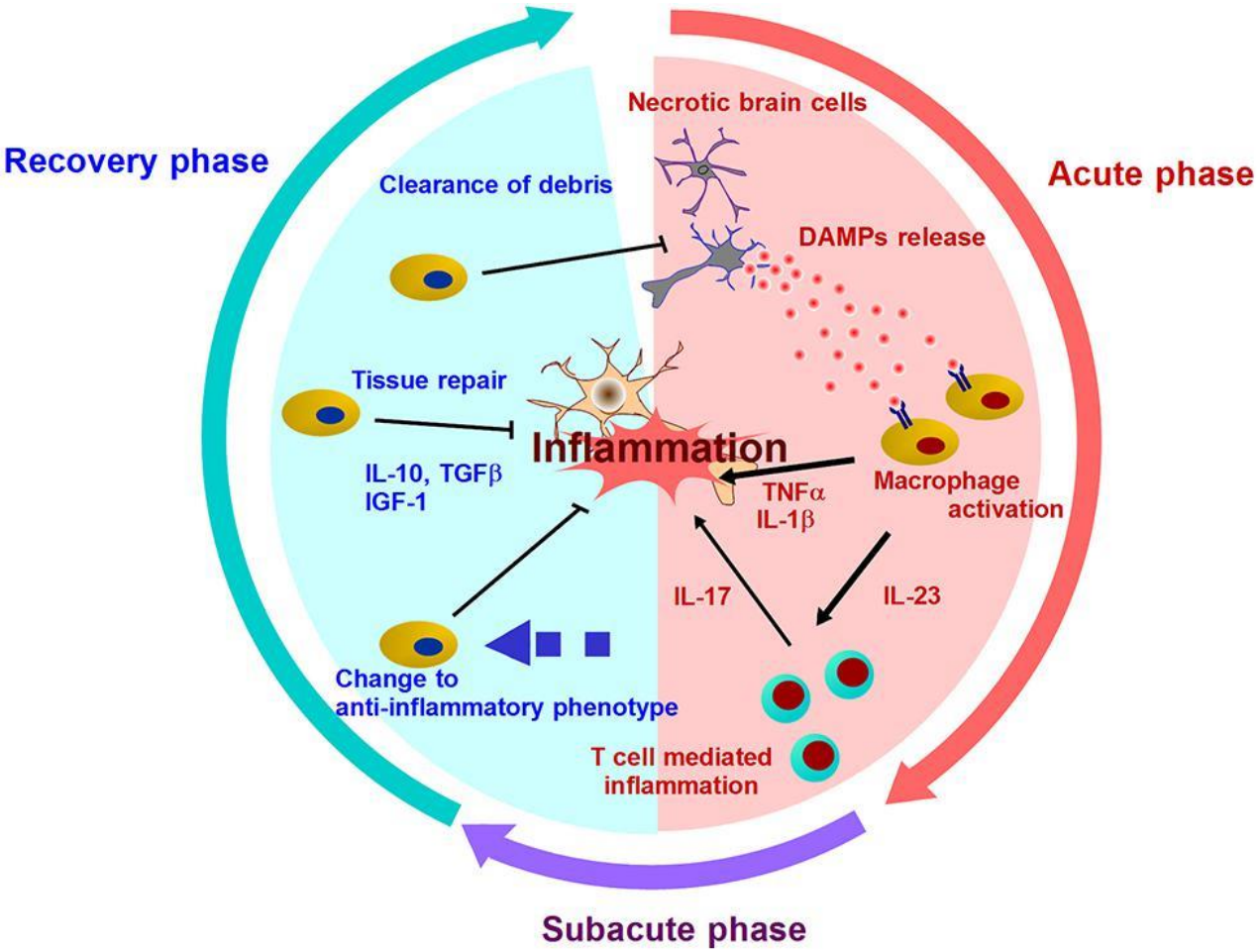

Figure 4. The activation of damage associated molecular patterns drives the immune and subsequent inflammatory pathways. In this model, it is assumed that the brain adopts an anti-inflammatory phenotype in the recovery phase that promotes clearance of debris and tissue repair in the days and weeks post-stroke. Reused from Shichita, T., Ito, M., & Yoshimura, A. (2014). Post-ischemic inflammation regulates neural damage and protection. *Frontiers in Cellular Neuroscience*, 8 (Article 319), used under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Inflammation is an immunovascular response, whereby vasoactive factors encourage vessel dilation and increased fluid flow and therefore increased attraction of immune cells into the damaged area (Gonzalez-Rey & Delgado, 2005). The endothelial tight junctions of cerebral microvessels and microglial endfeet on the blood brain barrier (BBB) are responsible for maintaining the fluid and chemical homeostasis in the brain. In the case of stroke, DAMPs, oxidative stress and inflammatory signalling, the tight junctions undergo many temporal phasic periods of assembly-disassembly and thereby contribute to the leaky BBB permeability phenotype (Sandoval & Witt, 2008). Assuming that reperfusion is achieved by medical treatment, the reflow effect is accompanied by a

short period of increased permeability (Figure 5). Given that most stroke patients exhibit already compromised vasculature, the sudden shear forces of reintroduced blood flow may initiate haemorrhagic transformations by rupturing already weakened blood vessels (Merali, Huang, Mikulis, Silver, & Kassner, 2017). Finally, the final phase of BBB permeability (18-96 hours after reperfusion) leads to an increase in vasogenic edema, the accumulation of vascular fluid in the interstitial space of cells in the ischemic penumbra (Sandoval & Witt, 2008). It is during this period that the greatest amount of peripheral blood-borne immune cells such as neutrophils, lymphocytes and monocytes are able to cross the BBB, a phenomenon known as 'leukocyte infiltration' (Kim, Park, Chang, Kim, & Lee, 2016). Leukocyte adhesion, binding and release of pro-inflammatory factors greatly increases BBB permeability during this phase and decreases further after 2-4 days (Gelderblom et al., 2009). Gender differences in BBB permeability and therefore cerebral blood flow can be attributed to the vasodilative effects of estrogen and progesterone on nitric oxide production (Haast, Gustafson, & Kiliaan, 2012).

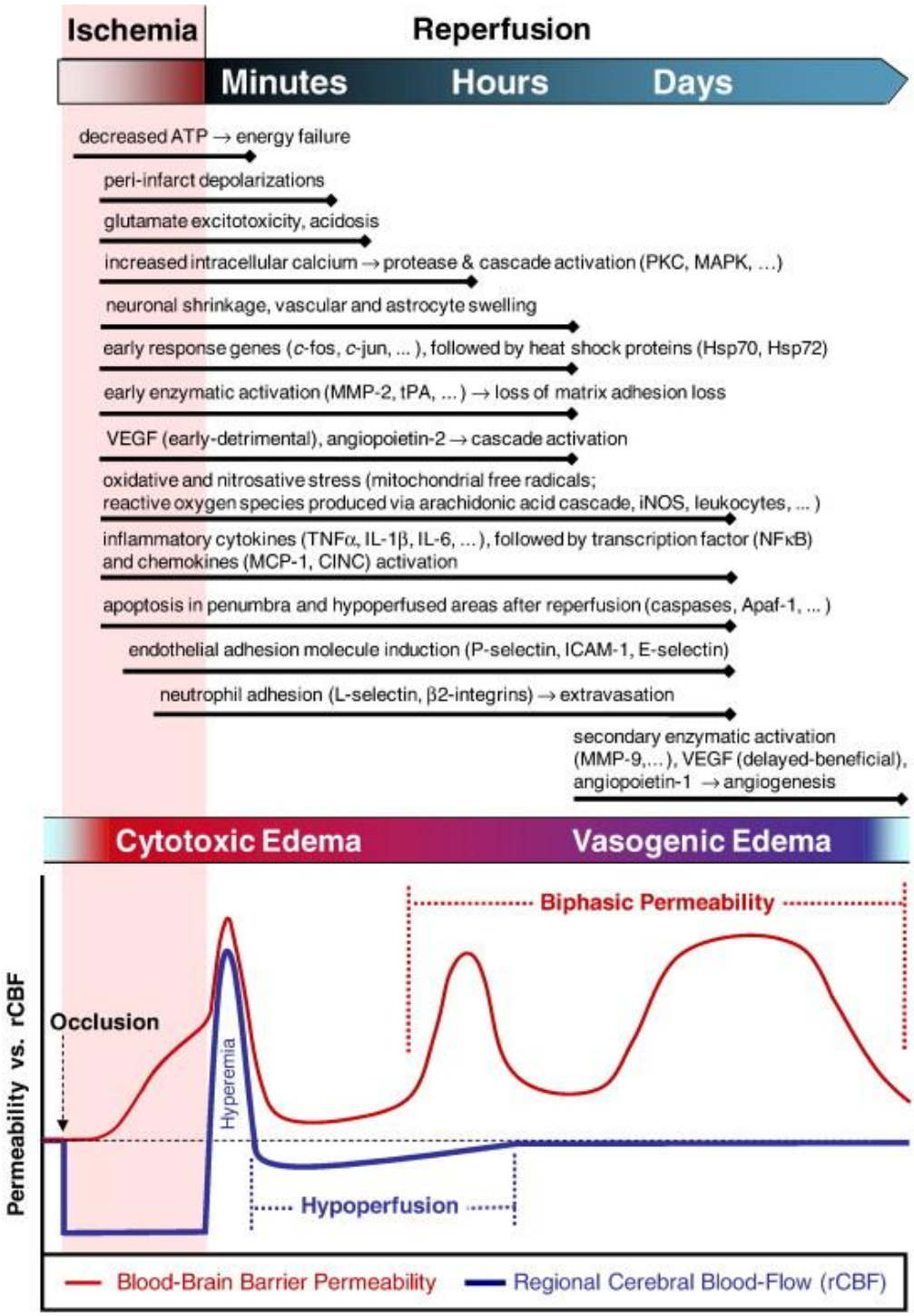

Figure 5. The relationship between blood brain barrier (BBB) permeability in the post-stroke period. Ischemic insult initiates a series of biochemical pathways that have a profound effect on the BBB within the first minutes to days post-stroke. Reprinted from *Neurobiology of Disease*, 32(2), Sandoval, K. E., Witt, K. A. Blood-brain barrier tight junction permeability and ischemic stroke, pp. 206, Copyright (2008), with permission from Elsevier.

Stroke recovery.

While most of the biochemical cascades in IS damage occur in the first week, the period of recovery extends well beyond the initial stroke event until most of the extracellular edema is dissipated (Gülke et al., 2018). As mentioned above, immune and inflammatory pathways comprise the central signalling mechanisms after a stroke and have long-lasting downstream effects in both the cerebral parenchyma and peripheral systems. In the recovery phase, these systems play a direct role in both further exacerbating damage or initiating repair mechanisms such as angiogenesis and neurogenesis. Heterogeneity of damage and function is the dominant theme in the clinical landscape of post-stroke recovery and rehabilitation (Knecht, Hesse, & Oster, 2011). Recovery in stroke is indicated by overlapping and interrelated trajectories of the biological cascades initiated by stroke damage and compounded by a complicated set of patient factors including pharmacological intervention, psychosocial factors, state of cardio-vasculature and rehabilitation programs. Although much is still currently unknown about this phase, this next section will briefly explore the biological processes likely to be involved in stroke recovery extending from the first days of damage.

The overactivation and recruitment of immune cells in the initial stroke damage is a natural response to cell damage and cell death (Gülke et al., 2018). Astrocytes undergo reactive gliosis and form glial scars to effectively seal off the lesion site, isolating damaged tissue from viable tissue (Huang et al., 2014). Thereafter, the waste products and cellular debris left following necrotic and apoptotic cell death processes need to be removed from the infarcted areas. Similar to lymphatic clearance in peripheral immune systems, resident glial cells provide the functional infrastructure for interaction with the glymphatic system, the main drainage pathway for interstitial fluid and immune waste in the mammalian brain (Benveniste et al., 2019). Microglia receive chemokine and cytokine spatial signalling from damaged cells and proceed to engage in clearance of

cellular debris by phagocytosis (Kim et al., 2016). Incidentally, anti-inflammatory M2-like microglia phenotypes dominate the infarcted area in the early hours of stroke and promote immune clearance, thereby improving neuronal survival (Hu et al., 2012). However, if the M1-like pro-inflammatory microglia response continues to be upregulated after 7 days post-stroke they become associated with deleterious levels of reactive oxygen species (ROS) and tumor necrosis factor alpha (TNF-a) accumulation. Furthermore, microglia are also capable of phagocytosing peripheral leukocytes after infiltration and prevent neutrophil aggregation in infarcted brain tissue (Otxoa-de-Amezaga et al., 2019).

The removal of cellular debris is crucial for the initiation of repair procedures, including angiogenesis and neurogenesis (Zhao & Willing, 2018). Indeed, these processes are responsible for the remodelling of the neurovascular unit and appear to continue weeks after IS (Strbian et al., 2008). Metabolic and hypoxic stress in cells induces the secretion of hypoxia inducible factor-1 (HIF-1) and following erythropoietin expression (Talwar & Srivastava, 2014). Aggregation and activation of vascular endothelial growth factors (VEGF) induces breakdown and remodelling of the damaged vasculature and stimulates revascularization (Talwar & Srivastava, 2014). However, the therapeutic administration of VEGF should consider the timing and status of vascular permeability and edema recovery to avoid further exaggerating stroke damage (Adamczak & Hoehn, 2015) and increasing the risk of further haemorrhages (Kanazawa et al., 2011).

Ongoing inflammatory processes are largely mediated by immune cells, of which there are many functional subtypes. Neutrophils are quickly implicated in initial cerebral ischemia and contribute to BBB permeability by releasing factors such as cytokines, chemokines, matrix metalloproteinase 9 and reactive oxygen species (Jickling et al., 2015). Lymphocytes such as T-cells and natural killer cells assist the movement and adhesion of neutrophils and platelets to cerebral damage, promoting thrombo-

inflammation and cytotoxicity (Planas, 2018). Towards the recovery phase, immune cells continue to play an active role as a mediator of inflammation and resolution of cellular damage. In animal models, macrophages have been shown to significantly contribute to long term recovery by resolving inflammation and vascular protection (Wattananit et al., 2016). Leukocyte counts of patients in recovery from stroke have been shown to rise rapidly after stroke but are not correlated with age, with evidence to suggest that aging contributes to maintenance of ongoing chronic pro-inflammatory responses (Roy-O'Reilly et al., 2020). Overall, the role of immune cell function towards long-term stroke recovery has not been well studied, despite evidence that suggests that recovery is largely mediated by immune processes.

Neurogenesis is spatially consistent with angiogenesis and occurs after the revascularization process provides scaffolding to deliver the nutrient supply required for neuronal repair in adult humans (Lindvall & Kokaia, 2015). Initially, elements such as chrondroitin sulfate proteoglycans in the glial scar, ongoing neuroinflammation and ROS production are inhibitory to axonal growth (Huang et al., 2014). In the days to weeks following stroke, growth-inhibitory proteoglycans are gradually reduced in an area surrounding the glial scar thus creating an environment inducive to axonal sprouting (Carmichael, 2006). Twenty-eight days post stroke, axonal sprouting forms complete anatomical connections in rodent models (Carmichael, Wei, Rovainen, & Woolsey, 2001). Post-stroke neurogenesis requires that neural stem cells (neuroblasts) proliferate into the areas of ischemic damage. Similar to the processes that promote angiogenesis following HIF-1 expression, the upregulation of cytokine erythropoietin enables the continual migration of neuroblasts into the area surrounding the glial scar (Sharp, Lu, Tang, & Millhorn, 2000).

Inflammation after ischemic damage has profound downstream effects on peripheral biological systems in animals and possibly exacerbates any already established

neuroinflammation in older humans (Gülke et al., 2018). As both glymphatic (glial) (Gaberel et al., 2014) and BBB perfusion (Sandoval & Witt, 2008) are perturbed immediately post-stroke, there are also measurable inflammatory and immune changes in other organs (Jin, Yang, & Li, 2010). Given that stroke is caused by clotting thromboembolisms, coagulation factors are also linked to immune activation by activated platelets that could be expected in any organ (Koupenova, Clancy, Corkrey, & Freedman, 2018; Zarbock, Polanowska-Grabowska, & Ley, 2007). The spleen is the largest peripheral reservoir for immune cells and participates in immune insults by rapidly releasing monocytes and macrophages factors into circulation (Bronte & Pittet, 2013). In experimental stroke models, splenocytes have been shown to upregulate pro-inflammatory cytokines and chemokines as early as 6 and 22 hours post-stroke (Offner, et al., 2005). By 96 hours, splenocyte counts and spleen mass were significantly reduced (Offner, Vandenbark & Hurn, 2009). In conjunction with sympathetic signalling and other circulating immune factors, host immunity is eventually depleted and is marked by a period extending from the first week of stroke where immune complications such as urinary tract infections or pneumonia are incredibly common and can lead to further mortality (Westendorp, Nederkoorn, Vermeij, Dijkgraaf, & van de Beek, 2011).

Table1
Brief Summary of Biosystems Affected by Ischemic Stroke

Biosystem	Post-stroke effects
Bioenergetics and Ionic Hemostasis	**Immediate – 48 hours** The most important biological function of cells and fails immediately with initial stroke hypoxia. Can switch to glycogen metabolism to maintain cellular functions. Cellular metabolism recovers somewhat with resolution of cerebral blood flow but has downstream effect on all following biological functions. Results in necrosis.
Oxidative Stress	**Immediate – Persistent** Fails with bioenergetics and inability of cells to resolve free radical accumulation. Has downstream and long-term effect on inflammation and immunity.
Blood Circulation	**Immediate – 48 hours (with treatment)**

	Loss of blood flow is the hallmark of stroke and can be resolved with removal of obstructive material reperfusion of blood vessels.
Blood Brain Barrier Permeability	**Immediate – 1 week** Permeability is vastly increased with stroke damage and shows oscillating patterns of opening and closing throughout the first week of stroke. Slightly raised even throughout long-term post-stroke.
Immunity & Inflammation	**2-12 hours – Persistent** Involved in responding to and initially propagating stroke damage. Astrocytes and microglia form a physical barrier (glial scar) around damaged area. Regulates blood brain barrier permeability with cytokines and chemokines. Peripheral leukocytes infiltrate CNS due to increased blood brain barrier permeability. Removes cellular debris and other products of cell death via apoptosis and autophagy.
Neurogenesis and Angiogenesis	**3-14 days – Persistent** Biologically occurs in most surviving patients but depends greatly on the extent of damage and possibly initial immune response. Requires the resolution of bioenergetics and oxidative stress to provide biochemical scaffolds for blood vessel growth and neuronal repair.

Although it has long been hoped that biological understanding (briefly summarized on Table 1) of the post-damage timeline leading to neural repair would provide the aetiological knowledge to be able to accurately predict rehabilitation outcomes, it is well known that post-stroke recovery is extremely heterogeneous, with each patient presenting with often unique combinations of complications including age, genetic makeup, life style risk factors, location and size of lesion (Kent, Soukup, & Fabian, 2001), cardiovascular integrity, immunoreactivity and psychosocial factors such as anxiety, depression and cognition. Therefore, it is difficult to predict which patients will assume positive post-stroke trajectories (Knecht et al., 2011; Sun, Tan, & Yu, 2014). As discussed, angiogenesis and neurogenesis are biological functions that primarily rely on age, previous vascular compromise and the ability to resolve initial stroke damage, thereby initiating the switching of a pro-inflammatory environment to an anti-inflammatory one (Kamel & Iadecola, 2012). These processes are reliant on the clearance of cellular debris, non-functioning cells and the restoration of cellular metabolism and ionic homeostasis, functions that remain difficult to assess in surviving stroke patients.

Furthermore in some cases the prevailing models for biological mechanisms in stroke recovery are derived from animal studies (Carmichael, 2008). Indeed, animal models of stroke typically ignore significant vascular risk factors and use young adult animals rather than older more senescent models and thus decrease their prognostic validity given the importance of age (Horvath & Raj, 2018) as a defining factor in predicting rehabilitation success of human stroke patients, (Bagg, Pombo, & Hopman, 2002). Therefore, there is a clinical need to address the biological processes involved in human stroke recovery, especially extending the timeline to months and years post-stroke.

References

Adamczak, J., & Hoehn, M. (2015). Poststroke angiogenesis (conference proceeding). *Stroke, 46*(5), e103-e104. doi:doi:10.1161/STROKEAHA.114.007642

Andersen, K. K., Olsen, T. S., Dehlendorff, C., & Kammersgaard, L. P. (2009). Hemorrhagic and ischemic strokes compared. *Stroke, 40*(6), 2068-2072. doi:doi:10.1161/STROKEAHA.108.540112

Anderson, W. D., DeCicco, D., Schwaber, J. S., & Vadigepalli, R. (2017). A data-driven modeling approach to identify disease-specific multi-organ networks driving physiological dysregulation. *PLoS Computational Biology, 13*(7), e1005627. doi:10.1371/journal.pcbi.1005627

Appelros, P., Stegmayr, B., & Terént, A. (2009). Sex differences in stroke epidemiology. *Stroke, 40*(4), 1082-1090. doi:doi:10.1161/STROKEAHA.108.540781

Bagg, S., Pombo, A. P., & Hopman, W. (2002). Effect of age on functional outcomes after stroke rehabilitation. *Stroke, 33*(1), 179-185. doi:doi:10.1161/hs0102.101224

Barrington, J., Lemarchand, E., & Allan, S. M. (2017). A brain in flame; do inflammasomes and pyroptosis influence stroke pathology? *Brain Pathology, 27*(2), 205-212. doi:10.1111/bpa.12476

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., . . . Muntner, P. (2017). Heart disease and stroke statistics—2017 update: A report from the american heart association. *Circulation, 135*(10), e146.

Benveniste, H., Liu, X., Koundal, S., Sanggaard, S., Lee, H., & Wardlaw, J. (2019). The glymphatic system and waste clearance with brain aging: A review. *Gerontology, 65*(2), 106-119. doi:10.1159/000490349

Bigourdan, A., Munsch, F., Coupé, P., Guttmann, C. R., Sagnier, S., Renou, P., . . . Sibon, I. (2016). Early fiber number ratio is a surrogate of corticospinal tract integrity and predicts motor recovery after stroke. *Stroke, 47*(4), 1053-1059.

Boehme, A. K., Esenwa, C., & Elkind, M. S. V. (2017). Stroke risk factors, genetics, and prevention. *Circulation Research, 120*(3), 472-495. doi:doi:10.1161/CIRCRESAHA.116.308398

Boyd, L. A., Hayward, K. S., Ward, N. S., Stinear, C. M., Rosso, C., Fisher, R. J., . . . Cramer, S. C. (2017). Biomarkers of stroke recovery: Consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *International Journal of Stroke, 12*(5), 480-493. doi:10.1177/1747493017714176

Brennan, A. M., Suh, S. W., Won, S. J., Narasimhan, P., Kauppinen, T. M., Lee, H., . . . Swanson, R. A. (2009). NAPDH oxidase is the primary source of superoxide induced by nmda receptor activation. *Nature Neuroscience, 12*(7), 857-863. doi:10.1038/nn.2334

Bronte, V., & Pittet, M. J. (2013). The spleen in local and systemic regulation of immunity. *Immunity, 39*(5), 806-818. doi:10.1016/j.immuni.2013.10.010

Brouns, R., & De Deyn, P. P. (2009). The complexity of neurobiological processes in acute ischemic stroke. *Clinical Neurology and Neurosurgery, 111*(6), 483-495. doi:https://doi.org/10.1016/j.clineuro.2009.04.001

Carmichael, S. T. (2006). Cellular and molecular mechanisms of neural repair after stroke: Making waves. *Annals of Neurology, 59*(5), 735-742. doi:10.1002/ana.20845

Carmichael, S. T. (2008). Themes and strategies for studying the biology of stroke recovery in the poststroke epoch. *Stroke, 39*(4), 1380-1388. doi:10.1161/STROKEAHA.107.499962

Carmichael, S. T., Wei, L., Rovainen, C. M., & Woolsey, T. A. (2001). New patterns of intracortical projections after focal cortical stroke. *Neurobiology of Disease, 8*(5), 910-922.

Cauwels, A., Rogge, E., Vandendriessche, B., Shiva, S., & Brouckaert, P. (2014). Extracellular ATP drives systemic inflammation, tissue damage and mortality. *Cell Death & Disease, 5*, e1102. doi:10.1038/cddis.2014.70

Ceruti, S., Villa, G., Genovese, T., Mazzon, E., Longhi, R., Rosa, P., . . . Abbracchio, M. P. (2009). The P2Y-like receptor PGR17 as a sensor of damage and a new potential target in spinal cord injury. *Brain, 132*(8), 2206-2218.

De Marchis, G. M., Katan, M., Weck, A., Fluri, F., Foerch, C., Findling, O., . . . Gensicke, H. (2013). Copeptin adds prognostic information after ischemic stroke: Results from the corisk study. *Neurology, 80*(14), 1278-1286.

Denes, A., Lopez-Castejon, G., & Brough, D. (2012). Caspase-1: Is IL-1 just the tip of the iceberg? *Cell Death & Disease, 3*(7), e338-e338. doi:10.1038/cddis.2012.86

Dirnagl, U., Iadecola, C., & Moskowitz, M. A. (1999). Pathobiology of ischaemic stroke: An integrated view. *Trends in Neurosciences, 22*(9), 391-397. doi:10.1016/S0166-2236(99)01401-0

Engl, E., & Attwell, D. (2015). Non-signalling energy use in the brain. *The Journal of Physiology, 593*(16), 3417-3429. doi:10.1113/jphysiol.2014.282517

Fugate, J. E., & Rabinstein, A. A. (2015). Absolute and relative contraindications to IV rt-PA for acute ischemic stroke. *The Neurohospitalist*, *5*(3), 110–121. https://doi.org/10.1177/1941874415578532

Gaberel, T., Gakuba, C., Goulay, R., De Lizarrondo, S. M., Hanouz, J.-L., Emery, E., . . . Gauberti, M. (2014). Impaired glymphatic perfusion after strokes revealed by contrast-enhanced mri. *Stroke, 45*(10), 3092.

García-Berrocoso, T., Giralt, D., Bustamante, A., Etgen, T., Jensen, J. K., Sharma, J. C., . . . Montaner, J. (2013). B-type natriuretic peptides and mortality after stroke: A systematic review and meta-analysis. *Neurology, 81*(23), 1976-1985. doi:10.1212/01.wnl.0000436937.32410.32

Gelderblom, M., Leypoldt, F., Steinbach, K., Behrens, D., Choe, C.-U., Siler, D. A., . . . Magnus, T. (2009). Temporal and spatial dynamics of cerebral immune cell accumulation in stroke. *Stroke, 40*(5), 1849-1857. doi:doi:10.1161/STROKEAHA.108.534503

Giralt, D., Domingues-Montanari, S., Mendioroz, M., Ortega, L., Maisterra, O., Perea-Gainza, M., . . . Montaner, J. (2012). The gender gap in stroke: A meta-analysis. *Acta Neurologica Scandinavica, 125*(2), 83-90. doi:https://doi.org/10.1111/j.1600-0404.2011.01514.x

Girijala, R. L., Sohrabji, F., & Bush, R. L. (2017). Sex differences in stroke: Review of current knowledge and evidence. *Vascular Medicine, 22*(2), 135-145. doi:10.1177/1358863x16668263

Gonzalez-Rey, E., & Delgado, M. (2005). Role of vasoactive intestinal peptide in inflammation and autoimmunity. *Current Opinion in Investigational Drugs, 6*(11), 1116-1123.

Gülke, E., Gelderblom, M., & Magnus, T. (2018). Danger signals in stroke and their role on microglia activation after ischemia. *Therapeutic Advances in Neurological Disorders, 11*, 1756286418774254-1756286418774254. doi:10.1177/1756286418774254

Haast, R. A. M., Gustafson, D. R., & Kiliaan, A. J. (2012). Sex differences in stroke. *Journal of Cerebral Blood Flow and Metabolism, 32*(12), 2100-2107. doi:10.1038/jcbfm.2012.141

Ham, P. B., & Raju, R. (2017). Mitochondrial function in hypoxic ischemic injury and influence of aging. *Progress in Neurobiology, 157*, 92-116. doi:https://doi.org/10.1016/j.pneurobio.2016.06.006

Harris, J. J., & Attwell, D. (2012). The energetics of CNS white matter. *The Journal of Neuroscience, 32*(1), 356-371. doi:10.1523/JNEUROSCI.3430-11.2012

Harston, G. W. J., Rane, N., Shaya, G., Thandeswaran, S., Cellerini, M., Sheerin, F., & Kennedy, J. (2015). Imaging biomarkers in acute ischemic stroke trials: A systematic review. *American Journal of Neuroradiology, 36*(5), 839-843. doi:10.3174/ajnr.A4208

Heo, J. H., Han, S. W., & Lee, S. K. (2005). Free radicals as triggers of brain edema formation after stroke. *Free Radical Biology and Medicine, 39*(1), 51-70. doi:https://doi.org/10.1016/j.freeradbiomed.2005.03.035

Heuser, R. R. (2017). The role for cardiologists in stroke intervention. *Progress in Cardiovascular Diseases, 59*(6), 549-554. doi:https://doi.org/10.1016/j.pcad.2017.05.002

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics, 19*(6), 371-384. doi:10.1038/s41576-018-0004-3

Howard, G., Banach, M., Cushman, M., Goff, D. C., Howard, V. J., Lackland, D. T., . . . Taylor, H. A. (2015). Is blood pressure control for stroke prevention the correct goal? *Stroke, 46*(6), 1595-1600. doi:doi:10.1161/STROKEAHA.115.009128

Hu, X., Li, P., Guo, Y., Wang, H., Leak, R. K., Chen, S., . . . Chen, J. (2012). Microglia/macrophage polarization dynamics reveal novel mechanism of injury expansion after focal cerebral ischemia. *Stroke, 43*(11), 3063-3070. doi:10.1161/STROKEAHA.112.659656

Huang, L., Wu, Z.-B., Zhuge, Q., Zheng, W., Shao, B., Wang, B., . . . Jin, K. (2014). Glial scar formation occurs in the human brain after ischemic stroke. *International Journal of Medical Sciences, 11*(4), 344-348. doi:10.7150/ijms.8140

Jickling, G. C., Liu, D., Ander, B. P., Stamova, B., Zhan, X., & Sharp, F. R. (2015). Targeting neutrophils in ischemic stroke: Translational insights from experimental

studies. Journal of Cerebral Blood Flow and Metabolism, 35(6), 888-901. doi:10.1038/jcbfm.2015.45

Jin, R., Yang, G., & Li, G. (2010). Inflammatory mechanisms in ischemic stroke: Role of inflammatory cells. *Journal of Leukocyte Biology, 87*(5), 779-789. doi:10.1189/jlb.1109766

Jung, S., Wiest, R., Gralla, J., McKinley, R., Mattle, H., & Liebeskind, D. (2017). Relevance of the cerebral collateral circulation in ischaemic stroke: Time is brain, but collaterals set the pace. *Swiss Medical Weekly, 147*(w14538), w14538.

Kamel, H., & Iadecola, C. (2012). Brain-immune interactions and ischemic stroke: Clinical implications. *Archives of Neurology, 69*(5), 576-581. doi:10.1001/archneurol.2011.3590

Kamel, H., Okin, P. M., Elkind, M. S. V., & Iadecola, C. (2016). Atrial fibrillation and mechanisms of stroke: Time for a new model. *Stroke, 47*(3), 895-900. doi:10.1161/STROKEAHA.115.012004

Kanazawa, M., Igarashi, H., Kawamura, K., Takahashi, T., Kakita, A., Takahashi, H., . . . Shimohata, T. (2011). Inhibition of VEGF signaling pathway attenuates hemorrhage after tpa treatment. *Journal of Cerebral Blood Flow and Metabolism, 31*(6), 1461-1474. doi:10.1038/jcbfm.2011.9

Kent, T. A., Soukup, V. M., & Fabian, R. H. (2001). Heterogeneity affecting outcome from acute stroke therapy. *Stroke, 32*(10), 2318-2327. doi:10.1161/hs1001.096588

Kim, J. Y., Park, J., Chang, J. Y., Kim, S.-H., & Lee, J. E. (2016). Inflammation after ischemic stroke: The role of leukocytes and glial cells. *Experimental Neurobiology, 25*(5), 241-251. doi:10.5607/en.2016.25.5.241

Kisler, K., Nelson, A. R., Montagne, A., & Zlokovic, B. V. (2017). Cerebral blood flow regulation and neurovascular dysfunction in alzheimer disease. *Nature Reviews. Neuroscience, 18*(7), 419-434. doi:10.1038/nrn.2017.48

Knecht, S., Hesse, S., & Oster, P. (2011). Rehabilitation after stroke. *Deutsches Arzteblatt International, 108*(36), 600-606. doi:10.3238/arztebl.2011.0600

Koupenova, M., Clancy, L., Corkrey, H. A., & Freedman, J. E. (2018). Circulating platelets as mediators of immunity, inflammation, and thrombosis. *122*(2), 337-351. doi:10.1161/CIRCRESAHA.117.310795

Laredo, C., Zhao, Y., Rudilosso, S., Renú, A., Pariente, J. C., Chamorro, Á., & Urra, X. (2018). Prognostic significance of infarct size and location: The case of insular stroke. *Scientific Reports, 8*(1), 9498-9498. doi:10.1038/s41598-018-27883-3

Lindvall, O., & Kokaia, Z. (2015). Neurogenesis following stroke affecting the adult brain. *Cold Spring Harbor Perspectives in Biology, 7*(11). doi:10.1101/cshperspect.a019034

Lisabeth, L., & Bushnell, C. (2012). Stroke risk in women: The role of menopause and hormone therapy. *The Lancet. Neurology, 11*(1), 82-91. doi:10.1016/S1474-4422(11)70269-1

Loh, K. Y., Wang, Z., & Liao, P. (2018). Oncotic cell death in stroke. In *Reviews of physiology, biochemistry and pharmacology*. Berlin, Heidelberg: Springer.

Majno, G., & Joris, I. (1995). Apoptosis, oncosis, and necrosis. An overview of cell death. *The American Journal of Pathology, 146*(1), 3-15.

Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., . . . Attia, J. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics, 50*(4), 524-537. doi:10.1038/s41588-018-0058-3

Manning, N. W., Campbell, B. C. V., Oxley, T. J., & Chapot, R. (2014). Acute ischemic stroke. *Stroke, 45*(2), 640-644. doi:10.1161/STROKEAHA.113.003798

Marie-Héléne, M., & Cramer, S. C. (2008). Biomarkers of recovery after stroke. *Current Opinion in Neurology, 21*(6), 654-659. doi:10.1097/WCO.0b013e3283186f96

Markus, H. S. (2004). Cerebral perfusion and stroke. *Journal of Neurology, Neurosurgery & Psychiatry, 75*(3), 353. doi:10.1136/jnnp.2003.025825

Massaro, A. R., Sacco, R. L., Scaff, M., & Mohr, J. P. (2002). Clinical discriminators between acute brain hemorrhage and infarction: A practical score for early patient identification. *Arquivos de Neuro-Psiquiatria, 60*, 185-191.

Menon, B., Ramalingam, K., Conjeevaram, J., & Munisusmitha, K. (2016). Role of brain natriuretic peptide as a novel prognostic biomarker in acute ischemic stroke. *Annals of Indian Academy of Neurology, 19*(4), 462-466. doi:10.4103/0972-2327.194422

Merali, Z., Huang, K., Mikulis, D., Silver, F., & Kassner, A. (2017). Evolution of blood-brain-barrier permeability after acute ischemic stroke. *PloS One, 12*(2), e0171558. doi:10.1371/journal.pone.0171558

Mergenthaler, P., Lindauer, U., Dienel, G. A., & Meisel, A. (2013). Sugar for the brain: The role of glucose in physiological and pathological brain function. *Trends in Neurosciences, 36*(10), 587-597. doi:10.1016/j.tins.2013.07.001

Misra, S., Kumar, A., Kumar, P., Yadav, A. K., Mohania, D., Pandit, A. K., . . . Vibha, D. (2017). Blood-based protein biomarkers for stroke differentiation: A systematic review. *Proteomics – Clinical Applications, 11*(9-10), 1700007. doi:10.1002/prca.201700007

O'Donnell, M. J., Xavier, D., Liu, L., Zhang, H., Chin, S. L., Rao-Melacini, P., . . . Yusuf, S. (2010). Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): A case-control study. *The Lancet, 376*(9735), 112-123. doi:10.1016/S0140-6736(10)60834-3

Offner, H., Subramanian, S., Parker, S. M., Afentoulis, M. E., Vandenbark, A. A., & Hurn, P. D. (2005). Experimental stroke induces massive, rapid activation of the

peripheral immune system. *Journal of Cerebral Blood Flow and Metabolism, 26*(5), 654-665. doi:10.1038/sj.jcbfm.9600217

Offner, H., Vandenbark, A. A., & Hurn, P. D. (2009). Effect of experimental stroke on peripheral immunity: CNS ischemia induces profound immunosuppression. *Neuroscience, 158*(3), 1098-1111. doi:10.1016/j.neuroscience.2008.05.033

Otxoa-de-Amezaga, A., Miró-Mur, F., Pedragosa, J., Gallizioli, M., Justicia, C., Gaja-Capdevila, N., . . . Planas, A. M. (2019). Microglial cell loss after ischemic stroke favors brain neutrophil accumulation. *Acta Neuropathologica, 137*(2), 321-341. doi:10.1007/s00401-018-1954-4

Pandey, S. P., Madhukar, P., Dev, P., Joshi, D., Mishra, V. N., Chaurasia, R. N., & Pathak, A. (2018). Blood biomarkers for ischemic stroke subtype differentiation: A systematic review. *Cardiovascular & Hematological Disorders Drug Targets*. doi:10.2174/1871529x18666180829142354

Paneni, F., Diaz Cañestro, C., Libby, P., Lüscher, T. F., & Camici, G. G. (2017). The aging cardiovascular system: Understanding it at the cellular and clinical levels. *Journal of the American College of Cardiology, 69*(15), 1952-1967. doi:https://doi.org/10.1016/j.jacc.2017.01.064

Planas, A. M. (2018). Role of immune cells migrating to the ischemic brain. *Stroke, 49*(9), 2261-2267. doi:10.1161/STROKEAHA.118.021474

Prentice, H., Modi, J. P., & Wu, J.-Y. (2015). Mechanisms of neuronal protection against excitotoxicity, endoplasmic reticulum stress, and mitochondrial dysfunction in stroke and neurodegenerative diseases. *Oxidative Medicine and Cellular Longevity, 2015*, article:964518-964518. doi:10.1155/2015/964518

Rai, A. T., Domico, J. R., Buseman, C., Tarabishy, A. R., Fulks, D., Lucke-Wold, N., . . . Carpenter, J. S. (2018). A population-based incidence of M2 strokes indicates potential expansion of large vessel occlusions amenable to endovascular therapy.

Journal of Neurointerventional Surgery, 10(6), 510. doi:10.1136/neurintsurg-2017-013371

Rama, R., & García, J. C. (2016). Excitotoxicity and oxidative stress in acute stroke. In *Ischemic Stroke*: IntechOpen.

Robbins, N. M., & Swanson, R. A. (2014). Opposing effects of glucose on stroke and reperfusion injury: Acidosis, oxidative stress, and energy metabolism. *Stroke, 45*(6), 1881-1886. doi:10.1161/STROKEAHA.114.004889

Roy-O'Reilly, M. A., Ahnstedt, H., Spychala, M. S., Munshi, Y., Aronowski, J., Sansing, L. H., & McCullough, L. D. (2020). Aging exacerbates neutrophil pathogenicity in ischemic stroke. Aging, 12(1), 436-461. doi:10.18632/aging.102632

Rymer, M. M. (2011). Hemorrhagic stroke: Intracerebral hemorrhage. *Missouri Medicine, 108*(1), 50-54.

Sachet, M., Liang, Y. Y., & Oehler, R. (2017). The immune response to secondary necrotic cells. *Apoptosis, 22*(10), 1189-1204. doi:10.1007/s10495-017-1413-z

Sandoval, K. E., & Witt, K. A. (2008). Blood-brain barrier tight junction permeability and ischemic stroke. *Neurobiology of Disease, 32*(2), 200-219. doi:https://doi.org/10.1016/j.nbd.2008.08.005

Sharp, F. R., Lu, A., Tang, Y., & Millhorn, D. E. (2000). Multiple molecular penumbras after focal cerebral ischemia. *Journal of Cerebral Blood Flow and Metabolism, 20*(7), 1011-1032.

Shichita, T., Hasegawa, E., Kimura, A., Morita, R., Sakaguchi, R., Takada, I., . . . Yanagawa, T. (2012). Peroxiredoxin family proteins are key initiators of post-ischemic inflammation in the brain. *Nature Medicine, 18*(6), 911.

Sporns, P. B., Hanning, U., Schwindt, W., Velasco, A., Minnerup, J., Zoubi, T., . . . Niederstadt, T. U. (2017). Ischemic stroke. *Stroke, 48*(8), 2206-2210. doi:10.1161/STROKEAHA.117.016590

Stokum, J. A., Gerzanich, V., & Simard, J. M. (2016). Molecular pathophysiology of cerebral edema. *Journal of Cerebral Blood Flow and Metabolism, 36*(3), 513-538. doi:10.1177/0271678X15617172

Stoll, G., & Bendszus, M. (2006). Inflammation and atherosclerosis. *Stroke, 37*(7), 1923-1932. doi:10.1161/01.STR.0000226901.34927.10

Stoll, G., Kleinschnitz, C., & Nieswandt, B. (2008). Molecular mechanisms of thrombus formation in ischemic stroke: Novel insights and targets for treatment. *Blood, 112*(9), 3555. doi:10.1182/blood-2008-04-144758

Strbian, D., Durukan, A., Pitkonen, M., Marinkovic, I., Tatlisumak, E., Pedrono, E., . . . Tatlisumak, T. (2008). The blood–brain barrier is continuously open for several weeks following transient focal cerebral ischemia. *Neuroscience, 153*(1), 175-181. doi:https://doi.org/10.1016/j.neuroscience.2008.02.012

Sun, J.-H., Tan, L., & Yu, J.-T. (2014). Post-stroke cognitive impairment: Epidemiology, mechanisms and management. *Annals of Translational Medicine, 2*(8), 80. doi:10.3978/j.issn.2305-5839.2014.08.05

Talwar, T., & Srivastava, M. V. P. (2014). Role of vascular endothelial growth factor and other growth factors in post-stroke recovery. *Annals of Indian Academy of Neurology, 17*(1), 1-6. doi:10.4103/0972-2327.128519

Villanueva, C., Kross, R. D., & Pérez-Astudillo, L. (2016). Free radicals and neuronal recovery from an ischaemic penumbra: A review. In *Free radicals and diseases*: IntechOpen.

Walko, T. D., Bola, R. A., Hong, J. D., Au, A. K., Bell, M. J., Kochanek, P. M., . . . Aneja, R. K. (2014). Cerebrospinal fluid mitochondrial DNA–a novel damp in pediatric traumatic brain injury. *Shock, 41*(6), 499.

Wattananit, S., Tornero, D., Graubardt, N., Memanishvili, T., Monni, E., Tatarishvili, J., . . . Kokaia, Z. (2016). Monocyte-derived macrophages contribute to

spontaneous long-term functional recovery after stroke in mice. *The Journal of Neuroscience*, 36(15), 4182. doi:10.1523/JNEUROSCI.4317-15.2016

Westendorp, W. F., Nederkoorn, P. J., Vermeij, J.-D., Dijkgraaf, M. G., & van de Beek, D. (2011). Post-stroke infection: A systematic review and meta-analysis. *BMC Neurology, 11*, 110-110. doi:10.1186/1471-2377-11-110

Whiteley, W., Jackson, C., Lewis, S., Lowe, G., Rumley, A., Sandercock, P., . . . Sudlow, C. (2009). Inflammatory markers and poor outcome after stroke: A prospective cohort study and systematic review of interleukin-6. *PLoS Medicine, 6*(9), e1000145. doi:10.1371/journal.pmed.1000145

Wintermark, M., Sanelli, P. C., Albers, G. W., Bello, J., Derdeyn, C., Hetts, S. W., . . . Meltzer, C. C. (2013). Imaging recommendations for acute stroke and transient ischemic attack patients: A joint statement by the American Society of Neuroradiology, the American College of Radiology, and the Society of Neurointerventional Surgery. *American Journal of Neuroradiology, 34*(11), E117-E127. doi:10.3174/ajnr.A3690

Xing, C.-Y., Tarumi, T., Liu, J., Zhang, Y., Turner, M., Riley, J., . . . Zhang, R. (2017). Distribution of cardiac output to the brain across the adult lifespan. *Journal of Cerebral Blood Flow and Metabolism, 37*(8), 2848-2856. doi:10.1177/0271678x16676826

Yusuf, S., Hawken, S., Ôunpuu, S., Dans, T., Avezum, A., Lanas, F., . . . Lisheng, L. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *The Lancet, 364*(9438), 937-952. doi:10.1016/S0140-6736(04)17018-9

Zarbock, A., Polanowska-Grabowska, R. K., & Ley, K. (2007). Platelet-neutrophil-interactions: Linking hemostasis and inflammation. *Blood Reviews, 21*(2), 99-111. doi:https://doi.org/10.1016/j.blre.2006.06.001

Zhao, L.-R., & Willing, A. (2018). Enhancing endogenous capacity to repair a stroke-damaged brain: An evolving field for stroke research. *Progress in Neurobiology, 163-164*, 5-26. doi:https://doi.org/10.1016/j.pneurobio.2018.01.004

Chapter 3

An Integrative Multi-Omics Systematic Review of Acute Ischemic and Hemorrhagic Stroke Systems Biology: Genetics, Transcriptomics and Proteomics

Vinh Nguyen

Nina Riddell

Leeanne Carey

Sheila Crewther

Preface

Perhaps the most powerful aspect of systems biology is the functionality to display biological systems and constituent molecules in a manner that is both graphically appealing and educational. Exploring resources such as Reactome and the Kyoto Encylopedia of Genes and Genomes makes it immediately apparent to the reader that biological molecules and systems are interconnected at the gene and protein expression levels. These tools can facilitate understanding of biology by presenting functional connections between molecules and systems that are well understood (e.g. Krebs or tricarboxylic acid (TCA) cycle) to biological systems that are less well known.

Chapter 3 involves a systematic review of untargeted systems biology of stroke recovery using novel bioinformatics approaches to present a comprehensive overview of the biological processes involved. The bioinformatics method utilized is a simplification of previously integrated multi-omics systematic review methods and does not use the raw quantitative data (i.e. fold change values) and only gathers the qualitative data (i.e. significant gene and protein names) from the reviewed studies. This approach was chosen to account for differences in methodologies and quantitative methods utilized by each study. Furthermore, genome wide association studies (GWAS) do not provide clear indication of whether or not genomic mutations are associated with downstream quantitative expression of messenger ribonucleic acid (mRNA) and proteins. Nevertheless, given that genetics, transcriptomics and proteomics studies all provide common data at the qualitative level, and can be expressed as significantly differentiated molecules between two conditions (stroke *vs* control), we can compare these studies across multiple omics levels, taking an integrative approach.

The current paper is being edited for journal submission in neurology and/or systems biology journals.

An Integrative Multi-Omics Systematic Review of Acute Ischemic and Hemorrhagic Stroke Systems Biology: Genetics, Transcriptomics and Proteomics

Stroke contributes to a large proportion of global mortality and disability (Powers et al., 2018). Lifestyle factors such as obesity and smoking are major risk factors for cardiovascular health while non-modifiable risk factors such as age and genetics (Boehme, Esenwa, & Elkind, 2017) also play a major role in the probability of stroke. Despite the vast literature associated with these well-established risk factors, the link between the predisposing stroke genome and the biological mechanisms that follow in stroke is more uncertain. To address these issues, this review aimed to characterize the relationships between human biological expression of acute ischemic stroke (IS) and intracerebral hemorrhage (ICH) by comparing the differentially expressed genome, transcriptome and proteome of individuals who had recently experienced first ever episode of stroke.

The biological determinants of first ever stroke can be inferred from established stroke risk factors. Non-modifiable risk factors for both IS and ICH share a large overlap, with atrial fibrillation, previous myocardial infarction, smoking status, hyperlipidemia, diabetes, hypertension and psychosocial factors such as depression and stress linked to stroke presentation (Boehme et al., 2017). Age, sex and genetics are often included as non-modifiable risk factors (Boehme et al., 2017). Although the causes of both IS and ICH are varied, these risk factors suggest that an abnormal cardiovascular substrate (Kamel, Okin, Elkind, & Iadecola, 2016) greatly increases the risk of stroke due to the acquired physiological stressors over the lifetime, compounded by the individual's biological makeup. Genome wide association studies (GWAS) of stroke largely support the overview that a combination of stroke-specific genetics in addition to cardiovascular dysfunction precedes stroke presentation. Meta analyses of GWAS studies have found that amongst 32 loci, genomic regions encoding for castor zinc finger 1 (CASZ1) and

paired like homeodomain 2 (PITX) overlap with previous associations in atrial fibrillation and blood pressure, while genes such as cyclin dependent kinase 6 (CDK6) are specific to stroke (Malik et al., 2018). As demonstrated by numerous genetics studies, stroke does not appear to be a monogenic disease that can be treated or diagnosed with a single molecular marker. Instead, these studies suggest that a wide range of biological processes that are linked with cardiovascular function influence stroke presentation.

One of the major causes of IS are cardioembolic clot formations that are closely linked to hemostasis and coagulation (Stoll, Kleinschnitz, & Nieswandt, 2008). The causes of ICH are more related to blood pressure and compounded by a compromised vascular system (K. K. Andersen, Olsen, Dehlendorff, & Kammersgaard, 2009). It is well accepted that inflammation and immunity are the most actively involved biological processes in mediating post-stroke damage and recovery (Kamel & Iadecola, 2012). In response to vascular-induced neurological damage, responses such as brain leukocyte infiltration, resolution of edema, clearance of cellular debris, scavenging of reactive oxygen species (ROS) are all mechanisms that are influenced by immune cells (Rayasam et al., 2018). Understanding the timing and degree of involvement of these biosystems in relation to stroke will allow for greater therapeutic influence over the trajectory of stroke recovery.

To achieve a greater understanding of stroke biology, current knowledge in stroke risk genetics should be further linked to studies of post-stroke biology. GWAS is an untargeted systems biology approach where analytical methods are used to quantify as many molecular elements in a particular sample as possible (Witte, 2010). In addition to GWAS, transcriptomic studies utilize microarray and next-gen sequencing approaches to examine messenger ribonucleic acid (mRNA) gene expression, often for the whole protein coding genome (Sharp et al., 2011). Untargeted proteomics studies differ in relation to GWAS and transcriptomics studies in the analytical sense as there are often

significant limitations to sample type and analytical methods that restrict overall protein coverage (Dayon & Kussmann, 2013). Furthermore, these 'hypothesis free' or 'agnostic' approaches provide a method to address disease pathophysiology without relying on past knowledge (Kitsios & Zintzaras, 2009). Indeed, these tools can be utilized to confirm and conceptualize current biological understanding into the context of biological processes with a systematic approach.

Recently, combining multiple untargeted omics approaches or 'multi-omics' has proven to be promising in providing a more comprehensive understanding in multiple diseases (Dimitrakopoulos et al., 2018). Such approaches involve the reduction of data to common names and values that can be standardized across genetics, transcriptomics and proteomics studies. As the genome is more stable over time, stroke GWAS have mostly been used to describe predispositions. In contrast, transcriptomic and proteomic studies examining messenger ribonucleic acid (mRNA) and proteins describe biological systems involved at that point in relation to the timing of stroke onset. Due to the organization of systems biology data, the name of identified genomic loci, transcript mRNA and proteins are conserved across genetics, transcriptomics and proteomics studies and thus can be compared. Thus, biosystems that are shown to overlap across multiple levels of systems biology would suggest a strong involvement of the gene-to-environment paradigm in response to stroke induced central nervous system damage.

In this review, we aim to employ a multi-omics systematically identify the overlapping relationship between the identified human stroke genome, transcriptome and proteome. Multi-omics approaches have been shown to be successful in providing a comprehensive overview of disease processes (Leon-Mimila et al., 2019) and have been used to describe shared disease processes between cardiovascular disease and type 2 diabetes (Shu et al., 2017). The significantly associated single nucleotide polymorphisms (SNPs) were obtained from a previously completed meta-analysis in stroke genetics, with

genes identified from the downstream versatile gene-based association study (VEGAS) in the published paper by Malik and colleagues (Malik et al., 2018). Building upon the traditional individual SNP to gene approach, VEGAS considers significant associations between trait expression (such as incidence of stroke) and all SNPs, correcting for linkage disequilibrium and gene size (Liu et al., 2010). As this review is qualitative in nature, only differentiated gene or protein names identified by the study as significant were obtained from studies. Based on a simplified view of the sequential transfer of genetic information, it was expected that there would be significant overlap between the stroke genome, transcriptome and proteome. It was further hypothesized that based on past research in stroke pathophysiology and the findings of individual omics studies, it is likely that the biological pathways and molecules that influence cardiovascular function and inflammatory responses will be shown to overlap between genetics, transcriptomics and proteomics studies.

Method

Genetics studies

The genetics data obtained for this multi-omics review were derived from a large international collaborative effort to conduct a meta-analysis of stroke GWAS studies (Malik et al., 2018). Recently, the MEGASTROKE consortium has published a comprehensive meta-analysis of 29 GWAS studies across different ethnic ancestries and stroke types, comparing the genomes of 67,162 stroke and 454,450 control cases (Malik et al., 2018). This analysis found significant associations for 32 SNPs and incidence stroke amongst European, African, East Asian, South Asian, mixed Asian and Latin American ancestry.

After a GWAS, gene-based tests such as VEGAS provide an analytical approach that assesses for enrichment of SNPs and the surrounding kilobases for associations with

diseases or traits (Liu et al., 2010; Mishra & Macgregor, 2015). Briefly, the VEGAS test improves upon standard per-SNP tests by utilizing permutations that account for linkage disequilibrium and genome size (Mishra & Macgregor, 2015). This data was used for the current analysis. For the gene boundary settings for SNP selection, SNPs were chosen within ± 10 kb, relative to the 5' and 3' untranslated region (Malik et al., 2018). Available in the supplementary data for the MEGASTROKE study, the VEGAS test has been conducted for all ethnicities in each stroke type and combined using Stouffer's *z* as a meta-analytical method to produce a final gene list. A total of 29 genes were related to all stroke subtypes, using a conservative statistical cut-off of Bonferroni corrected $p < 2.02 \times 10^{-6}$. Based on the statistical cut-offs utilized in the original VEGAS paper (Liu et al., 2010), our review has used the same cut-off of $p < .05 \times 10^{-4}$ to expand the list of genes to 158 (Appendix 2, Chapter 3, Supplementary Data 1).

Proteomics and Transcriptomics Studies

A similar search strategy was employed for both transcriptomics studies and proteomics studies, constituting two separate systematic searches. Stroke search terms were deliberately set as broad, with the intention to capture both ischemic and hemorrhagic stroke types. Transcriptomics studies were obtained by conducting a systematic search of PubMed incorporating studies utilizing transcriptome wide microarray and RNA sequencing platforms targeting mRNA (Figure 1). Proteomics studies were obtained by conducting a systematic search of PubMed (January 2019) incorporating studies utilizing techniques such as 2D electrophoresis and mass spectrometry in an untargeted format (Figure 2). The search strings used are available in Supplementary Methods (S1), with human selected as the search species in PubMed. To follow a similar study design to GWAS, only studies comparing the expressed mRNA transcriptome or proteome between stroke cases and controls were selected. Studies were only included if the time of stroke onset relative to blood collection was in the acute

period of <2 weeks (Roth et al., 2016). All articles available in English and published after the first draft of the human genome (Wright et al., 2001) from 2001 to September 2019 were included.

After performing the search, the PubMed search results were exported to the systematic review management web software, Covidence. VN performed the abstract screening process. For the transcriptomics articles, studies were excluded if there was no mention of 'genes' or 'gene expression' in the overall abstract and reference to 'mRNA', 'transcriptomics', 'microarray' or 'next generation sequencing' in the methodology. For the proteomics studies, studies were excluded if there was no mention of 'proteins' or protein expression in the overall abstract and reference to 'mass spectrometry', 'LC-MS', 'GC-MS', '2D-DIGE', 'SDS-PAGE'. VN and SGC performed the full-text review, examining articles for inclusion of a control group, usage of blood samples, untargeted methodology with compatible formatting of gene and protein identifiers. Studies must also have conducted their own transcriptomic or proteomic analysis and not utilized previously published data. Records were excluded if they did not utilize an untargeted transcriptomics or proteomics workflow, did not have blood as the sample type or did not have data in the published article or supplementary data. NR acted as a third reviewer to resolve disagreements on article inclusion.

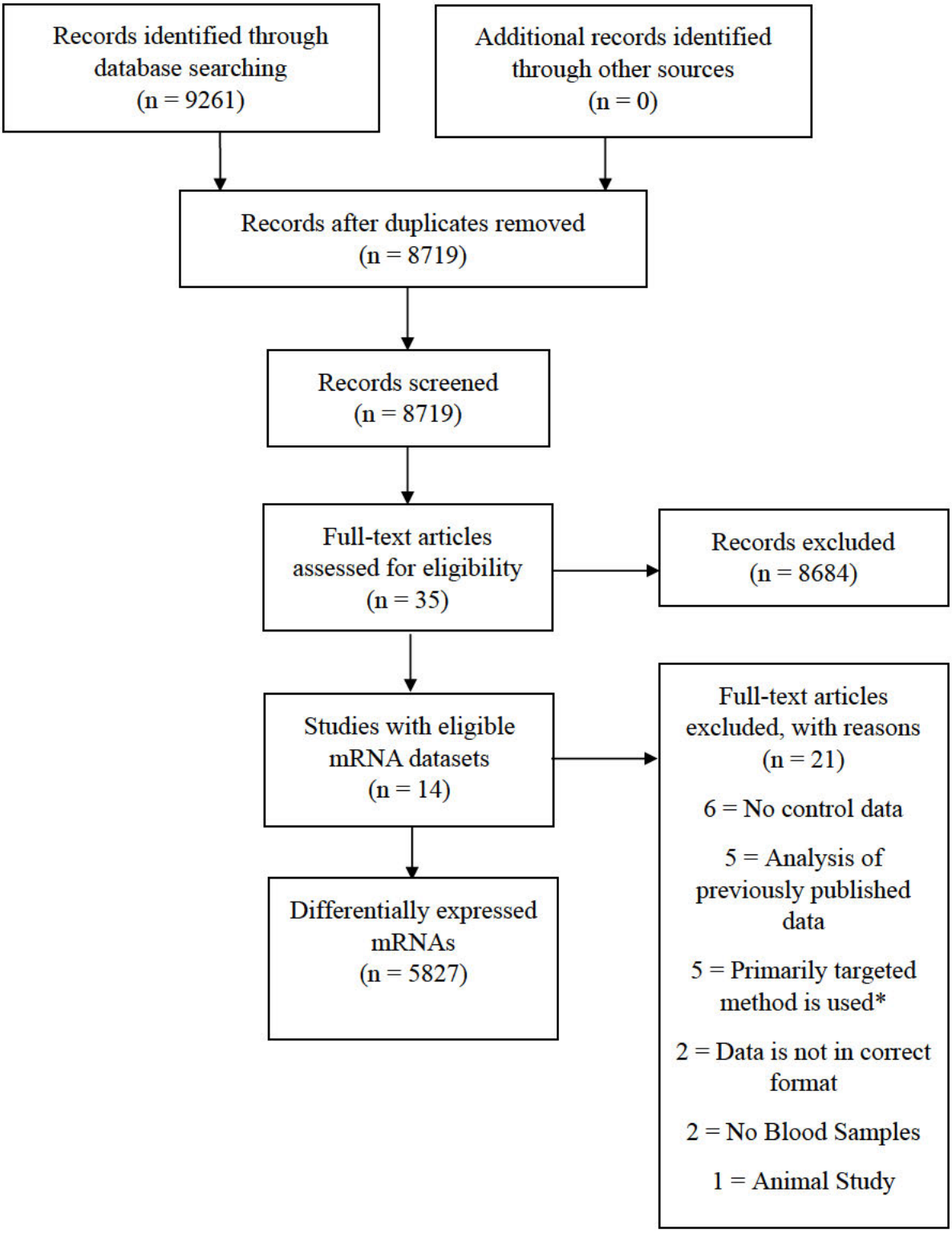

Figure 1. PRISMA flow-chart for stroke transcriptomics studies

*Note: A targeted transcriptomics method such as quantitative polymerase chain reaction (QT-PCR) or targeted microarrays is used where authors choose to assess selected mRNA only

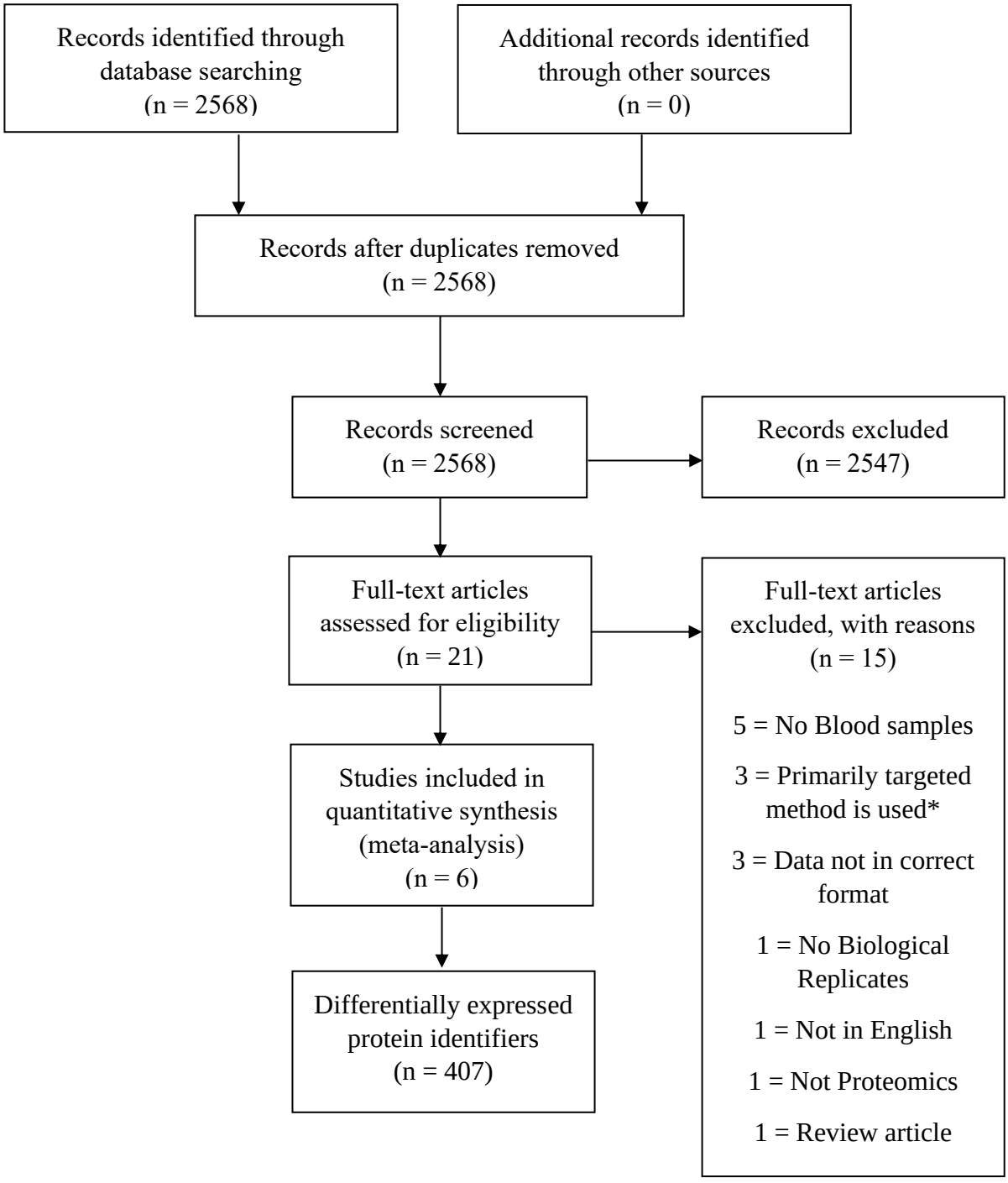

Figure 2. PRISMA flow-chart for stroke proteomics studies

*Note: A primarily targeted transcriptomics method such as quantitative targeted mass spectrometry is used where authors choose to assess selected proteins.

Data Extraction

The identifiers of differentiated mRNA transcripts (probe identifiers, gene symbol; Appendix 3, Chapter 3, Supplementary Data 2) or proteins (Uniprot ID, gene symbol; Appendix 4, Chapter 3, Supplementary Data 3) were obtained from published results or supplementary data, based on the author's statistical criteria. This criteria included data that was deemed as the first level of inferential statistics where single identifiers were significantly differentially expressed between stroke and controls, before the addition of advanced statistical methods such as the prediction analysis of microarray models (Dabney, 2005). Data was extracted based on the differential comparison group i.e. ischemic stroke vs control or hemorrhagic stroke vs control, per study. Studies containing findings with more than a single timepoint within the acute stroke window were collected and each dataset was extracted independently for each individual timepoint (i.e. 12 hours, 24 hours and 48 hours). Therefore, data was extracted for each differential comparison (ischemic stroke or hemorrhagic stroke), timepoint (e.g. 12 hours, 24 hours, 48 hours…) and also a combination of differential comparison and timepoints (e.g. hemorrhagic stroke at 12 hours or ischemic stroke at 24 hours). The list of included transcriptomics studies is included as Supplementary Table 1 and the list of included proteomics studies is included as Supplementary Table 2.

Multi-omics comparisons

To conduct a comparison over multiple omics approaches, molecular identifiers were transcribed from gene or protein accession numbers to the National Center for Biotechnology Information's (NCBI) Entrez gene identifiers using the gene mapping function in the org.Hs.eg.db R package (ver 3.7.0). The resulting lists were amalgamated into a large single list, with duplicates and missing values removed. This resulted in a single list of associated genes from the MEGASTROKE study, a single combined list of transcripts and a single combined list of proteins (Supplementary data 4). This method

provides a basic approach to overcome heterogeneity in sampling, methodology and stroke onset times with the noted disadvantages of reduced specificity. This step ensures similar formatting of data across the omics levels from transcriptomics probe identifiers, proteomics Uniprot identifiers and gene symbols, as well as allowing ease of access to bioinformatics tools that prefer the NCBI database (Maglott, Ostell, Pruitt, & Tatusova, 2011).

The GeneOverlap R package (ver 1.18.0) was employed to examine the degree of overlap between significant findings in the lists of literature defined significantly differentiated genes, transcripts and proteins (Shen, 2016). This package compares lists of identifiers utilizing Fisher's exact test in contingency tables against a defined number of background genes. In this case, the chosen background was the current NCBI defined protein coding genes with Entrez identifiers (n = 20,238). In a pairwise fashion, GeneOverlap provides the intersecting list of overlapping identifiers, the expected odds ratio, Jaccard index for evaluating the similarity or diversity between lists and a p-value. Benjamin-Hochberg corrections to the resulting contingency table were conducted using base R functions (ver 3.5.2). Sets of identifiers were analysed using all possible pairwise comparisons between the pooled genetics list, IS, ICH and a combination of both IS and IS.

Functional Gene Annotation

To further characterize the lists of significantly overlapping candidates into biological processes and pathways, functional annotations were conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID) classification tool (Huang, Sherman, & Lempicki, 2009; Huang et al., 2007). DAVID is a popular bioinformatics approach that can further classify significant findings from genetics, transcriptomics or proteomics studies into functional biological categories such as coagulation using Fisher exact tests in the Expression Analysis Systematic Explorer

(EASE) algorithm (Hosack, Dennis, Sherman, Lane, & Lempicki, 2003). The DAVID tool with Gene Ontology (GO) was used as the classification database (Mi et al., 2017). We utilized the GO Direct Biological Process (BP) categories where the terms are based on closest annotation to the source GO database, compared to base GO functions. The background list utilized the complete NCBI protein coding *Homo sapiens* genes (n = 20,238) with a minimum threshold of 3 candidates identified in the category against an FDR p of .05. Venn diagrams were constructed to visualize the amount of overlap between omics levels, using the VennDiagram R package (ver 1.6.20) (Chen, 2018). Chord diagrams were generated to display the categorical relationship between genes and their biological processes using the GOplot R package (ver 1.0.2) (Walter, Sánchez-Cabo, & Ricote, 2015).

Results

Systematic Review Overview

The systematic search of blood-based systems biology approaches in ischemic and hemorrhagic stroke identified 14 mRNA transcriptomics studies and 6 proteomics studies. Within the 22 available discrete mRNA transcriptomics datasets, there were 20 ischemic stroke datasets and 2 hemorrhagic stroke datasets available for extraction (Figure 3). Within the 6 available discrete proteomics datasets, there were 5 IS datasets and one dataset from Bian et al. (2014) incorporating both IS and ICH cases in their differentially expressed protein list. There were no proteomic comparisons of ICH cases to controls. As such, no comparisons are able to be conducted between differentially expressed ICH mRNA and proteins.

For stroke transcriptomics studies, studies utilized a range of sample types. These included QIAGEN PaxGene pre-prepared blood tubes (n = 9), whole blood with post-processing with reagents such as TRIzol (n = 3) and peripheral blood mononuclear cells (n = 2). For proteomics studies, studies have utilized serum (n = 2),

ethylenediaminetetraacetic acid plasma (n = 2), peripheral blood mononuclear cells (n = 1) and serum (n = 1).

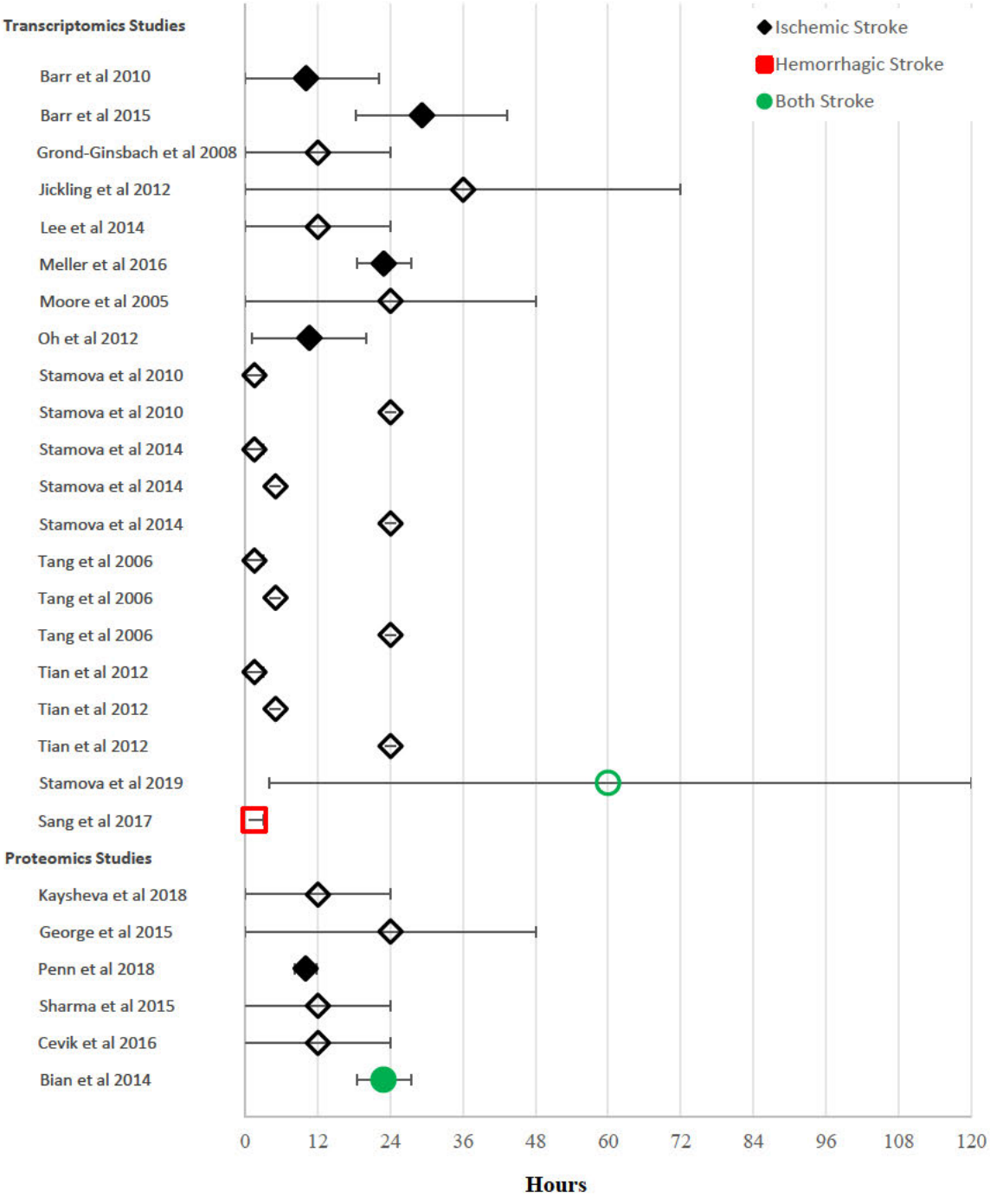

Figure 3. Maximum time to blood draw relative to stroke onset for 22 datasets from 14 transcriptomics studies and 6 datasets from 6 proteomics studies.
Note: Open markers with error bars represent range given (i.e <24 hours), closed markers with error bars represent mean and SD.

Overlapping Molecules

The differentially expressed results from past stroke genetics, transcriptomics and proteomics studies were compared. A total of 159 GWAS based candidates were identified from the MEGASTROKE VEGAS gene-based tests reaching $p < .05 \times 10^{-4}$. Once the studies were pooled for differentially expressed transcripts, there were 2747 mRNA identifiers from IS studies (Supplementary Data 2.1), 4292 mRNA identifiers from ICH studies (Supplementary Data 2.2), and a total combined 5827 mRNA identifiers from both IS and ICH studies (Supplementary Data 2.3). Pooled protein identifiers totalled 130 (Supplementary Data 3.1) for IS studies and 407 for the combination of IS and ICH studies (Supplementary Data 3.2). Overall, the proportion of the known human protein coding genome differentially expressed in all stroke transcriptomes and in all stroke proteomes was 28.79% and 6.98% respectively.

Overlap analysis revealed no significant overlap between the GWAS candidates and expressed transcriptome nor expressed proteome from the reviewed studies (Table 1). There was only significant overlap between the pooled IS and ICH mRNA list and pooled IS and ICH proteome ($p = 6.62 \times 10^{-11}$). A total of 178 molecules were found to significantly overlap between combined IS and ICH transcriptomics and proteomics studies. The significantly overlapping lists are available in the online only Supplementary Table 3.

Table 1

Table representing the *p*-value of all meaningful pairwise comparisons using GeneOverlap.

	Genetics	Transcriptomics			Proteomics	
	All Stroke (*n* = 159)	IS (*n* = 2747)	ICH (*n* = 4292)	IS + ICH (*n* = 5827)	IS (*n* = 130)	IS + ICH (*n* = 407)
Genetics						
All Stroke	-	-	-	-	-	-
IS	-	-	-	-	-	-
Transcriptomics						
IS	0.67	-	-	-	-	-
HS	0.63	0.89	-	-	-	-
IS + ICH	0.63	-	-	-	-	-
Proteomics						
IS	0.27	.60	-	-	-	-
IS + ICH	0.10	-	-	*6.62×10^{-11}	-	-

* Statistical significance at $p < .05$
n represents the size of final pooled identifier lists for each omics type.

To assess the categorization of significantly overlapping elements between omics into biological systems and pathways, we conducted further functional annotation analyses utilizing DAVID (Supplementary Data 4) and organized the top functional categories into a chord diagram (Figure 4). Using the GO Direct BP database through the DAVID bioinformatics platform (Huang et al., 2007), the relationship between mRNA and protein studies showed significant enrichment for 178 molecules in relation to 18 well defined biological processes. Given the time of stroke onset to blood collection in the reviewed studies, these results represent the context of molecules and their constituent biological processes in the acute period, as identified by untargeted omics methods. The significant overlap between stroke transcriptomics and proteomics studies resolved into pathways appear to cluster into wider biological functions including platelet function, response to oxidative stress and leukocyte chemotaxis.

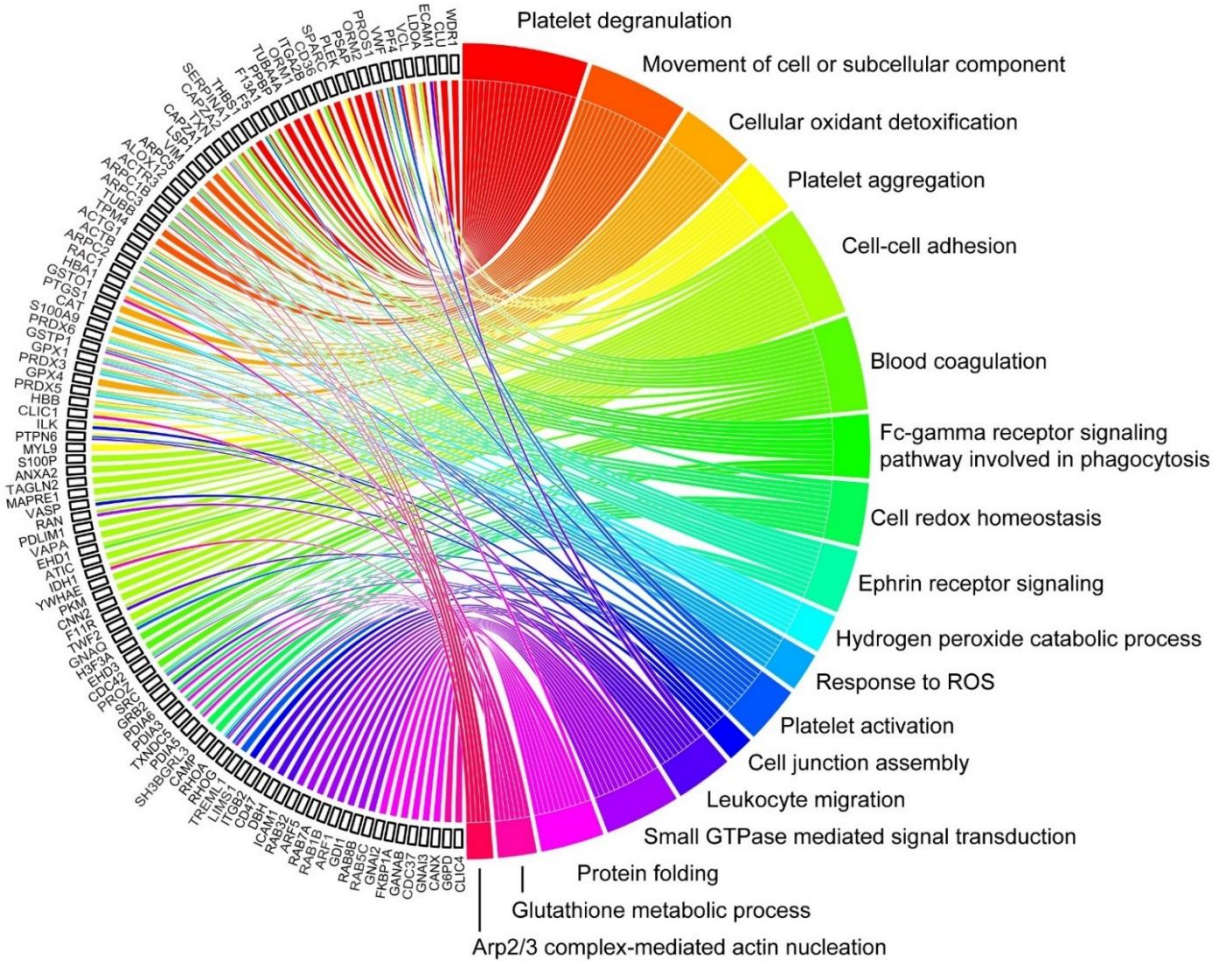

Figure 4. Chord diagram displaying the relationships between significantly enriched biological functions and their constituent molecules. The contributing molecules are shown on the left, with the lines depicting membership to the GO biological processes on the right. Molecules may be related to more than one biological process.

Discussion

Our study examined the overlapping molecules and biosystems identified in genetics, transcriptomics and proteomics studies in human stroke research. This was achieved by comparing the significantly identified molecules in previous studies, with genetics data from the MEGASTROKE meta-analysis (Malik et al., 2018) and transcriptomics and proteomics data compiled from published data. Our approach aimed to provide a descriptive and more broad overview of the molecules and mechanisms involved in the acute post-stroke timeline, with the goal to shift understanding to a 'candidate biosystems' approach as an augmentation to the traditional biomarker

approach in medical research (Feala et al., 2013). To our knowledge this is the first attempt to provide a unified comparison of the systems biology findings in stroke research. Our most important result shows that the unification of multiple studies across transcriptomics and proteomics provides a strong overview of the biological processes in stroke that largely agrees with the findings of the current literature.

At an overall 'omics' level, the lack of overlap between the genome and downstream expression in proteomics and transcriptomics is counterintuitive to the expected model of sequential transfer of genomic information, stating that DNA transfers the information for protein synthesis by mRNA expression (Cobb, 2017). This is also surprising given that we did not find a significant overlap between the relatively large portion of studies identifying significantly differential overlap in mRNA and proteins when compared to the overlap between mRNA or proteins to the stroke genome. However, the central dogma of molecular biology does not represent a complete picture of transcriptomics expression, demonstrating that factors such as epigenetics, microRNA and long non-coding RNA regulation add to the complexity of DNA to mRNA transcription (Gomes, Nolasco, & Soares, 2013). Indeed, the conceptual and statistical relationship of genetics to trait expression is typically low and it is implied that in many models of genetic risk, the non-coding regions of DNA act as regulatory mechanisms of the transcriptome as opposed to straightforward downstream expression (Qin, Niu, & Zhao, 2019).

The overlapping relationship between transcriptome and proteome has been shown to be strong in bacteria and eukaryotes, suggesting that a significant portion of protein abundance can be explained in part by transcriptome abundance (Qin et al., 2019). There was a relatively small number of proteomics studies incorporated here that have collectively analysed plasma samples and platelets, using mass spectrometry methods. Although blood samples are commonly collected in the medical setting, mass

spectrometry proteomics analysis of human plasma remains difficult at best, with the abundance of large proteins such as albumin dominating the protein content of the sample and thus restricting proteome coverage (Dayon & Kussmann, 2013). This issue is compounded in discovery proteomics studies where techniques such as reduction of mass range on the spectrometer or immunodepletion methods (Tu et al., 2010) introduce additional machine time and costs. Overall, this contributes to lower than expected numbers of identified proteins in these studies and explains the large discrepancy between identification of unique mRNA and proteins.

The biosystems identified as conserved between the stroke transcriptome and proteome are closely related to platelet function (Kim et al., 2018), response to oxidative stress (Allen & Bayraktutan, 2009), and adjacent inflammatory responses, such as cell-to-cell adhesion (Yilmaz & Granger, 2008), and leukocyte migration (Yilmaz & Granger, 2010). Given the average time of stroke onset to sample collection in the reviewed transcriptomics and proteomics studies was 24-48 hours, our analyses suggests that the overlapping biosystems are closely related to the timeline of haemostatic activation and post-ischemic inflammation pathways suggested in established theories of post-stroke biosystems involvement (Kamel & Iadecola, 2012; Saenger & Christenson, 2010). Therefore, the biosystems identified in this study are likely to be representative of the response to blood early within the acute stroke timeline.

Platelets are involved in a variety of roles to maintain hemostasis and immune function (Koupenova, Clancy, Corkrey, & Freedman, 2018). Our study has identified a range of mRNA transcripts and proteins related to platelet activation, aggregation and degranulation. Platelet activation begins after adhesion processes to activated sites of tissue injury expressing proteins such as fibrinogen, collagen and von Willebrand factor (Ruggeri & Mendolicchio, 2007). Some of the molecules identified in this study that share dual membership in coagulation and platelet pathways such as Ras-related C3

botulinum toxin substrate 1 (RAC1) (Dwivedi, Pandey, Khandoga, Brandl, & Siess, 2010), glycoprotein 9 (GP9) and glycoprotein Ib platelet beta subunit (GP1BB) (Israels & Rand, 2013) have some evidence for regulation of surface adhesion properties. The majority of platelet secretory proteins are stored within the cytosol in vesicles or granules (Koupenova, et al., 2018). The activated platelet phenotype begins a process of degranulation that is characterized by the exocytosis of these granules into the plasma and can provide chemotaxic localization of hemostasis, inflammation and immune responses (Yun, Sim, Goh, Park, & Han, 2016). Platelet factor 4 (PF4) is the one protein contained within platelet alpha granules and its secretion into plasma during platelet degranulation acts as a signalling molecule for platelet aggregation for immunological targets (Kowalska, Rauova, & Poncz, 2010). Lastly, aggregation is a process that involves recruitment of more platelets to assist in the formation of a haemostatic plug against vascular and endothelial damage, with proteins such as integrin-linked kinase shown to be central to focal adhesion and aggregation (Jones et al., 2014). Although there was no significant enrichment of immune related pathways between the transcriptomics and proteomics studies, possibly due to the lack of proteomics studies using peripheral blood mononuclear cells, platelet-leukocyte aggregates can typically form where coagulation and inflammatory events are interrelated such as stroke (Franks, Campbell, Weyrich, & Rondina, 2010).

Cellular redox homeostasis refers to a resolution of damaging intracellular accumulation of reactive oxygen species (ROS). Oxidative stress is a potent factor in stroke risk that exacerbates vascular damage and atherosclerotic plaque formation, a common issue with aging (Allen & Bayraktutan, 2009). It is well known that during stroke, rapid increases in oxidative stress contributes greatly to ischemic damage and can cause further tissue damage in the penumbra (Ramon et al., 2013) via damage-associated molecular pattern interactions with resident immune cells (Garcia-Bonilla & Iadecola,

2012). Indeed, free radical damage is suggested to be highest within 24 hours but is reduced at 3 months (Žitňanová et al., 2016). Identified in our study, haptoglobin (HP) is an abundant circulatory protein that participates in downstream haemoglobin antioxidation and may be neuroprotective of ischemic damage post-stroke (Andersen et al., 2016; Zhao et al., 2009). Similarly, thioredoxin refers to a family of mitochondrial enzymes that facilitates intracellular hydrogen peroxide buffering by reducing oxidation of peroxiredoxin 3 (PRDX3) to peroxiredoxin 5 (PRDX3) (Holzerova et al., 2015), all of which are molecules that are collected under oxidative stress pathways in our analyses. Clinical trials into pharmacological approaches targeting oxidative stress pathways such as edaravone and cyclosporine A are undergoing investigation (Matsumoto et al., 2018).

Limitations and Conclusions

The multi-omics review approach utilized in the current study has some limitations to interpretability. In an effort to encompass a wide selection of studies utilizing different methodological approaches and data reporting formats, we have chosen to collect the most superficial level of bioinformatics data at the molecular identifier only. Unfortunately, this limits the quantitative richness of bioinformatics data that includes the direction and intensity of transcript and protein expression that would have allowed us to describe if a biological process was up or down-regulated in relation to stroke damage. As such, it is difficult to suggest how these pathways would be activated or suppressed in the post-stroke context. Regardless, this level of descriptive data may be useful in providing understanding of the relationships between the ischemic stroke genome, transcriptome and proteome at an entry level. Ultimately, our study represents an initial attempt at unifying studies using discovery-based systems biology approaches. This unifying approach could be improved with studies adhering to universal standards of systems biology data sharing by providing clear molecular identifiers and quantitative information within data repositories or supplementary data. Furthermore, studies examining

differential protein or transcriptomic profiles could expand their investigations towards the longitudinal recovery phase in the months to years post-stroke. Future studies can utilize more powerful multi-omics workflows (Hu et al., 2018) on single cells such as leukocytes and platelets to better understand the peripheral biological processes involved in stroke recovery.

In this review, we have combined and compared multiple levels of omics data in stroke research, including available genetics, transcriptomics and proteomics studies. Overall, these studies have shown a small to moderate degree of overlap between omics levels, with the greatest degree shown between the stroke transcriptome and proteome. Given that the average time samples were taken relative to stroke onset in the reviewed studies was within the acute stroke phase, the identified biological processes of platelet function, oxidative stress and immune system signalling are in concordance with the current stroke literature. To improve future stroke clinical care, a more complete pathophysiological understanding of post-stroke biology from the acute phase to the longitudinal period of post-stroke recovery is needed to develop effective personalized treatment regimes.

References

Allen, C. L., & Bayraktutan, U. (2009). Oxidative stress and its role in the pathogenesis of ischaemic stroke. *International Journal of Stroke, 4*(6), 461-470. doi:10.1111/j.1747-4949.2009.00387.x

Andersen, C. B. F., Stødkilde, K., Sæderup, K. L., Kuhlee, A., Raunser, S., Graversen, J. H., & Moestrup, S. K. (2016). Haptoglobin. *Antioxidants & Redox Signaling, 26*(14), 814-831. doi:10.1089/ars.2016.6793

Andersen, K. K., Olsen, T. S., Dehlendorff, C., & Kammersgaard, L. P. (2009). Hemorrhagic and ischemic strokes compared. *Stroke, 40*(6), 2068-2072. doi:doi:10.1161/STROKEAHA.108.540112

Barr, T. L., Conley, Y., Ding, J., Dillman, A., Warach, S., Singleton, A., & Matarin, M. (2010). Genomic biomarkers and cellular pathways of ischemic stroke by rna gene expression profiling. *Neurology, 75*(11), 1009-1014. doi:10.1212/WNL.0b013e3181f2b37f

Barr, T. L., VanGilder, R., Rellick, S., Brooks, S. D., Doll, D. N., Lucke-Wold, A. N., . . . Matarin, M. (2015). A genomic profile of the immune response to stroke with implications for stroke recovery. *Biological Research for Nursing, 17*(3), 248-256. doi:10.1177/1099800414546492

Bian, F., Simon, R. P., Li, Y., David, L., Wainwright, J., Hall, C. L., . . . Zhou, A. (2014). Nascent proteomes in peripheral blood mononuclear cells as a novel source for biomarker discovery in human stroke. *Stroke, 45*(4), 1177-1179. doi:10.1161/STROKEAHA.113.004576

Boehme, A. K., Esenwa, C., & Elkind, M. S. V. (2017). Stroke risk factors, genetics, and prevention. *Circulation Research, 120*(3), 472-495. doi:doi:10.1161/CIRCRESAHA.116.308398

Cevik, O., Baykal, A. T., & Sener, A. (2016). Platelets proteomic profiles of acute ischemic stroke patients. *PloS One, 11*(6), e0158287. doi:10.1371/journal.pone.0158287

Chen, H. (2018). Venndiagram: Generate high-resolution venn and euler plots. Retrieved from https://CRAN.R-project.org/package=VennDiagram

Cobb, M. (2017). 60 years ago, Francis Crick changed the logic of biology. *PLoS Biology, 15*(9), e2003243. doi:10.1371/journal.pbio.2003243

Dabney, A. R. (2005). Classification of microarrays to nearest centroids. *Bioinformatics, 21*(22), 4148-4154. doi:10.1093/bioinformatics/bti681

Dayon, L., & Kussmann, M. (2013). Proteomics of human plasma: A critical comparison of analytical workflows in terms of effort, throughput and outcome. *EuPA Open Proteomics, 1,* 8-16. doi:https://doi.org/10.1016/j.euprot.2013.08.001

Dwivedi, S., Pandey, D., Khandoga, A. L., Brandl, R., & Siess, W. (2010). Rac1-mediated signaling plays a central role in secretion-dependent platelet aggregation in human blood stimulated by atherosclerotic plaque. *Journal of Translational Medicine, 8,* 128-128. doi:10.1186/1479-5876-8-128

Feala, J. D., Abdulhameed, M. D. M., Yu, C., Dutta, B., Yu, X., Schmid, K., . . . Reifman, J. (2013). Systems biology approaches for discovering biomarkers for traumatic brain injury. *Journal of Neurotrauma, 30*(13), 1101-1116. doi:10.1089/neu.2012.2631

Franks, Z. G., Campbell, R. A., Weyrich, A. S., & Rondina, M. T. (2010). Platelet-leukocyte interactions link inflammatory and thromboembolic events in ischemic stroke. *Annals of the New York Academy of Sciences, 1207,* 11-17. doi:10.1111/j.1749-6632.2010.05733.x

Garcia-Bonilla, L., & Iadecola, C. (2012). Peroxiredoxin sets the brain on fire after stroke. *Nature Medicine, 18*(6), 858-859. doi:10.1038/nm.2797

George, P. M., Mlynash, M., Adams, C. M., Kuo, C. J., Albers, G. W., & Olivot, J. M. (2015). Novel tia biomarkers identified by mass spectrometry-based proteomics. *International Journal of Stroke, 10*(8), 1204-1211. doi:10.1111/ijs.12603

Grond-Ginsbach, C., Hummel, M., Wiest, T., Horstmann, S., Pfleger, K., Hergenhahn, M., . . . Wagner, S. (2008). Gene expression in human peripheral blood mononuclear cells upon acute ischemic stroke. *Journal of Neurology, 255*(5), 723-731. doi:10.1007/s00415-008-0784-z

Gomes, A. Q., Nolasco, S., & Soares, H. (2013). Non-coding RNAs: Multi-tasking molecules in the cell. *International Journal of Molecular Sciences, 14*(8), 16010-16039. doi:10.3390/ijms140816010

Holzerova, E., Danhauser, K., Haack, T. B., Kremer, L. S., Melcher, M., Ingold, I., . . . Distelmaier, F. (2015). Human thioredoxin 2 deficiency impairs mitochondrial redox homeostasis and causes early-onset neurodegeneration. *Brain, 139*(2), 346-354. doi:10.1093/brain/awv350

Hosack, D. A., Dennis, G., Jr., Sherman, B. T., Lane, H. C., & Lempicki, R. A. (2003). Identifying biological themes within lists of genes with ease. *Genome Biology, 4*(10), R70-R70. doi:10.1186/gb-2003-4-10-r70

Hu, Y., An, Q., Sheu, K., Trejo, B., Fan, S., & Guo, Y. (2018). Single cell multi-omics technology: Methodology and application. *Frontiers in Cell and Developmental Biology, 6*(28). doi:10.3389/fcell.2018.00028

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research, 37*(1), 1-13. doi:10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., Tan, Q., Collins, J. R., Alvord, W. G., Roayaei, J., . . . Lempicki, R. A. (2007). The DAVID gene functional classification tool: A novel

biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology, 8*(9), R183-R183. doi:10.1186/gb-2007-8-9-r183

Israels, S. J., & Rand, M. L. (2013). What we have learned from inherited platelet disorders. *Pediatric Blood & Cancer, 60*(S1), S2-S7. doi:10.1002/pbc.24345

Jickling, G. C., Zhan, X., Stamova, B., Ander, B. P., Tian, Y., Liu, D., . . . Sharp, F. R. (2012). Ischemic transient neurological events identified by immune response to cerebral ischemia. *Stroke, 43*(4), 1006-1012. doi:10.1161/strokeaha.111.638577

Jones, C. I., Tucker, K. L., Sasikumar, P., Sage, T., Kaiser, W. J., Moore, C., . . . Gibbins, J. M. (2014). Integrin-linked kinase regulates the rate of platelet activation and is essential for the formation of stable thrombi. *Journal of Thrombosis and Haemostasis, 12*(8), 1342-1352. doi:10.1111/jth.12620

Kamel, H., & Iadecola, C. (2012). Brain-immune interactions and ischemic stroke: Clinical implications. *Archives of Neurology, 69*(5), 576-581. doi:10.1001/archneurol.2011.3590

Kamel, H., Okin, P. M., Elkind, M. S. V., & Iadecola, C. (2016). Atrial fibrillation and mechanisms of stroke: Time for a new model. *Stroke, 47*(3), 895-900. doi:10.1161/STROKEAHA.115.012004

Kaysheva, A. L., Kopylov, A. T., Ponomarenko, E. A., Kiseleva, O. I., Teryaeva, N. B., Potapov, A. A., . . . Archakov, A. I. (2018). Relative abundance of proteins in blood plasma samples from patients with chronic cerebral ischemia. *Journal of Molecular Neuroscience, 64*(3), 440-448. doi:10.1007/s12031-018-1040-3

Kim, J.-T., Choi, K.-H., Park, M.-S., Lee, J. S., Saver, J. L., & Cho, K.-H. (2018). Clinical significance of acute and serial platelet function testing in acute ischemic stroke. *Journal of the American Heart Association, 7*(11), e008313. doi:10.1161/JAHA.117.008313

Koupenova, M., Clancy, L., Corkrey, H. A., & Freedman, J. E. (2018). Circulating platelets as mediators of immunity, inflammation, and thrombosis. *122*(2), 337-351. doi:doi:10.1161/CIRCRESAHA.117.310795

Kowalska, M. A., Rauova, L., & Poncz, M. (2010). Role of the platelet chemokine platelet factor 4 (pf4) in hemostasis and thrombosis. *Thrombosis Research, 125*(4), 292-296. doi:10.1016/j.thromres.2009.11.023

Lee, H.-B., Kim, Y., Yoo, H., Lee, J.-M., Kim, Y.-K., Kim, N.-K., . . . Oh, S.-H. (2014). Blood genomic profiling in extracranial- and intracranial atherosclerosis in ischemic stroke patients. *Thrombosis Research, 134*(3), 686-692. doi:https://doi.org/10.1016/j.thromres.2014.06.025

Leon-Mimila, P., Wang, J., & Huertas-Vazquez, A. (2019). Relevance of multi-omics studies in cardiovascular diseases. *Frontiers in Cardiovascular Medicine, 6*(91). doi:10.3389/fcvm.2019.00091

Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., . . . Macgregor, S. (2010). A versatile gene-based test for genome-wide association studies. *American Journal of Human Genetics, 87*(1), 139-145. doi:10.1016/j.ajhg.2010.06.009

Maglott, D., Ostell, J., Pruitt, K. D., & Tatusova, T. (2011). Entrez gene: Gene-centered information at ncbi. *Nucleic Acids Research, 39*(Database issue), D52-D57. doi:10.1093/nar/gkq1237

Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., . . . Attia, J. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics, 50*(4), 524-537. doi:10.1038/s41588-018-0058-3

Matsumoto, S., Murozono, M., Kanazawa, M., Nara, T., Ozawa, T., & Watanabe, Y. (2018). Edaravone and cyclosporine a as neuroprotective agents for acute ischemic stroke. *Acute Medicine & Surgery, 5*(3), 213-221. doi:10.1002/ams2.343

Meller, R., Pearson, A. N., Hardy, J. J., Hall, C. L., McGuire, D., Frankel, M. R., & Simon, R. P. (2016). Blood transcriptome changes after stroke in an african american population. *Ann Clin Transl Neurol, 3*(2), 70-81. doi:10.1002/acn3.272

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. D. (2017). Panther version 11: Expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research, 45*(D1), D183-D189. doi:10.1093/nar/gkw1138

Mishra, A., & Macgregor, S. (2015). VEGAS2: Software for more flexible gene-based testing. *Twin Research and Human Genetics, 18*(1), 86-91. doi:10.1017/thg.2014.79

Moore, D. F., Li, H., Jeffries, N., Wright, V., Cooper, R. A., Elkahloun, A., . . . Baird, A. E. (2005). Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke. *Circulation, 111*(2), 212-221. doi:doi:10.1161/01.CIR.0000152105.79665.C6

Oh, S.-H., Kim, O.-J., Shin, D.-A., Song, J., Yoo, H., Kim, Y.-K., & Kim, J.-K. (2012). Alteration of immunologic responses on peripheral blood in the acute phase of ischemic stroke: Blood genomic profiling study. *Journal of Neuroimmunology, 249*(1), 60-65. doi:https://doi.org/10.1016/j.jneuroim.2012.04.005

Penn, A. M., Saly, V., Trivedi, A., Lesperance, M. L., Votova, K., Jackson, A. M., . . . Borchers, C. H. (2018). Differential proteomics for distinguishing ischemic stroke from controls: A pilot study of the spectra project. *Translational stroke research, 9*(6), 590-599. doi:10.1007/s12975-018-0609-z

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., Becker, K., . . . Tirschwell, D. L. (2018). 2018 guidelines for the early management of patients with acute ischemic stroke: A guideline for healthcare

professionals from the American Heart Association/American Stroke Association. *Stroke*.

Qin, H., Niu, T., & Zhao, J. (2019). Identifying multi-omics causers and causal pathways for complex traits. *Frontiers in Genetics, 10*(110). doi:10.3389/fgene.2019.00110

Ramon, R., Rodrigo, F.-G., Rodrigo, G., Jose Manuel, M., Rodrigo, C., Andres, M.-M., & Walter, F. (2013). Oxidative stress and pathophysiology of ischemic stroke: Novel therapeutic opportunities. *CNS & Neurological Disorders - Drug Targets, 12*(5), 698-714. doi:http://dx.doi.org/10.2174/18715273113120500 15

Rayasam, A., Hsu, M., Kijak, J. A., Kissel, L., Hernandez, G., Sandor, M., & Fabry, Z. (2018). Immune responses in stroke: How the immune system contributes to damage and healing after stroke and how this knowledge could be translated to better cures? *Immunology, 154*(3), 363-376. doi:10.1111/imm.12918

Roth, D., Sheehan, O., Huang, J., Rhodes, J., Judd, S., Kilgore, M., . . . Haley, W. (2016). Medicare claims indicators of healthcare utilization differences after hospitalization for ischemic stroke: Race, gender, and caregiving effects. *International Journal of Stroke, 11*. doi:10.1177/1747493016660095

Ruggeri, Z. M., & Mendolicchio, G. L. (2007). Adhesion mechanisms in platelet function. *Circulation Research, 100*(12), 1673-1685. doi:doi:10.1161/01.RES.0000267878.97021.ab

Saenger, A. K., & Christenson, R. H. (2010). Stroke biomarkers: Progress and challenges for diagnosis, prognosis, differentiation, and treatment. *Clinical Chemistry, 56*(1), 21. doi:10.1373/clinchem.2009.133801

Sang, M., Wang, X., Zhang, H., Sun, X., Ding, X., Wang, P., . . . Zhang, G. (2017). Gene expression profile of peripheral blood mononuclear cells in response to intracerebral hemorrhage. *DNA and Cell Biology, 36*(8), 647-654. doi:10.1089/dna.2017.3650

Sharp, F. R., Jickling, G. C., Stamova, B., Tian, Y., Zhan, X., Liu, D., . . . Ander, B. P. (2011). Molecular markers and mechanisms of stroke: RNA studies of blood in animals and humans. *Journal of Cerebral Blood Flow and Metabolism, 31*(7), 1513-1531. doi:10.1038/jcbfm.2011.45

Sharma, R., Gowda, H., Chavan, S., Advani, J., Kelkar, D., Kumar, G. S., . . . Christopher, R. (2015). Proteomic signature of endothelial dysfunction identified in the serum of acute ischemic stroke patients by the itraq-based lc-ms approach. *Journal of proteome research, 14*(6), 2466-2479. doi:10.1021/pr501324n

Shen, L. (2016). *Geneoverlap : An R package to test and visualize gene overlaps*.

Shu, L., Chan, K. H. K., Zhang, G., Huan, T., Kurt, Z., Zhao, Y., . . . Yang, X. (2017). Shared genetic regulatory networks for cardiovascular disease and type 2 diabetes in multiple populations of diverse ethnicities in the united states. *PLOS Genetics*, 13(9), e1007040. doi:10.1371/journal.pgen.1007040

Stamova, B., Ander, B. P., Jickling, G., Hamade, F., Durocher, M., Zhan, X., . . . Sharp, F. R. (2019). The intracerebral hemorrhage blood transcriptome in humans differs from the ischemic stroke and vascular risk factor control blood transcriptomes. *Journal of Cerebral Blood Flow and Metabolism, 39*(9), 1818-1835. doi:10.1177/0271678x18769513

Stamova, B., Jickling, G. C., Ander, B. P., Zhan, X., Liu, D., Turner, R., . . . Sharp, F. R. (2014). Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PloS One, 9*(7), e102550. doi:10.1371/journal.pone.0102550

Stamova, B., Xu, H., Jickling, G., Bushnell, C., Tian, Y., Ander, B. P., . . . Sharp, F. R. (2010). Gene expression profiling of blood for the prediction of ischemic stroke. *Stroke, 41*(10), 2171-2177. doi:doi:10.1161/STROKEAHA.110.588335

Stoll, G., Kleinschnitz, C., & Nieswandt, B. (2008). Molecular mechanisms of thrombus formation in ischemic stroke: Novel insights and targets for treatment. *Blood, 112*(9), 3555. doi:10.1182/blood-2008-04-144758

Tang, Y., Xu, H., Du, X., Lit, L., Walker, W., Lu, A., . . . Sharp, F. R. (2006). Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: A microarray study. *Journal of Cerebral Blood Flow and Metabolism, 26*(8), 1089-1102. doi:10.1038/sj.jcbfm.9600264

Tian, Y., Stamova, B., Jickling, G. C., Liu, D., Ander, B. P., Bushnell, C., . . . Sharp, F. R. (2012). Effects of gender on gene expression in the blood of ischemic stroke patients. *Journal of Cerebral Blood Flow and Metabolism, 32*(5), 780-791. doi:10.1038/jcbfm.2011.179

Tu, C., Rudnick, P. A., Martinez, M. Y., Cheek, K. L., Stein, S. E., Slebos, R. J. C., & Liebler, D. C. (2010). Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of Proteome Research, 9*(10), 4982-4991. doi:10.1021/pr100646w

Walter, W., Sánchez-Cabo, F., & Ricote, M. (2015). Goplot: An r package for visually combining expression data with functional analysis. *Bioinformatics, 31*(17), 2912-2914. doi:10.1093/bioinformatics/btv300

Wright, F. A., Lemon, W. J., Zhao, W. D., Sears, R., Zhuo, D., Wang, J.-P., . . . Yuan, B. (2001). A draft annotation and overview of the human genome. *Genome Biology, 2*(7), research0025.0021. doi:10.1186/gb-2001-2-7-research0025

Yilmaz, G., & Granger, D. N. (2008). Cell adhesion molecules and ischemic stroke. *Neurological Research, 30*(8), 783-793. doi:10.1179/174313208X341085

Yilmaz, G., & Granger, D. N. (2010). Leukocyte recruitment and ischemic brain injury. *Neuromolecular Medicine, 12*(2), 193-204. doi:10.1007/s12017-009-8074-1

Yun, S.-H., Sim, E.-H., Goh, R.-Y., Park, J.-I., & Han, J.-Y. (2016). Platelet activation: The mechanisms and potential biomarkers. *BioMed research international, 2016*, 9060143-9060143. doi:10.1155/2016/9060143

Zhao, X., Song, S., Sun, G., Strong, R., Zhang, J., Grotta, J. C., & Aronowski, J. (2009). Neuroprotective role of haptoglobin after intracerebral hemorrhage. *The Journal of Neuroscience, 29*(50), 15819-15827. doi:10.1523/JNEUROSCI.3776-09.2009

Žitňanová, I., Šiarnik, P., Kollár, B., Chomová, M., Pazderová, P., Andrezálová, L., . . . Turčáni, P. (2016). Oxidative stress markers and their dynamic changes in patients after acute ischemic stroke. *Oxidative Medicine and Cellular Longevity, 2016*, 9761697-9761697. doi:10.1155/2016/9761697

Supplementary Table 1. List of included transcriptomics studies.

Study	*n* stroke (m/f) per stroke type	*n* Control (m/f)	Sample Type	Platform	Time	*N* DE Genes	Statistical Criteria	Data Obtained From
Barr 2010	39 (17/22) IS	24 (10/14)	Whole Blood (Paxgene)	Illumina HumanRef-8 v2	10.06 Hours	9	t-test, >2 fold difference	Paper
Barr 2015	34 (14/20) IS	8 (5/3)	Whole Blood (Paxgene)	Illumina HumanRef-8 v2	24Hr - 48hr	3	t-test, >1.5 fold difference, FDR p <.05	Paper
Grond-Ginsbach 2008	20 (8/12) IS	15 (4/11)	Peripheral Blood Mononuclear cells	Affymetrix U133A GeneChips	24Hr	44	t-test, p < .001	Supplementary
Jickling 2012	94 (46/48) IS	44 (22/22)	Whole Blood (Paxgene)	Affymetrix U133 2 Plus GeneChips	<72 Hours	26	t-test, >1.2, FDR 0.5	Supplementary
Lee 2014	12 (7/5) IS	24 (13/11)	Whole Blood	IlluminaHT - 12 V3 expression bead chips	<24 Hours	166	t-test, >2 fold difference, p < .05	Supplementary
Meller 2016	17 (10/7) IS	28 (28/9)	Whole Blood (Paxgene)	Solid 5500XL	22.9 Hours	115	t-test, >2 fold difference	Supplementary
Moore 2005	19 (6/13) IS	10 (4/6)	Peripheral Blood Mononuclear cells	Affymetrix Human Genome U133A	<48 Hours	771	T-test, p < .05	Supplementary
Oh 2012	12 (7/5) IS	12 (7/5)	Whole Blood	Illumina HumanHT-12 v3	12.7 Hours	32	t-test, >2 fold difference, p < .05	Supplementary
Stamova 2010	70 (40/30) IS	17 (12/5)	Whole Blood (Paxgene)	Affymetrix Genome U133 Plus 2 GeneChips	<3 hours, <24 hours	36	PAM modelling	Supplementary
Stamova 2014	23 (12/11) IS	23 (12/11)	Whole Blood (Paxgene)	Affymetrix Genome U133 Plus 2 GeneChips	<3 hours, <5 hours, < 24hours	>1000	ANCOVA, >1.2 fold difference, p <.05	Supplementary

Tang 2006	15 (11/4) IS	8 (7/1)	Whole Blood (Paxgene)	Affymetrix Genome U133 Plus 2 GeneChips	<3 hours, <5 hours, <24hours	>1000	t-test, FDR < .05	Supplementary
Tian 2012	52 (28/24) IS	51 (27/24)	Whole Blood (Paxgene)	Affymetrix Genome U133 Plus 2 GeneChips	<3 hours, <5 hours, <24hours	>1000	t-test, >1.2 fold difference, FDR < .05	Supplementary
Sang 2017	4* (?) ICH	4* (?)	Peripheral Blood Mononuclear cells	Illumina HiSeq3000	<3 hours	4040	t-test, >1.5 fold difference, p <.05	Supplementary
Stamova 2019	33 (24/9) IS 33 (24/9) ICH	33 (24/9)	Whole Blood (Paxgene)	GeneChip HTA 2.0 Array	<24 hours, 24-48 hours, >48 hours	489	Mixed regression, FDR < 0.2, p < 0.005, fold difference >1.2	Supplementary

Note. * Sang 2017 analyzed a subgroup of samples but did not specify subgroup demographics for gender.

Supplementary Table 2. List of included proteomics studies.

Study	n stroke	n controls	Sample Type	Platform	Time	Depletion	*n* stroke	*n* controls	*n* Identified Proteins	*n* DE Proteins	Statistical Criteria	Data Obtained From
Bian 2014	7 (3/4)	8 (5/3)	Peripheral blood mononuclear cells	Waters Synapt G2S	<24 hours	-	7 (3/4)	8 (5/3)	-	428	Unclear	Supplementary
Kaysheva 2018	17 (13/4)	20 (10/10)	Plasma EDTA	Thermo Scientific Q Exactive	<24 Hours	MARS column	17 (13/4)	20 (10/10)	110	44	NSAF Criterion	Supplementary
George 2015	15 (10/5)	12 (6/6)	Serum	Waters NanoAcuity	<48 hours	-	15 (10/5)	12 (6/6)	141	3	T-test FDR	Paper
Penn 2018	20 (8/12)	20 (7/13)	Plasma EDTA	Analytical Technologies 4000 QTRAP	<10 hours	-	20 (8/12)	20 (7/13)	147	28	FDR <.20	Paper
Sharma 2015	20*	20*	Serum	LTQ Orbitrap	<24 Hours	MARS 14	20 (?)	20 (?)	329	60	1.5 FC	Paper
Cevik 2016	9*	9*	Platelets	Waters Synapt G1	<24 Hours	-	9 (?)	9 (?)	500	84	Mann-Whitney U	Supplementary

Note. * Sharma 2015 and Cevik 2016 analyzed a subgroup of samples but did not specify subgroup demographics for gender.

Supplementary Table 3. List of intersecting mRNAs and proteins for both reviewed ischemic and hemorrhagic stroke studies.

Gene Symbol	Full Gene Name	ENTREZID*
ORM1	orosomucoid 1	5004
S100A12	S100 calcium binding protein A12	6283
ZYX	zyxin	7791
F11R	F11 receptor	50848
RPN2	ribophorin II	6185
SH3BGRL3	SH3 domain binding glutamate rich protein like 3	83442
MDH1	malate dehydrogenase 1	4190
MAPRE2	microtubule associated protein RP/EB family member 2	10982
CAPZA1	capping actin protein of muscle Z-line subunit alpha 1	829
BASP1	brain abundant membrane attached signal protein 1	10409
CCL5	C-C motif chemokine ligand 5	6352
COTL1	coactosin like F-actin binding protein 1	23406
S100P	S100 calcium binding protein P	6286
HBA1	hemoglobin subunit alpha 1	3039
PTGS1	prostaglandin-endoperoxide synthase 1	5742
CD36	CD36 molecule	948
S100A9	S100 calcium binding protein A9	6280
H2AFY	H2A histone family member Y	9555
PGAM1	phosphoglycerate mutase 1	5223
ANXA2	annexin A2	302
CAT	catalase	847
VIM	vimentin	7431
H3F3A	H3 histone family member 3A	3020
SERPINB1	serpin family B member 1	1992
IDH1	isocitrate dehydrogenase (NADP(+)) 1, cytosolic	3417
S100A8	S100 calcium binding protein A8	6279
ITGB2	integrin subunit beta 2	3689
F5	coagulation factor V	2153
SQOR	sulfide quinone oxidoreductase	58472
SH3BGRL	SH3 domain binding glutamate rich protein like	6451
GNAI3	G protein subunit alpha i3	2773
MNDA	myeloid cell nuclear differentiation antigen	4332
CDA	cytidine deaminase	978
SLC2A3	solute carrier family 2 member 3	6515
PECAM1	platelet and endothelial cell adhesion molecule 1	5175
TALDO1	transaldolase 1	6888
PDIA5	protein disulfide isomerase family A member 5	10954
HSD17B4	hydroxysteroid 17-beta dehydrogenase 4	3295
STX11	syntaxin 11	8676
PLEK	pleckstrin	5341
PGD	phosphogluconate dehydrogenase	5226

PGK1	phosphoglycerate kinase 1	5230
LCP1	lymphocyte cytosolic protein 1	3936
FKBP1A	FK506 binding protein 1A	2280
CANX	calnexin	821
SNAP23	synaptosome associated protein 23	8773
PRDX3	peroxiredoxin 3	10935
SERPINA1	serpin family A member 1	5265
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	471
KPNB1	karyopherin subunit beta 1	3837
TUBB	tubulin beta class I	203068
SEPTIN6	septin 6	23157
ENO2	enolase 2	2026
THBS1	thrombospondin 1	7057
MYL9	myosin light chain 9	10398
HIST1H2BO	histone cluster 1 H2B family member o	8348
F12	coagulation factor XII	2161
PROZ	protein Z, vitamin K dependent plasma glycoprotein	8858
TREML1	triggering receptor expressed on myeloid cells like 1	340205
LCN2	lipocalin 2	3934
ITGA2B	integrin subunit alpha 2b	3674
HBB	hemoglobin subunit beta	3043
PROS1	protein S	5627
MGLL	monoglyceride lipase	11343
CAMP	cathelicidin antimicrobial peptide	820
PTPN6	protein tyrosine phosphatase, non-receptor type 6	5777
RAB32	RAB32, member RAS oncogene family	10981
VWF	von Willebrand factor	7450
GPX1	glutathione peroxidase 1	2876
MANF	mesencephalic astrocyte derived neurotrophic factor	7873
HIST1H2BD	histone cluster 1 H2B family member d	3017
ALDOA	aldolase, fructose-bisphosphate A	226
STXBP2	syntaxin binding protein 2	6813
SPARC	secreted protein acidic and cysteine rich	6678
LRG1	leucine rich alpha-2-glycoprotein 1	116844
VASP	vasodilator stimulated phosphoprotein	7408
CSTB	cystatin B	1476
HIST1H1C	histone cluster 1 H1 family member c	3006
EHD1	EH domain containing 1	10938
PPBP	pro-platelet basic protein	5473
PKM	pyruvate kinase M1/2	5315
TUBA1A	tubulin alpha 1a	7846
CLU	clusterin	1191
LIMS1	LIM zinc finger domain containing 1	3987
HSPA6	heat shock protein family A (Hsp70) member 6	3310

GSTO1	glutathione S-transferase omega 1	9446
RHOG	ras homolog family member G	391
PNP	purine nucleoside phosphorylase	4860
CLIC1	chloride intracellular channel 1	1192
TAGLN2	transgelin 2	8407
ANXA3	annexin A3	306
PF4	platelet factor 4	5196
TPM4	tropomyosin 4	7171
LSP1	lymphocyte specific protein 1	4046
PDIA3	protein disulfide isomerase family A member 3	2923
F13A1	coagulation factor XIII A chain	2162
ICAM1	intercellular adhesion molecule 1	3383
TKT	transketolase	7086
ARF1	ADP ribosylation factor 1	375
GDI1	GDP dissociation inhibitor 1	2664
ACTB	actin beta	60
SH3BGRL2	SH3 domain binding glutamate rich protein like 2	83699
YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon	7531
WDR1	WD repeat domain 1	9948
PDIA6	protein disulfide isomerase family A member 6	10130
PSAP	prosaposin	5660
PDLIM1	PDZ and LIM domain 1	9124
NAPA	NSF attachment protein alpha	8775
ANXA1	annexin A1	301
TXN	thioredoxin	7295
RAC1	Rac family small GTPase 1	5879
SRC	SRC proto-oncogene, non-receptor tyrosine kinase	6714
LGALS1	galectin 1	3956
LMNB1	lamin B1	4001
ORM2	orosomucoid 2	5005
CNN2	calponin 2	1265
NME1	NME/NM23 nucleoside diphosphate kinase 1	4830
RAB5C	RAB5C, member RAS oncogene family	5878
GPX4	glutathione peroxidase 4	2879
ARPC1B	actin related protein 2/3 complex subunit 1B	10095
TUBA4A	tubulin alpha 4a	7277
EHD3	EH domain containing 3	30845
ACTG1	actin gamma 1	71
TYMP	thymidine phosphorylase	1890
ARPC3	actin related protein 2/3 complex subunit 3	10094
FHL1	four and a half LIM domains 1	2273
PRDX5	peroxiredoxin 5	25824
RHOA	ras homolog family member A	387
VDAC1	voltage dependent anion channel 1	7416

TWF2	twinfilin actin binding protein 2	11344
CFL1	cofilin 1	1072
GSN	gelsolin	2934
SAR1A	secretion associated Ras related GTPase 1A	56681
ARF5	ADP ribosylation factor 5	381
ALOX12	arachidonate 12-lipoxygenase, 12S type	239
GRB2	growth factor receptor bound protein 2	2885
CLIC4	chloride intracellular channel 4	25932
VAPA	VAMP associated protein A	9218
GNAI2	G protein subunit alpha i2	2771
CDC37	cell division cycle 37	11140
TUBA1B	tubulin alpha 1b	10376
CAPNS1	calpain small subunit 1	826
RAB7A	RAB7A, member RAS oncogene family	7879
PAFAH1B2	platelet activating factor acetylhydrolase 1b catalytic subunit 2	5049
MAPRE1	microtubule associated protein RP/EB family member 1	22919
GSTP1	glutathione S-transferase pi 1	2950
RGS10	regulator of G protein signaling 10	6001
SLC25A3	solute carrier family 25 member 3	5250
CDC42	cell division cycle 42	998
ARPC2	actin related protein 2/3 complex subunit 2	10109
G6PD	glucose-6-phosphate dehydrogenase	2539
RAB1B	RAB1B, member RAS oncogene family	81876
RAN	RAN, member RAS oncogene family	5901
DBH	dopamine beta-hydroxylase	1621
FCGBP	Fc fragment of IgG binding protein	8857
PITPNM3	PITPNM family member 3	83394
ATP5F1B	ATP synthase F1 subunit beta	506
CD47	CD47 molecule	961
GNAQ	G protein subunit alpha q	2776
H2AFZ	H2A histone family member Z	3015
ILK	integrin linked kinase	3611
TPM2	tropomyosin 2	7169
VDAC3	voltage dependent anion channel 3	7419
TXNDC5	thioredoxin domain containing 5	81567
ARPC5	actin related protein 2/3 complex subunit 5	10092
B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2	10678
CAPZA2	capping actin protein of muscle Z-line subunit alpha 2	830
GANAB	glucosidase II alpha subunit	23193
RSU1	Ras suppressor protein 1	6251
SACM1L	SAC1 like phosphatidylinositide phosphatase	22908
TBC1D1	TBC1 domain family member 1	23216
VCL	vinculin	7414

ACTR3	ARP3 actin related protein 3 homolog	10096
TUBB1	tubulin beta 1 class VI	81027
PRDX6	peroxiredoxin 6	9588
PRL	prolactin	5617
RAB8B	RAB8B, member RAS oncogene family	51762
MYH10	myosin heavy chain 10	4628

*Note. The ENTREZID is the accession number utilized by the National Centre for Biotechnology Information.

Chapter 4

Methodological Issues and Considerations

Preface

Chapter 4 will detail the methodologies utilized in the experimental series, covering sample selection, protocols of blood storage, considerations for proteomics analyses and bioinformatics. The additional details within this chapter reflects the additional methodological considerations and choices before the commencement of each experiment that may not be discussed elsewhere.

Chapter Overview

Each experimental study of this thesis will detail the specific methodology used within the chapter as if written as a manuscript for publication. However, this necessarily means there are other background issues and considerations for sample selection and methodologies used in the experimental series that will not be detailed there, and thus require more detailed and theoretical explanation – hence the focus of this chapter.

The START Cohort

For the experimental series of this thesis, we have selected a longitudinal cohort of post-stroke patients from an existing cohort, the **ST**roke im**A**ging p**R**evention and **T**reatment (START) longitudinal stroke cohort (*n* = 219). The START cohort comprised two sub-cohorts that have been drawn from the large multi-centre cohort: i.e. **EX**tending the time for **T**hrombolysis in **E**mergency **N**eurological **D**eficits (START_EXTEND) (Ma et al., 2011) and the **PRE**diction and **P**revention to **A**chieve Optimal **R**ecovery **E**ndpoints after stroke (START_PrePARE) cohorts (Carey et al., 2013). Ethics for this project has been obtained from from La Trobe University (HEC10-071, Appendix 5) and Austin Hospital (H2010/03588). These cohort studies included multimodal investigations with the overall aim to characterise and better understand the biological changes underlying the longitudinal profile of clinical outcomes evolving over the first year post-stroke. A particular aim was to identify the blood based biomarkers and brain networks involved with the development of post-stroke depression (PSD) and recovery over time (Carey et al., 2013). For this purpose, multiple blood samples were taken in conjunction with a battery of clinical examinations at different timepoints within the first year of stroke, including demographics, diet, functional recovery, depressive symptoms and cognitive function. The START cohort is one of the few cohort studies with a longitudinal approach utilizing a wide range of assessments and

represents the best case for a prospective study of stroke recovery.

Study Participants

The inclusion criteria for the combined START cohort was separate for the START_EXTEND and START_PrePARE studies. The START_PrePARE study recruited patients that were aged ≥18 years with clinical signs of hemispheric infarction (Carey et al., 2013). Additionally, due START_EXTEND being an investigation of thrombolytic agent usage in acute stroke, the study had stricter inclusion criteria that included patients with an NIHSS score of ≥4–26 receiving thrombolytic treatment from 3-9.5 hours post stroke onset (Ma et al., 2011). The exclusion criteria for START_PrePARE included patients that present with contraindication to magnetic resonance imaging (MRI) imaging studies and evidence of previous disability (Carey et al., 2013). Further exclusion criteria for START_PrePARE included use of heparin within the previous 48 hours and use of glycoprotein IIb–IIIa inhibitors such as abciximab and eptifibatide within the previous 72 hours (Ma et al., 2011). Both studies had exclusion criteria for hemorrhagic stroke presentation. Overall, these studies aimed to sample the typical ischemic stroke population of elderly adults (Cadilhac et al., 2017). Table 1 details the baseline patient characteristics of the combined START_PrePARE and START_EXTEND cohorts.

Table 1

Baseline Stroke Patient Characteristics for the Total START Cohort (n = 219)

	M	*SD*	Median	*IQR*
Age (years)	69.85	13.20	72.37	13.54
NIHSS (baseline)	7.99	6.74	6.00	9.00
Heart Rate (per minute)	76.06	13.83	75.00	18.00
Systolic Blood Pressure (mm Hg)	144.87	23.04	142.00	26.50
Diastolic Blood Pressure (mm Hg)	79.37	12.97	78.00	16.00
3 Month				
NIHSS	2.54	4.90	2.00	2.00
mRS*	1.74	1.56	1.00	1.00
MoCA	24.61	5.92	26.00	5.75
MADRS	7.40	7.56	5.00	10.00
	Frequency	Percentage		
Death	21	9.59		
Lost to Follow-up	8	3.65		
Withdrawal	2	0.91		
Past Atrial Fibrillation	50	22.83		
Hypertension	124	56.62		
Lipid Disorder	101	46.12		
Ischemic Heart Disease	50	22.83		
Diabetes Mellitus	42	19.18		

Study Outcomes and Measures

The START cohort study utilized a wide range of assessments that aimed to understand the trajectory of post-stroke recovery and the biological correlates of post-stroke outcomes (Carey et al., 2013). Clinical recovery assessments were conducted at 3-7 days, 3 months and 12 months post-stroke, with bloods collected from patients at all available timepoints (see Table 2). As the aims of this thesis are to examine the biological correlates of stroke recovery, we examined the post-stroke clinical assessments from the START Cohort that measured stroke severity, functional recovery, cognition and depression symptoms. The National Institute of Health Stroke Scale (NIHSS) is a bedside measurement of stroke severity, acting as a composite of several domains including motor function, language, sensation and visual fields and gaze, where higher scores indicate an increased degree of neurological damage (Jauch et al., 2006). While the NIHSS only provides a rough estimate of

stroke severity (Hand, Page, & White, 2014; Jauch et al., 2006) compared to more comprehensive tests of each subdomain, it is still the most commonly used bedside measure of stroke severity in medical practice and research due to its ease of use (Lyden, 2017). Functional recovery post-stroke is assessed using the modified Rankin Scale (mRS), a broad measure of functioning in everyday activities and is most commonly used at longitudinal follow-up sessions to assess recovery (Banks & Marotta, 2007). The pre-stroke mRS is also available in the acute setting to determine pre-stroke disability (Quinn et al., 2017). Although the mRS has similar disadvantages to the NIHSS in terms of comprehensiveness of assessed domains and subjective determination by the examiner (Broderick, Adeoye, & Elm, 2017), its widespread usage in clinical stroke and research is important to consider for generalization of results to current practice. The Montreal Cognitive Asessment (MoCA) is a rapid measure of cognition that sees general use in stroke research alongside generalized measures of cognition such as the Mini-Mental State Examiantion (MMSE) (Chiti & Pantoni, 2014). Compared to the MMSE, the MoCA is a more sensitive predictor of post-stroke cognitive impairment if assessed in the acute phase and also has good correlation with more extensive neuropsychological batteries at the post-stroke phase (Chiti & Pantoni, 2014). The Montgomery–Åsberg Depression Rating Scale (MADRS) is a measure that was originally designed to measure depression episode severity in primary mood disorders (Montgomery & Åsberg, 1979). For the purposes of adapting the measure to stroke patients, the START cohort study utilized the structured interview format for the MADRS that could be delivered to aphasic stroke survivors (Williams & Kobak, 2008). Overall, the MADRS has been shown to have good reliability and symptom discrimination when compared with other measures of clinical depression in stroke research such as the Hospital Anxiety and Depression Scale (HADS) (Sagen et al., 2009) and the Beck Depression Inventory (BDI) (Kang et al., 2013).

Table 2

Schedule of Assessments for the START cohort.

Study outcome	Investigation	Baseline	12–24 h	3-7 Days	3 Months	12 Months
Patient Demographics	Age, gender Stroke History	X		X		
Clinical examinations	National Institute of Health Stroke Scale (NIHSS)	X	X	X	X	X
	Modified Rankin Scale (mRS)	Pre-stroke			X	X
	Physical examination	X			X	X
	Barthel Index				X	X
	Physical risk factors interview			X	X	X
	Montreal Cognitive Assessment (MoCA)			X	X	X
	Quality of life – Stroke Impact Scale (SIS)				X	X
Depression	Depression – Montgomery Asberg Depression Rating Scale (MADRS)			X	X	X
	Depression history and interview questions			X	X	X
Diet and lifestyle questionnaires	Cancer Council of Victoria (CCV) diet questionnaire (part)			X	X	X
	Rapid Assessment of Physical Activity (RAPA) questionnaire			X	X	X
Routine laboratory blood tests	Hematology, lipid profile, electrolytes, Complete Blood Count	X		X	X	X
Research bloods	Vascular, inflammatory, dietary	X	X	X	X	X

Overview of Data Samples Selected based on Study Aims

Details of the sample selection process for each experimental chapter is provided below.

The aim of Chapter 5, was to investigate the predictive ability of routine laboratory blood tests in relation to post-stroke clinical outcomes, including: neurological status, assessed using the NIHSS (Lyden et al., 2009); depressive symptoms based on the MADRS

(Williams & Kobak, 2008); cognitive function using the Montreal Cognitive Assessment (MoCA) (McLennan, Mathias, Brennan, & Stewart, 2011); and functional recovery based on the modified Rankin Scale (Banks & Marotta, 2007). Consistent with the protocol for the START cohort studies (Carey et al., 2013), routine laboratory blood tests and additional bloods were taken for later storage on hospital arrival, within 12-24 hours and 3-7 days, 3 months and 12 months post-stroke. This provided a unique opportunity to examine a large range of biomarkers such as white blood cell and neutrophil count that are commonly used in healthcare. As such, Chapter 5 utilized a smaller subset of the START cohort that had available data and complete data on routine blood biomarkers across the first week of stroke and clinical outcomes at 3 and 12 months (n = 154).

The aim of the second experimental study described in Chapter 6 was to better understand the molecular biological systems associated with the longitudinal profile of stroke recovery. This was achieved by an exploratory correlational proteomics approach in relation to post-stroke depression symptoms (Nguyen et al., 2016). As the analysis for this study was conducted during the timeframe that the START and START_PrePARE cohort data collection was ongoing, the sample bloods chosen for this study were the available plasma samples at 3 months post-stroke with MADRS scores (n = 44).

The third experimental study, described in Chapter 7, aimed to extend the information gained from the findings in Chapter 6 and further elaborate understanding of changes in biomarkers and biosystems involved over the first year of natural stroke recovery. Our focus here was on biomarkers of depression and associations over the first-year post-stroke. We therefore chose to conduct a pilot longitudinal proteomics analysis comparing differing timepoints at 3-7 days, 3 months and 12 months post-stroke. From the results of our pilot proteomics study (described in Chapter 6), we reviewed the complete START cohort in order

to conduct a sample selection process that would have the maximum chance of detecting biomarkers associated with changes in MADRS defined depression symptomology (Fantino & Moore, 2009), over the three timepoints using a 2-way analysis of variance (ANOVA). This included reassigning our selected stroke population on the basis of 3 month post-stroke MADRS scores into experimental groups, including no depressive symptoms (MADRS score 0-2), some depressive symptoms (MADRS score 4-7), and depressive symptoms warranting treatment (severe) (MADRS score 14+) based on the 3 month post-stroke time-point. These clinically defined groups were selected to maximise separation between groups and were informed by recommended guidelines for interpretation of MADRS scores (Montgomery & Åsberg, 1979; Sagen et al., 2009). An *a priori* power analysis for a 3x3 repeated measures design was conducted and found that 69 samples were needed per cell to detect a large effect size (Cohen's $f = 0.40$) at $\alpha = 0.05$ (Salkind, 2010). Within the START cohort, 107 patients had complete blood samples available at 3-7 days, 3 months and 12 months post-stroke, together with complete MADRS scores. Hence, a fully powered analysis was not possible. We therefore chose to conduct a pilot study involving subgroups with distinct groupings of MADRS scores and equal number of samples in each subgroup. Only 20 participants presented with severe depressive symptoms, thus defining the size of the subgroups. The selection process involved ranking the available sample of $n = 107$ according to MADRS scores then randomly selecting 20 stroke patients whose scores were within each of the 3 MADRS score depression bracket. This yielded a sample size of ($n = 60$) across 3 timepoints, with equal cell sizes for each depression category of 'no presence of depressive symptoms' ($n = 20$), 'mild depressive symptoms' ($n = 20$), and 'depressive symptoms warranting treatment' ($n = 20$).

Blood Samples and Analysis

Blood is measured routinely in emergency room clinical practice and hence is readily available. In any living body the vascular system ensures that blood perfuses all bodily tissue and is the medium for carrying essential nutrients, chemical signalling molecules and waste products to and from cells. Hence blood tests are collected routinely in Emergency Room clinical practice as the fastest expression of ongoing human biophysiology (Dayon & Kussmann, 2013) while simultaneously making an incredibly rich diverse line of cell types, genes, proteins, metabolites and plasma available for research. This thesis has utilized several blood based methodologies within this experimental series to examine the contents of blood for biomarkers and biosystems functioning.

Collection

Whole blood is a solution containing platelets, white and red blood cells suspended in fluid plasma with each element requiring specialized equipment, workflows and storage protocols. Current phlebotomy practices state that blood is most commonly collected via vacuum sealed storage containers or open systems such as needles and syringes (World Health Organization, 2010). Centrifugal force is typically applied to separate blood into its constituent compartments based on mass. From heaviest to lightest, this includes packed red blood cells or erythrocytes, a buffy coat layer containing leukocytes and a top layer of plasma (Figure 1) (Basu & Kulkarni, 2014). Serum is the liquid compartment after centrifuging with the addition of silica particles that encourage the blood to coagulate and is commonly used to analyse lipids (Nigam, 2011) and glucose levels (Kim, 2016). Conversely, the addition of anticoagulants such as ethylenediaminetetraacetic acid (EDTA), heparin and citrate halts the normal coagulation processes and results in liquid plasma where coagulation factors are mostly preserved (Bell, 2000). Citrate anticoagulated blood is often used for assessing

prothrombin times and EDTA plasma is useful for a wide range of applications including the determination of specific cell counts (Zini, 2014).

Within the START protocol, blood was collected by venepuncture into Becton Dickinson (BD) EDTA, serum, heparin and citrate tubes at admission, within 12-24 hours, 3-7 days, 3 months and 12 months. Samples were centrifuged at 1100-1300 *g* at room temperature, the resulting supernatant aliquoted into 1ml aliquots and cryogenically stored for later analysis.

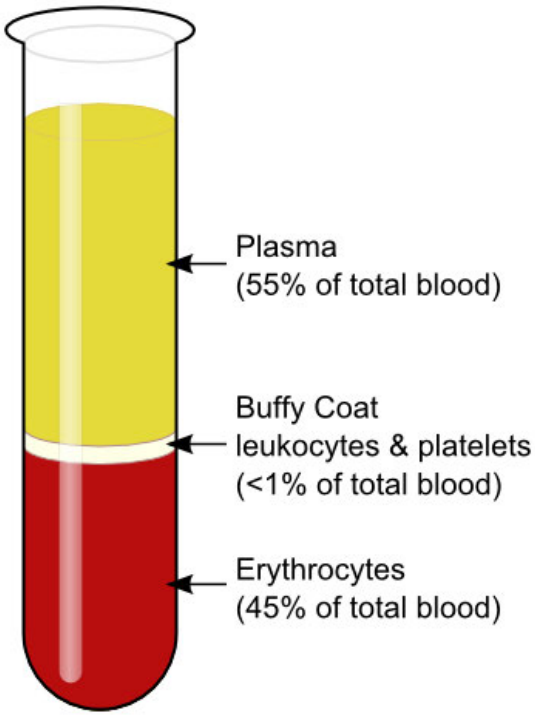

Figure 1. A diagram detailing the different compartments of blood after centrifuge. Due to differences in mass, red blood cells are packed at the bottom, leukocytes form a buffy coat layer with the top layer consisting of liquid plasma.

Storage

Time to beginning of analysis of any blood is time-critical, especially given the ongoing concern for stability of proteins (Dayon & Kussmann, 2013) and cellular

components (Zini, 2014). This necessitates a logistical workflow that can facilitate rapid analysis of blood or immediate long-term storage. This is especially important to consider in the hospital setting as turnaround times and specimen viability are important factors in the point of care (Chung, Lee, Chun, Park, & Min, 2009). In the case of lengthy cohort studies recruiting over many sites, many studies opt to separate and store blood compartments at the time of blood draw to maintain a biobank for future analyses (Mohamadkhani & Poustchi, 2015). Recommendations for storage of liquid serum or plasma state that centrifugal separation should occur within 2 hours and be freeze dried in liquid nitrogen cyro-stable plastic tubes before transferring into -80˚ C freezer (Mohamadkhani & Poustchi, 2015). This procedure ensures maximum stability for biomolecules such as metabolites and proteins (Breier et al., 2014). Long-term storage of leukocytes can be achieved with the addition of dimethyl sulfoxide ($DMSO_4$) and pentastarch to ensure cellular viability after thawing (Stroncek et al., 2011). If RNA or DNA extraction is intended for cryopreserved cells, cells can retain stable gene expression even after 5 years of storage in liquid nitrogen at -140˚ C (Stroncek et al., 2011).

Routine Laboratory Assessments

Routine standard care blood tests describe a set of patient variables that are taken as a clinically reproducible investigative measure in the clinical workflow (Janssen, van Hoeven, & Rautmann, 2015). Within Australian hospitals, laboratory blood tests are typically outsourced to commercial pathology services with median turnaround times of less than 2 hours for all commonly requested measures (Vecellio et al., 2015). Commercial laboratories have the advantage of accessing large scale platforms called hematology analyzers that are purpose built for high throughput and automation, often combining multiple analytical modalities on the same machine (Davis & Barnes, 2012).

The assessment of white blood cell counts (WBC) and neutrophils obtained from hospital data, as reported in Chapter 5, form the basic measurements of the full blood examination (FBE) or complete blood count (CBC). As this test is a routine medical procedure with different hospitals engaging in separate pathology services, it is difficult to unify the methodologies used. It is assumed that this test was conducted on medical analyzers using flow cytometry as the industry standard (Sullivan, 2006), although it is also acknowledged that other methods such as microscopic review (Vis & Huisman, 2016) and Coulter counting (Jog, 2013) remain relevant in hematology analysis. For the purposes of cell counting, flow cytometry can achieve a rate of up to thousands of cells per second, capable of providing leukocyte and red blood cell counts within minutes (Craig & Foon, 2008). Flow cytometry measures the physical and chemical characteristics of cells via optical and fluorescence modalities. Cells are drawn from a source into a stream of isotonic fluid that creates laminar flow of sequential single cells towards a laser. The light emitted from the interrogation point of the cell and laser is collected by optics and a series of filters for specific wavelengths (Brown & Wittwer, 2000). Although not utilized in the present thesis, flow cytometry in combination with fluorescence activated cell sorting can further provide further quantitative counts on differing WBC populations such as eosinophils, basophils, lymphocytes and monocytes (Roussel, Benard, Ly-Sunnaram, & Fest, 2010). This extended level of WBC differentiation is useful for determining characteristics such as the neutrophil-to-lymphocyte ratio, and is accepted as an established marker of systematic inflammation (Xue et al., 2017; Zhao et al., 2016).

Proteomics

Proteomics is a field of study combining large scale investigations of protein expression and structure with bioinformatics approaches (Aslam, Basit, Nisar, Rasool, &

Khurshid, 2017). Proteins are large complex molecules that are the final expressional endpoint of the genome and are involved in cellular composition, cell-to-cell signalling, transportation and enzymatic activities (Freeman & Hemby, 2004). Studies of proteins are more structurally and functionally related to phenotypic expressions of cells than genes and therefore proteomics studies can be interpreted as a closer definition of disease states that can be readily translated when compared to genomics studies (Graves & Haystead, 2002). Although there are many methods available to detect and quantify single proteins such as immunoblotting and enzyme linked immunosorbent assays, proteomic methodologies consist of a relatively new set of techniques with complex considerations for each component of the workflow. Proteomics workflows universally include steps to isolate and separate proteins or peptides from a complex sample, identify and quantify analytes and subsequent bioinformatics analysis. An advantage of proteomics is that it is theoretically able to assess a large number of proteins within a single sample when compared to transcriptomics (Manzoni et al., 2016). The addition of bioinformatic algorithms and associated methodologies allows the organization and explanation of these molecules in the context of biological pathways. Considering the experimental series, these systems biology methodologies are the only approaches that can address the collective aims of Chapters 3, 6 and 7 to identify the biological systems involved in stroke recovery from blood samples. This section will address the past and current proteomics methodologies, with a focus on proteomics analysis of blood plasma.

Gel Based Proteomics

Some of the earliest techniques that were aimed at separating and addressing whole protein expression were gel based methodologies such as 2-dimensional–difference in gel electrophoresis (2D-DIGE) and 2-dimensional–polyacrylamide gel electrophoresis (2D-

PAGE). Gel based electrophoresis techniques are also used in Western Blotting methods to assess proteins based on the single physiochemical properties of proteins isoelectric point or molecular weight (MW) to facilitate the identification and quantification of proteins. The 2-dimensional aspect of SD-DIGE and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is based on the combination of techniques utilizing an immobilized pH gradient to separate proteins based on isoelectric point and MW simultaneously on a 2D plane (Meleady, 2018). Samples can be differentiated by labelling with the addition of cyanine or fluorescent dyes. This produces a planar map of protein spots, whereby each separated spot can be identified as a protein and the intensity of the staining provides quantification relative to other proteins (Figure 2) (Carrette, Burkhard, Sanchez, & Hochstrasser, 2006). Image processing software is often used to resolve protein spots based on the axis coordinates and identification of proteins can be achieved by automated cross-referencing on protein databases such as Mascot (Braun & Senkler, 2012).

An advantage of 2D gel techniques is that they can detect miniscule changes in protein structure and can further examine post-translational modifications by detecting small shifts in the isoelectric point (Meleady, 2018). However, there are issues that are inherent to 2D-DIGE techniques that are prohibitive in assessing many blood samples in a timely matter. Severely limited throughput and by extension, gel related batch effects, are a notable limitation to conducting gel based proteomics on a large number of samples (Strohkamp, Gemoll, & Habermann, 2016). Although pooling samples can address this issue as well as reduce non-specific background noise (Weinkauf, Hiddemann, & Dreyling, 2006), it is often difficult to justify sample pooling unless there are clinically distinct groups to analyse. Another limitation associated with gel proteomics is the resolution of proteins when protein spots are found to visually overlap (Meleady, 2018). Although this issue exists in other

proteomics methods, there is a pronounced lack of sensitivity and resolution for low-abundance proteins even when compared to some mass-spectrometry (MS) techniques (Meleady, 2018). Overall, 2D electrophoresis techniques may be suitable for preclinical and low sample experiments but requires additional attention to reconcile issues with resolution. Considering these limitations, the author has chosen not to utilize 2D electrophoresis techniques due to the analytical requirements for resolution of many samples.

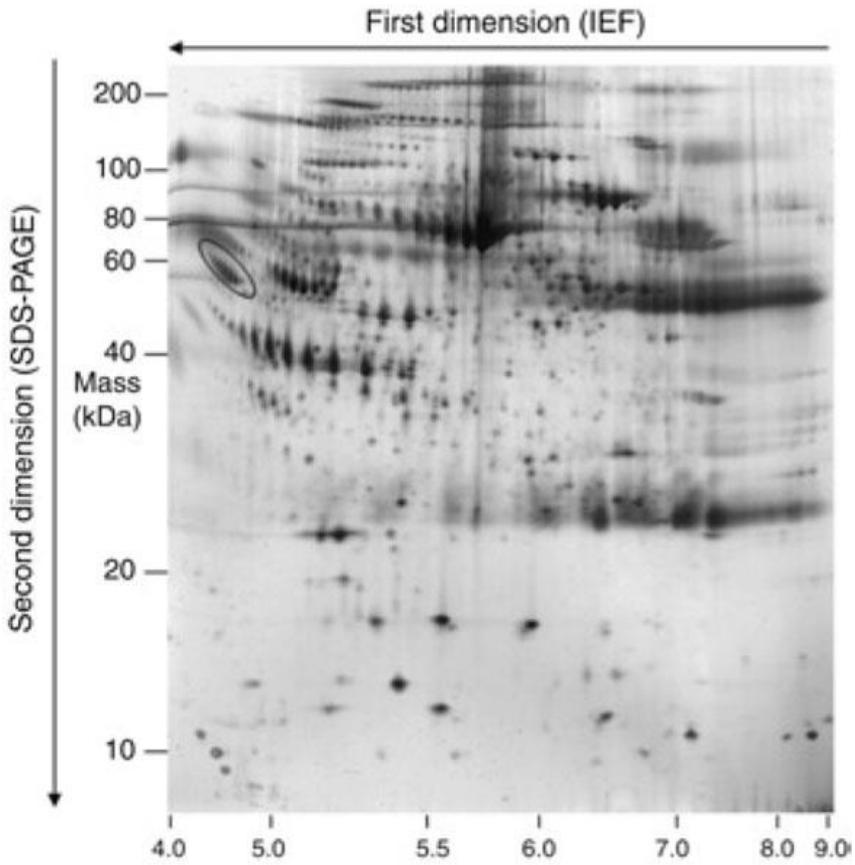

Figure 2. Whole cell proteome human plasma on sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and stained with sensitive silver. Reprinted by permission from the Licensor: Springer Nature, *Nature Protocols*, State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research, Carette, O., Burkhard, P. R., Sanchez, J., & Hochstrasser, D. F. COPYRIGHT, (2006).

Mass Spectrometry Proteomics

Mass spectrometry (MS) is a powerful analytical platform that has been utilized in resolving protein expression since the development of ionization techniques that could vaporize large complex molecules such as proteins in 1981 (Finehout & Lee, 2004). The MS proteomics workflow can address a wide range of sample types from plasma, cerebrospinal fluid, urine to whole cells and cellular components (Budnik, Levy, Harmange, & Slavov, 2018). Mass spectrometers are currently the most popular platform for proteomics, with high throughput and instrumentation detection range having the greatest advantages in discovery-based applications (Aslam et al., 2017). This section will detail the methodological considerations for MS based proteomics within the experimental Chapters 6 and 7 in this thesis.

Considerations for Plasma

Plasma is an extremely rich and complex sample, responsible for the transportation of tens of thousands of distinct protein species. As stated above, human blood samples are usually assessed utilizing automated workflows that address simple cell counts or single biomarkers. As the present thesis aims to comment on entire biosystems utilizing blood samples, the issues with proteomics analysis of blood must first be addressed.

A collaborative effort using liquid chromatography mass spectrometry (LC-MS) techniques in plasma profiling has found over 9000 proteins identified with one or more peptides, with 889 proteins identified after statistical adjustment for multiple hypotheses at 95% CI (States et al., 2006). However, due to the extremely wide dynamic range for protein abundance and high proteome complexity for plasma, more common LC-MS proteomics methods will report closer to 200 identified proteins (Tu et al., 2010). In case of MS based proteomics, dynamic range refers to the quantitative differences between the relative

concentrations least and most abundant molecules in the present sample. In plasma, this issue can be illustrated by the difference in the relatively high abundance human serum albumin (HSA) (50 mg/mL) compared to the relatively low abundance interleukin-6 (IL-6) (5 pg/mL), demonstrating a concentration range of over 10^{10} fold (Geyer, Holdt, Teupser, & Mann, 2017). Similar to visual protein spot resolution issues in 2D-DIGE proteomics, high abundance proteins present with more intense peak signals and can therefore overlap and obfuscate peaks of lower abundance molecules (Dayon & Kussmann, 2013). This leads to a greatly underestimated proteome coverage of plasma samples (Hortin & Sviridov, 2010). Several *a priori* considerations are available at multiple stages of the proteomics workflow to address plasma dynamic range.

Sample Preparation

Sample preparation for plasma includes preliminary steps to extract plasma samples from storage and isolate proteins for instrumentation. The plasma proteome is overall resilient to factors that would otherwise influence large shifts in mRNA or metabolite expression such as time from blood collection, storage temperature and freeze-thaw cycles (Zimmerman, Li, Yarbrough, Slebos, & Liebler, 2012). Regardless, the nature of clinical cohort often precludes that plasma collected for exploratory and novel biomarker studies are conducted after the completion of the study. The proteomics facility contracted throughout this experimental series requested plasma amounts of 10µl. To avoid additional freeze-thaw cycles, frozen plasma aliquots were carefully cut on a bed of dry ice and transferred to microtubes for transport, using sharp metal apparatus (Figure 3).

Figure 3. A collection of sharp metal micro-spatulas that were used to cut and extract frozen plasma samples.

Bottom-up proteomics involves steps to digest proteins into peptides. Pre-treatment of many sample types intended for MS proteomics includes a high concentration of urea for stabilization or denaturation (Chen, Cociorva, Norris, & Yates, 2007). Urea is a molecule that is useful for disrupting hydrogen bonds in intraprotein interactions and enhances digestion of proteins. Trypsin is the most used enzyme for in-sample digestion of proteins, contributing to usage in over 90% of published proteomics datasets (Tsiatsiani & Heck, 2015). Trypsin enzymatically cleaves proteins into peptides with a high level of specificity at the C-terminal between lysine and arginine residues, except when followed by proline (Switzar, Giera, & Niessen, 2013). This leads to better resolution for MS fragmentation spectra and confident peptide identification in peptide database searches that are often built on trypsin proteomics (Switzar et al., 2013). Although trypsin is often used in bottom-up up proteomics, other endoproteases such as Lys-C, Glu-C, Asp-N have specific considerations for stricter specificity for protein cleavage sites and protein mass ranges that may be useful after more generalized MS-based trypsin proteomics is conducted (Zhang, 2015).

Liquid Chromatography and Mass Spectrometry

To analyse digested peptide samples, an *apriori* separation step is required to ensure an orderly rate of flow and that the peptide mixture does not overwhelm the detection instrument. Liquid chromatography is currently the most widely used separation method in

proteomics laboratories and separates samples based on the chemical properties of the solvents in the mobile phase and solid beads in the stationary phase, with the physical properties of pressure and temperature (Karpievitch, Polpitiya, Anderson, Smith, & Dabney, 2010). As such, there is a large amount of variation available in selecting the chemical and physical components of the separation setup that contribute greatly to the resolution of MS spectra.

For the analysis of plasma, our proteomics studies utilized a gradient mobile phase of formic acid and acetonitrile, with a stationary phase using a C_{18} PepMap 300 µm trap column (Thermo-Fisher Scientific) and BioSphere C_{18} 75 µm analytical column (NanoSeparation). Acetonitrile is an organic solvent that dissolves a wide range of ionic and nonpolar components such as trypsin cleaved peptides and is commonly used as a mobile phase liquid chromatography (Fritz, Ruth, & Kragl, 2009). Trapping columns provide sample clean-up and retention of unwanted peptides, before eluting into an analytical column. Due to the small number of samples in our first proteomics study, the initial lengthy separation/fractionation step provided better resolution of low-abundance proteins and allowed for a greater amount of identified proteins (n = 475, across 44 samples) (Nguyen et al., 2016) compared to our second study (n = 163, across 180 samples). Other methods to address plasma dynamic range issues include usage of the Multiple Affinity Removal System (Agilent) columns or spin columns with methodology similar to DNA purification for the purposes of depleting the most abundant proteins (Tu et al., 2010)**.** However, addition of these extra steps to the proteomics workflow introduce increased funding and machine time costs.

The elute from the chemical separation workflows is ionized before entering the mass spectrometer. Within our studies, electrospray ionisation (ESI) was used to aerosolize the

sample and produce ions by subjecting the sample to extremely high voltages. As ESI is a soft ionization technique, it has advantages for plasma in trypsin based proteomics by not significantly fragmenting molecules in the ionization phase (Wang & Hanash, 2009). The mass spectrometer serves as the detector instrument and identifies ionized molecules based on their mass-to-charge (*m*/z) ratio. This is the physical principle whereby the mass atomic composition of molecules corresponds to its electrical and ionic properties (Saraswathy & Ramalingam, 2011). The most popular MS technique follows the time of flight (TOF) principle, where ions are accelerated by an electrical field in a vacuum prior to reaching the detector, with different *m/z* ratios resulting in distinct travel time. Recently, Thermo-Fisher Scientific has developed the Orbitrap series of MS, an analyser that utilizes electrical fields to trap ions in a central spindle, resulting in harmonic axial motions that can be detected with greater resolution than traditional TOF-MS (Zubarev & Makarov, 2013). This technique is employed in the two variants of MS used in our studies, the Orbitrap Elite (Thermo Fisher Scientific) and Orbitrap Q Exactive (Thermo Fisher Scientific), both of which have demonstrated resolution in plasma proteomics (Robinson, 2019). The output from MS based techniques is peak spectra, whereby peak horizontal position (*m/*z) is used to identify the molecule based on peak height and the relative intensity or abundance of that molecule present in the sample (Figure 4).

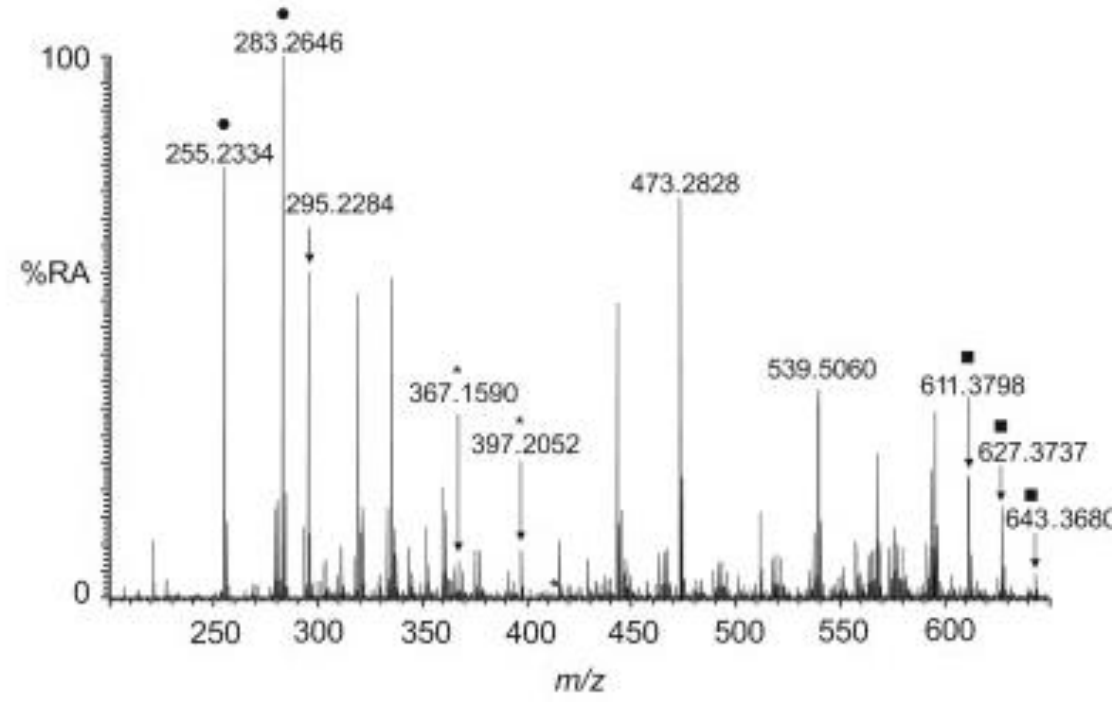

Figure 4. Example peak spectra of plasma samples. Reused from Griffiths et al. (2012). Analytical strategies for characterization of oxysterol lipidomes: Liver X receptor ligands in plasma, *Free Radical Biology and Medicine*, 59, pp. 81, under the Creative Commons Attribution License 3.0 Unported (CC BY 3.0) from Elseiver.

Bioinformatics

Bioinformatics refers to the computational aspects of omics workflows, especially the post-data acquisition phase. The several uses of bioinformatics in proteomics includes the automatization of identifying protein and peptides from gel and MS based proteomics and adding a layer of functional annotation of results to augment traditional statistical approaches. Within this experimental series, there has been a focus on utilizing bioinformatics approaches to further understand, organize and test the relationship of large proteomics datasets to functional biological processes. Situated at the end of the proteomics analytical chemistry workflow, bioinformatics approaches ultimately allow the research presented in this thesis to address the aims of identifying and quantifying the molecules and biosystems related to stroke recovery. This section will detail the considerations of the bioinformatics approaches used.

Peptide and Protein Matching

The overwhelming amount of peak level data from MS requires considerable computational time for peptide and protein identification. Bioinformatics in proteomics has largely replaced the traditional approach of visually matching each peak by hand against published examples of known spectra to automatically resolve peaks against curated protein spectra databases (Christin, Bischoff, & Horvatovich, 2011). This approach has allowed for a greater degree of consistency and speed in peptide identification in complex samples. The present thesis utilized the MaxQuant software package, allowing for a streamlined data processing workflow that interfaces with popular peptide databases such as Mascot and Andromeda (Cox & Mann, 2008; Cox et al., 2011). A false discovery rate of 1% for protein identification on Andromeda has been utilized in this thesis, as this allows for confident peptide identification over other approaches, including advanced proteomics methods examining modified peptides and post-translational modifications (Cox et al., 2011).

Set Based Bioinformatics Methods

Due to the large amount of data generated from standardized MS proteomics pipelines, traditional statistical approaches are only able to provide a limited interpretation of raw outputs. Functional annotation is a procedure that provides additional biological context to the data from omics experiments. This requires the interfacing of biochemical and pathway databases such as Gene Ontology (GO) (Ashburner et al., 2000), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa & Goto, 2000), Reactome (Fabregat et al., 2018) and Biocarta (Nishimura, 2001). These databases maintain a collection of molecules and their predetermined relationship to a higher biological category, with expected expression profiles curated from previous studies (C. Chen, Huang, & Wu, 2017). The majority of the databases are built upon knowledge from genomics studies, incorporating data from RNA sequencing

and microarray studies. As such, protein data is often arbitrarily translated into gene nomenclature to be able to access genomics databases (Manzoni et al., 2016). Although proteomics databases exist, their usage is more suited towards examining protein structure and protein-protein interactions, as opposed to describing systems level biological processes (Manzoni et al., 2016). Ultimately, clinical proteomics bioinformatics approaches greatly underestimate the complexity of the proteome (Pible & Armengaud, 2015). Nevertheless, advanced proteomics and bioinformatics approaches are a complementary addition that builds upon findings in discovery-based approaches, such as those employed in the current experimental series.

Overrepresentation analysis is a basic bioinformatics approach that attempts to categorize proteomics and transcriptomics data into databased defined categories. For example, lipopolysaccharide binding protein is categorized under the toll like receptor 4 cascade and signalling pathways in Reactome and KEGG databases respectively. This can be accomplished by simple web-based interfaces such as the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2007) and Protein Analysis THrough Evolutionary Relationships (PANTHER) (Mi et al., 2017). The DAVID approach employs a Fisher exact test for calculating the statistical likelihood of finding experimentally defined molecular identifiers as a proportion of those found in the GO database (Hosack, Dennis, Sherman, Lane, & Lempicki, 2003). These methods were utilized within the systematic review Chapter 3 of this thesis to determine if the lists of differentially expressed genes, mRNAs and proteins that were interrelated within stroke studies could be explained as functional biological processes.

The pilot proteomics study in this experimental series utilized the Gene Set Enrichment Analysis (GSEA) method, an analytical approach that provides a stronger

statistical rationale than basic overrepresentation methods (Subramanian et al., 2005). Furthermore, it is one of the only approaches that is able to perform a correlation test to explore the relationship between machine expression data and continuous variables, such as MADRS depression symptom scores (Liberzon et al., 2011). GSEA utilizes a weighted Kolmogorov-Smirnov-like statistic to determine an overrepresentation at the top or bottom of an experimental gene set in relation to a chosen database gene set (Figure 5) (Subramanian et al., 2005). In addition to providing categorical overrepresentation analysis, the 'leading edge' subset further defines which genes or proteins contribute most to the overall enrichment of the overrepresented gene set (Subramanian et al., 2005). In the pilot proteomics study, this method was used to identify complement system dysregulation in relation to increasing MADRS scores across 4 biochemical databases (Reactome, KEGG, Hallmark and Biocarta), with proteins such as complement components 3, 5, 9 (C3, C5, C9) and mannose binding lectin 2 contributing significantly to the overall enrichment of this biological category.

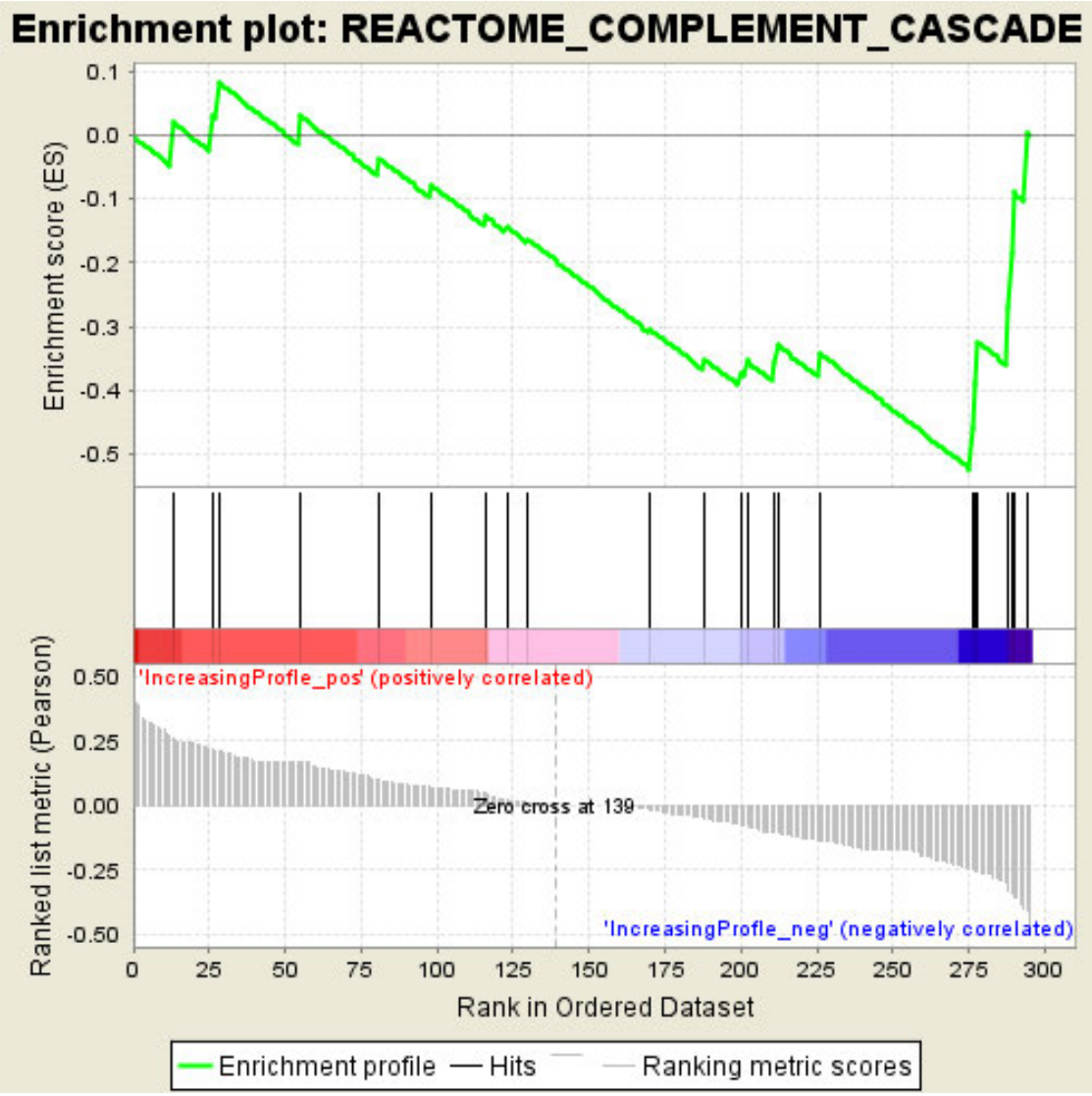

Figure 5. Gene Set Enrichment Analysis of the protein expression profile of *n* = 44 plasma samples, correlated with 3-month MADRS depression symptom scores from Chapter 5.

Topology Based Enrichment Methods. Biochemical relationships can also be better understood as pathways, with one molecule having a clear effect on the abundance of another molecule. The set-based bioinformatics analysis methods presented previously are not able to assess the relationship of one molecule to another. In addition to set based enrichment methods, topological methods are able to leverage expected molecule-to-molecule biochemical relationships to create hierarchical analyses and diagrams of gene set information, therefore providing a true representation of a biochemical pathway (Figure 6).

For example, an increase in complement component 3 is expected to lead to an increase in complement component 5, thus describing information about two separate nodes and the correlational hierarchy between them. Further development of biochemical databases such as KEGG and Reactome and the adoption of a recent developments in Systems Biology Markup Language (Hucka et al., 2018) have allowed for the additional representation of topological data, as well as facilitating bioinformatics workflows at the pathway level.

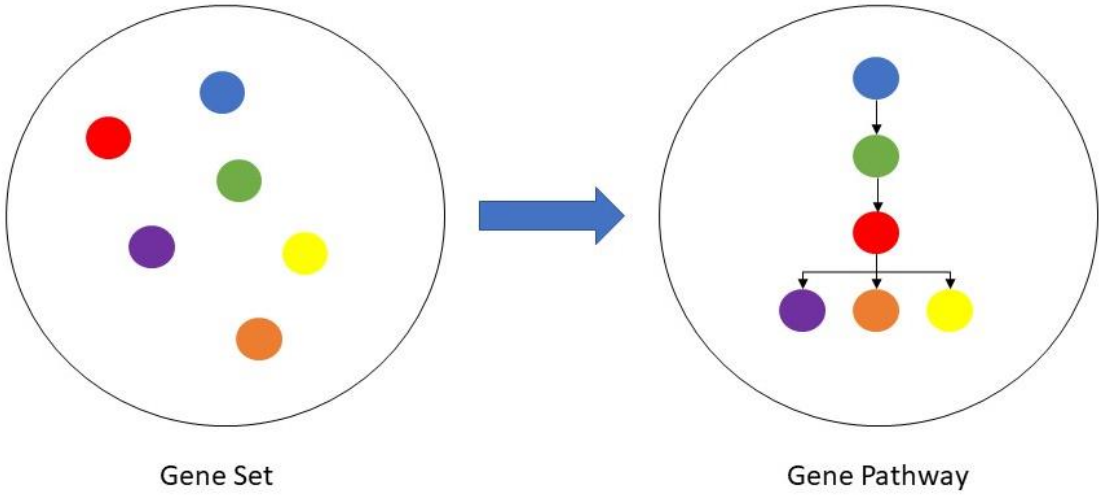

Figure 6. The aim of bioinformatics-based pathway analysis is to provide organizational structure for disparate molecules. This aids with the interpretability of biological data and can address both gene sets and whole pathways simultaneously.

The final Chapter 7 of the experimental series in this thesis additionally aimed to investigate the molecule-to-molecule biochemical relationships from molecule to molecule within pathways and hence chose to utilize the Gene Graph Enrichment Analysis approach to examine the plasma proteome of 60 post-stroke patients over 3 distinct timepoints. This is a complex statistical approach, involving resolution of gene sets from experimental and database-defined gene sets in classical overrepresentation analysis and a further test for overall map consistency whereby the majority of molecule-to-molecule relationships agrees

or disagrees with the expected pathway maps from biochemical databases (Geistlinger, Csaba, Küffner, Mulder, & Zimmer, 2011). As such, GGEA can provide a complete statistical and bioinformatics overview of pathway structure consistency in set level enrichment, as well as quantify the molecule-to-molecule relationships (Figure 7). Furthermore, this approach is able to identify significant gene sets, requiring a lower rate of additional single genes in a permutation approach before enrichment is established (Geistlinger et al., 2011). The inclusion of discernible statistical methods in the R package (Geistlinger, Csaba, & Zimmer, 2016), also allows for ability to test for changes in molecule-to-molecule or regulatory strength across conditions, such as timepoints.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. The Gene Ontology consortium. *Nature Genetics, 25*(1), 25-29. doi:10.1038/75556

Aslam, B., Basit, M., Nisar, M. A., Rasool, M. H., & Khurshid, M. (2017). Proteomics: Technologies and their applications. *Journal of Chromatographic Science, 55*(2), 182-196. doi:10.1093/chromsci/bmw167

Banks, J. L., & Marotta, C. A. (2007). Outcomes validity and reliability of the modified rankin scale: Implications for stroke clinical trials. *Stroke, 38*(3), 1091.

Basu, D., & Kulkarni, R. (2014). Overview of blood components and their preparation. *Indian Journal of Anaesthesia, 58*(5), 529-537. doi:10.4103/0019-5049.144647

Bell, W. R. (2000). Blood, coagulants and anticoagulants. In *Kirk-Othmer encyclopedia of chemical technology*.

Braun, H.-P., & Senkler, M. (2012). Functional annotation of 2D protein maps: The gelmap portal. *Frontiers in Plant Science, 3*(87). doi:10.3389/fpls.2012.00087

Breier, M., Wahl, S., Prehn, C., Fugmann, M., Ferrari, U., Weise, M., . . . Lechner, A. (2014). Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. *PloS One, 9*(2), e89728-e89728. doi:10.1371/journal.pone.0089728

Broderick, J. P., Adeoye, O., & Elm, J. (2017). Evolution of the modified rankin scale and its use in future stroke trials. *Stroke, 48*(7), 2007-2012. doi:doi:10.1161/STROKEAHA.117.017866

Brown, M., & Wittwer, C. (2000). Flow cytometry: Principles and clinical applications in hematology. *Clinical Chemistry, 46*(8), 1221.

Budnik, B., Levy, E., Harmange, G., & Slavov, N. (2018). Mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *bioRxiv*, 102681. doi:10.1101/102681

Cadilhac, D. A., Kilkenny, M. F., Levi, C. R., Lannin, N. A., Thrift, A. G., Kim, J., . . . Anderson, C. S. (2017). Risk-adjusted hospital mortality rates for stroke: Evidence from the australian stroke clinical registry (AUSCR). *Medical Journal of Australia, 206*(8), 345-350. doi:10.5694/mja16.00525

Carey, L. M., Crewther, S., Salvado, O., Lindén, T., Connelly, A., Wilson, W., . . . Donnan, G. A. (2013). Stroke imaging prevention and treatment (START): A longitudinal stroke cohort study: Clinical trials protocol. *International Journal of Stroke*, *10*(4), 636-644. doi:10.1111/ijs.12190

Carrette, O., Burkhard, P. R., Sanchez, J.-C., & Hochstrasser, D. F. (2006). State-of-the-art two-dimensional gel electrophoresis: A key tool of proteomics research. *Nature Protocols*, *1*(2), 812-823. doi:10.1038/nprot.2006.104

Chen, C., Huang, H., & Wu, C. H. (2017). Protein bioinformatics databases and resources. *Methods in Molecular Biology, 1558*, 3-39. doi:10.1007/978-1-4939-6783-4_1

Chen, E. I., Cociorva, D., Norris, J. L., & Yates, J. R., 3rd. (2007). Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *Journal of Proteome Research*, *6*(7), 2529-2538. doi:10.1021/pr060682a

Chiti, G., & Pantoni, L. (2014). Use of montreal cognitive assessment in patients with stroke. *Stroke*, *45*(10), 3135.

Christin, C., Bischoff, R., & Horvatovich, P. (2011). Data processing pipelines for comprehensive profiling of proteomics samples by label-free LC–MS for biomarker

discovery. *Talanta, 83*(4), 1209-1224. Doi:https://doi.org/10.1016/j.talanta.2010.10.029

Chung, H.-J., Lee, W., Chun, S., Park, H.-I., & Min, W.-K. (2009). Analysis of turnaround time by subdividing three phases for outpatient chemistry specimens. *Annals of Clinical and Laboratory Science, 39*(2), 144-149.

Cox, J., & Mann, M. (2008). Maxquant enables high peptide identification rates, individualized p.P.B.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology, 26,* 1367. doi:10.1038/nbt.1511

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: A peptide search engine integrated into the maxquant environment. *Journal of Proteome Research, 10*(4), 1794-1805. doi:10.1021/pr101065j

Craig, F. E., & Foon, K. A. (2008). Flow cytometric immunophenotyping for hematologic neoplasms. *Blood, 111*(8), 3941. doi:10.1182/blood-2007-11-120535

Davis, B. H., & Barnes, P. W. (2012). Automated cell analysis: Principles. In *Laboratory hematology practice.*

Dayon, L., & Kussmann, M. (2013). Proteomics of human plasma: A critical comparison of analytical workflows in terms of effort, throughput and outcome. *EuPA Open Proteomics, 1,* 8-16. doi:https://doi.org/10.1016/j.euprot.2013.08.001

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., . . . D'Eustachio, P. (2018). The Reactome pathway knowledgebase. *Nucleic Acids Research, 46,* 649-655. doi:10.1093/nar/gkx1132

Fantino, B., & Moore, N. (2009). The self-reported Montgomery-Åsberg Depression Rating Scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry, 9,* 26-26. doi:10.1186/1471-244X-9-26

Finehout, E. J., & Lee, K. H. (2004). An introduction to mass spectrometry applications in biological research. *Biochemistry and Molecular Biology Education, 32*(2), 93-100. doi:10.1002/bmb.2004.494032020331

Freeman, W. M., & Hemby, S. E. (2004). Proteomics for protein expression profiling in neuroscience. *Neurochemical Research, 29*(6), 1065-1081. doi:10.1023/b:nere.0000023594.21352.17

Fritz, R., Ruth, W., & Kragl, U. (2009). Assessment of acetone as an alternative to acetonitrile in peptide analysis by liquid chromatography/mass spectrometry. *Rapid Communications in Mass Spectrometry, 23*(14), 2139-2145. doi:10.1002/rcm.4122

Geistlinger, L., Csaba, G., Küffner, R., Mulder, N., & Zimmer, R. (2011). From sets to graphs: Towards a realistic enrichment analysis of transcriptomic systems. *Bioinformatics, 27*(13), 366-373. doi:10.1093/bioinformatics/btr228

Geistlinger, L., Csaba, G., & Zimmer, R. (2016). Bioconductor's enrichmentbrowser: Seamless navigation through combined results of set- & network-based enrichment analysis. *BMC Bioinformatics, 17*, 45. doi:10.1186/s12859-016-0884-1

Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. (2017). Revisiting biomarker discovery by plasma proteomics. *Molecular Systems Biology, 13*(9), 942-942. doi:10.15252/msb.20156297

Graves, P. R., & Haystead, T. A. J. (2002). Molecular biologist's guide to proteomics. *Microbiology and molecular biology reviews : MMBR, 66*(1), 39-63. doi:10.1128/mmbr.66.1.39-63.2002

Griffiths, W., Crick, P., Wang, Y., Ogundare, M., Tuschl, K., Morris, A., . . . Wang, Y. (2012). *Analytical strategies for characterization of oxysterol lipidomes: Liver x receptor ligands in plasma* (Vol. 59).

Hand, B., Page, S. J., & White, S. (2014). Stroke survivors scoring zero on the nih stroke scale score still exhibit significant motor impairment and functional limitation. *Stroke Research and Treatment, 2014*, 462681. doi:10.1155/2014/462681

Hortin, G. L., & Sviridov, D. (2010). The dynamic range problem in the analysis of the plasma proteome. *Journal of Proteomics, 73*(3), 629-636. doi:https://doi.org/10.1016/j.jprot.2009.07.001

Hosack, D. A., Dennis, G., Jr., Sherman, B. T., Lane, H. C., & Lempicki, R. A. (2003). Identifying biological themes within lists of genes with ease. *Genome Biology, 4*(10), R70-R70. doi:10.1186/gb-2003-4-10-r70

Huang, D. W., Sherman, B. T., Tan, Q., Collins, J. R., Alvord, W. G., Roayaei, J., . . . Lempicki, R. A. (2007). The DAVID gene functional classification tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology, 8*(9), R183-R183. doi:10.1186/gb-2007-8-9-r183

Hucka, M., Bergmann, F. T., Dräger, A., Hoops, S., Keating, S. M., Le Novère, N., . . . Wilkinson, D. J. (2018). The systems biology markup language (sbml): Language specification for level 3 version 2 core. *Journal of Integrative Bioinformatics, 15*(1), 20170081. doi:10.1515/jib-2017-0081

Janssen, M., van Hoeven, L., & Rautmann, G. (2015). Trends and observations on the collection, testing and use of blood and blood components in europe.

Jauch, E. C., Lindsell, C., Broderick, J., Fagan, S. C., Tilley, B. C., & Levine, S. R. (2006). Association of serial biochemical markers with acute ischemic stroke. *Stroke, 37*(10), 2508-2513. doi:doi:10.1161/01.STR.0000242290.01174.9e

Jog, N. K. (2013). *Electronics in medicine and biomedical instrumentation*: PHI Learning Pvt. Ltd.

Kanehisa, M., & Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27-30.

Kang, H. J., Stewart, R., Kim, J. M., Jang, J. E., Kim, S. Y., Bae, K. Y., . . . Yoon, J. S. (2013). Comparative validity of depression assessment scales for screening poststroke depression. *Journal of Affective Disorders, 147*(1-3), 186-191. doi:10.1016/j.jad.2012.10.035

Karpievitch, Y. V., Polpitiya, A. D., Anderson, G. A., Smith, R. D., & Dabney, A. R. (2010). Liquid chromatography mass spectrometry-based proteomics: Biological and technological aspects. *The annals of applied statistics, 4*(4), 1797-1823. doi:10.1214/10-AOAS341

Kim, H. S. (2016). Blood glucose measurement: Is serum equal to plasma? *Diabetes & Metabolism Journal, 40*(5), 365-366. doi:10.4093/dmj.2016.40.5.365

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., & Mesirov, J. P. (2011). Molecular signatures database (msigdb) 3.0. *Bioinformatics, 27*(12), 1739-1740. doi:10.1093/bioinformatics/btr260

Lyden, P. (2017). Using the national institutes of health stroke scale. *Stroke, 48*(2), 513-519. doi:doi:10.1161/STROKEAHA.116.015434

Lyden, P., Raman, R., Liu, L., Emr, M., Warren, M., & Marler, J. (2009). NIHSS certification is reliable across multiple venues. *Stroke, 40*(7), 2507-2511. doi:10.1161/STROKEAHA.108.532069

Ma, H., Parsons, M. W., Christensen, S., Campbell, B. C. V., Churilov, L., Connelly, A., . . . Donnan, G. A. (2011). A multicentre, randomized, double-blinded, placebo-controlled phase iii study to investigate extending the time for thrombolysis in

emergency neurological deficits (EXTEND). *International Journal of Stroke, 7*(1), 74-80. doi:10.1111/j.1747-4949.2011.00730.x

Manzoni, C., Kia, D. A., Vandrovcova, J., Hardy, J., Wood, N. W., Lewis, P. A., & Ferrari, R. (2016). Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences. *Briefings in Bioinformatics, 19*(2), 286-302. doi:10.1093/bib/bbw114

McLennan, S. N., Mathias, J. L., Brennan, L. C., & Stewart, S. (2011). Validity of the montreal cognitive assessment (MoCA) as a screening test for mild cognitive impairment (MCI) in a cardiovascular population. *Journal of Geriatric Psychiatry and Neurology, 24*(1), 33-38. doi:10.1177/0891988710390813

Meleady, P. (2018). Two-dimensional gel electrophoresis and 2d-dige. In K. Ohlendieck (Ed.), *Difference gel electrophoresis: Methods and protocols* (pp. 3-14). New York, NY: Springer New York.

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. D. (2017). Panther version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research, 45*(D1), D183-D189. doi:10.1093/nar/gkw1138

Mohamadkhani, A., & Poustchi, H. (2015). Repository of human blood derivative biospecimens in biobank: Technical implications. *Middle East Journal of Digestive Diseases, 7*(2), 61-68.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry, 134*(4), 382-389. doi:10.1192/bjp.134.4.382

Nguyen, V. A., Carey, L. M., Giummarra, L., Faou, P., Cooke, I., Howells, D. W., . . . Crewther, S. G. (2016). A pathway proteomic profile of ischemic stroke survivors

reveals innate immune dysfunction in association with mild symptoms of depression – a pilot study. *Frontiers in Neurology, 7*, 85. doi:10.3389/fneur.2016.00085

Nigam, P. K. (2011). Serum lipid profile: Fasting or non-fasting? *Indian Journal of Clinical Biochemistry, 26*(1), 96-97. doi:10.1007/s12291-010-0095-x

Nishimura, D. (2001). Biocarta. *Biotech Software & Internet Report, 2*(3), 117-120. doi:10.1089/152791601750294344

Organization, W. H. (2010). Who guidelines on drawing blood: Best practices in phlebotomy.

Pible, O., & Armengaud, J. (2015). Improving the quality of genome, protein sequence, and taxonomy databases: A prerequisite for microbiome meta-omics 2.0. *Proteomics, 15*(20), 3418-3423. doi:10.1002/pmic.201500104

Quinn, T. J., Taylor-Rowan, M., Coyte, A., Clark, A. B., Musgrave, S. D., Metcalf, A. K., . . . Myint, P. K. (2017). Pre-stroke modified Rankin Scale: Evaluation of validity, prognostic accuracy, and association with treatment. *Frontiers in Neurology, 8*, 275-275. doi:10.3389/fneur.2017.00275

Robinson, C. V. (2019). Mass spectrometry: From plasma proteins to mitochondrial membranes. *Proceedings of the National Academy of Sciences, 116*(8), 2814. doi:10.1073/pnas.1820450116

Roussel, M., Benard, C., Ly-Sunnaram, B., & Fest, T. (2010). Refining the white blood cell differential: The first flow cytometry routine application. *Cytometry Part A, 77A*(6), 552-563. doi:10.1002/cyto.a.20893

Sagen, U., Vik, T. G., Moum, T., Mørland, T., Finset, A., & Dammen, T. (2009). Screening for anxiety and depression after stroke: Comparison of the hospital anxiety and depression scale and the montgomery and åsberg depression rating scale. *Journal of*

Psychosomatic Research, 67(4), 325-332. doi:https://doi.org/10.1016/j.jpsychores.2009.03.007

Salkind, N. (2010). Encyclopedia of research design. doi:10.4135/9781412961288

Saraswathy, N., & Ramalingam, P. (2011). Mass spectrometry for proteomics. In N. Saraswathy & P. Ramalingam (Eds.), *Concepts and techniques in genomics and proteomics* (pp. 171-183): Woodhead Publishing.

States, D. J., Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., & Hanash, S. M. (2006). Challenges in deriving high-confidence protein identifications from data gathered by a hupo plasma proteome collaborative study. *Nature Biotechnology, 24*(3), 333-338. doi:10.1038/nbt1183

Strohkamp, S., Gemoll, T., & Habermann, J. K. (2016). Possibilities and limitations of 2DE-based analyses for identifying low-abundant tumor markers in human serum and plasma. *Proteomics 16*(19), 2519-2532. doi:10.1002/pmic.201600154

Stroncek, D. F., Xing, L., Chau, Q., Zia, N., McKelvy, A., Pracht, L., . . . Jin, P. (2011). Stability of cryopreserved white blood cells (WBCs) prepared for donor WBC infusions. *Transfusion, 51*(12), 2647-2655. doi:10.1111/j.1537-2995.2011.03210.x

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), 15545.

Sullivan, E. (2006). Hematology analyzer: From workhorse to thoroughbred. *Laboratory Medicine, 37*(5), 273-278. doi:10.1309/tmq6t4cbcg408141

Switzar, L., Giera, M., & Niessen, W. M. A. (2013). Protein digestion: An overview of the available techniques and recent developments. *Journal of Proteome Research, 12*(3), 1067-1077. doi:10.1021/pr301201x

Tsiatsiani, L., & Heck, A. J. R. (2015). Proteomics beyond trypsin. *The FEBS Journal, 282*(14), 2612-2626. doi:10.1111/febs.13287

Tu, C., Rudnick, P. A., Martinez, M. Y., Cheek, K. L., Stein, S. E., Slebos, R. J. C., & Liebler, D. C. (2010). Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of Proteome Research, 9*(10), 4982-4991. doi:10.1021/pr100646w

Vecellio, E., Li, L., Westbrook, J. I., Xiong, J., Georgiou, A., Eigenstetter, A., . . . Wilson, R. (2015). Examination of variation in hospital pathology investigations by diagnosis-related groups and associations with outcomes and costs, *Australian Department of Health*.

Vis, J. Y., & Huisman, A. (2016). Verification and quality control of routine hematology analyzers. *International Journal of Laboratory Hematology, 38*(S1), 100-109. doi:10.1111/ijlh.12503

Wang, H., & Hanash, S. (2009). Electrospray mass spectrometry for quantitative plasma proteome analysis. *Methods in Molecular Biology, 564*, 227-242. doi:10.1007/978-1-60761-157-8_13

Weinkauf, M., Hiddemann, W., & Dreyling, M. (2006). Sample pooling in 2-D gel electrophoresis: A new approach to reduce nonspecific expression background. *Electrophoresis 27*(22), 4555-4558. doi:10.1002/elps.200600207

Williams, J. B., & Kobak, K. A. (2008). Development and reliability of a Structured Interview guide for the Montgomery-Åsberg Depression Rating Scale (SIGMA). *The British Journal of Psychiatry, 192*(1), 52-58.

Xue, J., Huang, W., Chen, X., Li, Q., Cai, Z., Yu, T., & Shao, B. (2017). Neutrophil-to-lymphocyte ratio is a prognostic marker in acute ischemic stroke. *Journal of Stroke and Cerebrovascular Diseases, 26*(3), 650-657. doi:10.1016/j.jstrokecerebrovasdis.2016.11.010

Zhang, X. (2015). Less is more: Membrane protein digestion beyond urea-trypsin solution for next-level proteomics. *Molecular & Cellular Proteomics, 14*(9), 2441-2453. doi:10.1074/mcp.R114.042572

Zhao, L., Dai, Q., Chen, X., Li, S., Shi, R., Yu, S., . . . Zhang, R. (2016). Neutrophil-to-lymphocyte ratio predicts length of stay and acute hospital cost in patients with acute ischemic stroke. *Journal of Stroke and Cerebrovascular Diseases, 25*(4), 739-744. doi:10.1016/j.jstrokecerebrovasdis.2015.11.012

Zimmerman, L. J., Li, M., Yarbrough, W. G., Slebos, R. J. C., & Liebler, D. C. (2012). Global stability of plasma proteomes for mass spectrometry-based analyses. *Molecular & Cellular Proteomics, 11*(6), M111.014340-M014111.014340. doi:10.1074/mcp.M111.014340

Zini, G. (2014). Stability of complete blood count parameters with storage: Toward defined specifications for different diagnostic applications. *International Journal of Laboratory Hematology, 36*(2), 111-113. doi:10.1111/ijlh.12181

Zubarev, R. A., & Makarov, A. (2013). Orbitrap mass spectrometry. *Analytical Chemistry, 85*(11), 5288-5296. doi:10.1021/ac4001223

Chapter 5

Experimental Study 1

Acute routine leukocyte and neutrophil counts are predictive of post-stroke recovery at 3- and 12-months post-stroke: An exploratory study

Preface

Clinical management of stroke begins with the patient visit to an emergency department. Within this period, many investigational objective physiological measures and in particular blood tests are taken to confirm if the symptoms are specific to stroke. Common symptoms that often ‘mimic’ stroke include infections and hypoglycemia and hence require differential diagnosis usually via blood tests.

Blood tests generate a great amount of patient data that can have additional uses beyond the acute stroke phase though their usefulness in the context of longitudinal stroke recovery has not yet been fully explored. Immune cell counts are non-specific measures of immunity whereby greater white blood cell (or leukocyte) and neutrophil counts are often interpreted as increased immune challenge or presence of trauma and lessened leukocyte or neutrophil counts are interpreted as immunosuppression. Clinically, these tests are often interpreted based on normative reference ranges that are currently more useful as a measure of infection. As such, these measures are often not used interpreted as a measure of longitudinal recovery from stroke. Thus, the overall aim of Chapter 4 was to address immune cell counts in relation to measures of longitudinal stroke outcomes to link acute measures of peripheral immunity to aspects of stroke recovery.

The manuscript presented here has been accepted for publication into *Neurorehabilitation and Neural Repair* on the 25th of June, 2020.

Acute routine leukocyte and neutrophil counts are predictive of post-stroke recovery at 3 and 12 months post-stroke: An exploratory study

Vinh A Nguyen, BPsych (Hons)[1, 2, 3]

Sheila G Crewther, PhD[2]

David W Howells, PhD[23, 4]

Tissa Wijeratne, MD[25]

Henry Ma, MD PhD[26]

Graeme J Hankey, MD PhD[27]

Stephen Davis, MD PhD[23,8]

Geoffrey A Donnan, MD PhD[23,8]

Leeanne M Carey, PhD[21, 3]

and the START research team

Affiliations

[1]Department of Occupational Therapy, Social Work and Social Policy, La Trobe University, Bundoora, Australia

[2]Department of Psychology and Counselling, La Trobe University, Bundoora, Australia

[3]Neurorehabilitation and Recovery, Stroke Theme, The Florey Institute of Neuroscience and Mental Health, Heidelberg, Australia

[4]School of Medicine, University of Tasmania, Hobart, Australia

[5]Neurology, Western Health, Sunshine, Australia

[6]Stroke and Ageing Research Group, Monash University, Clayton, Australia

[7]Medical School, University of Western Australia, Perth, Australia

[8]Melbourne Brain Centre, Royal Melbourne Hospital, Parkville, Australia

Word count: 4050 (Body)

Tables: 7

Figures: 1

Abstract

Background and Aims

White blood cell (WBC) and neutrophil counts (NC) are common markers of inflammation and neurological stroke damage and could be expected to predict post-stroke outcomes.

Objective

The aim of this study was to explore the prognostic value of early post-stroke WBC and NC to predict cognition, mood and disability outcomes at 3- and 12-months post-stroke.

Methods

Routine clinical analyses WBC and NC were collected at 3 timepoints in the first 4 days of hospitalization from 156 acute stroke patients. Correlations with additional hierarchical or ordinal regressions were explored between acute WBC and NC and functional recovery, depression and cognition at 3- and 12-months post-stroke; after covarying for age and baseline stroke severity.

Results

We found significant increases in NC between <12 hours and 24-48 hour timepoints (p = .05). Hierarchical regressions, covaried for age and baseline stroke severity, found that 24-48 hour WBC (p = .05) and NC (p = .04) significantly predicted 3-month cognition scores. Similarly, 24-48 hour WBC (p = .05) and NC (p = .02) predicted cognition scores at 12-months. Increases in WBC and NC were predictive of increased cognition scores at both 3 and 12 months (positive recovery) though there were no significant associations between WBC and NC and disability or depression scores.

Conclusions

Routine acute stroke clinical laboratory tests such as WBC and NC taken between 24-48 hours post-stroke are predictive of cognition post-stroke. It is interpreted that higher rapid immunological activation in the acute phase is an indicator for the trajectory of positive stroke recovery.

Keywords

Blood biomarkers, Post-stroke recovery, cognition, white blood cell, neutrophil, leukocytes

Introduction

Acute focal brain ischemia is characterized by rapid decrease in oxygen and adenosine triphosphate (ATP) availability leading to impairment of ionic pumps compromise of basement membranes and cellular swelling[1]. Such neuronal excitotoxicity immediately induces a sterile inflammatory response mediated by damage associated molecular patterns (DAMPs), increased permeability of the blood-brain barrier[2], and local infiltration by peripheral blood leukocytes[3]. Thus the interaction of immune and inflammatory pathways have long been conceptualized as one of the mechanisms underlying long term stroke recovery[4,5]. Indeed it has also been proposed that recovery of neurological function after stroke may be mediated by the initial infiltration of leukocytes into the infarcted area and/or with the associated immune and inflammatory response within the first to second week post stroke[5,6]. However, general clinical blood tests such as the complete blood count (CBC) or full blood examination (FBE)[7] and component leukocyte or white blood cell (WBC) count that are prioritized following suspected stroke admission[8] are predominantly utilized to investigate the differential diagnoses presenting as stroke mimics such as the presence of infections or hypoglycemia[9], rather than as a marker of tissue injury, recovery and prognosis.

Previous studies in post-stroke populations, have shown that persistent leukocytosis (WBC > 11 x 10^9 cells/L in adults) 48 hours after admission is associated with higher baseline scores on the National Institute of Health stroke scale (NIHSS), and worse functional recovery at discharge and longer lengths of stay (M = 10.4 days *vs* M = 6.6 days)[10]. Other studies have also shown that even after controlling for age and baseline stroke severity, relatively elevated leukocyte count (regardless of reference range defined leukocytosis [11]), is a predictor of post-stroke NI qS severity, modified Rankin scale scores (mRS) and stroke volume[12,13]. Indeed, controlling for age and baseline stroke severity is

extremely important to reveal changes in post-stroke immune responses whilst addressing the natural progression of immunosenescence in elderly adults[14] and the large effect of acute stroke on the peripheral immune system[15]. However, knowing when the most clinically and scientifically valuable-time points post-stroke onset to measure blood products such as WBC and NC for the prediction of post-stroke recovery, has not yet been established. Studies examining the acute trajectory of post-stroke WBC[16] and NC[17] have shown that cell counts increase rapidly within the first 0-2 days post-stroke and then fall between 3-7 days. Indeed, both the timing and amplitude of acute post-stroke immune responses have been found to correlate with stroke severity within 24 hours of symptom onset[18] and also recently been reported to correlate with improvement on NIHSS scores over the first week[19].

A recent 2017 review and international consensus recommendation concludes that there is a relative lack of evidence for blood-based biomarkers to predict stroke outcomes and recovery[20], when post-stroke recovery is defined using the mRS disability and functional recovery scale[21]. Yet, this measure does not include mood or cognitive outcomes, despite the fact that they are commonly reported as unmet needs by stroke survivors[22]. Recent studies examining the association of WBCs, neutrophils and lymphocytes have determined that increased acute neutrophil-to-lymphocyte ratio is often predictive of post-stroke depression[23]. Indeed, despite accumulating evidence for the use of acute routine investigations of blood products in the association and prediction of stroke outcomes, these measures have yet to be adopted by clinical guidelines[24].

The role of immune markers in depression and cognition regardless of stroke, involves a complex pathophysiological system linking immune dysregulation, chronic inflammation and neuro-behavioural changes[25]. Previous meta analyses have identified that increased levels of inflammatory mediators such as interleukin 1, interleukin 6 and C-reactive

protein are associated with increased symptoms of depression[26]. Indeed, chronic inflammation has the potential to mediate neurotoxic response dysregulation of the glutamatergic system and upregulation of reactive oxygen and nitrogen species, with structural changes such as hippocampal and prefontal cortex matter loss[25]. Furthermore, there are sickness behaviour similarities between subjects facing immune challenge (such as infection) and depression symptoms in many mammalian species[27]. The link between cognition and immunity has not been as well established as the physiological basis of immunity in depression[27]. Multiple large prospective studies examining elderly adults found both positive and non-significant associations between cognitive outcomes and leukocyte count[28,29] or inflammatory markers[29,30]. Studies in stroke immunity and cognition have identified that increases in myelin basic protein antibodies that respond to central nervous system (CNS) antigens are associated with longitudinal cognitive decline[4,31].

We aimed to explore the associations between measures of acute post-stroke WBC and NC acquired within the first week of hospitalization, and longer-term post-stroke recovery at 3- and 12-months defined by clinical scales for functional recovery (modified Rankin Score (mRS)), depression (Montgomery-Åsberg Depression Rating Scale (MADRS))[32] and cognition (Montreal Cognitive Assessment (MoCA))[33]. Investigation of predictive associations between routine blood tests conducted in the acute phase to clinical outcomes at 3- and 12-months post-stroke was possible using the longitudinal START stroke cohort[34]. It was hypothesized that increases in hematological markers of immune function within 0-12 hours, 24-48 hours and 3-7 days post-stroke would be associated with poorer post-stroke functional, cognitive and depressive outcomes at 3- and 12-months post-stroke. Furthermore, we investigated from the acute samples available, the optimal timing for blood draw and analysis that would be the most useful predictor of outcomes. Knowledge of an

association between routine bloods tests, such as WBC and NC, combined with measures of recovery in the first year post-stroke may assist with longer-term prediction and highlight the importance of treating early immunological dysfunction to reestablish immune and inflammatory homeostasis for better long-term outcome.

Method

Participants

Patients were recruited as part of START-EXTEND (Neuroscience Trials Australia: NTA 0901, NTA 0901, NCT01580839)[35] and START-PrePARE (NTA 0902)[34] cohort studies. These studies recruited 219 patients from a group of participating hospitals in Australia between 2011 – 2015. Briefly, the START-EXTEND study investigated the clinical characteristics of extension of the thrombolysis treatment window and included patients within 9 hours of stroke onset, and the START-PrePARE cohort allowed for recruitment of patients with stroke onset up to 3 days with the purpose of examining predictors of recovery, including the functional brain networks associated with longitudinal stroke recovery[16]. A subset of 156 participants (104 male and 41 female) aged 71 (SD = 12.78) from whom clinical scores and blood test results (Table 3) were available, were selected for the current study.

Blood Tests

Routine blood tests for a range of hematological markers were conducted at participating hospital sites by onsite commercial pathology services. Venipuncture was conducted at baseline (<12 hours), 24-48 hours and 3-7 days post-stroke, as outlined in the START study protocols[34]. Data presented are drawn from electronic patient records and standardized to units recommended by the Royal College of Pathologists Australia[11]. Only total WBC and NC were consistently available from data obtained from the START cohort.

Baseline Demographics and Clinical Characteristics

Patient demographics were obtained by semi-structured interview at baseline admission as part of inclusion in the START cohort studies[34]. Routine demographic and clinical assessments also included age, gender, stroke risk factors and regular blood pressure monitoring. Stroke severity was assessed at baseline using the NIHSS according to clinical trial protocols[36]. Baseline patient characteristics including risk factors and physiological measures were obtained.

Clinical Measures

A series of standardized assessments of neurological function, disability status, cognition and depression were conducted at 3-months and 12-months post-stroke by double-blinded trained health professionals and encompassed several domains of post-stroke recovery. Descriptive details of clinical measures and scoring criteria are presented in Table 1.

The MADRS[37] is a measure of depressive symptoms and has been widely used in major depressive disorders[38] and post-stroke research[39]. The MADRS was administered using the structured interview guide, which is reported to have higher reliability than the self-report method[37]. A score above 8 on the MADRS for depressive symptom screening has been proposed as the normative cut-off in stroke cases[32] for increased sensitivity (>0.80) and acceptable specificity (0.74) compared to the original recommended score of 12[37]. The mRS was used to describe functional disability of patients[21].

The MoCA[33] is a screening measure of multiple cognitive domains and has been shown to be more sensitive to post-stroke cognitive impairment than the Mini-Mental State Examination[40]. Normative cut-offs to determine cognitive impairment in post-stroke research have been suggested at 19-22 for acute administration and 20-27 in the chronic phase[41].

Table 1

Clinical Assessments administered at 3-Months and 12-Months Post-Stroke

Clinical Scale	Purpose	No. of Items	Test scoring and interpretation [a]
NIHSS National Institute of Health Stroke Scale[36]	Quantification of the degree of stroke severity by assessment of multiple neurological domains including motor, vision, speech and cognition	11	0: No Stroke Symptoms 1-4: Mild Stroke severity 5-15: Moderate Stroke severity 16-20: Moderate to Severe Stroke 21-42: Severe Stroke
mRS Modified Rankin Scale[21]	Assessment of degree of disability or dependence in daily activities	Structured Interview	0 - No symptoms. 1 - No significant disability. 2 - Slight disability. 3 - Moderate disability. 4 - Moderately severe disability. 5 - Severe disability. 6 - Dead.
MADRS Montgomery-Åsberg Depression Rating Scale – structured interview[37]	Assessment of presence and severity of depressive symptoms in the past week	10	0-6: Normal, no/minimal depressive symptoms 7-19: Mild Depressive symptoms 20-34: Moderate Depressive symptoms > 34: Severe Depressive symptoms
MoCA Montreal Cognitive Assessment[33]	Rapid screening assessment of generalized cognitive performance	30	≥ 26: Normal 17 - 25: Mild Cognitive Impairment ≤ 16: Cognitive Impairment

[a] Scoring and interpretation relates to total score obtained for each of the assessments. Total score is typically obtained following addition of item scores, as recommended for each of the tools. Details are provided to facilitate interpretation of test scores reported in this manuscript.

Data Analyses

A Pearson's correlation analysis was first conducted to independently assess the associations between acute immune cell counts and longitudinal functional recovery, mood and cognition status. A post-hoc power analysis showed that our sample size (n = 156) was adequate to detect a medium correlation (r = .30) with a power of 0.8 and α level of 0.05

(n_{pwr} = 84). Normality assessment using the Kolmogorov-Smirnov (K-S) test was conducted after outlier removal at +/- >2.5 interquartile range (IQR). The K-S demonstrated that normality was within acceptable bounds for parametric tests. Repeated measure ANCOVAs were used to test for significant differences between immune cell counts at 0-12 hours, 24-48 hours and 3-7 days timepoints post-stroke and T-tests were used to test changes between clinical scores at 3- and 12-months. Post-hoc ANCOVA analyses were used to determine changes between blood test timepoints.

Pearson's partial correlations were used to explore the associations between acute laboratory blood tests (i.e. <12 hours, 24-48 hours and 3-7 days) and mRS, MADRS and MoCA at 3- and 12-months. To control for confounding factors such as age and baseline stroke severity (previously identified as being associated with WBC count and/or recovery[42,43]), these factors were entered as covariates in the partial correlational matrix to allow for examination of the unique contribution of WBC and NC to post-stroke clinical recovery. Hierarchical regressions were conducted following significant partial correlations to determine the changes in explained variance from the addition of immune cell counts to stroke and severity predicting MADRS and MoCA scores at 3 and 12 months. As this was an exploratory study, an α level of 0.05 was used to detect statistical significance after adjusting for covariates, for each of the three outcomes.

Results

Baseline patient characteristics including risk factors are reported in Table 2. Based on individual patient[44] NIHSS scores (mean of 8.47, range = 0 – 32, *SD* = 6.49, median = 6, IQR = 3 to 6) the neurological severity of ischemic stroke in this cohort at the baseline assessment (<12 hours) was primarily mild to moderate. At 3- and 12-months post-stroke, mean stroke severity was still in the mild range (Table 3). Quantification of recovery

outcomes showed relatively mild to moderate clinical scores on the mRS[21] and MADRS[45] at 3 and 12 months (Table 3). However, the trajectory of recovery for cognitive outcomes post-stroke was poor, with a significant decrease of MoCA scores from 3-months to 12-months post stroke (M_{diff} = 0.96, t = 2.72, df = 137, p = .001), with a small effect size (d = 0.23). Clinically, this represents a trend from cognitive functioning in the unimpaired range towards mild cognitive impairment[46]. Stroke severity symptoms assessed using the NIHSS were significantly increased between 3-months and 12-months post-stroke (M_{diff} = 0.64, t = 3.34, df = 180, p = .001, d = 0.21), indicating deterioration in neurological symptoms.

Table 2

Patient Baseline Characteristics (n = 157)

	Frequency (%)	Mean (SD)
Diabetes	27 (17.2%)	
Lipid Disorder	76 (51.3%)	
Ischemic Heart Disease	41 (26.1%)	
Hypertension	91 (57.6%)	
Atrial Fibrillation	41 (25.9%)	
Height		162.56 (25.24)
Weight		78.76 (16.84)
Temperature		36.40 (0.65)
Heart Rate		76.10 (13.90)
Systolic Blood Pressure		144.71 (23.07)
Diastolic Pressure		79.44 (12.98)
Baseline NIHSS		8.47 (6.49)

Note. NIHSS: National Institute of Health Stroke scale

Table 3

Clinical Scores at 3-Months and 12-Months Post-stroke (n = 157)

	3 Months				12 Months					
	M (SD)	Range	Median (IQR)	25th, 75th percentiles	M (SD)	Range	Median (IQR)	25th, 75th percentiles	M_{diff}	p
NIHSS	1.33 (2.41) *None-very mild stroke symptoms*	0-17	1 (2)	0-2	1.29 (2.78)	0-18	0 (1.75)	0-1	0.64	>.01**
mRS	1.34 (1.26) *None- slight disability*	0-5	1 (2)	1-3	1.25 (1.21)	0-4	1 (2)	0-2	0.14	.08
MADRS	7.86 (7.94) *Mild depressive symptoms*	0-33	5 (10)	1-11	6.50 (6.91)	0-29	4 (10)	0-10.25	0.83	.08
MoCA	24.61 (5.39) *Normal to MCI*	6-31	26 (5.25)	23-28.25	21.13 (8.98)	0-31	26 (7)	21-28	0.96	>.01**

Note. NIHSS: National Institute of Health Stroke scale, mRS: Modified Rankin Scale, MADRS: Montgomery-Åsberg Depression Rating Scale, MoCA: Montreal Cognitive Assessment. Scoring comments are based on descriptors given for scoring ranges based on original assessment.

Paired samples t-tests were conducted on all clinical scores between 3 and 12 months.

** $p < .01$

Table 4

Routine Blood Tests: Details of Acquisition and Counts at 0-12 hours, 12-48 hours and 3-7 days.

		0 -12 hours			24-48 Hours			3-7 Days		
	Reference Ranges[a]	M (SD)	Range	Median (IQR)	M (SD)	Range	Median (IQR)	M (SD)	Range	Median (IQR)
Time from Stroke to Blood Draw		5.12 (1.63) Hrs	1.42-9.18	5.13 (1.97)	29.62 (3.15)	21.25 - 35.75	30.23 (3.50)	3.29 (0.69) Days	1.92 - 5.33	3.28 (0.55)
Immune Function										
White Blood Cell Count (WBC)	4.3 – 10.3 x 10^9/L	8.50 (2.85)	3.60 - 19.40	8.05 (3.28)	9.06 (3.37)	4.10 – 23.00	8.96 (4.08)	8.64 (2.65)	4.10 – 23	8.35 (3.33)
Neutrophil Count (NC)	1.5 – 8.0 x 10^9/L	5.96 (2.68)	1.87 - 15.30	5.3 (2.97)	6.48 (3.30)	1.60 – 21.10	6 (3.45)	6.18 (2.56)	2.20 – 16.40	5.80 (2.62)

[a]Reference ranges are obtained for adults but are slightly lower on average for elderly patients[75].

Details of acquisition, including the average time from stroke onset to blood draw for each timepoint, and counts of WBC and neutrophils at <12 hours, 24-48 hours and 3-7 days are provided in Table 4. WBC and NC showed an increasing pattern from <12 hours to 24-48 hours post-stroke, followed by a trend back towards <12 hour levels at 3-7 days (Figure 1). A repeated measures ANCOVA for NC at <12 hours, 24-48 hours and 3-7 days, with age and stroke severity as covariates, was significant $F(2, 76) = 4.52$, $p = .01$, with post-hocs showing that changes between <12 hours and 24-48 hours were significantly different ($M_{diff} = 0.88$, 95% CI = [-0.38, 1.78], $p = .05$).

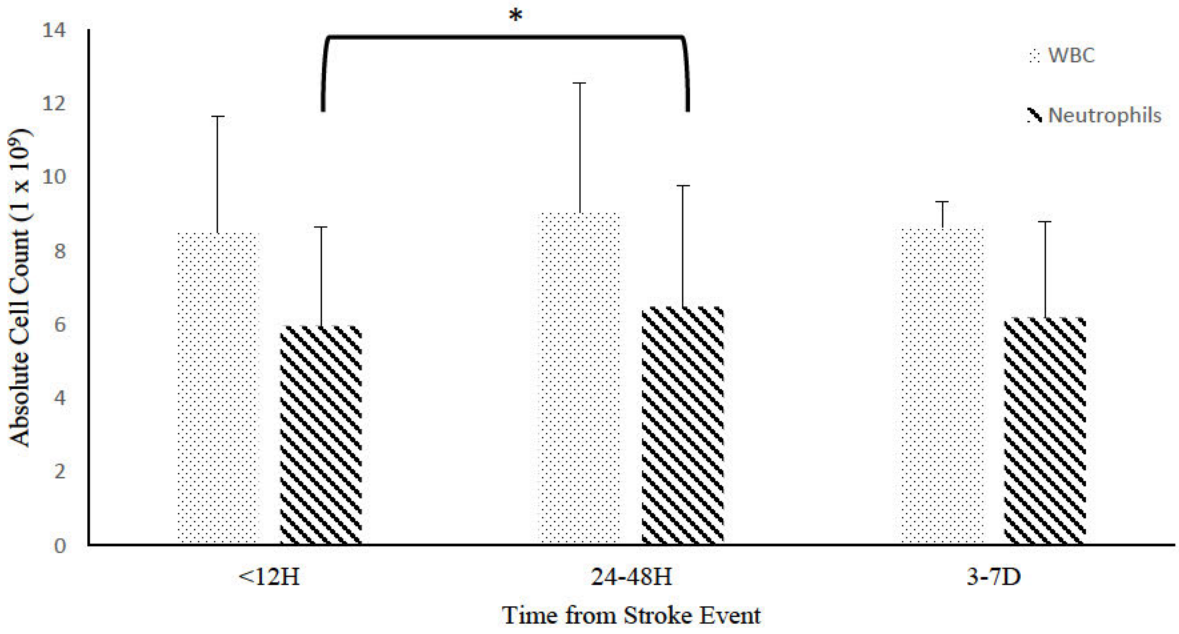

Figure 1. Trend of mean White Blood Cell Count and Neutrophil Count at <12 hours, 24-48 hours and 3-7 days. Error bars represent SD.
*$p < .05$ ANCOVA, covaried for age and baseline stroke severity.

A partial correlation matrix was used to explore the relationship between the early blood tests at <12 hours, 24-48 hours and 3-7 days and clinical measures at 3-months and 12-months, adjusted for age and baseline NIHSS severity score (Table 5). Depression symptoms as measured by the MADRS were not significantly associated with any acute blood WBC or NC counts.

Table 5

Correlation Between Immune Cell Counts (Multiple Acute Times) and Clinical Outcome Scores at 3-Months and 12-Months Post-Stroke (Covaried For Age and Baseline Stroke Severity).

Immune Cell Counts	3 Months			12 Months		
	mRS	MADRS	MoCA	mRS	MADRS	MoCA
White Blood Cell Count – 0-12 hours	0.09	0.18	-0.03	0.12	0.16	0.06
White Blood Cell Count – 24-48 Hours	0.13	0.04	0.26*	0.09	-0.10	0.26*
White Blood Cell Count - 3-7 Days	0.09	0.08	-0.15	0.20*	-0.11	0.01
Neutrophil Count – 0-12 hours	0.09	0.15	-0.02	0.17*	0.13	-0.01
Neutrophil Count – 24-48 Hours	0.08	0.10	0.29*	0.07	-0.09	0.31*
Neutrophil Count - 3-7 Days	-0.03	0.05	-0.18	0.09	-0.07	-0.01

Pearson correlations covaried for age and NIHSS score.
*$p < .05$

As follow-up analyses to the significant correlations, hierarchical regressions were conducted between WBC and NC counts at <12 hours, 24-48 hours and 3-7 days with 3-month (Table 6) and 12-month (Table 7) post-stroke MoCA scores, with age and baseline NIHSS entered as covariates for all models in Step 1. In prediction of 3-month MoCA scores, the regression model with age and baseline NIHSS entered as Step 1, showed medium effect sizes ($R^2 = .20$, $F = (2, 54) = 6.84$ $p = .002$). The addition of 24-48 hour WBC counts ($B = 0.49$, $p = .05$) to Step 1 improved explained variance with a small effect size (R^2 change = .06, $F = (1, 53) = 3.88$, $p = .05$). Similarly, the addition of 24-48 hour NC counts ($B = 0.45$, $p = .04$) to Step 1 improved explained variance with a small effect size (R^2 change = .06, F =

(1, 53) = 4.71, p = .04). In prediction of 12-month MoCA scores, the regression model with baseline NIHSS entered as Step 1 showed large effect sizes (R^2 = .34, F = (2, 56) = 14.70, p = > .001). The addition of 24-48 hour WBC counts (B = 0.74, p = .05) to Step 1 improved explained variance for 12-month MoCA scores with a small effect size (R^2 change = .05, F = (1, 55) = 4.08, p = .048). The addition of 24-48 hour NC counts (B = 0.45, p = .02) to Step 1 improved explained variance with a small effect size (R^2 change = .06, F = (1, 55) = 5.87, p = .02). As expected, age and baseline stroke severity predict a significant portion of 3- and 12-month MoCA scores, generally in the expected direction of cognitive decline. In contrast, given the unstandardized B of regressions models with the addition of immune cell counts, an increase of 1 x 10^9 WBC count and NC at 24-48 hours post-stroke is predictive of an approximate 0.53 increase in MoCA scores at 3- and 12-months.

Table 6

Hierarchical Regressions for White Blood Cell Count and Neutrophil Count at 24-48 hours with 3-month MoCA scores, controlled for by Age and Baseline National Institute of Stroke Scale.

Explanatory Variable	R^2	R^2 Δ	p	*B*	*SE B*	*β*	P
Step 1	.20	-	>.01*				
Age				-0.10	0.05	-0.26	.04*
Baseline NIHSS				-0.23	0.09	-0.31	.02*
Step 2	.26	.06	>.01*				
Age				-0.10	0.05	-0.25	.04*
Baseline NIHSS				-0.31	0.10	-0.42	>.01**
White Blood Cell Count 24-48 hours				0.49	0.25	0.26	>.05*
Step 3	.27	.06	>.01*				
Age				-0.11	0.05	-0.27	.03*
Baseline NIHSS				-0.33	0.10	-0.45	>.00*
Neutrophil Count 24-48 Hours				0.45	0.21	0.29	.04*

*p < .05, **p < .01

Note: *p value is less than 0.05 but has been rounded up for reporting in the table.

Table 7

Hierarchical Regressions for White Blood Cell Count and Neutrophil Count at 24-48 hours with 12-month MoCA scores, controlled for by Age and Baseline National Institute of Stroke Scale.

Explanatory Variable	R^2	$R^2\Delta$	p	*B*	*SE B*	*β*	P
Step 1	0.34	-	>.01**				
Age				-0.14	0.07	-0.21	.06
Baseline NIHSS				-0.61	0.14	-0.50	>.00**
Step 2	0.39	.05	>.01**				
Age				-0.13	0.07	-0.20	.07
Baseline NIHSS				-0.73	0.15	-0.60	>.00**
White Blood Cell Count 24-48 hours				0.74	0.37	0.23	>.05*
Step 3	0.41	.06	>.01**				
Age				-0.14	0.07	-0.22	.04*
Baseline NIHSS				-0.78	0.15	-0.64	>.01**
Neutrophil Count 24-48 Hours				0.73	0.30	0.29	.02*

*$p < .05$, **$p < .01$

Note: **p* value is less than 0.05 but has been rounded up for reporting in the table.

An ordinal regression, with age and baseline NIHSS as covariates, between WBC count at 3-7 days and mRS scores at 12-months found an odds ratio of 1.02, 95% CI = [0.89, 1.18], Wald $\chi2(1) = .09$, which was not significant ($p = .76$). An ordinal regression was conducted with age and baseline NIHSS entered as covariates with NC at <12 hours predicting mRS scores at 12-months found a significant odds ratio of 1.13, 95% CI = [1.00, 1.27], Wald $\chi2(1) = 3.91$, $p = .048$.

Discussion

In this longitudinal cohort study, we have investigated the associations between hospital-based hematological markers of immune function taken within 1 week post-stroke and outcomes of clinical functioning and recovery at 3-months and 12-months post vascular event. A significant increase in NC was found between <12 hours and 24-48 hours. The partial correlations and hierarchical regressions revealed that elevations in post-stroke WBC and NC at 24-48 hours are associated with better cognitive outcome at 3- and 12-months

post-stroke, independent of age and baseline stroke severity. Although the convergence of regression models showed that both WBC and NC are predictive of increased MoCA scores at two separate time points, these data should be interpreted with caution given the relatively weak but consistent effect sizes and small sample sizes. To date, there has only been a single study showing that WBC and/or NC early (within 24-48 hours) are correlated or predictive of better cognition at longitudinal timepoints such as 3- or 12- months post-stroke[47]. Rather, as recently indicated, the majority of research has examined novel biomarkers such as interleukins and apolipoproteins in relation to post-stroke cognitive impairment and recovery (see review[48]).

The significant regression of immune-related white blood cells to cognitive outcomes can be interpreted as evidence for the involvement of immune system activation following cerebral ischemia in acute post-stroke having long-term effects on functional recovery[2]. As shown in the hierarchical regression models, baseline stroke severity and age account for a large amount of variance in relation to longitudinal MoCA scores and are in the expected direction of predicting cognitive decline. Indeed, demographics, functional factors[49] and lesion characteristics[50] often predict worsened cognitive outcomes post-stroke. Interestingly, the addition of 24-48 hour WBC and NC to these models significantly explained additional variance and while the regression models overall were predictive of cognitive decline, WBC and NC as individual predictors were found to be in the unexpected direction of predicting positive cognitive outcomes. Previous studies examining the link between blood-based biomarkers and post-stroke cognition have largely focused on the association of inflammatory markers to impairment[48]. Increases in serum C-reactive protein have been reported to predict lowered global cognitive function within the first month of stroke[51]. Furthermore, prolonged erythrocyte sedimentation rates have been associated with cognitive

decline[52]. However, few studies have described a blood-based biomarker for positive cognitive outcomes; especially where blood-based measures are acquired early in admission and investigated in association with long term cognition.

The biophysiological correlates of worsened post-stroke outcome are usually understood in the context of ongoing peripheral and CNS inflammation[53] and improved outcomes, including cognition, are understood in the context of neural repair and plasticity[54]. Although it is difficult to speculate on the functional neuro-immunomodulation of inflammation and neural repair in the context of gross peripheral leukocyte counts, acute innate immune responses in stroke have been suggested to be both inflammatory and neuroprotective[55] Indeed, more rigorous studies with better defined timing for blood draw and increased sample size representative of more severe stroke cases are required to replicate the positive associations between immune cell counts and cognitive function found here.

Our study did not find any significant correlations between acute immune cell count and depressive symptoms, when adjusted for age and stroke severity, despite previous research suggesting an inflammation and immune based pathogenesis for post-stroke depression[56,57]. The significant prediction of NC at <12 hours to mRS at 12 months is in agreement with previous studies that show an association between increased acute leukocyte counts and poor mRS functional recovery outcomes at both 3 months[58] and 12 months[18]. These findings may appear inconsistent with our results that found an association with increased WBC and NC with positive cognitive outcomes in the same cohort. Although functional recovery might be expected to be associated with cognition, the mRS is perhaps not sensitive enough to adequately capture cognitive function or neuropsychological impairments in stroke patients in comparison to the MoCA[59].

Our results also found that an increase in NC at <12 hours post-stroke, is predictive of poorer functional recovery at 12 months. An increasing NC from 0 to 48 hours post-stroke has been previously reported [17] and may reflect changes in blood-brain barrier (BBB) permeability due to disassembly of the tight junctions by hypoxia-reoxygenation[2] and disruptions to both peripheral and central nervous system immune/inflammatory pathways (see review[60]). Although blood tests at 12 hours post-stroke could not be compared to pre-stroke reference levels, this systemic change is confirmation of several known physiological responses to stroke such as peumbral excitotoxicity and perivascular edema[61] and resulting increased glymphatic perfusion[62]. Indeed, these results may be related to the molecular signaling timeline following neuronal oncosis driven by mitochondrial and associated bioenergetic impairment (ie reduced availability of ATP[63,64]). Indeed acute edema and inflammation following stroke-induced excitotoxity, would be expected to lead to previously activated autophagy and necrosis pathways[65], and be associated with the abrupt upregulation of immune activity that has been predicted to occur around 24-48 hours post-stroke in conjunction with blood-brain barrier (BBB) permeability and be related to peripheral innate immunity as pro-inflammatory cytokines and other endocrine responses influence leukocyte recruitment across a vulnerable BBB[66]. The timeline of leukocyte levels presented here closely resembles dynamic contrast-enhanced MRI defined BBB permeability pattern within the first 3 days of stroke[2]. Previous studies within this acute stage have shown a phasic increase in permeability-surface area product from 6 to 48 hours while subsequently returning to 6 hours levels beyond 48 hours post-stroke[2]. Similarly, the return of WBC and NC to <12 hour levels at 3-7 days may suggest a shift in the pathophysiological cell death cascade, including functions such as cell necrosis, vascular leakage, cerebral edema and maintenance processes such as the clearance of cellular debris, ionic osmoregulation and

restoration of energy processes; processes that have been well described in animal models[5,63]. The apparent reduction in WBC and NC may also be neuroprotective, as persistent antigen-specific immune cell recruitment and activity in the brain is predictive of poorer outcome and mortality[4].

Considering the ionic, biochemical and molecular cascades accompanying stroke damage[67], the leukocyte pattern observed may be reflective of the mild nature of stroke in this cohort and an ongoing trajectory to positive post-stroke recovery that has not been observed in other studies[42]. Our regression analyses of acute immune cells counts predicting increased MoCA scores also suggest that trajectories of later post-stroke recovery could be based on the extent of the early leukocyte response and the subsequent attempt to restore immune cell balance within the first week. The timing of the leukocyte infiltration response may also be indicative of the duration of BBB permeability, vessel occlusion and therefore increased tissue damage, with longer duration strokes showing an earlier leukocyte infiltration[68]. On the contrary, poor post-stroke recovery is commonly followed by continuing immunosuppression that may continue beyond 3 months[5,69]. This trend towards immunosuppression supports the prevalence and timeline of post-stroke infections such as upper respiratory tract infection, pneumonia and urinary tract infections that commonly present towards the end of the first week[70,71]. Indeed, the timeline for immunoreactivity seems to precede optimal trajectories for patients undergoing early physiotherapy rehabilitation at 15.6 days post-stroke, followed by a period of 4 weeks of rapid functional recovery[72].

There are several limitations associated with the methodologies in this study. This study has examined a cross section of stroke patients obtained from a longitudinal stroke cohort study examining thrombolysis outcomes and stroke recovery outcomes: it is not

inclusive of all patients as may be the case in national level data registries[34,35]. The nature of the original inclusion criteria and ability to participate in a year-long study may bias the characteristics of the sample, in particular the mild stroke severity, depressive symptoms and cognitive outcomes. Age and baseline NIHSS are factors that have been shown in previous studies to correlate highly with clinical measures at 3 months[42,73] and thus were selected to adjust for in these analyses. It is acknowledged however that there are other baseline measures such as infection status or atherosclerosis in relationship to WBC and NC that could be addressed as covariates. Furthermore, there may be large variability in the amplitude of immune responses[17] given the wide acute blood sampling times achieved in this study. It is also noted that this study did not collect lymphocyte cell counts that could be used to explore the neutrophil to lymphocyte ratio which has been suggested to be a better measure of acute inflammation than immune cell counts alone[74]. In order to provide more robust evidence for predictive value of immune cell counts in stroke recovery, future studies should examine national or hospital level data registries with similar blood test protocols. Further, sample sizes should be greatly increased, in contrast to the relatively low sample size used in this exploratory study.

Conclusion

Blood tests are routinely prescribed as an investigative tool in acute stroke management but are seldom used as clinical tools to predict post-stroke recovery in the longer term. We have found that increased WBC and neutrophils count within the first week of stroke are consistently independent predictors of better cognitive outcomes at 3- and 12-months. In conjunction with the observed trends in immune cell counts, this data suggests that the degree and timing of the peripheral immune response in the acute phase is likely to play a role in setting the trajectory for long-term recovery. Further studies should be

conducted on acute blood tests and their role in similar long-term clinical outcomes, with the goal of eventual integration into stroke guidelines and best practice.

The Authors declare that there are no conflicts of interest.

References

1. Weerasinghe P, Buja L. Oncosis: An important non-apoptotic mode of cell death. *Exp Mol Pathol.* 2012;93(3):302-308.
2. Merali Z, Huang K, Mikulis D, Silver F, Kassner A. Evolution of blood-brain-barrier permeability after acute ischemic stroke. *PLoS One.* 2017;12(2):e0171558.
3. McCombe PA, Read SJ. Immune and inflammatory responses to stroke: Good or bad? *Int J Stroke.* 2008;3(4):254-265.
4. Becker KJ, Kalil AJ, Tanzi P, et al. Autoimmune responses to brain following stroke are associated with worse outcome. *Stroke.* 2011;42(10):2763-2769.
5. Kamel H, Iadecola C. Brain-immune interactions and ischemic stroke: clinical implications. *Arch Neurol.* 2012;69(5):576-581.
6. Planas AM. Role of immune cells migrating to the ischemic brain. *Stroke.* 2018;49(9):2261-2267.
7. Lappé JM, Horne BD, Shah SH, et al. Red cell distribution width, C-reactive protein, the complete blood count, and mortality in patients with coronary disease and a normal comparison population. *Clin Chim Acta.* 2011;412(23-24):2094-2099.
8. Kristensen M, Iversen AKS, Gerds TA, et al. Routine blood tests are associated with short term mortality and can improve emergency department triage: a cohort study of >12,000 patients. *Scand J Trauma Resusc Emerg Med.* 2017;25:115.
9. Hosseininezhad M, Sohrabnejad R. Stroke mimics in patients with clinical signs of stroke. *Caspian journal of internal medicine.* 2017;8(3):213-216.
10. Boehme AK, Kumar AD, Lyerly MJ, et al. Persistent leukocytosis—is this a persistent problem for patients with acute ischemic stroke? *J Stroke Cerebrovasc Dis.* 2014;23(7):1939-1943.

11. Hawkins RC, Badrick T. Reporting unit size and measurement uncertainty: current Australian practice in clinical chemistry and haematology. *Pathology.* 2015;47(5):462-465.

12. Nardi K, Milia P, Eusebi P, Paciaroni M, Caso V, Agnelli G. Admission leukocytosis in acute cerebral ischemia: influence on early outcome. *J Stroke Cerebrovasc Dis.* 2012;21(8):819-824.

13. Zhou J, Wu J, Zhang J, et al. Association of stroke clinical outcomes with coexistence of hyperglycemia and biomarkers of inflammation. *J Stroke Cerebrovasc Dis.* 2015;24(6):1250-1255.

14. Fulop T, Larbi A, Dupuis G, et al. Immunosenescence and inflamm-aging as two sides of the same coin: friends or foes? *Front Immunol.* 2018;8:1960.

15. Kammersgaard LP, Jørgensen HS, Nakayama H, Reith J, Raaschou HO, Olsen TS. Leukocytosis in acute stroke: Relation to initial stroke severity, infarct size, and outcome: The Copenhagen stroke study. *J Stroke Cerebrovasc Dis.* 1999;8(4):259-263.

16. Haeusler KG, Schmidt WUH, Föhring F, et al. Cellular immunodepression preceding infectious complications after acute ischemic stroke in humans. *Cerebrovasc Dis.* 2008;25(1-2):50-58.

17. Petrone AB, Eisenman RD, Steele KN, Mosmiller LT, Urhie O, Zdilla MJ. Temporal dynamics of peripheral neutrophil and lymphocytes following acute ischemic stroke. *Neurol Sci.* 2019;40(9):1877-1885.

18. Quan K, Wang A, Zhang X, Wang Y. Leukocyte count and adverse clinical outcomes in acute ischemic stroke patients. *Front Neurol.* 2019;10(1240).

19. Vogelgesang A, Witt C, Heuer C, et al. Clinical improvement following stroke promptly reverses post-stroke cellular immune alterations. *Front Neurol.* 2019;10(414).

20. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: Consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Neurorehabilitation and Neural Repair.* 2017;31(10-11):864-876.

21. Banks JL, Marotta CA. outcomes validity and reliability of the modified rankin scale: implications for stroke clinical trials. *Stroke.* 2007;38(3):1091.

22. Rohde D, Gaynor E, Large M, et al. Stroke survivor cognitive decline and psychological wellbeing of family caregivers five years post-stroke: a cross-sectional analysis. *Top Stroke Rehabil.* 2019;26(3):180-186.

23. Chen H, Luan X, Zhao K, et al. The association between neutrophil-to-lymphocyte ratio and post-stroke depression. *Clin Chim Acta.* 2018;486:298-302.

24. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 guidelines for the early management of patients with acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* 2018.

25. Leonard BE. The concept of depression as a dysfunction of the immune system. *Curr Immunol Rev.* 2010;6(3):205-212.

26. Howren MB, Lamkin DM, Suls J. Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosom Med.* 2009;71(2):171-186.

27. Dantzer R. Cytokine, sickness behavior, and depression. *Immunol Allergy Clin North Am.* 2009;29(2):247-264.

28. Taniguchi Y, Shinkai S, Nishi M, et al. Nutritional biomarkers and subsequent cognitive decline among community-dwelling older japanese: A prospective study. *The Journals of Gerontology: Series A.* 2014;69(10):1276-1283.

29. Kao T-W, Chang Y-W, Chou C-C, Hu J, Yu Y-H, Kuo H-K. White blood cell count and psychomotor cognitive performance in the elderly. *Eur J Clin Invest.* 2011;41(5):513-520.

30. Todd MA. Inflammation and cognition in older adults: Evidence from Taiwan. *Biodemography Soc Biol.* 2017;63(4):309-323.

31. Becker KJ, Tanzi P, Zierath D, Buckwalter MS. Antibodies to myelin basic protein are associated with cognitive decline after stroke. *J Neuroimmunol.* 2016;295-296:9-11.

32. Sagen U, Vik TG, Moum T, Mørland T, Finset A, Dammen T. Screening for anxiety and depression after stroke: Comparison of the Hospital Anxiety and Depression Scale and the Montgomery and Åsberg Depression Rating Scale. *J Psychosom Res.* 2009;67(4):325-332.

33. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53(4):695-699.

34. Carey LM, Crewther S, Salvado O, et al. STroke imAging pRevention and Treatment (START): A longitudinal stroke cohort study: Clinical trials protocol. *Int J Stroke.* 2013;10(4):636-644.

35. Ma H, Parsons MW, Christensen S, et al. A multicentre, randomized, double-blinded, placebo-controlled phase III study to investigate Extending the Time for

Thrombolysis in Emergency Neurological Deficits (EXTEND). *Int J Stroke.* 2011;7(1):74-80.

36. Lyden P, Raman R, Liu L, Emr M, Warren M, Marler J. NIHSS certification is reliable across multiple venues. *Stroke.* 2009;40(7):2507-2511.

37. Williams JB, Kobak KA. Development and reliability of a structured interview guide for the Montgomery-Åsberg Depression Rating Scale (SIGMA). *The British Journal of Psychiatry.* 2008;192(1):52-58.

38. Fantino B, Moore N. The self-reported Montgomery-Åsberg depression rating scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry.* 2009;9:26-26.

39. Welin L, Bjälkefur K, Roland I. Open, randomized pilot study after first stroke. *Stroke.* 2010;41(7):1555-1557.

40. Blackburn DJ, Bafadhel L, Randall M, Harkness KA. Cognitive screening in the acute stroke setting. *Age and Ageing.* 2013;42(1):113-116.

41. Chiti G, Pantoni L. Use of Montreal Cognitive Assessment in patients with stroke. *Stroke.* 2014;45(10):3135.

42. Furlan JC, Vergouwen MDI, Fang J, Silver FL. White blood cell count is an independent predictor of outcomes after acute ischaemic stroke. *Eur J Neurol.* 2014;21(2):215-222.

43. Wu T-H, Chien K-L, Lin H-J, et al. Total white blood cell count or neutrophil count predict ischemic stroke events among adult Taiwanese: report from a community-based cohort study. *BMC Neurology.* 2013;13:7-7.

44. Fonarow GC, Saver JL, Smith EE, et al. Relationship of national institutes of health stroke scale to 30-day mortality in medicare beneficiaries with acute ischemic stroke. *Journal of the American Heart Association.* 2012;1(1):42-50.

45. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;134(4):382-389.

46. McLennan SN, Mathias JL, Brennan LC, Stewart S. Validity of the Montreal Cognitive Assessment (MoCA) as a screening test for mild cognitive impairment (MCI) in a cardiovascular population. *J Geriatr Psychiatry Neurol.* 2011;24(1):33-38.

47. Tsai AS, Berry K, Beneyto MM, et al. A year-long immune profile of the systemic response in acute stroke survivors. *Brain.* 2019;142(4):978-991.

48. Wahul AJ, Pranav, Kumar A, Chakravarty S. Association of diagnostic stroke biomarkers with post stroke cognitive impairment. *Journal of Neurological Disorders & Stroke.* 2018;6(1):1134-1140.

49. Zulkifly M, Faizal M, Ghazali SE, Che Din N, Singh DKA, Subramaniam P. A review of risk factors for cognitive impairment in stroke survivors. *The scientific world journal.* 2016;2016.

50. Yatawara C, Ng KP, Chander R, Kandiah N. Associations between lesions and domain-specific cognitive decline in poststroke dementia. *Neurology.* 2018;91(1):e45.

51. Rothenburg LS, Herrmann N, Swardfager W, et al. The relationship between inflammatory markers and post stroke cognitive impairment. *Journal of Geriatric Psychiatry and Neurology.* 2010;23(3):199-205.

52. Kliper E, Bashat DB, Bornstein NM, et al. Cognitive decline after stroke: relation to inflammatory biomarkers and hippocampal volume. *Stroke.* 2013;44(5):1433-1435.

53. Barrington J, Lemarchand E, Allan SM. A brain in flame; do inflammasomes and pyroptosis influence stroke pathology? *Brain Pathol.* 2017;27(2):205-212.

54. Carey L, Walsh A, Adikari A, et al. Finding the intersection of neuroplasticity, stroke recovery, and learning: scope and contributions to stroke rehabilitation. *Neural Plast.* 2019;2019.

55. Amantea D, Greco R, Micieli G, Bagetta G. Paradigm shift to neuroimmunomodulation for translational neuroprotection in stroke. *Front Neurosci.* 2018;12(241).

56. Feng C, Fang M, Liu X-Y. The neurobiological pathogenesis of poststroke depression. *The Scientific World Journal.* 2014;2014:521349.

57. Pascoe MC, Crewther SG, Carey LM, Crewther DP. Inflammation and depression: why poststroke depression may be the norm and not the exception. *Int J Stroke.* 2011;6(2):128-135.

58. Pagram H, Bivard A, Lincz LF, Levi C. Peripheral immune cell counts and advanced imaging as biomarkers of stroke outcome. *Cerebrovascular diseases extra.* 2016;6(3):120-128.

59. Rautalin IM, Sebök M, Germans MR, et al. Screening tools for early neuropsychological impairment after aneurysmal subarachnoid hemorrhage. *Neurol Sci.* 2019.

60. Brouns R, De Deyn PP. The complexity of neurobiological processes in acute ischemic stroke. *Clin Neurol Neurosurg.* 2009;111(6):483-495.

61. Zador Z, Stiver S, Wang V, Manley GT. Role of aquaporin-4 in cerebral edema and stroke. *Handb Exp Pharmacol.* 2009(190):159-170.

62. Gaberel T, Gakuba C, Goulay R, et al. Impaired glymphatic perfusion after strokes revealed by contrast-enhanced MRI. *Stroke.* 2014;45(10):3092.

63. Santos Samary C, Pelosi P, Leme Silva P, Rieken Macedo Rocco P. Immunomodulation after ischemic stroke: potential mechanisms and implications for therapy. *Critical Care.* 2016;20(1):391.

64. Liesz A, Dalpke A, Mracsko E, et al. DAMP signaling is a key pathway inducing immune modulation after brain injury. *The Journal of Neuroscience.* 2015;35(2):583-598.

65. Chen W, Sun Y, Liu K, Sun X. Autophagy: a double-edged sword for neuronal survival after cerebral ischemia. *Neural Regeneration Research.* 2014;9(12):1210-1216.

66. Yilmaz G, Granger DN. Leukocyte recruitment and ischemic brain injury. *Neuromolecular Med.* 2010;12(2):193-204.

67. Xing C, Arai K, Lo EH, Hommel M. Pathophysiologic cascades in ischemic stroke. *International journal of stroke : official journal of the International Stroke Society.* 2012;7(5):378-385.

68. Grønberg NV, Johansen FF, Kristiansen U, Hasseldam H. Leukocyte infiltration in experimental stroke. *J Neuroinflammation.* 2013;10:115-115.

69. Nguyen VA, Carey LM, Giummarra L, et al. A pathway proteomic profile of ischemic stroke survivors reveals innate immune dysfunction in association with mild symptoms of depression – a pilot study. *Front Neurol.* 2016;7:85.

70. Emsley HCA, Hopkins SJ. Acute ischaemic stroke and infection: recent and emerging concepts. *The Lancet Neurology.* 2008;7(4):341-353.

71. Grabska K, Gromadzka G, Członkowska A. Infections and ischemic stroke outcome. *Neurol Res Int.* 2011;2011.

72. Lee KB, Lim SH, Kim KH, et al. Six-month functional recovery of stroke patients: a multi-time-point study. *International journal of rehabilitation research Internationale Zeitschrift fur Rehabilitationsforschung Revue internationale de recherches de readaptation.* 2015;38(2):173-180.

73. Huber T, Kleine JF, Kaesmacher J, et al. Blood leukocytes as prognostic parameter in stroke thrombectomy. *Cerebrovasc Dis.* 2016;42(1-2):32-40.

74. Xue J, Huang W, Chen X, et al. Neutrophil-to-lymphocyte ratio is a prognostic marker in acute ischemic stroke. *J Stroke Cerebrovasc Dis.* 2017;26(3):650-657.

75. Aminzadeh Z, Parsa E. Relationship between age and peripheral white blood cell count in patients with sepsis. *Int J Prev Med.* 2011;2(4):238-242.

Chapter 6

Experimental Chapter 2

A Pathway Proteomic Profile of Ischemic Stroke Survivors Reveals Innate Immune Dysfunction in Association with Mild Symptoms of Depression – A Pilot Study

Preface

The study reported in Chapter 6 aimed to identify the link between biological systems in the acute period and measures of longitudinal clinical outcome, specifically depression. Depression was chosen as the outcome given the clinical significance of depression and the sizable literature that links the behavioural and neurological presentations of depression to immune functioning. Furthermore, due to the chronic perturbations in immune system functioning follow focal neurological and incipient neurodegeneration in stroke, post-stroke depression is a likely occurrence (Pascoe et. al 2012).

This study utilized an untargeted 'hypothesis-free' discovery approach that has shown to be useful where there is a large amount of data available and the biological context of the disease being studied is not well understood. Firstly, the untargeted proteomic approach employed here analyzed plasma samples based on the ability of the chemistry approach (including sample preparation methods and chromatography) to separate the samples into molecules and the resolution of the mass spectrometer to detect those molecules. This is in contrast to traditional biology techniques for protein elucidation that are generally aimed at quantifying certain molecules by use of antibodies.

Secondly, the Gene Set Enrichment Analysis statistical analysis provides a link between the data obtained in this study to multiple well defined databases of biological systems and their constituent molecules while concurrently limiting the ability to examine the molecule-to-molecule interactions as would be expected in previous biochemical pathway technology. Given the relatively new use of the technique at the time, this study was deemed as a pilot study for both the mass spectrometry and bioinformatics approaches.

This chapter was published as Nguyen, V. A., Carey, L. M., Giummarra, L., Faou, P., Cooke, I., Howells, D. W., … Crewther, S. G. (2016). A Pathway Proteomic Profile of Ischemic Stroke Survivors Reveals Innate Immune Dysfunction in Association with Mild Symptoms of Depression - A Pilot Study. Frontiers in Neurology, 7, 85.

A Pathway Proteomic Profile of Ischemic Stroke Survivors Reveals Innate Immune Dysfunction in Association with Mild Symptoms of Depression – A Pilot Study

Nguyen, V. A[1,2,3]., Carey, L. M[1,2*]., Giummarrra, L[3]., Faou, P[4]., Cooke, I[4]., Howells, D. W[5]., Tse, T[1,2]., Macaulay, S. L[6]., Ma, H[7,8]., Davis, S[7,8]., Donnan, G[7,10]., Crewther, S. G[2,3*].

[1]Occupational Therapy, School of Allied Health, College of Science health and Engineering, La Trobe University, Melbourne, Victoria, Australia
[2]Neurorehabilitation and Recovery, Stroke, The Florey Institute of Neuroscience and Mental Health, Melbourne, Victoria, Australia
[3]School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia
[4]School of Molecular Sciences, La Trobe University, Melbourne
[5]School of Medicine, University of Tasmania, Hobart, Australia
[6]Commonwealth Science and Industrial Research Organisation (CSIRO), Melbourne, Victoria, Australia
[7]The Florey Institute of Neuroscience and Mental Health, Parkville, Victoria, Australia
[8]Monash University, Clayton, Victoria, Australia
[9]Department of Medicine, Melbourne Brain Centre, Royal Melbourne Hospital, Melbourne, Victoria, Australia
[10]University of Melbourne, Parkville, Victoria, Australia

*Joint senior authors

Acknowledgements

We acknowledge financial support for conduct of the research from the Commonwealth Scientific and Industrial Research Organization (CSIRO) of Australia, Flagship Collaboration Fund through the Preventative Health Flagship; data collection and project planning by the Stroke Imaging, Prevention and Treatment research team (www.START.csiro.au); support for analysis, write up and/or researchers from the James S. McDonnell Foundation 21st Century Science Initiative in Cognitive Rehabilitation - Collaborative Award (# 220020413); Victorian Government's Operational Infrastructure Support Program; an Australian Research Council Future Fellowship awarded to LMC [#FT0992299]; and a La Trobe University Post Graduate Scholarship supported by the Understanding Diseases Research Focus Group at La Trobe University.

Correspondence: Professor Sheila Crewther, School of Public Health and Psychology, La Trobe University, Bundoora, Victoria, Australia. s.crewther@latrobe.edu.au

Abstract

Depression after stroke is a common occurrence, raising questions as to whether depression could be a long term biological and immunological sequela of stroke. Early explanations for post stroke depression (PSD) focused on the neuropsychological/psychosocial effects of stroke on mobility and quality of life. Recent investigations however have revealed imbalances of inflammatory cytokine levels in association with PSD, though to date there is only one published proteomic pathway analysis testing this hypothesis. Thus we examined the serum proteome of stroke patients (n = 44, mean age = 63.62 years) and correlated these with the Montgomery-Asberg Depression Rating Scale (MADRS) scores at 3 months post-stroke. Overall, the patients presented with mild depression symptoms on the MADRS, M = 6.40 (SD = 7.42). A discovery approach utilizing label-free relative quantification was employed utilizing an LC-ESI-MS/MS coupled to a LTQ-Orbitrap Elite (Thermo-Scientific). Identified peptides were analyzed using the Gene Set Enrichment (GSEA) approach on several different genomic databases that all indicated significant downregulation of the complement and coagulation systems with increasing MADRS scores. Complement and coagulation systems are traditionally thought to play a key role in the innate immune system and are established precursors to the adaptive immune system through pro-inflammatory cytokine signaling. Both systems are known to be globally affected after ischemic or hemorrhagic stroke. Thus our results suggest that lowered complement expression in the periphery in conjunction with depressive symptoms post stroke may be a biomarker for incomplete recovery of brain metabolic needs, homeostasis and inflammation following ischemic stroke damage. Further proteomic investigations are now required to construct the temporal profile leading from acute lesion damage to manifestation of depressive symptoms. Overall, the findings provide support for the involvement of inflammatory and

immune mechanisms in post stroke depression symptoms and further demonstrates the value and feasibility of the proteomic approach in stroke research.

Keywords

Ischemic stroke, proteomics, post stroke depression, complement and coagulation, immunity and inflammation, stroke neurological recovery, blood biomarkers

Glossary

ApoA4, ApoC2, APOE: Apolipoproteins A4, C2, E

BBB: Blood brain barrier

C6, C3, C3a, C5a: Complement component 6, 3, 3 a, 5a

EDTA: ethylenediaminetetraacetic acid

ESI: Electrospray Ionization

FDR: False Discovery Rate

GO: Gene Ontology

GSEA: Gene Set Enrichment Analysis

GSN: Gelsolin

ITRAQ: Isobaric tagging for relative and absolute quantification

KEGG: Kyoto Encyclopedia of Genes and Genomes

LC: Liquid Chromatography

LFQ: Label-free quantification

MADRS: Montgomery-Asberg Depression Rating Scale

MAC: Membrane attack complex

MS, MS/MS: Mass spectrometry, tandem mass spectrometry

MBL2: Mannan-binding lectin 2

MoCA: Montreal Cognitive Assessment

mRS: modified Rankin Scale

NES: Normalized enrichment score

NIHSS: National Institute of Health Stroke Scale

PHQ-9, PHQ-2: Patient Health Questionnaire 9, 2

PrePARE: Prediction and Prevention to Achieve optimal Recovery Endpoints

PSD: Post stroke depression

RNA: Ribonucleic acid

S100A12: S100 calcium binding protein A12

SIGMA: Structured Interview Guide for the Montgomery-Asberg Depression Rating Scale

SLE: Systemic lupus erythematous

START: Stroke Imaging, Prevention and Treatment longitudinal stroke cohort study

TF: Transferrin

TIA: Transient ischemic attack

TOAST: Trial of ORG 10172 in acute treatment

tPA: Tissue plasminogen activator

TLR4: Toll-like receptor 4

Introduction

Over 15 million people worldwide experience a stroke each year; 5 million of those events are fatal and another 5 million people are left with a permanent disability (Corbyn, 2014). Previous epidemiological reviews have concluded that approximately 30% of stroke survivors are likely to experience post stroke depression (PSD) (Hackett, Yapa, Parag, & Anderson, 2005; Whyte & Mulsant, 2002). Prevalence of depression is stroke survivors is reported to peak at 3 months post-stroke based on testing using the Diagnostic and Statistical Manual V (DSM-V; American Psychiatric Association) (Paolucci, 2008). Studies have also characterized depression symptoms on the same criteria as early as 15 days and as late as 12 months post-stroke (Paolucci, 2008). Stroke patients with PSD show poorer functional and recovery outcomes compared to patients not suffering depression (Ahn, Lee, Jeong, Kim, & Park, 2015; Schmid et al., 2011). PSD is also found to contribute to poorer quality of life and increased mortality rates (Ahn et al., 2015; Corbyn, 2014) highlighting the need for patient care that extends beyond that of physical and cognitive rehabilitation. Investigation of biomarkers linked with underlying mechanisms has potential to guide the targeting of therapy, both prevention and treatment.

Currently, there is no consensus regarding the aetiology of PSD with much debate as to the extent to which PSD stems from a purely biological origin and/or more likely incorporates elements of psychosocial response (Pascoe, Crewther, Carey, & Crewther, 2011; Schmid et al., 2011). Following the biological argument for PSD progression, it was first proposed that the location of the stroke lesion could predict PSD presentation (Robinson & Szetela, 1981). This has been long debated without resolution (Vataja et al., 2014). More recently, basal ganglia or frontal lobe lesions (Schmid et al., 2011), white matter hyperintensities (Tang et al., 2010) and interruption of connecting pathways (Mitra et al., 2014) have been linked to PSD. Furthermore, a recent review suggests that right

hemisphere stroke predicts PSD incidence in the sub-acute, 1-6 month post-stroke period (Zubarev, 2013). Other explanations for correlations between PSD symptoms and stroke lesions have suggested that the overall balance of monoamines such as serotonin, dopamine and norepinephrine are disrupted following cerebrovascular damage (Loubinoux et al., 2012). These same amines have previously been associated with depression, consistent with a link between stroke and depressive symptoms (Robinson, Shoemaker, Schlumpf, Valk, & Bloom, 1975).

Other reviews and meta-analyses have failed to find clear evidence for lesion location as a risk factor for PSD (Paolucci, 2008; Whyte & Mulsant, 2002). For instance, similar incidence rates of comorbid depression can be seen in other cerebrovascular diseases such as vascular dementia in the elderly where brain lesions are common (Perini et al., 2001; Vu & Aizenstein, 2013). Surprisingly, prevalence rates of depression are also similar in patients who have been subject to transient ischemic attacks (TIA) (Broomfield, Quinn, Abdul-Rahim, Walters, & Evans, 2014; El Husseini et al., 2012; Luijendijk et al., 2011) or carotid artery stenosis patients (Rao, Jackson, & Howard, 2001) where there is little to no presentation of lesions visible on computerized axial tomography or magnetic resonance imaging. In TIA especially, prevalence of PSD is comparable to stroke even at 12 months post-stroke (El Husseini et al., 2012). Thus, while currently detectable brain lesions may contribute to the etiology of PSD, they are unlikely to be the primary cause with evidence suggesting other pathogenesis involving immune disruption and circulating cytokines long after the stroke (Spalletta et al., 2006).

It is possible that PSD could be viewed as depression resulting firstly from damage caused by the initial ischemia and reperfusion injury (Vogelgesang et al., 2010) and exacerbated by psychosocial issues such as anxiety, loss of confidence and apathy associated with even mild loss of mobility or functioning (Hama, Yamashita, Yamawaki, & Kurisu, 2011). In line with this view, the cytokine hypothesis proposes a role for pro-

inflammatory cytokines such as tumor necrosis factor α, interleukin-1 (IL-1), IL-6, IL-8 and anti-inflammatory IL-10. This is consistent with evidence that pro-inflammatory cytokines have been found to be increased in both depressive mood (Dowlati et al., 2010; Galea & Brough, 2013) and cerebrovascular events (J.-M. Kim et al., 2011; Noonan, Carey, & Crewther, 2013; Pascoe et al., 2011; Perini et al., 2001). Pro-inflammatory cytokine levels are elevated during and long after the therapeutic window of typical tissue plasminogen activator treatment (tPA; <4.5h) (Lambertsen, Biber, & Finsen, 2012; Perini et al., 2001; Worthmann et al., 2010). Recent immunoassay research has established that cytokine levels often remained significantly elevated at one year post stroke, leading the authors to suggest that the homeostatic balance of pro and anti-inflammatory cytokines may be disrupted in the long term (Su, Chou, Tsai, & Hung, 2012). Elevated pro-inflammatory cytokine expression is also likely to influence glucocorticoid resistance and sensitivity which is reported to lead to overactivation of the hypothalamus-pituitary-adrenal (HPA) axis thereby inducing depressive mood (Leonard & Maes, 2012; Maes, 2011; Zunszain, Anacker, Cattaneo, Carvalho, & Pariante, 2011). Furthermore, persistently high concentrations of pro-inflammatory cytokine levels may provide functional markers for long term impaired immunological responses as the body attempts to resolve the lesion induced neurodegeneration.

Given the multifaceted nature of PSD, it is difficult to ascertain the underlying molecular basis of stroke damage associated with depressive symptoms at the 3 month peak prevalence period in human patients. In recent years, proteomics has provided a powerful platform for pathology research. Downstream from transcriptomic processes, the proteome describes the totality of proteins that can be produced by the organism's genome. The complexity of the proteome accounts for the biological functioning of an organism and as such includes not only the primary function of the protein itself but also how it fits into the biological environment that is further determined by protein-protein

interactions, post translational modifications (Wetie, Woods, & Darie, 2014) and protein degradation. With the addition of alternative splicing to these processes, the 25,000 genes in the human genome can theoretically encode up to 1,000,000 different proteins (P. Sharma, Cosme, & Gramolini, 2013). To this end, proteomics approaches coupled with bioinformatics have proven useful in identifying potential therapeutic biomarkers in pathologies including cancer (Koomen et al., 2008) and cardiovascular disease (P. Sharma et al., 2013). In the case of proteomic disease research, biomarkers refer to a subset of proteins that are either upregulated, downregulated, activated or deactivated and are detectable using a particular methodology in a disease phenotype.

The proteomic composition of a biological organism can reflect not only rapid changes in response to external challenges (Giardina, Stanley, & Chiang, 2014; Pavlov & Ehrenberg, 2013) but also can be studied to understand disease progression and biological changes in the long term (Kumar et al., 2012; Wildsmith et al., 2014). In stroke, the damage caused by the event would be expected to show both acute and later chronic consequences. A discovery proteomics approach utilizing a mass spectrometer with high dynamic range in order to assay as much of the human proteome as possible is therefore warranted in the context of proteome complexity and the multifactorial nature of PSD aetiology. Previous studies adopting this approach have been successful in providing new insights and support for the interpretation of molecular biological mechanisms following a stroke (Ning, Lopez, Cao, Buonanno, & Lo, 2012). The majority of these studies focused on revealing biomarkers that might be useful aids to rapid diagnosis, for example in the case of differentiation between hemorrhagic stroke versus ischemic stroke (Lopez et al., 2012) and ischemic stroke versus healthy controls (R. Sharma et al., 2015). Research has also been successful in characterizing the treatment effects of tPA (Ning et al., 2010) and electroacupuncture (S. Pan et al., 2011) on the human proteome,

demonstrating that proteomics approaches can objectively characterize the molecular changes associated with certain treatments.

A review of the literature up to December 2015 revealed only a single published study examining the proteomic profile of PSD. Isobaric tagging for relative and absolute quantification (iTRAQ) was employed by Zhan and colleagues (2014) to compare the ethylenediaminetetraacetic acid (EDTA) anticoagulated blood of patients with stroke and PSD, stroke without PSD, and healthy controls. The iTRAQ approach involves isobaric tagging of tryptic digested samples based on phenotypes of interest for comparison through commercially available iTRAQ reagents, and thus facilitates proteome to proteome comparisons of disease versus control groups (Pottiez, Wiederin, Fox, & Ciborowski, 2012). This is conducted by introducing a pooled mixture of all the samples to the mass spectrometer whereby relative quantitation can be assessed by comparing differences in the peak intensity between labeled peptides (Ross et al., 2004); an approach well suited for investigations of disease phenotypes. The Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics tool for gene enrichment revealed that of the peptides that statistically differentiated stroke patients with and without PSD, only the complement and coagulation cascade pathway accounted for the clustering of identified proteins. Given that this research was conducted on patients at one month post stroke, it was suggested that these findings indicate a homeostatic imbalance of pro-inflammatory and anti-inflammatory processes in PSD (Zhan et al., 2014). These findings are in line with the principles of the cytokine hypothesis of PSD albeit complement pathways are more reliably identifiable in blood as biomarkers of immune disruption (Janeway, et al., 2001;Pekna & Pekny, 2012) . Zhan and colleagues (2014) also further examined the top protein candidates by protein immunoblotting of apolipoproteins A IV (ApoA-4) and C-II (ApoC-2), C-reactive protein (CRP), gelsolin, haptoglobin and leucine rich glycoprotein (LRG). All of these proteins were significantly altered in PSD

patients compared to stroke without depression. As there are limited data on the biological processes directly involved in PSD, Zhan et al.'s (2014) study represents a successful first step in applying proteomics approaches to PSD, a disorder that has been previously difficult to characterize in molecular terms.

As the molecular basis of depression after stroke is not well understood, we aimed to take a discovery approach to identifying molecular pathways impacted by depression in stroke survivors. We sought to use a proteomics technique that is capable of identifying and quantifying a large number of biological entities. Our research aim was to identify and quantify peptides to be used to understand the biological mechanisms associated with post stroke depression symptoms. Thus, a data-driven discovery proteomics methodology was used to investigate the biological mechanisms associated with depression aetiology 3 months post stroke. Given the recent finding of Zahn et al (2014) we also sought to investigate whether this finding could be replicated using a different proteomics and analytic approach to that previously reported, in order to strengthen the validity of the proteomic profile of PSD. Our analysis was conducted at 3 months post stroke.

Firstly, a label free quantitation (LFQ) technique was employed to examine the blood samples of stroke patients. The LFQ technically differs from the iTRAQ approach in that the protein samples are not pooled or tagged for analysis and each sample generates its own proteomic profile (Neilson et al., 2011). The relationship between protein expression and various clinical measurements can then be explored instead of comparisons being planned *a priori,* thus conforming to a discovery approach better suiting our aim. Furthermore, it is debatable that iTRAQ provides better quantitation than LFQ approaches, with LFQ shown to perform similarly (Wang, Alvarez, & Hicks, 2012) or more accurately (Trinh et al., 2013) than iTRAQ approaches.

Secondly, blood serum was chosen compared to EDTA anticoagulated plasma. To date, there has been little or no research published comparing the differences between serum and plasma blood samples in stroke proteomics. EDTA has been shown to be the least proteolytically active of the plasma samples and has previously been shown to reveal low abundance proteins in healthy patients (Jambunathan & Galande, 2014). However, the proteomic and secretomic profile of serum may be more suitable for profiling of coagulation and complement systems (Atala, 2002). Furthermore, in a small batch optimization analysis of these bloods, we found overall increased expression of coagulation and complement in serum compared to EDTA, as expected (Yi, Kim, & Gelfand, 2007).

Thirdly, biological pathways and mechanisms were ascertained using the Gene Set Enrichment Analysis (GSEA) statistical and bioinformatics tool. This tool compares the current protein expression data from our study to those from the Molecular Signatures Database (MSIGDB) (Subramanian et al., 2005). The approach of DAVID considers gene annotation clustering after statistical procedures have identified significant differences in protein expression (Huang, Sherman, & Lempicki, 2008). The GSEA approach however develops an enriched gene set from the original expression data and compares them to sets from the databases on MSIGDB (Subramanian et al., 2005). Using pathway analysis affords greater explanatory power in highlighting relationships between gene sets and phenotypes that might go unnoticed in a comparison of individual proteins. While this is inherently a technique for genes and genomic profiling, the statistical approach of database list ranking in GSEA has been demonstrated to be applicable to mass spectrometry (MS) based proteomics (Käll & Vitek, 2011; Schmidt, Forne, & Imhof, 2014; Sridhar et al., 2008). The databases cannot yet interpret the full complexity of the molecular proteome as they do not consider protein/protein interactions and post-translational modifications. However, despite these limitations, GSEA represents one of

the best tools currently available for understanding cellular pathways (Wu, Hasan, & Chen, 2014). Although the sample sizes in this study are small, this exploratory study aims to show that the proteomics approach is a reliable and viable method that can generate hypotheses and increase understanding of PSD and stroke pathophysiology.

Subjects and Method

Subjects

A subset of 44 stroke patients (30 female, 14 male) were obtained consecutively and prospectively from the START_PrePARE (STroke imAging pRevention and Treatment_ Prediction and Prevention to Achieve optimal Recovery Endpoints) cohort: a longitudinal stroke cohort study with advanced clinical and neuroimaging data conducted in Australia (Carey et al., 2015). The participants were recruited following a first ischemic stroke and were over 18 years of age. A diagnosis of ischemic stroke was determined by an experienced neurologist using clinical assessments supplemented with computerized tomography (CT) or magnetic resonance imaging (MRI). Further information on inclusion and exclusion criteria for this study can be found in protocol papers for START_PrePARE (Carey et al., 2015) and START_EXTEND (Ma et al, 2012). Patients included were selected consecutively from the prospectively collected START_PrePARE cohort as patients were being recruited. Patients were not excluded if they had a prior history of depression, as our focus was to identify biological factors associated with the presence of depression in stroke survivors at a particular point in time, irrespective of whether depression was present prior to or post stroke (Carey et al., 2015). Inclusion of patients with prior history of depression was also important to improve ecological validity and permit generalization of findings to clinical populations, given prior history of depression in stroke patients.

Design

The data presented in this manuscript was obtained primarily at 3 months post-stroke (± 7 days), when clinical data and blood were collected. Ethics was approved by the Human Research Ethics Committee of Austin Hospital, Heidelberg (HREC code: H2010/03588) and relevant university and hospital sites. Further details on study design are provided in related protocol papers for START_PrePARE (Carey et al., 2015) and START_EXTEND (Ma et al, 2012).

Clinical Assessments

The National Health Institute Stroke Severity Scale (NIHSS) was administered by a trained neurologist or health care professional. This measure is designed to be a test of patient neurological status and correlates highly with stroke severity (Brott et al., 1989). Prior history of depression was obtained using the 2 item Patient Health Questionnaire (PHQ-2) (Kroenke, Spitzer, & Williams, 2003) at 3-7 (±1) days post-stroke and again at 12 months (±7 days) post-stroke. The PHQ-2 is a brief screening tool that employs the first two questions of the 9 item Patient Health Questionnaire (PHQ-9) (Kroenke, Spitzer, & Williams, 2001), and shares high correlation, interrater reliability and internal consistency with the PHQ-9 (Janneke et al., 2012).

Depressive symptoms were assessed at 3 months post-stroke using the Montgomery-Asberg Depression Rating Scale (MADRS). The MADRS is a validated measure in clinical depression research (Fantino & Moore, 2009; Montgomery & Asberg, 1979), and was delivered using the Structured Interview Guide (SIGMA) (Williams & Kobak, 2008). This is a standardized interview format for the MADRS, providing the clinician with more versatility to probe the circumstances surrounding depression symptomology compared to a self-report method. The SIGMA format is reported to have higher reliability than the self-report MADRS (Williams & Kobak, 2008). The

MADRS_SIGMA was selected as it is a standardised and validated observer-rating of depression at a point in time (Williams & Kobak, 2008). It was administered by a health care professional (stroke nurse, occupational therapist or doctor) specifically trained in the administration of this assessment. Assessors used structured interview guides, detailed protocol manuals and training videos to enhance standardization in the administration of the tool. Non-verbal supports were available to patients that were aphasic. Higher scores on the MADRS indicate more depressive symptoms.

Global cognitive impairment was screened using the Montreal Cognitive Assessment (MoCA), a validated screening measure of cognitive impairment in dementia and post-stroke recovery (Dong et al., 2010). While positive scores on the MoCA indicate optimal cognitive functioning, an adjusted cut off score of <23 in stroke populations suggests cognitive impairment (Cumming, Churilov, Linden, & Bernhardt, 2013). The modified Rankin Scale (mRS) was used to measure functional disability. It is an interview based measure, with lower scores indicating lesser levels of observable functional impairment in daily life (Rankin, 1957). All assessments were administered by a health care professional specifically trained in the administration of these assessments. Background details on age, gender, subtype of acute ischaemic stroke, and thrombolysis were also obtained.

Blood Collection and Serum Separation

Blood samples were obtained by venipuncture at the 3 month follow-up assessment. All samples were collected in plastic serum separating tube (SST) vacutainers and were allowed to clot at ambient temperate for 30 minutes. The tubes were then centrifuged at 1100-1300g at room temperature and the resulting serum was aliquoted into Eppendorf 4 x 2.0ml tubes and immediately stored at -80 degrees C. Upon moving

blood samples from hospital sites to the central laboratory, temperature was kept at -20° C prior to transfer into a -80° C freezer.

Sample Preprocessing and Trypsination

Ten microliters of serum from each patient was first stabilised in 100 µL of 8 M urea pH=8.3 and stored at -80°C until used. For proteomic analysis, the stabilised samples were processed as follows: 20 µL protein solution was added to 90 µL of 8 M urea pH=8.3 and reduced for 5 hours with 1 µL of 200 mM tris(2-carboxyethyl)phosphine (TCEP). After this, samples were alkylated for 1 hour at 25°C in the dark with 4 µL of 1 M iodoacetamide (IAA). In sample-digests were performed overnight (37°C) by addition of 1 µg of trypsin (Promega, Madison WI, USA) and 900 µL of 50 mM Tris pH=8.3 followed by a second digestion step with 1 µg trypsin and an additional incubation of 4 hours at 37°C. Two hundred microliters of the digested solution were collected and dried by SpeedVac centrifugation. The digested proteins were resuspended in 100 µl of 1% (v/v) formic acid and centrifuged at 14,000rpm for 2 minutes. The solid phase extraction was performed with Empore reversed-phase extraction disks (SDB-XC reversed-phase material, 3M) according to (Ishihama, Rappsilber et al. 2006) with the following modifications: the membrane was conditioned with 50 µL of 80 % (v/v) acetonitrile, 0.1 % (w/v) trifluroacetic acid, then washed with 50 µL of 0.1% trifluroacetic acid before the tryptic peptides were bound to the membrane. The bound peptides were eluted by 50 µL 80% (v/v) acetonitrile, 0.1% (w/v) trifluroacetic acid, and dried in a SpeedVac centrifuge.

Mass Spectrometry

Tryptic peptides reconstituted in 0.1% formic acid and 2% acetonitrile (buffer A) were analysed by LC-ESI-MS/MS on a LTQ-Orbitrap Elite (*Thermo-Fisher Scientific*). Peptides were loaded onto a trap column (C18 PepMap 100 µm i.d. × 2 cm trapping

column, *Thermo-Fisher Scientific*) at 5 µL/min for 6 min before switching the precolumn in line with the analytical column (Easy-Spray 75 µm i.d. × 50 cm, *Thermo-Fisher Scientific*). The separation of peptides was performed at 250 nl/min using a linear acetonitrile gradient of buffer A and buffer B (0.1% formic acid, 80% acetonitrile), starting from 5% buffer B to 60% over 300 min. This final separation step is equivalent to fractionation and was conducted in order to avoid potential biases and increased sample variability due to depletion. Although it is impossible to eliminate dynamic range issues in serum samples this technique greatly increases the dynamic range detectable in our experiment.

Data was collected in Data Dependent Acquisition mode using *m*/*z* 300–1500 as MS scan range, CID MS/MS spectra were collected for the 20 most intense ions. Dynamic exclusion parameters were set as follows; repeat count 1, duration 90s, the exclusion list size was set at 500 with early expiration disabled. Other instrument parameters for the Orbitrap were the following: MS scan at 120 000 resolution, maximum injection time 150 ms, AGC target 1 × 106, CID at 35% energy for a maximum injection time of 150ms with AGT target of 5000. The Orbitrap Elite was operated in dual analyzer mode with the Orbitrap analyzer being used for MS and the linear trap being used for MS/MS. This procedure was performed on 2 technical replicates. The samples were then analysed using the in house Mascot server for protein identification. MaxQuant (Max-Planck Institute for Biochemistry, Martinsried, Germany) was used to obtain the relative quantification of identified proteins in the samples. Relative intensity or quantification is a measurement of peak height in a single sample that is compared to the same measurement in other samples. The absolute concentrations of the proteins in the sample are not known and require other methodologies to obtain and thus relative protein expression data cannot be generalized to other protein assays. Absolute quantification is

possible on a mass spectrometer but requires prior knowledge of target proteins and extensive methodologies.

Protein identification and Label-free quantitation

Identification and label free quantitation of obtained spectra across all 44 samples was performed using MaxQuant version 1.4.1.2 to obtain identified proteins (Cox and Mann 2008). All raw data and complete details of MaxQuant parameters and result files have been deposited in ProteomeXChange and are available with accession number PDX003494. Identification of peptides and proteins was performed internally by MaxQuant using the Andromeda (Cox and Mann 2008) search engine to search against all reviewed and unreviewed human proteins in the Uniprot database (August 2013; 133798 entries in total). Common contaminants and decoys (reversed sequences) were included automatically by Andromeda. Prior to searching, MS/MS spectra were filtered according to MaxQuant default settings for ion trap MS/MS spectra by retaining only the top eight peaks per 100Da. Main search, precursor mass tolerance was set to 4.5ppm and MS/MS tolerance to 0.5 Da. Carbamidomethylation of cysteines was set as a fixed modification and N-term acetylation and Oxidation of methionine were included as variable modifications. Up to two missed cleavages were allowed and peptides were required to be at least 7 amino acids in length. False discovery rate (FDR) cut-offs for both peptides and proteins in the database search were set to 1%. Both unique and razor peptides were used for quantitation with a minimum of 2 peptides including at least one unique peptide required to calculate a protein quantitative value. The “match between runs” setting in MaxQuant was used to transfer peptide identifications from one run to another on the basis of matching retention time and mass to charge ratio.

Data Pre-processing and Analysis

The initial output from MaxQuant consisted of 515 protein groups. After removal of contaminants, the list was shortened to 475 protein groups. The LFQ signal intensity was $\log_2$ normalized to account for naturally skewed intensity values (Cho, Smalley, Theodorescu, Ley, & Lee, 2007) and averaged over technical replicates. The discovery approach here employs relative quantification instead of absolute and as such, it was better suited to examine a collection of genes/proteins as ontologies of biological structures or pathway processes to extract biological significance from the proteomics data and understanding of the systems involved in depression symptoms post stroke (Wu et al., 2014). As relative protein expression is limited to comparison within the study or similar MS based studies, pathway analysis provides greater explanatory power than traditional statistical biomarker approaches. Thus, the resulting protein group expression data were reduced to gene names then paired with the continuous MADRS scores to prepare for GSEA (Broad Institute, MIT). GSEA allows for a robust comparison of continuous phenotypes to gene expression with the selection of the 'Pearson' metric. All analyses in GSEA were conducted with the GENE_SYMBOL chip and default number of permutations (1000) on full gene sets from Hallmark, Gene Ontology (GO), the Kyoto Encyclopedia of Genes and Genomes (KEGG), Biocarta, Reactome Positional and Immunologic Signatures, acquired from MSIGDB v5.00. The Hallmark database is recommended as an entry to GSEA analysis, as it collects gene sets that represent well defined biological states and processes with expression scores computed from many existing gene sets to reduce noise and redundancy, acting as the searchable 'meta-analysis' of gene sets. GO is the earliest but most up to date functional gene annotation database and encompasses the largest variety of annotations under 3 headings: biological processes, molecular function and cellular component (Ashburner et al., 2000). KEGG (Kanehisa, Goto, Kawashima, & Nakaya, 2002), Biocarta (Nishimura, 2001) and Reactome (Croft et al., 2014) gene sets are curated from external databases, with each

database representing different approaches to compiling such as genome sequencing, microarray profiling and computational methods to build complex networks. Positional gene sets represent the locations of genes on chromosomes and cytogenic bands (Eyre et al., 2006) while the Immunologic Signatures database is a collection of immune responses curated from separate microarray studies. Included gene sets were no smaller than 15 and no larger than 500. As per recommendation by the GSEA program for discovery experiments, a FDR of 25% was deemed acceptable for statistical significance of enriched gene sets where there is a 75% chance of rejecting a false positive (Subramanian et al., 2005). A nominal *p* (nom-*p*) value based on the statistical significance of each individual database is also included for reference. This statistic is not adjusted for multiple testing whereas the FDR value is.

An enrichment score is a Kolmogorov—Smirnov-like ranking statistic that reflects the degree to which a set is overrepresented at the top (positive enrichment) or bottom (negative enrichment) of the list when compared to another list (Subramanian et al., 2005). For example, when adjusted for correlation with increasing score on a clinical scale such as the MADRS, positive enrichment suggests biological up-regulation of a given gene list whilst negative enrichment suggests down-regulation of a gene list with increasing presentation of depression symptoms. The normalized enrichment score (NES) is the primary statistic used for evaluating and comparing gene sets and is understood as (Subramanian et al., 2005):

$$\mathrm{NES} = \frac{\text{actual Enrichment Score}}{mean(\text{Enrichment Scores against all permutations of the dataset})}$$

This analysis yields the ‘leading edge subset’ that refers to the cluster of proteins that contribute most to the enrichment score and can be interpreted as the genes that are most likely to affect change in in complex pathway function.

To minimize the possible confound of prior history of depression on the analysis, the appropriate T-test or non-parametric (Mann-Whitney) comparisons were conducted and compared with group comparisons with Monte Carlo simulations of the MADRS scores in subgroups with and without prior history of depression.

Results

Patient demographics and clinical status

The average age of ischemic stroke patients in the sample was 63.62 years, range 34-87 (*SD* = 13.52); within the range of previous epidemiological findings, although younger than the mean age of 74.40 years (Lee, Shafe, & Cowie, 2011). Thirty were males and 14 females. Based on the TOAST classification (Adams et al., 1993), 11 patients had large artery atherosclerosis, 7 had cardioembolism and 11 had small vessel occlusion. The remaining 15 patients were unclassified. Nine patients were administered thrombolytic tPA within the 4.5 hour time window and 2 were unknown. Patient demographics and NIHSS score at time of recruitment (i.e baseline *M* = 1.39, *SD* = 1.26 days post-stroke) and clinical characteristics at 3-7 days and at 3 months post-stroke are presented in Table 1.

Table 1
Patient characteristics at baseline, 3 -7 days and 3 months post stroke.

Patient Characteristic	Baseline*	3-7 days (±1 day) post-stroke	3 months post-stroke (± 7 days)
Marital status, n (%)			
Single	2 (4.5)		
Married	30 (68.2)	-	-
Widowed	4 (9.1)		
Separated or divorced	6 (13.6)		
Other	2 (4.5)		
Working status, n (%)			
Employed	25 (56.8)	25 (56.8)	14 (32)
Given up work	-	-	7 (16)
Returning to work	-	-	4 (9)
Homemaker	1 (2.3)	1 (2.3)	1 (2)
Retired	18 (40.9)	18 (40.9)	18 (41)
Living situation, n (%)			
Home alone	9 (20)	9 (20)	8 (18.2)
Home with others	34 (77)	34 (77)	34 (77.3)
Retirement home	1 (2)	1 (2)	1 (2.3)
High level care nursing home	-	-	1 (2.3)
Schooling completed, (*years*), Mean (SD)	11.81 (3.66)	-	-
NIHSS (*1-42*), mean (SD)	3.84 (4.03)	1.68 (0.33)	0.36 (0.12)
mRS (*0-6*), median (IQR)	0.00 (0.00)**	-	1.00 (2.00)
MoCA (0-30), mean (SD)		25.70 (0.43)	26.70 (0.45)
MADRS (*0-60*), mean (SD)		7.00 (5.72)	6.40 (7.42)

* Baseline measurements were taken at $M = 1.39$, $SD = 1.26$ days, i.e at time of recruitment to the study.
** This mRS score was taken at recruitment and assesses pre-stroke functioning.
NIHSS: National Institute of Health Stroke Severity Scale; 1-42; < 5 suggests mild stroke severity, with 42 being most impaired; mRS: modified Rankin Scale; range: 0-6, with a score of 1 indicating no significant disability and able to carry out all daily activities; MoCA: Montreal Cognitive Assessment; range: 0-30, <23 mild global cognitive impairment; MADRS: Montgomery-Asberg Depression Rating Scale; 0-60; > 6 suggests

presence of depressive symptoms, > 17 indicates experience of a major depressive episode within 2 weeks.

The average MADRS score for the sample was 6.40 (*SD* = 7.42, 95% CI = [4.17, 8.63]). Within a range of 0-26; 12 (27.27%) of patients had a score of > 6, indicating that this sample mostly had mild depression symptoms. Only 5 (11.36%) patients had scores within the severe symptoms range. However, a person scoring 12 in major depressive disorder or 8 post-stroke on the MADRS can be considered for treatment (Sagen et al., 2009) The sample presented with mild stroke severity at 3 months, with a median NIHSS score of 0, and range 0-5. Similarly, scores on the MoCA and mRS suggest that this sample had relatively few cognitive problems and had recovered in daily activities at 3 months from the stroke. A correlation analysis of the MADRS with these clinical measures and patient age revealed no significant results at 3 months post-stroke.

Screening on the PHQ-2 revealed that 11 of the 44 patients had a prior history of depression. As it was possible that patients with a pre-stroke history of depression may be predisposed to more severe post stroke depression symptoms, a Mann-Whitney comparison was conducted to test heterogeneity of MADRS scores between those with and without prior history of depression. Preliminary analysis suggested that the distributions of both groups were similar and thus the median metric is used (Hart, 2001). After running a Monte Carlo simulation for 1,000,000 samples at 99% CI to account for low sample sizes, it was found that patients with a previous history of depression (*Mdn* = 5.00) did not significantly differ from patients without any history (*Mdn* = 2.50), *U* = 165.00, *z* = -0.591, Monte Carlo *p* = .565 (99% CI = [.563, .566]). Following this validation, it was possible to continue with the available range of MADRS scores without having to categorize based on prior history of depression.

Gene Set Enrichment Analysis

Peptides (n=475) from the proteomic analysis were analyzed to reveal underlying molecular pathways and gene ontologies associated with depressive symptoms. Five different data sets (Hallmark, Gene Ontology, KEGG and Reactome) were interrogated to identify enriched data sets. Multiple databases were used in our discovery approach to not only explore a comprehensive range of possible pathways (given the fact that each database is constructed differently) but also to identify commonalities across databases; thus strengthening the robustness and generalizability of our findings.

Of the gene set databases that were entered in GSEA (Table 2), Positional and Immunologic Signatures did not return enriched sets. With the exception of GO, all databases showed significantly enriched gene sets pertaining to complement cascade activation or general immune upregulation/downregulation. All sets were negatively enriched when compared with increasing MADRS score, suggesting significant protein down-regulation with increasing level of depressive symptoms. There was also no clear enrichment pattern that would associate individual genes with MADRS scores. However, there are slight variations in set size and top gene contributors that reflect the differences in the compiling of the databases itself. Thus, although the statistics and bioinformatics analysis show that these pathways are associated with depression, biological interpretation is dependent upon further information from the individual databases. Positional gene sets collect data about chromosomal cytogenic band positioning of the genes involved while Immnuological Signatures collect published examples of specific immune activity against an immune challenge such as dendritic cell activity in human immunodeficiency virus.

Table 2

Size of Gene Set Databases Searched in GSEA and Corresponding Number of Enriched Sets Adjusted for FDR

Gene set database	Gene sets	Enriched gene sets	Significantly enriched gene sets (FDR <25%)
Hallmark	50	3	2
Gene Ontology	1454	25	0
KEGG	186	2	2
Biocarta	217	1	1
Reactome	674	7	3
Positional	326	0	0
Immunologic Signatures	1910	0	0

Note. These gene sets are publicly available at http://software.broadinstitute.org/gsea/msigdb/collections.jsp

Hallmark

As the Hallmark gene sets are computed from a collection of similar biological processes, this set provides strong support for the negative enrichment of this set and associated gene expressions (Table 3). A negatively skewed NES for both coagulation and complement Hallmark gene sets correlated against increasing MADRS scores indicates that there is an overexpression of down-regulated genes in these sets. These genes can be seen in the leading edge subset (highlighted in bold) whilst the other genes comprise the structure of the set and are important in affirming the construct validity of the obtained sets, they do not significantly contribute to the overall magnitude and direction of the enrichment score.

Table 3

Significantly Enriched Gene Sets and Individual Gene Contributors from the Hallmark Database

Name	NES	Individual Gene Contributors	FDR	Nom *p*
Coagulation	-1.268	**COMP PROC MBL2 TF C8B C9 APOC1 F2 PROZ FN1 PLG GSN MST1 F11** GP1BA KLKB1 CFI APOC3 C2 ITIH1SPARC HRG CLU CPB2 F10 SERPINC1 PF4 SERPING1 F13B F12 APOA1 A2M C1R FGA APOC2 VWF C1QA CFH PROS1 SERPINA1 CFB CPN1 ANG C8G C1S C8A C3	.218	.157
Complement	-1.662	**CP S100A12 C9 C1QC APOC1 F2 FN1 PLG** LTF GP1BA KLKB1 C2 ITIH1 CLU CA2 F10 SERPINC1 F5 SERPING1 C1R C4BPB C1QA CFH SERPINA1 CFB APOA4 ANG C1S C3	.157	.013

Note. The genes in **bold** represent the leading edge subset of genes (see methods for overview) that contribute most to the enrichment score of the set. NES = Normalized enrichment score. FDR = False discovery rate, Nom *p* = Nominal p-value

Kyoto Encyclopedia of Genes

Coagulation and complement cascades and systemic lupus erythematous (SLE) were implicated in the KEGG database (Table 4). In this database, the SLE set is defined by the antigen-activated complement pathway, demonstrating the characteristic molecular cascades of immune dysfunction. Even though SLE is a significant risk factor of ischemic stroke (Chiu et al., 2012), as seen from the leading edge (Table 4), it is more likely that this set was merely expressing statistical enrichment of a gene set that is largely comprised of complement. Central to all complement pathways, complement component 3 (C3) is one of the main activators and mediators of this pathway and was found to contribute highly to the negative enrichment score.

Table 4

Significantly Enriched Gene Sets and Individual Gene Contributors from the KEGG Database

Name	NES	Individual Gene Contributors	FDR	Nom *p*
Complement and Coagulation Cascades	-1.377	**C5 C4B C8G C1S KNG1 C8A C3 PROC MBL2 C8B C9 C1QC F2 C4BPA C1QB SERPINF2 PLG C6 F11** FGB SERPINA5 KLKB1 CFI C2 SERPIND1 CPB2 F10 SERPINC1 F5 MASP1 SERPING1 F13B F12 A2M C1R FGA VWF C4BPB C1QA CFH C4A PROS1 SERPINA1 CFB C7	.083	.070
Systemic Lupus Erythematous	-1.731	**CP S100A12 C9 C1QC APOC1 F2 FN1 PLG** C7 C5 C4B C8G C1S C8A C3 C8B C9 C1QC C1QB C6	.019	.016

Note. The genes in **bold** represent the leading edge subset of genes that contribute most to the enrichment score of the set.

Biocarta

The set size of Biocarta complement was only 17 with the majority of genes contributing to the various complement pathways and formation of the membrane attack complex (MAC) that is involved in attacking target cells. This gene list suggests that both mannan-binding lectin 2 (MBL2) and complement component 1 (C1Q), precursors of both the lectin and classical pathways are down-regulated in association with depressive symptoms (Table 5). C3 is central to the entire cascade as its activation is required for both lectin and classical pathways and attraction of the adaptive immune system.

Table 5

Significantly Enriched Gene Sets and Individual Gene Contributors from the Biocarta Database

Name	NES	Individual Gene Contributors	FDR	Nom *p*
Complement	-1.626	**CFB C7 C5 C4B C1S C8A C3 MBL2 C9 C1QC C1QB C6** C2 MASP1 C1R C1QA C4A	.029	.029

Note. The genes in **bold** represent the leading edge subset of genes that contribute most to the enrichment score of the set.

Reactome

The Reactome database implicated 3 gene sets with significant negative enrichment in association with MADRS scores, all with similar individual gene contributions (Table 6). This database consists of a large pathway map of which complement cascades are categorized under the innate immune system. The results for the complement pathways were similar to those obtained previously; both classical and lectin pathways are down-regulated and MAC activation genes such as C9 and C6 are underrepresented. These genes are essential to the overall function of the complement system (Figure 1).

Table 6

Significantly Enriched Gene Sets and Individual Gene Contributors from the Reactome Database

Name	NES	Individual Gene Contributors	FDR	Nom *p*
Immune System	-1.651	**C7 C5 C8G C1S CARD9 C8A C3 ICAM1 MBL2 C8B S100A12 C9 C1QC C4BPA C1QB C6 LBP** CRP CFI C2 MASP1 SAA1 SELL CD14 C4BPB C1QA CFH C4A PROS1 B2M CFB	.075	.029
Innate Immune System	-1.639	**C7 C5 C8G C1S CARD9 C8A C3 MBL2 C8B S100A12 C9 C1QC C4BPA C1QB C6 LBP** CRP CFI C2 MASP1 SAA1 CD14 C4BPB C1QA CFH C4A PROS1 CFB	.039	.018
Complement	-1.334	**C7 C5 C8G C1S C8A C3 MBL2 C8B C9 C1QC C4BPA C1QB C6** CRP CFI C2 MASP1 C4BPB C1QA CFH C4A PROS1 CFB	.194	.126

Note. The genes in **bold** represent the leading edge subset of genes that contribute most to the enrichment score of the set.

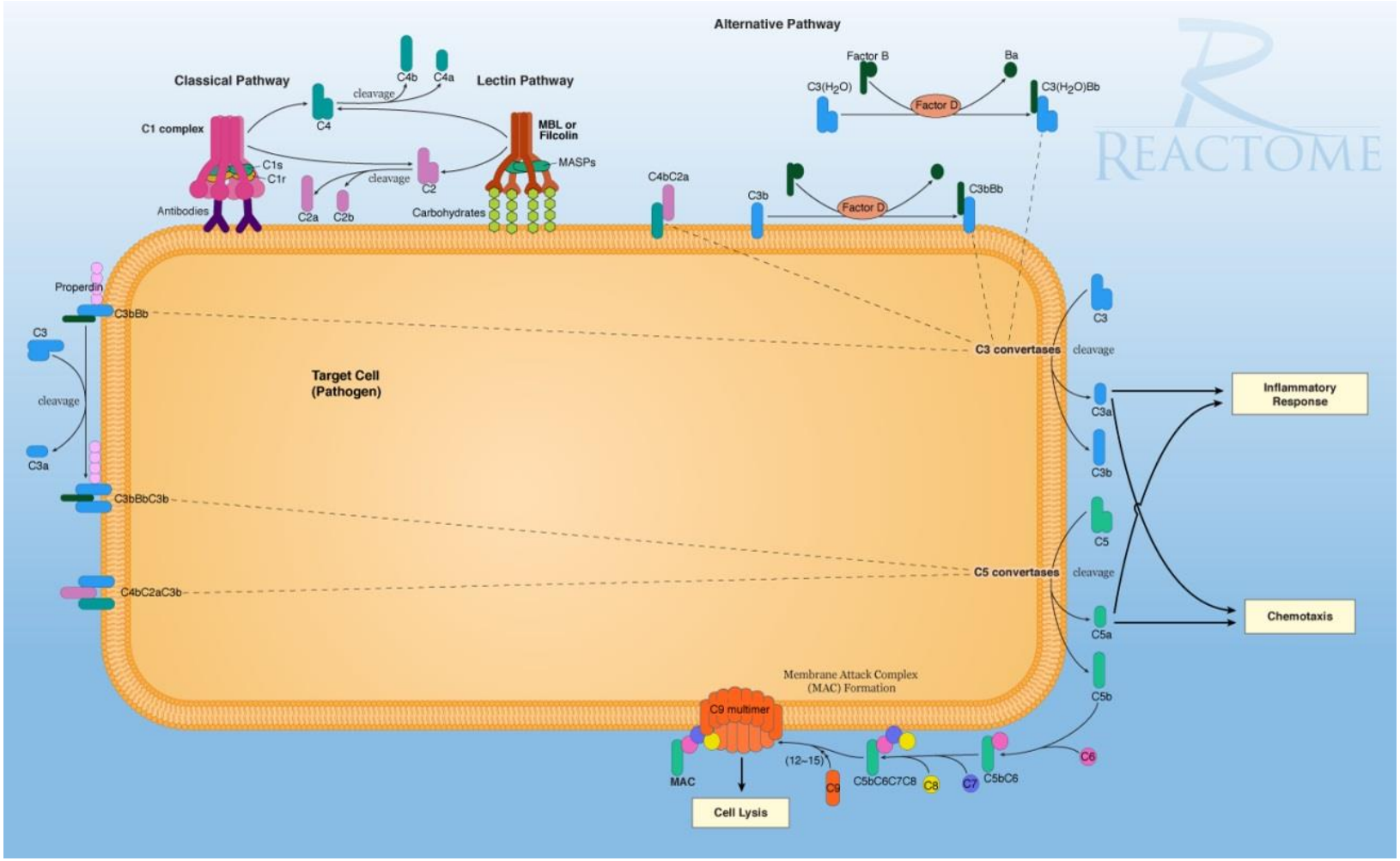

Figure 1. The Reactome pathway diagram, demonstrating the primary signaling and functional molecular pathways of the complement system on a target cell. In this diagram, multiple activation pathways are shown, eventually effecting cell lysis and breakdown of the target cell. This also shows secondary signaling pathways for C3a signaling for pro-inflammation and cytokine activity with C5a promoting chemotaxic recruitment of inflammatory cells. Retrieved from http://www.reactome.org/figures/Pathway_Illustrations/Complement_Cascade_72_oicr.png

Discussion

Depression after stroke is a common occurrence, negatively impacting functional outcome, response to rehabilitation and quality of life. Utilizing a discovery proteomics approach, this study was aimed to find biological associates of depression symptoms (MADRS scores) post stroke in human serum. This was achieved by employing a label free proteomics workflow utilizing an LTQ Orbitrap with high dynamic range to maximize the range of protein identification. Statistical and bioinformatics analysis was completed on the GSEA platform on a wide range of curated gene set databases. This approach revealed consensus among well-developed databases for decreased gene expression that was associated with the complement pathways, and partial support for precursors of the coagulation pathway.

The most important findings of this study are down-regulation of coagulation and complement cascades in serum bloods from all stroke patients with increasing level of MADRS defined depressive symptoms, as depicted by Hallmark, KEGG, Reactome and Biocarta databases. Most PSD studies examine a severe depression phenotype. However, we have shown here that even patients that have very mild depression symptoms and have recovered well can still exhibit biological changes that can be characterized by a proteomics approach. Furthermore, even though the analytical approach was different, our findings provide additional support for those obtained by Zhan et al. (2014). Our findings also support the dysregulation of the coagulation and complement cascade pathway that was identified in Zhan et al. (2014), in serum as opposed to EDTA blood samples. The contributing differences in individual gene expression are likely due to a number of factors such as stroke timeline, study design and statistical approaches but most importantly, the use of serum contrasted with EDTA for blood preparation. Considering these methodological differences

between the two studies, finding similar results only provides further confirmation for the proteomic profile of stroke and strengthens biological understanding of PSD.

From a molecular standpoint, the primary function of the complement system can be summarized as formation of the membrane attack complex on target cells while secondary functions include signaling for toll like receptor 4 (TLR4) mediated inflammatory response (Bradley, Zipfel, & Hughes, 2011), leading to increased pro-inflammatory cytokine levels (Zhang, Ramesh, Uematsu, Akira, & Reeves, 2008). Complement has been traditionally viewed as a versatile system, conforming to 3 distinct activation pathways, leading to similar outcomes (Fujita, 2002). Of the leading edge subsets of the databases searched, it is clear that a majority of the molecular determinants of the complement system are down-regulated in relation to mild depression symptoms. Of note, complement component 1 q (C1q) and complement component 3 (C3) are both integral to the activation of this system while downstream products that contribute to inflammatory signaling and formation on the MAC (Figure 1) are all down-regulated. This may indicate perturbation in brain homeostasis following stroke, characterized by an ongoing state of lowered resistance to oxidative stress (Collard et al., 2000) and immune related changes (Roos et al., 2001) in maintaining depression symptoms post stroke. This is further implied by the Reactome database analysis, demonstrating negative enrichment of calcium binding protein A12 (S100A12) and lipopolysaccharide binding protein (LBP), attributed to innate immune functioning. S100A12 is an established cytoskeletal protein involved in the signaling of neutrophil response and a candidate biomarker of inflammation (Meijer, Gearry, & Day, 2012).

Coagulation is a key process in vascular diseases, especially in ischemic stroke where coagulation factors have been identified as pre-stroke risk factors (Stankovic & Majkic-Singh, 2010). This process is heavily involved in the stroke itself and treatment with tPA

(Fassbender et al., 1999; Uchiyama, Yamazaki, Hara, & Iwata, 1997). Coagulation is an ongoing process in serum and deficiencies of several proteins such as protein Z (PROZ) (Sofi et al., 2010), fibronectin 1 (FN1) (Uchiyama et al., 1997) and gelsolin (GSN) (Chou et al., 2011) have been identified to be involved in cerebrovascular and cardiovascular diseases. The overlap of genes between coagulation and complement sets is expected as the coagulation processes have been identified to cleave into the central components of the complement system (Amara et al., 2008). The inclusion of both complement and coagulation cascades into one gene set in KEGG compared to the presentation of 2 separate sets in Hallmark illustrates a previous debate as to whether the pathways should be considered separately (Delvaeye & Conway, 2009; Markiewski, Nilsson, Nilsson Ekdahl, Mollnes, & Lambris, 2007). However, further studies have shown that while complement deficiencies alone do not increase bleeding frequency and coagulation deficiencies alone do not impair immune responses, the functions of the two pathways are linked, and indeed in the case of innate immunity (Agarwal, Talens, Grandits, & Blom, 2015; Amara et al., 2010). It has been suggested that coagulation and complement are involved in a feedback loop where complement activation increases platelet activation area on the target cell that in turn augments complement activity (Esmon, Xu, & Lupu, 2011). Coagulation has independently been shown to be involved in the process of cell death and is associated with activation of the kinin-kallikrein pathway and facilitates defensive inflammatory responses by enhancing leukocyte activity (Chen & Nuñez, 2010). Finally, binding platelets also have the ability to enhance neutrophil activity via TLR4 signaling, a process that is vital to phagocytosis (Esmon et al., 2011).

Although a robust association between relative presence of depressive symptoms and downregulation of complement and coagulation pathway was found in our necessarily small

sample of mild stroke survivors living in the community, our findings need to be interpreted with care especially as some patients had a prior history of depression. We did use a group comparison utilizing Monte Carlo simulations that indicated that prior depression did not significantly impact our findings. However, due to the discovery approach and the patient characteristics of this sample, the generalizability of the findings is limited to survivors with mild stroke severity.

Complement Post Stroke: From Acute Cell Death to Immunodepression And Depression

Cell death in the brain is the inevitable consequence of any stroke damage. In stroke, the first wave of cell death occurs as a result of hypoxia. This releases damage associated molecular patterns (DAMPs) that can begin and perpetuate apoptotic and necrotic cell death cascades in neighboring neurons. Such molecules can include but are not limited to intracellular adenosine triphosphate and uridine triphosphate that has been leaked into the extracellular space, nitric oxide, heat shock proteins, S100 proteins, extracellular calcium ion levels and cytokines (Gelderblom, Sobey, Kleinschnitz, & Magnus, 2015; Kono & Rock, 2008). The presence of DAMPs on the central side of the blood brain barrier (BBB) are also associated with endothelial damage which further contributes to increased permeability and infiltration of immune cells from the periphery (Shichita, Ito, & Yoshimura, 2014).

Transcriptomic research has not yet identified pathways involved in the timeline of PSD pathogenesis in bloods, however, an overview of studies examining the whole blood ribonucleic acid (RNA) profiles of different stroke subtypes including TIA in the <1 week have yielded interesting results (Sharp et al., 2011). Functional analysis of the genes in these studies have shown that immune and homeostasis pathway expression can differentiate between cardioembolic and atherosclerosis stroke (Xu et al., 2008) and Gene Ontology

clustering in cardioembolic ischemic stroke suggests gene expression indicative of cell death, lipid metabolism and metal ion transport. Thus here we propose that the proteomically detectable state of the peripheral immune system may be indicative of unresolved ischemic damage in the central parenchyma, resulting in disruptions to inflammatory, metabolic and homeostatic balance that is related to the transcriptomic profile in the early stages of stroke damage (Sharp et al., 2011). While there is no neurobiological evidence presented in our study, current knowledge of the aetiology of PSD suggests that persistent neuroinflammation, driven by an increase in pro-inflammatory but also a decrease in anti-inflammatory cytokine signaling, is responsible for depressive symptoms post-stroke (Fang & Cheng, 2009; Spalletta et al., 2006). Indeed, it is possible that peripheral immunodepression may be caused by central and upstream molecular cascades that include cross BBB cytokine signaling (W. Pan et al., 2011) or bioavailability of immunoglobulin antibodies (St-Amour et al., 2013) and leukocyte immune cells (Famakin, 2014).Thus, PSD can be considered as a natural sequelae of incomplete recovery from stroke that can be further exacerbated by anxiety from psychosocial issues (Pascoe et al., 2011).

We have shown that two innate immune pathways in peripheral bloods, complement and coagulation, trend towards downregulation at the 3 month phase in correlation with mild symptoms of depression. The complement system is largely involved in an array of normal and immunoregulatory functions, although its role in immunodepression post stroke has not been explored (Figure 2). As a whole however, immunodepression after stroke is a well-documented and a natural consequence of ischemia but is poorly understood in functional terms (Macrez et al., 2011). Initially, it appears that immunodepression is counterproductive after stroke as it may increase the chances of commonly reported secondary infections such as urinary tract infection and pneumonia (Dirnagl et al., 2007; Offner, Vandenbark, & Hurn,

2009). Additionally, immune suppression is thought to be an adaptive response to central inflammation as an autoimmune response against the brain would be detrimental to recovery outcomes and possibly exacerbate damage (Kamel & Iadecola, 2012; Macrez et al., 2011). There are many factors that may maintain poor immune responses and recovery in animal models such as induced psychosocial stress (Kubera, Obuchowicz, Goehler, Brzeszcz, & Maes, 2011), abnormal blood brain barrier (BBB) permeability (Abbott, 2000), and possible ongoing antigen related responses that have yet to be fully characterized (Emsley, Hedley, & Hopkins, 2010; Urra, Miró, Chamorro, & Planas, 2014).

There has been little research exploring the role of complement and indeed immunity post stroke in relation to depression. The results here suggest that there is ongoing immunosuppression in the periphery with mild depressive symptoms, even at 3 months post stroke. Previous research into antigen presenting cells such as macrophages and dendritic cells has established that their expression is elevated centrally but downregulated peripherally post stroke (Gelderblom et al., 2009; Iadecola & Anrather, 2011). A trend towards immune recovery would stipulate that a balance has been reached in these cell levels for both central and periphery. Indeed, it has also been theorized that this discrepancy accounts for the recruitment of peripheral dendritic cells into the brain to maintain a central inflammatory and immune response (Yilmaz et al., 2009). Thus given these lines of evidence, it is possible that our current findings reflect an ongoing sub-acute state of pro-inflammation that has not yet transitioned to anti-inflammation (Shichita et al., 2014), manifesting behaviorally as depression symptoms and molecularly as peripheral immunodepression of the complement and coagulation systems.

	Pre-stroke function and risk factors	Stroke (0-48h)	Stroke Outcomes (48h+)
Central	+ C3a regulates neurogenesis in vitro by determining differentiation and migration of neural progenitor cells (Shinjyo, Ståhlberg, Dragunow, Pekny, & Pekna, 2009). + Complement system involved in synaptic pruning by opsonization involving recruitment of microglia (Stephan, Barres, & Stevens, 2012).	+ Complement inhibition by various methods such as C5a antagonists (G. H. Kim et al., 2008) and intravenous immunoglobulin (Arumugam et al., 2009) are neuroprotective. – Enhances neutrophil adhesion and leukocyte activity which leads to further tissue damage (Atkinson et al., 2006). – Propagation of neuroinflammation and apotopic cell death (Alawieh et al., 2015; Elvington et al., 2012).	+ Astrocyte modulated complement responses enhance the ability of microglia to remove cell death debris (Van Beek et al., 2000). + Complement promotes neurogenesis post cerebral ischemia (Rahpeymai et al., 2006).
Peripheral	+ Constitutes part of the immune system, involved in clearance of pathogens via opsonization and cell lysis. Can initiate a local inflammatory response (Janeway, Travers, Walport, & Mark, 2001). – Increased serum C4 levels in patients with coronary artery disease predicts stroke risk (Cavusoglu et al., 2007). – C5a induces vasodilation independent of histamine (José, Forrest, & Williams, 1981) and increased central venous pressure (Lundberg, Marceau, & Hugli, 1987). – Serum C3 levels independently associated with myocardial infarctions and ischemic events,		– Early elevation of plasma levels of complement predict negative functional outcome following aneurysmal subarachnoid hemorrhage (Mack et al., 2007) – Reduced complement protein expression in blood associated with mild depression symptoms (current study).

including TIA (Muscari et al., 1995).

Figure 2. A brief overview of the complement system in both central and peripheral functioning related to stroke risk, the stroke and post stroke outcomes as understood from both human and animal studies. The + denoted functions indicate normal or positive functioning while the – denoted functions indicate counterproductive functioning. Overall, the complement system is involved in regulation of immune function either by localized cell lysis cascades or targeting for removal by phagocytosis. The same is true in both central parenchyma and periphery, where microglia are the primary phagocytes in the brain while neutrophils, dendritic cells and monocytes are the primary phagocytes in the periphery. However, this distinction no longer holds in the event of a stroke, where BBB disruption allows some immune cells to infiltrate from the periphery. The complement system exhibits dual roles in the brain, propagating cell death by encouraging the DAMP initiated signaling cascades of apoptosis but also neuroprotection by removing cellular debris and engaging in synaptic pruning, contributing to post stroke recovery.

Limitations and future studies

The patients in this study were not severely depressed and had recovered well after their stroke, with some patients recording a history of pre-stroke depression. Furthermore, there was no age-matched control group present in this study. The current study was a pilot to investigate the feasibility of the proteomics approach in stroke. While we found a significant association even in a sample with mild depressive symptoms and mild stroke severity, future studies that employ larger sample sizes with depressive symptoms and neurological severity that represent the range commonly experienced, are recommended. Future studies may consider basic or laboratory assessment of immunological condition of patients post stroke as well as relationship between clinical and/or psychosocial factors on PSD. We used a discovery approach where the findings can stand alone for the stroke cohort; although it is recognized that this approach is not as robust as with age matched controls. We therefore recommend future comparison with age-matched non-stroke controls, with and without depression, to enable a more comprehensive interpretation of findings. In addition, future studies may compare the proteomes of other biofluids such as CSF and urine of the same patients to develop a better understanding of the compartmentalization or relationships of BBB and kidney physiology post stroke. Complement and coagulation functioning can also be assayed in blood in the traditional hematological laboratory setting. From a technical perspective, further studies could optimize the comparison of ionic chelation properties of anticoagulants as this is not well understood in whole blood samples. It is also feasible to employ immunodepletion or different fractionation techniques on the mass spectrometer in the sample preparation stages to resolve dynamic range and peak detection issues for low abundance compounds.

Conclusion

This study examined the serum proteomic profile of stroke survivors at 3 months post stroke using a label free approach. The findings here and in Zhan et al. (2014) are complementary and provide a basis for further research into blood proteins recently identified to be involved in the pathophysiology of PSD and possibly in other cerebrovascular diseases with comorbid anxiety. Analysis by GSEA on various databases has revealed enriched gene sets that are identifiable as complement and innate immune processes. As all of these gene sets were negatively enriched when correlated with increasing depressive symptoms of ischemic stroke survivors, this was interpreted as peripheral immunodepression, indicative of unresolved ongoing inflammatory processes in the brain. This interpretation is supported by the substantial body of literature that has linked PSD aetiology to overactive immunologic processes, leading to increased inflammatory processes in both peripheral and central compartments. While these findings add to the growing body of evidence for differentially expressed proteins in PSD, more research is needed to characterize their molecular processes and how their expressions may change as result of the stroke and in development of PSD.

References

Abbott, N. J. (2000). Inflammatory mediators and modulation of blood–brain barrier permeability. *Cellular and Molecular Neurobiology, 20*(2), 131-147.

Adams, H. P., Bendixen, B. H., Kappelle, L. J., Biller, J., Love, B. B., Gordon, D. L., & Marsh, E. E. (1993). Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke, 24*(1), 35-41. doi: 10.1161/01.str.24.1.35

Agarwal, V., Talens, S., Grandits, A. M., & Blom, A. M. (2015). A Novel interaction between complement inhibitor C4b-binding protein and plasminogen that enhances plasminogen activation. *Journal of Biological Chemistry, 290*(30), 18333-18342. doi: 10.1074/jbc.M114.619494

Ahn, D.-H., Lee, Y.-J., Jeong, J.-H., Kim, Y.-R., & Park, J.-B. (2015). The Effect of post-stroke depression on rehabilitation outcome and the impact of caregiver type as a factor of post-stroke depression. *Annals of Rehabilitation Medicine, 39*(1), 74-80. doi: 10.5535/arm.2015.39.1.74

Alawieh, A., Elvington, A., Zhu, H., Yu, J., Kindy, M. S., Atkinson, C., & Tomlinson, S. (2015). Modulation of post-stroke degenerative and regenerative processes and subacute protection by site-targeted inhibition of the alternative pathway of complement. *Journal of neuroinflammation, 12*(1), 247.

Amara, U., Flierl, M. A., Rittirsch, D., Klos, A., Chen, H., Acker, B., . . . Huber-Lang, M. (2010). Molecular intercommunication between the complement and coagulation systems. *Journal of Immunology, 185*(9), 5628-5636. doi: 10.4049/jimmunol.0903678

Amara, U., Rittirsch, D., Flierl, M., Bruckner, U., Klos, A., Gebhard, F., . . . Huber-Lang, M. (2008). Interaction between the coagulation and complement system. *Advances in Experimental Medicine and Biology, 632*, 71-79.

Arumugam, T. V., Woodruff, T. M., Lathia, J. D., Selvaraj, P. K., Mattson, M. P., & Taylor, S. M. (2009). Neuroprotection in stroke by complement inhibition and immunoglobulin therapy. *Neuroscience, 158*(3), 1074-1089. doi: 10.1016/j.neuroscience.2008.07.015

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics, 25*(1), 25-29. doi: 10.1038/75556

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Atala, A. (2002). *Methods of tissue engineering*: Gulf Professional Publishing.

Atkinson, C., Zhu, H., Qiao, F., Varela, J. C., Yu, J., Song, H., . . . Tomlinson, S. (2006). Complement-Dependent P-Selectin Expression and Injury following Ischemic Stroke. *The Journal of Immunology, 177*(10), 7266-7274. doi: 10.4049/jimmunol.177.10.7266

Bradley, D. T., Zipfel, P. F., & Hughes, A. E. (2011). Complement in age-related macular degeneration: a focus on function. *Eye, 25*(6), 683-693. doi: 10.1038/eye.2011.37

Broomfield, N. M., Quinn, T. J., Abdul-Rahim, A. H., Walters, M. R., & Evans, J. J. (2014). Depression and anxiety symptoms post-stroke/TIA: prevalence and associations in cross-sectional data from a regional stroke registry. *BMC Neurology, 14*, 198. doi: 10.1186/s12883-014-0198-8

Brott, T., Adams, H. P., Olinger, C. P., Marler, J. R., Barsan, W. G., Biller, J., . . . Hertzberg, V. (1989). Measurements of acute cerebral infarction: a clinical examination scale. *Stroke, 20*(7), 864-870. doi: 10.1161/01.str.20.7.864

Carey, L. M., Crewther, S., Salvado, O., Lindén, T., Connelly, A., Wilson, W., . . . the, S. R. T. (2015). STroke imAging pRevention and treatment (START): A

longitudinal stroke cohort study: Clinical trials protocol. *International Journal of Stroke, 10*(4), 636-644. doi: 10.1111/ijs.12190

Cavusoglu, E., Eng, C., Chopra, V., Ruwende, C., Yanamadala, S., Clark, L. T., . . . Marmur, J. D. (2007). Usefulness of the serum complement component C4 as a predictor of stroke in patients with known or suspected coronary artery disease referred for coronary angiography. *The American Journal of Cardiology, 100*(2), 164-168. doi: http://dx.doi.org/10.1016/j.amjcard.2007.02.075

Chen, G. Y., & Nuñez, G. (2010). Sterile inflammation: sensing and reacting to damage. *Nature Reviews Immunology, 10*(12), 826-837. doi: 10.1038/nri2873

Chiu, C.-C., Huang, C.-C., Chan, W.-L., Chung, C.-M., Huang, P.-H., Lin, S.-J., . . . Leu, H.-B. (2012). Increased risk of ischemic stroke in patients with systemic lupus erythematosus: A nationwide population-based study. *Internal Medicine, 51*(1), 17-21. doi: 10.2169/internalmedicine.51.6154

Cho, H., Smalley, D. M., Theodorescu, D., Ley, K., & Lee, J. K. (2007). Statistical identification of differentially labeled peptides from liquid chromatography tandem mass spectrometry. *Proteomics, 7*(20), 3681-3692.

Chou, S. H.-Y., Lee, P.-S., Konigsberg, R. G., Gallacci, D., Chiou, T., Arai, K., . . . Ning, M. (2011). Plasma-Type gelsolin is decreased in human blood and cerebrospinal fluid after subarachnoid hemorrhage. *Stroke, 42*(12), 3624-3627. doi: 10.1161/strokeaha.111.631135

Collard, C. D., Väkevä, A., Morrissey, M. A., Agah, A., Rollins, S. A., Reenstra, W. R., . . . Stahl, G. L. (2000). Complement activation after oxidative stress: role of the lectin complement pathway. *The American Journal of Pathology, 156*(5), 1549-1556. doi: http://dx.doi.org/10.1016/S0002-9440(10)65026-2

Corbyn, Z. (2014). Stroke: A growing global burden. *Nature, 510*(7506), S2-S3. doi: 10.1038/510S2a

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., . . . D'Eustachio, P. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Research, 42*, D472-D477. doi: 10.1093/nar/gkt1102

Cumming, T. B., Churilov, L., Linden, T., & Bernhardt, J. (2013). Montreal Cognitive Assessment and Mini–Mental State Examination are both valid cognitive tools in stroke. *Acta Neurologica Scandinavica, 128*(2), 122-129. doi: 10.1111/ane.12084

Delvaeye, M., & Conway, E. M. (2009). Coagulation and innate immune responses: can we view them separately? *Blood, 114*(12), 2367-2374.

Dirnagl, U., Klehmet, J., Braun, J. S., Harms, H., Meisel, C., Ziemssen, T., . . . Meisel, A. (2007). Stroke-Induced immunodepression: Experimental evidence and clinical relevance. *Stroke, 38*(2), 770-773. doi: 10.1161/01.STR.0000251441.89665.bc

Dong, Y., Sharma, V. K., Chan, B. P.-L., Venketasubramanian, N., Teoh, H. L., Seet, R. C. S., . . . Chen, C. (2010). The Montreal Cognitive Assessment (MoCA) is superior to the Mini-Mental State Examination (MMSE) for the detection of vascular cognitive impairment after acute stroke. *Journal of the Neurological Sciences, 299*(1), 15-18. doi: 10.1016/j.jns.2010.08.051

Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E. K., & Lanctôt, K. L. (2010). A meta-analysis of cytokines in major depression. *Biological Psychiatry, 67*(5), 446-457. doi: http://dx.doi.org/10.1016/j.biopsych.2009.09.033

El Husseini, N., Goldstein, L. B., Peterson, E. D., Zhao, X., Pan, W., Olson, D. M., . . . Laskowitz, D. T. (2012). Depression and antidepressant use after stroke and transient ischemic attack. *Stroke*. doi: 10.1161/strokeaha.111.643130

Elvington, A., Atkinson, C., Zhu, H., Yu, J., Takahashi, K., Stahl, G. L., . . . Tomlinson, S. (2012). The alternative complement pathway propagates inflammation and injury in murine ischemic stroke. *Journal of Immunology, 189*(9), 4640-4647. doi: 10.4049/jimmunol.1201904

Emsley, A., Hedley, C., & Hopkins, S. J. (2010). Post-stroke immunodepression and infection: an emerging concept. *Infectious Disorders-Drug Targets, 10*(2), 91-97.

Esmon, C. T., Xu, J., & Lupu, F. (2011). Innate Immunity and Coagulation. *Journal of Thrombosis and Haemostasis, 9*(Suppl 1), 182-188. doi: 10.1111/j.1538-7836.2011.04323.x

Eyre, T. A., Ducluzeau, F., Sneddon, T. P., Povey, S., Bruford, E. A., & Lush, M. J. (2006). The HUGO Gene Nomenclature Database, 2006 updates. *Nucleic Acids Research, 34*(suppl 1), D319-D321. doi: 10.1093/nar/gkj147

Famakin, B. M. (2014). The immune response to acute focal cerebral ischemia and associated post-stroke immunodepression: A Focused Review. *Aging and Disease, 5*(5), 307-326. doi: 10.14336/AD.2014.0500307

Fang, J., & Cheng, Q. (2009). Etiological mechanisms of post-stroke depression: a review. *Neurological Research, 31*(9), 904-909. doi: doi:10.1179/174313209X385752

Fantino, B., & Moore, N. (2009). The self-reported Montgomery-Åsberg depression rating scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry, 9*(1), 26.

Fassbender, K., Dempfle, C.-E., Mielke, O., Schwartz, A., Daffertshofer, M., Eschenfelder, C., . . . Hennerici, M. (1999). Changes in coagulation and fibrinolysis markers in acute ischemic stroke treated with recombinant tissue plasminogen activator. *Stroke, 30*(10), 2101-2104. doi: 10.1161/01.str.30.10.2101

Fujita, T. (2002). Evolution of the lectin-complement pathway and its role in innate immunity. *Nature Reviews Immunology, 2*(5), 346-353.

Galea, J., & Brough, D. (2013). The role of inflammation and interleukin-1 in acute cerebrovascular disease. *Journal of Inflammation Research, 6*, 121-128. doi: 10.2147/JIR.S35629

Gelderblom, M., Leypoldt, F., Steinbach, K., Behrens, D., Choe, C.-U., Siler, D. A., . . . Magnus, T. (2009). Temporal and spatial dynamics of cerebral immune cell accumulation in Stroke. *Stroke, 40*(5), 1849-1857. doi: 10.1161/strokeaha.108.534503

Gelderblom, M., Sobey, C. G., Kleinschnitz, C., & Magnus, T. (2015). Danger signals in stroke. *Ageing Research Reviews*. doi: http://dx.doi.org/10.1016/j.arr.2015.07.004

Giardina, B. J., Stanley, B. A., & Chiang, H.-L. (2014). Glucose induces rapid changes in the secretome of Saccharomyces cerevisiae. *Proteome Science, 12*(9).

Hackett, M. L., Yapa, C., Parag, V., & Anderson, C. S. (2005). Frequency of depression after stroke: a systematic review of observational studies. *Stroke, 36*(6), 1330-1340. doi: 10.1161/01.str.0000165928.19135.35

Hama, S., Yamashita, H., Yamawaki, S., & Kurisu, K. (2011). Post-stroke depression and apathy: Interactions between functional recovery, lesion location, and emotional response. *Psychogeriatrics, 11*(1), 68-76. doi: 10.1111/j.1479-8301.2011.00358.x

Hart, A. (2001). Mann-Whitney test is not just a test of medians: differences in spread can be important. *British Medical Journal, 323*(7309), 391-393.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols, 4*(1), 44-57. doi: http://www.nature.com/nprot/journal/v4/n1/suppinfo/nprot.2008.211_S1.html

Iadecola, C., & Anrather, J. (2011). The immunology of stroke: from mechanisms to translation. *Nature Medicine, 17*(7), 796-808. doi: http://www.nature.com/nm/journal/v17/n7/abs/nm.2399.html#supplementary-information

Ishihama, Y., Rappsilber, J., & Mann, M. (2006). Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in

proteomics. *Journal of Proteome Research, 5*(4), 988-994. doi: 10.1021/pr050385q

Jambunathan, K., & Galande, A. K. (2014). Sample collection in clinical proteomics—Proteolytic activity profile of serum and plasma. *Proteomics – Clinical Applications, 8*(5-6), 299-307. doi: 10.1002/prca.201300037

Janeway, C. A., Travers, P., Walport, M., & Mark, S. (2001). *Immunobiology: The Immune System in Health and Disease* (5 ed.). New York: Garland Science.

Janneke, M., Gooskens, F., Schepers, V. P., Schuurmans, M. J., Lindeman, E., & Hafsteinsdóttir, T. B. (2012). Screening for poststroke depression using the patient health questionnaire. *Nursing Research, 61*(5), 333-341.

José, P. J., Forrest, M. J., & Williams, T. J. (1981). Human C5a des Arg increases vascular permeability. *The Journal of Immunology, 127*(6), 2376-2380.

Käll, L., & Vitek, O. (2011). Computational mass spectrometry–based proteomics. *PLoS Computional Biology, 7*(12), e1002277. doi: 10.1371/journal.pcbi.1002277

Kamel, H., & Iadecola, C. (2012). Brain-Immune Interactions and Ischemic Stroke: Clinical Implications. *Archives of Neurology, 69*(5), 576-581. doi: 10.1001/archneurol.2011.3590

Kanehisa, M., Goto, S., Kawashima, S., & Nakaya, A. (2002). The KEGG databases at GenomeNet. *Nucleic Acids Research, 30*(1), 42-46.

Kim, G. H., Mocco, J., Hahn, D. K., Kellner, C. P., Komotar, R. J., Ducruet, A. F., . . . Connolly, E. S. (2008). Protective effect of C5a receptor inhibition after murine reperfused stroke. *Neurosurgery, 63*(1), 122-126. doi: 10.1227/01.NEU.0000335079.70222.8D

Kim, J.-M., Stewart, R., Kim, S.-W., Shin, I.-S., Kim, J.-T., Park, M.-S., . . . Yoon, J.-S. (2011). Associations of cytokine gene polymorphisms with post-stroke

depression. *The World Journal of Biological Psychiatry, 13*(8), 579-587. doi: 10.3109/15622975.2011.588247

Kono, H., & Rock, K. L. (2008). How dying cells alert the immune system to danger. *Nature Reviews Immunology, 8*(4), 279-289. doi: 10.1038/nri2215

Koomen, J. M., Haura, E. B., Bepler, G., Sutphen, R., Remily-Wood, E. R., Benson, K., . . . Dalton, W. S. (2008). Proteomic contributions to personalized cancer care. *Molecular & Cellular Proteomics, 7*(10), 1780-1794. doi: 10.1074/mcp.R800002-MCP200

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. *Medical care, 41*(11), 1284-1292.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. doi: 10.1046/j.1525-1497.2001.016009606.x

Kubera, M., Obuchowicz, E., Goehler, L., Brzeszcz, J., & Maes, M. (2011). In animal models, psychosocial stress-induced (neuro)inflammation, apoptosis and reduced neurogenesis are associated to the onset of depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(3), 744-759. doi: http://dx.doi.org/10.1016/j.pnpbp.2010.08.026

Kumar, Y., Liang, C., Bo, Z., Rajapakse, J. C., Ooi, E. E., & Tannenbaum, S. R. (2012). Serum proteome and cytokine analysis in a longitudinal cohort of adults with primary dengue infection reveals predictive markers of DHF. *PLoS Neglected Tropical Diseases, 6*(11), e1887. doi: 10.1371/journal.pntd.0001887

Lambertsen, K. L., Biber, K., & Finsen, B. (2012). Inflammatory cytokines in experimental and human stroke. *Journal of Cerebral Blood Flow and Metabolism, 32*(9), 1677-1698. doi: http://www.nature.com/jcbfm/journal/v32/n9/suppinfo/jcbfm201288s1.html

Lee, S., Shafe, A. C. E., & Cowie, M. R. (2011). UK stroke incidence, mortality and cardiovascular risk management 1999–2008: time-trend analysis from the General Practice Research Database. *British Medical Journal Open, 1*(2). doi: 10.1136/bmjopen-2011-000269

Leonard, B., & Maes, M. (2012). Mechanistic explanations how cell-mediated immune activation, inflammation and oxidative and nitrosative stress pathways and their sequels and concomitants play a role in the pathophysiology of unipolar depression. *Neuroscience & Biobehavioral Reviews, 36*(2), 764-785. doi: http://dx.doi.org/10.1016/j.neubiorev.2011.12.005

Lopez, M. F., Sarracino, D. A., Prakash, A., Athanas, M., Krastins, B., Rezai, T., . . . Ning, M. (2012). Discrimination of ischemic and hemorrhagic strokes using a multiplexed, mass spectrometry-based assay for serum apolipoproteins coupled to multi-marker ROC algorithm. *Proteomics – Clinical Applications, 6*(3-4), 190-200. doi: 10.1002/prca.201100041

Luijendijk, H. J., Stricker, B. H. C., Wieberdink, R. G., Koudstaal, P. J., Hofman, A., Breteler, M. M., & Tiemeier, H. (2011). Transient ischemic attack and incident depression. *Stroke, 42*(7), 1857-1861. doi: 10.1161/strokeaha.110.604405

Lundberg, C., Marceau, F., & Hugli, T. (1987). C5a-induced hemodynamic and hematologic changes in the rabbit. Role of cyclooxygenase products and polymorphonuclear leukocytes. *The American Journal of Pathology, 128*(3), 471.

Ma, H., Parsons, M. W., Christensen, S., Campbell, B. C. V., Churilov, L., Connelly, A., . . . Donnan, G. A. (2012). A multicentre, randomized, double-blinded, placebo-controlled phase III study to investigate Extending the Time for Thrombolysis in Emergency Neurological Deficits (EXTEND). *International Journal of Stroke, 7*(1), 74-80. doi: 10.1111/j.1747-4949.2011.00730.x

Mack, W. J., Ducruet, A. F., Hickman, Z. L., Garrett, M. C., Albert, E. J., Kellner, C. P., . . . Connolly Jr, E. S. (2007). Early plasma complement C3a levels correlate with functional outcome after aneurysmal subarachnoid hemorrhage. *Neurosurgery, 61*(2), 255-261.

Macrez, R., Ali, C., Toutirais, O., Le Mauff, B., Defer, G., Dirnagl, U., & Vivien, D. (2011). Stroke and the immune system: from pathophysiology to new therapeutic strategies. *The Lancet Neurology, 10*(5), 471-480.

Maes, M. (2011). Depression is an inflammatory disease, but cell-mediated immune activation is the key component of depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(3), 664-675. doi: http://dx.doi.org/10.1016/j.pnpbp.2010.06.014

Markiewski, M. M., Nilsson, B., Nilsson Ekdahl, K., Mollnes, T. E., & Lambris, J. D. (2007). Complement and coagulation: strangers or partners in crime? *Trends in Immunology, 28*(4), 184-192. doi: http://dx.doi.org/10.1016/j.it.2007.02.006

Meijer, B., Gearry, R., & Day, A. (2012). The role of S100A12 as a systemic marker of inflammation. *International Journal of Inflammation, 2012,* article 907078.

Mitra, J., Shen, K.-K., Ghose, S., Bourgeat, P., Fripp, J., Salvado, O., . . . Rose, S. (2014). Predicting poststroke depression from brain connectivity in *Computational Diffusion MRI* (pp. 89-99): Springer.

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry, 134*(4), 382-389.

Muscari, A., Bozzoli, C., Puddu, G. M., Sangiorgi, Z., Dormi, A., Rovinetti, C., . . . Puddu, P. (1995). Association of serum C3 levels with the risk of myocardial infarction. *The American Journal of Medicine, 98*(4), 357-364. doi: http://dx.doi.org/10.1016/S0002-9343(99)80314-3

Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., . . . Haynes, P. A. (2011). Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics,, 11*(4), 535-553. doi: 10.1002/pmic.201000553

Ning, M., Lopez, M., Cao, J., Buonanno, F. S., & Lo, E. H. (2012). Application of proteomics to cerebrovascular disease. *Electrophoresis, 33*(24), 10.1002/elps.201200481. doi: 10.1002/elps.201200481

Ning, M., Sarracino, D. A., Buonanno, F. S., Krastins, B., Chou, S., McMullin, D., . . . Lo, E. H. (2010). Proteomic protease substrate profiling of tPA treatment in acute ischemic stroke patients: a step toward individualizing thrombolytic therapy at the bedside. *Translational Stroke Research, 1*(4), 268-275. doi: 10.1007/s12975-010-0047-z

Nishimura, D. (2001). BioCarta. *Biotech Software & Internet Report, 2*(3), 117-120. doi: 10.1089/152791601750294344

Noonan, K., Carey, L. M., & Crewther, S. G. (2013). Meta-analyses Indicate Associations between Neuroendocrine Activation, Deactivation in Neurotrophic and Neuroimaging Markers in Depression after Stroke. *Journal of Stroke and Cerebrovascular Diseases, 22*(7), e124-e135. doi: http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2012.09.008

Offner, H., Vandenbark, A. A., & Hurn, P. D. (2009). effect of experimental stroke on peripheral immunity: CNS ischemia induces profound immunosuppression. *Neuroscience, 158*(3), 1098-1111. doi: 10.1016/j.neuroscience.2008.05.033

Pan, S., Zhan, X., Su, X., Guo, L., Lv, L., & Su, B. (2011). Proteomic analysis of serum proteins in acute ischemic stroke patients treated with acupuncture. *Experimental Biology and Medicine, 236*(3), 325-333. doi: 10.1258/ebm.2011.010041

Pan, W., Stone, K. P., Hsuchou, H., Manda, V. K., Zhang, Y., & Kastin, A. J. (2011). Cytokine signaling modulates blood-brain barrier function. *Current Pharmaceutical Design, 17*(33), 3729-3740.

Paolucci, S. (2008). Epidemiology and treatment of post-stroke depression. *Neuropsychiatric Disease and Treatment, 4*(1), 145-154.

Pascoe, M. C., Crewther, S. G., Carey, L. M., & Crewther, D. P. (2011). Inflammation and depression: why poststroke depression may be the norm and not the exception. *International Journal of Stroke, 6*(2), 128-135. doi: 10.1111/j.1747-4949.2010.00565.x

Pavlov, M. Y., & Ehrenberg, M. (2013). Optimal control of gene expression for fast proteome adaptation to environmental change. *Proceedings of the National Academy of Sciences of the United States of America, 110*(51), 20527-20532. doi: 10.1073/pnas.1309356110

Pekna, M., & Pekny, M. (2012). The neurobiology of brain injury. *Cerebrum: the Dana Forum on Brain Science, 2012*, 9.

Perini, F., Morra, M., Alecci, M., Galloni, E., Marchi, M., & Toso, V. (2001). Temporal profile of serum anti-inflammatory and pro-inflammatory interleukins in acute ischemic stroke patients. *Neurological Sciences, 22*(4), 289-296. doi: http://dx.doi.org/10.1007/s10072-001-8170-y

Pottiez, G., Wiederin, J., Fox, H. S., & Ciborowski, P. (2012). Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *Journal of Proteome Research, 11*(7), 3774-3781. doi: 10.1021/pr300414z

Rahpeymai, Y., Hietala, M. A., Wilhelmsson, U., Fotheringham, A., Davies, I., Nilsson, A.-K., . . . Pekna, M. (2006). Complement: a novel factor in basal and ischemia-induced neurogenesis. *The EMBO Journal, 25*(6), 1364-1374. doi: 10.1038/sj.emboj.7601004

Rankin, J. (1957). Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scottish Medical Journal, 2*(5), 200.

Rao, R., Jackson, S., & Howard, R. (2001). Depression in older people with mild stroke, carotid stenosis and peripheral vascular disease: a comparison with healthy controls. *International Journal of Geriatric Psychiatry, 16*(2), 175-183. doi: 10.1002/1099-1166(200102)16:2<175::AID-GPS298>3.0.CO;2-0

Robinson, R. G., Shoemaker, W. J., Schlumpf, M., Valk, T., & Bloom, F. E. (1975). Effect of experimental cerebral infarction in rat brain on catecholamine and behaviour. *Nature*, 255, 332-334.

Robinson, R. G., & Szetela, B. (1981). Mood change following left hemispheric brain injury. *Annals of Neurology, 9*(5), 447-453.

Roos, A., Bouwman, L. H., van Gijlswijk-Janssen, D. J., Faber-Krol, M. C., Stahl, G. L., & Daha, M. R. (2001). Human IgA activates the complement system via the mannan-binding lectin pathway. *The Journal of Immunology, 167*(5), 2861-2868. doi: 10.4049/jimmunol.167.5.2861

Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., . . . Pappin, D. J. (2004). Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics, 3*(12), 1154-1169. doi: 10.1074/mcp.M400129-MCP200

Sagen, U., Vik, T. G., Moum, T., Mørland, T., Finset, A., & Dammen, T. (2009). Screening for anxiety and depression after stroke: Comparison of the Hospital Anxiety and Depression Scale and the Montgomery and Åsberg Depression Rating Scale. *Journal of Psychosomatic Research, 67*(4), 325-332. doi: http://dx.doi.org/10.1016/j.jpsychores.2009.03.007

Schmid, A. A., Kroenke, K., Hendrie, H. C., Bakas, T., Sutherland, J. M., & Williams, L. S. (2011). Poststroke depression and treatment effects on functional outcomes. *Neurology, 76*(11), 1000-1005. doi: 10.1212/WNL.0b013e318210435e

Schmidt, A., Forne, I., & Imhof, A. (2014). Bioinformatic analysis of proteomics data. *BMC Systems Biology, 8*(Suppl 2), S3.

Sharma, P., Cosme, J., & Gramolini, A. O. (2013). Recent proteomic advances in cardiac cells. *Journal of Proteomics, 81*, 3-14. doi: 10.1016/j.jprot.2012.10.026

Sharma, R., Gowda, H., Chavan, S., Advani, J., Kelkar, D., Kumar, G. S. S., . . . Christopher, R. (2015). Proteomic signature of endothelial dysfunction identified in the serum of acute ischemic stroke patients by the iTRAQ-Based LC–MS approach. *Journal of Proteome Research, 14*(6), 2466-2479. doi: 10.1021/pr501324n

Sharp, F. R., Jickling, G. C., Stamova, B., Tian, Y., Zhan, X., Liu, D., . . . Ander, B. P. (2011). Molecular markers and mechanisms of stroke: RNA studies of blood in animals and humans. *Journal of Cerebral Blood Flow & Metabolism, 31*(7), 1513-1531. doi: 10.1038/jcbfm.2011.45

Shichita, T., Ito, M., & Yoshimura, A. (2014). Post-ischemic inflammation regulates neural damage and protection. *Frontiers in Cellular Neuroscience, 8*. doi: 10.3389/fncel.2014.00319

Shinjyo, N., Ståhlberg, A., Dragunow, M., Pekny, M., & Pekna, M. (2009). Complement-derived anaphylatoxin C3a Regulates in vitro differentiation and migration of neural progenitor cells. *Stem Cells, 27*(11), 2824-2832. doi: 10.1002/stem.225

Sofi, F., Cesari, F., Abbate, R., Gensini, G. F., Broze, G., & Fedi, S. (2010). A meta-analysis of potential risks of low levels of protein Z for diseases related to vascular thrombosis. *Thrombosis and Haemostasis, 103*(4), 749-756. doi: 10.1160/TH09-09-0645

Spalletta, G., Bossu, P., Ciaramella, A., Bria, P., Caltagirone, C., & Robinson, R. G. (2006). The etiology of poststroke depression: a review of the literature and a new hypothesis involving inflammatory cytokines. *Molecular Psychiatry*, *11*(11), 984-991.

Sridhar, S., Schembri, F., Zeskind, J., Shah, V., Gustafson, A. M., Steiling, K., . . . Brody, J. S. (2008). Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics*, *9*(1), 259.

St-Amour, I., Paré, I., Alata, W., Coulombe, K., Ringuette-Goulet, C., Drouin-Ouellet, J., . . . Calon, F. (2013). Brain bioavailability of human intravenous immunoglobulin and its transport through the murine blood–brain barrier. *Journal of Cerebral Blood Flow & Metabolism*, *33*(12), 1983-1992. doi: 10.1038/jcbfm.2013.160

Stankovic, S., & Majkic-Singh, N. (2010). Genetic aspects of ischemic stroke: coagulation, homocysteine, and lipoprotein metabolism as potential risk factors. *Critical Reviews in Clinical Laboratory Sciences*, *47*(2), 72-123. doi: 10.3109/10408361003791520

Stephan, A. H., Barres, B. A., & Stevens, B. (2012). The complement system: an unexpected role in synaptic pruning during development and disease. *Annual Review of Neuroscience*, *35*(1), 369-389. doi: doi:10.1146/annurev-neuro-061010-113810

Su, J.-A., Chou, S.-Y., Tsai, C.-S., & Hung, T.-H. (2012). Cytokine changes in the pathophysiology of poststroke depression. *General Hospital Psychiatry*, *34*(1), 35-39. doi: http://dx.doi.org/10.1016/j.genhosppsych.2011.09.020

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the*

National Academy of Sciences of the United States of America, 102(43), 15545-15550. doi: 10.1073/pnas.0506580102

Tang, W., Chen, Y., Lu, J., Chu, W. C., Mok, V., Ungvari, G. S., & Wong, K. (2010). White matter hyperintensities in post-stroke depression: a case control study. *Journal of Neurology, Neurosurgery & Psychiatry, 81*(12), 1312-1315.

Trinh, H. V., Grossmann, J., Gehrig, P., Roschitzki, B., Schlapbach, R., Greber, U. F., & Hemmi, S. (2013). iTRAQ-Based and label-free proteomics approaches for studies of human adenovirus infections. *International Journal of Proteomics, 2013*, 16. doi: 10.1155/2013/581862

Uchiyama, S., Yamazaki, M., Hara, Y., & Iwata, M. (1997). Alterations of platelet, coagulation, and fibrinolysis markers in patients with acute ischemic stroke. *Seminars in Thrombosis and Hemostasis, 23*(6), 535-541.

Urra, X., Miró, F., Chamorro, A., & Planas, A. M. (2014). Antigen-specific immune reactions to ischemic stroke. *Frontiers in Cellular Neuroscience, 8*, 278. doi: 10.3389/fncel.2014.00278

Van Beek, J., Bernaudin, M., Petit, E., Gasque, P., Nouvelot, A., MacKenzie, E. T., & Fontaine, M. (2000). Expression of receptors for complement anaphylatoxins C3a and C5a following permanent focal cerebral ischemia in the mouse. *Experimental Neurology, 161*(1), 373-382. doi: http://dx.doi.org/10.1006/exnr.1999.7273

Vataja, R., Leppävuori, A., Pohjasvaara, T., Mäntylä, R., Aronen, H. J., Salonen, O., . . . Erkinjuntti, T. (2014). Poststroke depression and lesion location revisited. *The Journal of Neuropsychiatry and Clinical Neurosciences*.

Vogelgesang, A., May, V. E. L., Grunwald, U., Bakkeboe, M., Langner, S., Wallaschofski, H., . . . Dressel, A. (2010). Functional status of peripheral blood T-Cells in ischemic stroke patients. *PLoS ONE, 5*(1), e8718. doi: 10.1371/journal.pone.0008718

Vu, N. Q., & Aizenstein, H. J. (2013). Depression in the elderly: brain correlates, neuropsychological findings, and role of vascular lesion load. *Current Opinion in Neurology, 26*(6), 656-661. doi: 10.1097/wco.0000000000000028

Wang, H., Alvarez, S., & Hicks, L. M. (2012). Comprehensive comparison of iTRAQ and label-free LC-Based quantitative proteomics approaches using two chlamydomonas reinhardtii strains of interest for biofuels engineering. *Journal of Proteome Research, 11*(1), 487-501. doi: 10.1021/pr2008225

Wetie, A. N., Woods, A., & Darie, C. (2014). Mass spectrometric analysis of post-translational modifications (PTMs) and protein–protein interactions (PPIs). In A. G. Woods & C. C. Darie (Eds.), *Advancements of Mass Spectrometry in Biomedical Research* (Vol. 806, pp. 205-235): Springer International Publishing.

Whyte, E. M., & Mulsant, B. H. (2002). Post stroke depression: epidemiology, pathophysiology, and biological treatment. *Biological Psychiatry, 52*(3), 253-264. doi: 10.1016/S0006-3223(02)01424-5

Wildsmith, K. R., Schauer, S. P., Smith, A. M., Arnott, D., Zhu, Y., Haznedar, J., . . . Honigberg, L. A. (2014). Identification of longitudinally dynamic biomarkers in Alzheimer's disease cerebrospinal fluid by targeted proteomics. *Molecular neurodegeneration, 9*(1), 1-14.

Williams, J. B. W., & Kobak, K. A. (2008). Development and reliability of a structured interview guide for the Montgomery–Åsberg Depression Rating Scale (SIGMA). *British Journal of Psychiatry, 192*(1), 52-58.

Worthmann, H., Tryc, A. B., Goldbecker, A., Ma, Y. T., Tountopoulou, A., Hahn, A., . . . Weissenborn, K. (2010). The temporal profile of inflammatory markers and mediators in blood after acute ischemic stroke differs depending on stroke outcome. *Cerebrovascular Diseases, 30*(1), 85-92.

Wu, X., Hasan, M. A., & Chen, J. Y. (2014). Pathway and network analysis in proteomics. *Journal of Theoretical Biology, 362*, 44-52. doi: http://dx.doi.org/10.1016/j.jtbi.2014.05.031

Xu, H., Tang, Y., Liu, D.-Z., Ran, R., Ander, B. P., Apperson, M., . . . Sharp, F. R. (2008). Gene expression in peripheral blood differs after cardioembolic compared with large-vessel atherosclerotic stroke: biomarkers for the etiology of ischemic stroke. *Journal of Cerebral Blood Flow & Metabolism, 28*(7), 1320-1328. doi: 10.1038/jcbfm.2008.22

Yi, J., Kim, C., & Gelfand, C. A. (2007). Inhibition of intrinsic proteolytic activities moderates preanalytical variability and instability of human plasma. *Journal of Proteome Research, 6*(5), 1768-1781. doi: 10.1021/pr060550h

Yilmaz, A., Fuchs, T., Dietel, B., Altendorf, R., Cicha, I., Stumpf, C., . . . Kollmar, R. (2009). Transient decrease in circulating dendritic cell precursors after acute stroke: potential recruitment into the brain. *Clinical Science, 118*(2), 147-157. doi: 10.1042/cs20090154

Zhan, Y., Yang, Y. T., You, H. M., Cao, D., Liu, C. Y., Zhou, C. J., . . . Xie, P. (2014). Plasma-based proteomics reveals lipid metabolic and immunoregulatory dysregulation in post-stroke depression. *European Psychiatry, 29*(5), 307-315. doi: http://dx.doi.org/10.1016/j.eurpsy.2014.03.004

Zhang, B., Ramesh, G., Uematsu, S., Akira, S., & Reeves, W. B. (2008). TLR4 signaling mediates inflammation and tissue injury in nephrotoxicity. *Journal of the American Society of Nephrology, 19*(5), 923-932. doi: 10.1681/asn.2007090982

Zubarev, R. A. (2013). The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics, 13*(5), 723-726. doi: 10.1002/pmic.201200451

Zunszain, P. A., Anacker, C., Cattaneo, A., Carvalho, L. A., & Pariante, C. M. (2011). Glucocorticoids, cytokines and brain abnormalities in depression. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 35*(3), 722-729. doi: 10.1016/j.pnpbp.2010.04.011

Chapter 7

Experimental Study 3

Longitudinal Stroke Recovery Dominated by Dysregulation of Complement System – A Proteomics Pathway Analysis

Preface

Following the pilot proteomics study conducted in Chapter 6, untargeted proteomics techniques were applied to investigations of the longitudinal recovery in stroke survivors. The unique characteristic of the present cohort had both clinical and blood samples available at longitudinal timepoints 3 months and 12 months, in addition to the acute 3-7 days.

More recent Advances in Systems Biology Markup Language allow databases to include intensity and expected direction of change (i.e positive or negative) in one molecule relative to another and thus enhance ability to codify molecule-to-molecule interactions on a biology wide scale. In addition to being able to present molecules as part of their constituent biological categories, databases can now host information such as the positive and negative interaction between two molecules. In the context of biological systems, regulation as a molecule-to-molecule interaction can now be studied on a systems-wide scale. Hence Gene Graph Enrichment Analysis was employed in this study to address the limitations of Gene Set Enrichment Analysis and was able to provide a much deeper level of explanation to the protein expression profile of stroke survivors with e.g. in regard to the complement system. Although both Chapters 5 and 6 found that the complement system was related to the post-stroke timeline Chapter 5 was only able to comment on complement system dysregulation whereas Chapter 6 has allowed quantification of both differential expression profiles and regulation effects over time.

This chapter has been published as ‘Longitudinal Stroke Recovery Associated With Dysregulation of Complement System—A Proteomics Pathway Analysis’ in *Frontiers in Neurology* on 28/07/2020, https://doi.org/10.3389/fneur.2020.00692.

Longitudinal Stroke Recovery Associated with Dysregulation of Complement System – A Proteomics Pathway Analysis

Vinh A. Nguyen[1,2,3,4]*, Nina Riddell[2], Sheila G. Crewther[2], Pierre Faou[5], Harinda Rajapaksha[5], David W. Howells[6], Graeme J. Hankey[7,8], Tissa Wijeratne[3,9], Henry Ma[10], Stephen Davis[11], Geoffrey A. Donnan[11] and Leeanne M. Carey[1,3]

[1] Department of Occupational Therapy, La Trobe University, Bundoora, VIC, Australia,
[2] Department of Psychology and Counselling, La Trobe University, Bundoora, VIC, Australia,
[3] Neurorehabilitation and Recovery, Stroke, The Florey Institute of Neuroscience and Mental Health, Heidelberg, VIC, Australia, [4]Western Health, Department of Neurology, Sunshine, VIC, Australia,
[5] Department of Biochemistry and Genetics, La Trobe University, Bundoora, VIC, Australia,
[6] Medical Sciences Precinct, University of Tasmania, Hobart, TAS, Australia,
[7] Faculty of Health and Medical Sciences, Internal Medicine,
University of Western Australia, Perth, WA, Australia,
[8] Clinical Research, Harry Perkins Institute of Medical Research, Perth, WA, Australia,
[9] Department of Medicine, The University of Melbourne, Sunshine, VIC, Australia,
[10] Monash Health, Neurology and Stroke, Clayton, VIC, Australia,
[11] Royal Me bourne Hospital, Department of Neurology, Parkville, VIC, Australia.

Corresponding author:
Mr. Vinh Nguyen

v.nguyen@latrobe.edu.au

Abstract

Currently the longitudinal proteomic profile of post-ischemic stroke recovery is relatively unknown with few well accepted biomarkers or understanding of the biological systems that underpin recovery. We aimed to characterise plasma derived biological pathways associated with recovery during the first year post event using a discovery proteomics workflow coupled with a topological pathway systems biology approach. Blood samples (n = 180, ethylenediaminetetraacetic acid plasma) were collected from a subgroup of 60 first episode stroke survivors from the Australian START study at 3 timepoints: 3-7 days (T1), 3-months (T2) and 12-months (T3) post-stroke. Samples were analysed by liquid chromatography mass spectrometry using label-free quantification (data available at ProteomeXchange with identifier PXD015006). Differential expression analysis revealed that 29 proteins between T1 and T2, and 33 proteins between T1 and T3 were significantly different, with 18 proteins commonly differentially expressed across the two time periods. Pathway analysis was conducted using Gene Graph Enrichment Analysis on both the Kyoto Encyclopedia of Genes and Genomes and Reactome databases. Pathway analysis revealed that the significantly differentiated proteins between T1 and T2 were consistently found to belong to the complement pathway. Further correlational analyses were utilized to examine the changes in regulatory effects of proteins over time identified significant inhibitory regulation of clusterin on complement component 9. Longitudinal post-stroke blood proteomics profiles suggest that the alternative pathway of complement activation remains in a state of higher activation from 3-7 days to 3 months post-stroke, while simultaneously being regulated by clusterin and vitronectin. These findings also suggest that post-stroke induced sterile inflammation and immunosuppression could inhibit recovery within the 3-month window post-stroke.

Keywords:
Longitudinal, stroke, proteomics, immune system, complement system

Introduction

Ischemic stroke covers a variety of cerebrovascular events that affect up to 800,000 people in the United States every year, with 133,000 deaths reported in 2017 (16.74%) (Benjamin et al., 2017). Of the survivors, 30% are reported to experience prolonged cognitive impairment (Sun, Tan, & Yu, 2014) and depressive symptoms at any point 5 years post-stroke (Robinson & Jorge, 2016). Currently there are few well accepted biomarkers for recovery and comparatively little literature exploring the biological systems that drive recovery or even the most optimal times for monitoring biological and behavioural recovery. Evidence from stroke rehabilitation studies suggest the greatest efficacy for motor-based rehabilitation is within this 3 month time window (Hatem et al., 2016), though recovery may continue at a slower rate over subsequent months and years. Although there has been increasing research examining the blood biomarkers of stroke recovery (Bang, 2017), the additional linking of biomarkers to biological systems remains speculative. Hence, we aimed to investigate the changes in the molecular profile of proteins in plasma samples via a mass spectrometry (MS) based discovery proteomics approach (Aslam, Basit, Nisar, Rasool, & Khurshid, 2017). Mass spectrometry and nuclear magnetic resonance (NMR) based techniques examining protein expression are among the most versatile techniques for protein identification and quantification, with the ability to address a wide range of biological samples, especially plasma and serum (Gowda & Djukovic, 2014; Xue et al., 2017). Proteomics utilizes the advantage of systems biology techniques to quantify a large number of analytes in an exploratory fashion, with a computational bioinformatics approach to further categorize biomarkers into biosystems (Chandramouli & Qian, 2009).

Proteomics have recently been used to pursue multiple clinical questions within stroke research, relating to differentiation of ischemic from hemorrhagic stroke (Montaner et al., 2012; Penn et al., 2018a; Penn et al., 2018b) and investigations of potential

biomarkers involved in post-stroke recovery (Nguyen et al., 2016; Zhou et al., 2017). Although traditional bioinformatics methods were originally developed to accommodate gene expression data, proteomics studies can utilize these methods to organize and visualize findings by adopting standardized change scores and using annotations that are common across proteomics and genomics (Schmidt, Forne, & Imhof, 2014). Indeed, our laboratory has previously used proteomic methods and Gene Set Enrichment Analysis (GSEA) to investigate the relationship between to investigate protein changes in plasma at 3-months post-stroke and affective (depression) outcomes (Subramanian et al., 2005).The results indicated that proteins involved in the complement but not the coagulation pathway of the immune system are likely to be associated with post-stroke depressive symptoms (Nguyen et al., 2016). The complement system is recognized as an innate immune pathway that contributes to primary host defence by encouraging phagocytosis of unwanted cells. This new study aims to expand upon our earlier single time point study by using discovery proteomics to identify longitudinal changes in blood plasma protein expression over the post-stroke timeline of recovery; specifically 3 timepoints post ischemic stroke: 3-7 days (T1), 3-months (T2) and 12-months (T3). We also aimed to improve upon our previous set-based functional annotation methods by utilizing Gene Graph Enrichment Analysis (GGEA). The GGEA approach differs from the GSEA approach by further addition to traditional set-based functional annotations by incorporation of *networks* of established biochemical relationships (gene to gene) using topological omics databases such as Reactome (Fabregat et al., 2018) and KEGG (Kanehisa & Goto, 2000) for the analysis of structural representation of biological pathways in the analytical workflow. This more novel approach further addresses the regulatory mechanisms in gene and protein pathways by examining co-expression and co-regulation networks using correlation analyses (Geistlinger, Csaba, Küffner, Mulder, & Zimmer, 2011). Examining the changes in correlational strength also allows for

quantification of the changes in the regulatory effect of proteins between timepoints. Although discovery approaches are hypothesis-free by nature, based on our previous study (Nguyen et al., 2016) and others (Shichita, Ito, & Yoshimura, 2014) that suggest inflammatory and immune homeostasis will be disrupted in the post-stroke recovery timeline, we hypothesise that the complement system will be dysregulated when comparing early 3-7 days post-stroke to later 3-month and 12-month post-stroke timepoints.

Materials and Methods

Subjects

Data from a subset of ischemic stroke patients was obtained from the longitudinal stroke cohort known as START, which comprised participants from START_PrePARE (STroke imAging pRevention and Treatment: Prediction and Prevention to Achieve optimal Recovery Endpoints; (20) (Neuroscience Trials Australia: NTA 0902) and START_EXTEND (STroke imAging pRevention and Treatment: EXtending the time for Thrombolysis in Emergency Neurological Deficits (21) (Neuroscience Trials Australia: NTA 0901, Clinicaltrials.gov number: NCT01580839)) studies. These prospective, integrated studies were longitudinal and provided long term plasma and serum samples at 3 time-points post-stroke. Selection of the subset of 106 ischemic stroke patients for the current study was initially based on the availability of ethylenediaminetetraacetic acid (EDTA) treated plasma samples at each of the 3 timepoints: 3-7 days (T1), 3-months (T2) and 12 months (T3). Laboratory batch processing limitations and time/cost considerations allowed for 180 EDTA plasma samples to be processed, i.e. samples from 60 stroke patients (43 male) across 3 timepoints. We therefore conducted a stratified selection process involving patients with complete blood samples at the 3 repeated times and full clinical scores, to ensure a spread of scores across clinical outcomes of stroke severity, mood and cognition. Although the present subset contained only participants who

completed the 3 sessions in the 12 months post first acute stroke event, 21 other deaths occurred, 8 others were lost to follow-up and 2 withdrew among the original complete START_PrePARE and START_EXTEND cohorts. See Table 1 below for baseline characteristics for patients and Table 2 for clinical characteristics at T1, T2 and T3 and Supplementary Table 1 for a comparison between the selected subset and the complete cohort. Healthy control data was not available as part of the original protocols for the START_PrePARE and START EXTEND studies. Ethics was approved by the Human Research Ethics Committee of Austin Hospital, Heidelberg (HREC code: H2010/03588), and relevant university and hospital sites.

Table 1
Baseline Sample Characteristics (*n* = 60)

Baseline	Mean	SD	Median	IQR
Age (years)	68.00	14.60	69.00	18.00
NIHSS	4.70	4.76	2.50	3.00
Height (cm)	160.41	26.80	170.00	20.00
Weight (kg)	74.66	19.80	80.00	26.75
Heart Rate (per minute)	74.52	11.06	74.50	14.50
Systolic Blood Pressure (mm Hg)	141.47	23.27	134.50	29.75
Diastolic Blood Pressure (mm Hg)	78.50	11.77	78.00	14.00
TOAST Criteria	Frequency	Percentage		
Large artery atherosclerosis	14	24.14%		
Cardioembolism	6	10.34%		
Small vessel occlusion	15	25.86%		
Stroke of other determined aetiology	3	5.17%		
Stroke or undetermined aetiology, two or more causes identified	3	5.17%		
Stroke of undetermined aetiology, negative evaluation	3	5.17%		
Stroke of undetermined aetiology, incomplete evaluation	14	24.13%		
Comorbidities				
Past Atrial Fibrillation	4	6.7%		
Hypertension	26	73.3%		
Lipid Disorder	24	40.0%		
Ischemic Heart Disease	11	18.3%		
Diabetes Mellitus	9	15.0%		

Note. *n* = 60, NIHSS: National Institute of Health Stroke Scale, 0: No Stroke symptoms, 1-4: Minor Stroke, 5-15: Moderate Stroke, 16-20: Moderate to Severe Stroke, 21-42: Severe Stroke (Lyden *et al.*, 2009)
TOAST: Trial of Org 10172 in Acute Stroke Treatment classification for ischemic stroke subtype

Table 2
Sample Clinical Characteristics at 3-7 days, 3 months and 12 months (n = 60)

	3-7 Days (T1)				**3 Months (T2)**				**12 Months (T3)**			
	Mean	SD	Median	IQR	Mean	SD	Median	IQR	Mean	SD	Median	IQR
Weight (kg)					76.30	15.30	78.00	19.00	78.75	13.57	79.50	20.00
Heart Rate (per minute)					63.02	21.03	65.00	20.00	61.50	21.01	70.00	19.00
Systolic Blood Pressure (mm Hg)					125.76	16.92	125.50	21.00	124.28	15.34	124.00	17.25
Diastolic Blood Pressure (mm Hg)					74.78	10.17	76.00	11.75	72.47	9.78	70.00	15.00
NIHSS	2.35	3.05	1.00	3.00	1.18	2.31	0.00	1.00	2.45	12.87	0.00	1.00
mRS*					1.25	1.31	1.00	2.00	1.12	1..22	1.00	2.00
MoCA	24.21	5.00	26.00	5.00	25.93	4.57	28.00	5.00	25.07	5.02	26.00	4.00
MADRS	6.93	7.31	4.00	11.00	8.72	8.81	5.50	13.75	7.13	7.41	5.00	10.75

NIHSS: National Institute of Health Stroke scale, mRS: modified Rankin Scale, MADRS: Montgomery-Åsberg Depression Rating Scale, MoCA: Montreal Cognitive Assessment.

* The mRS is not conducted at 3-7 days as it is a measure of post-stroke disability within previous 30 days.

Blood Collection and Storage

All samples were collected in Benton-Dickson (BD) EDTA-coated 4ml vacutainers and were mixed and left to stand in ambient room temperature for 30 minutes. Average time to blood draw was 3.45 days (range = 2.03 – 7.58 days, SD = 1.174) after stroke onset for T1. The tubes were then centrifuged at 1100-1300 *g* at room temperature and the resulting plasma was aliquoted into cryotubes and immediately stored in a -80^{o-}C freezer. For transport from the central study freezer to the analysis site, temperature was kept at -70^{o-}C to -80^{o-}C on dry ice before transfer into a -80^{o-}C freezer. This procedure was consistent for the 3-month and 12-month follow up periods.

Mass Spectrometry

Label-free quantitation (LFQ) proteomics was conducted on the Q Exactive HF Orbitrap instrument (Thermo-Fisher Scientific). Details of the sample preparation, instrument parameters and protein identification are available in the Supplementary Methods.

Bioinformatics and Statistical Analysis

The initial MaxQuant output consisted of 358 identified proteins across 180 samples. After removal of site-identified contaminants and filtering for at least 80% of data present across the protein, 163 proteins remained (Supplementary Data 1). This procedure regarding missing values is an expected step in addressing proteomics data as the protein identification search on MaxQuant (utilizing an FDR of 1%) will not output as an identified protein unless there is a single unique peptide, regardless of peptide length. The Limma R package (version 3.34.9) with Benjamini-Hochberg (BH) adjustments for familywise error rate was used to compare which proteins were differentially expressed across T1, T2 and T3. To further understand the functional organisation of significant protein sets, the EnrichmentBrowser R package (version 2.09.17) was employed with GGEA as the network-based enrichment method across the KEGG and Reactome

databases using default settings (minimum set size = 3, maximum set size = 500, permutations = 1000). Correlations with BH corrections were used to further explore the changes in proteins that regulate identified pathways, with Fisher *r* to *z* transformations conducted to explore changes in regulatory effect across timepoints. Statistical significance levels were set to $a = .05$ after multiple comparisons adjustment.

Data Availability

Due to the requirements of ethics and the nature of ongoing clinical trials with the START cohorts, unidentified patient clinical data may only be made available upon request. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD015006.

Results

The differential expression (DE) analysis revealed that 29 proteins significantly differed between T1 and T2, and that 33 proteins significantly differed between T1 and T3 (Table 3) with 8 proteins at FDR of 0.05 indicating that the effects are higher than would occur due to chance alone. The changes between T1 to T2 and T1 to T3 constitute 17.79% and 20.25% respectively in proportion to the total number of proteins identified. Eleven proteins were uniquely expressed between T1 and T2 and 15 proteins were uniquely expressed between T1 and T3, with 18 of the same proteins expressed both between T1 and T2 and T1 and T3. There were no significant differences in protein expression between T2 and T3, potentially suggesting that the currently identified proteome in post-stroke survivors does not change significantly between 3-month and 12-month times. See Supplementary Data 1 for a full list of proteins and fold change values across all comparisons.

Table 3
Differentially expressed proteins (BH $p < .05$) detected between T1, T2 and T1, T3

3-7 Days (T1) to 3 Months (T2)			3-7 Days (T1) to 12 Months (T3)		
Gene Symbol	log2 Fold Change	adj *p*	Gene Symbol	log2 Fold Change	adj *p*
C8A	0.6657	0.0000	C8A	0.7394	0.0000
ACTB	0.3766	0.0000	APOA4	0.3585	0.0000
APOA4	0.3169	0.0000	ACTB	0.3712	0.0000
A1BG	-0.3210	0.0004	CFI	0.3346	0.0000
C9	-0.2814	0.0027	APOA2	0.2349	0.0003
CLEC3B	0.1814	0.0027	C9	-0.3231	0.0003
FBLN1	0.2953	0.0044	PGLYRP2	0.1863	0.0006
CFI	0.2417	0.0086	CLEC3B	0.1992	0.0006
APOD	0.1584	0.0104	APOB	-0.1754	0.0049
PGLYRP2	0.1634	0.0104	CFD	0.2459	0.0065
APOF	-0.2793	0.0109	CPN1	-0.1435	0.0065
IGFALS	0.1975	0.0131	TF	0.2243	0.0065
F9	-0.2180	0.0133	LRG1	-0.2285	0.0066
SERPINA3	0.1332	0.0133	RBP4	0.2153	0.0067
SAA2-SAA4	-0.2462	0.0141	FGA	0.1504	0.0068
RBP4	0.1836	0.0146	APOD	0.1399	0.0068
TF	0.2026	0.0146	A1BG	-0.2353	0.0073
APOC1	-0.1699	0.0286	IGFALS	0.1906	0.0073
FETUB	0.1995	0.0286	SAA2-SAA4	-0.2728	0.0081
HPR	-0.1673	0.0295	A2M	-0.1860	0.0095
PLTP	0.3322	0.0295	AFM	0.1487	0.0160
SAA1	-0.9248	0.0295	C4BPB	-0.2011	0.0160
FGB	-0.2347	0.0308	CSPG4	0.2386	0.0160
APOB	-0.1347	0.0389	C4BPA	0.1206	0.0187
C8G	-0.1511	0.0389	APCS	-0.2772	0.0198
VTN	0.0723	0.0413	VTN	0.0737	0.0244
SERPINA10	-0.2143	0.0420	GPLD1	0.2335	0.0274
HPX	0.0918	0.0433	ORM1	-0.3497	0.0274
C6	0.0980	0.0493	LBP	-0.2497	0.0282
			FBLN1	0.2117	0.0296
			SAA1	-0.9380	0.0334
			SERPINA3	0.1000	0.0374
			HPX	0.0898	0.0440

Note. $p < 0.05$ adjusted by the Benjamini-Hochberg procedure was considered statistically significant. Lists are ordered by *p* value.

The list of DE proteins and the full expression matrix were submitted to EnrichmentBrowser using GGEA as the network enrichment method (1000 permutations and a = .05). This revealed significant functional annotation of our set of proteins in both

KEGG and Reactome databases only between T1 and T2, after BH adjustment of *p* values (Table 4) (see Supplementary Data 2 a full list of nominally significant sets from KEGG and Reactome). In the case of the Reactome database, the nested structures in the ontological organisation of the pathways can lead to redundancy resulting in identification of multiple pathways, especially as the 'Complement Cascade' is located as a subset of the 'Innate Immune System' and the 'Immune System.' Further interpretation is needed to account for the number and type of identified proteins between the reference pathways and experimental data and should occur at the level closest to the relevant biological processes. Both KEGG and Reactome databases concurred on the identification of the complement system as the significantly enriched network in these plasma samples.

Table 4
Significantly Enriched Gene Pathways between T1 (3-7 days) and T2 (3-month) timepoints in Stroke Survivors

Gene Set	Identifier	Normed Score	adj *p*
KEGG			
Complement and Coagulation Cascades	hsa04610	1.91	0.003
Reactome			
Immune System	R-HSA-168256	0.497	0.0009
Innate Immune System	R-HSA-168249	0.497	0.0009
Complement Cascade	R-HSA-166658	0.497	0.0009

Note. $p < 0.05$ adjusted by the Benjamini-Hochberg procedure was considered statistically significant.

A visual illustration of this enriched pathway on the Reactome database reveals the changes and interactions of the detected molecules on the complement pathway (Supplementary Figure 1), displaying information on nodes (proteins) and edges (the correlation between nodes). Similarly, the results from the KEGG database includes elements of the complement pathway but also the coagulation cascades (Supplementary Figure 2). As the complexity of information is difficult to interpret without a significant degree of system specific understanding in these pathway diagrams, a composite network

pathway was created to amalgamate the statistical relationships found in both diagrams pertaining to the complement system as shown in Figure 1.

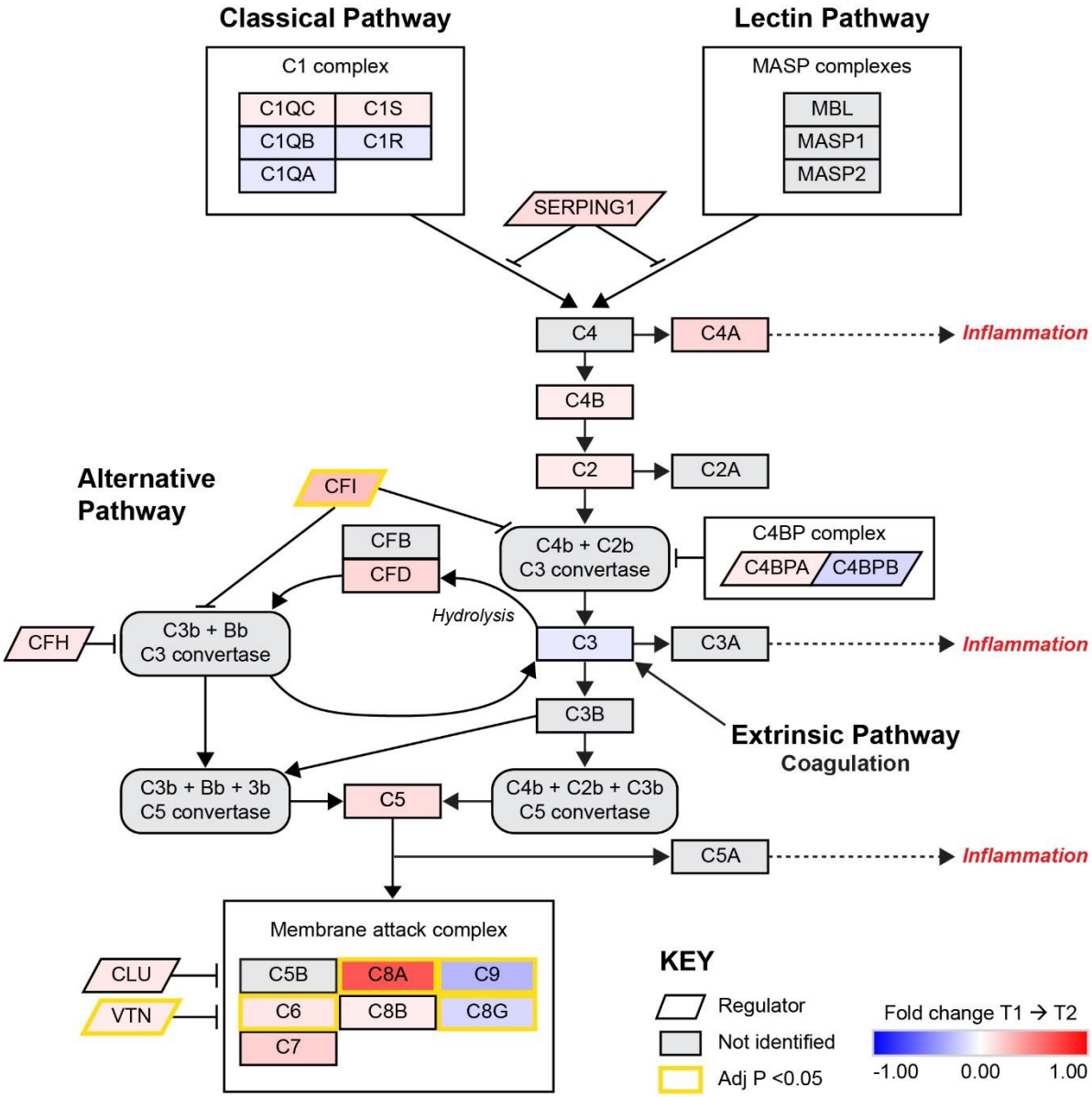

Figure 1. A pathway diagram contrasting information from differential proteomic analysis of KEGG and Reactome GGEA complement system pathways at 3-7 days (T1) and 3-months (T2) post-stroke. This pathway is activated by the C1q complex antigen linked immunoglobulin through the classical pathway and pathogen surface linked mannose binding linked (MBL) proteins through the MASP complex. The alternative pathway of complement activation follows a constant and low state of activation by a feedback loop that is heavily regulated by complement factor I (CFI) and complement factor H (CFH). Simultaneous upregulation of complement factor D (CFD) and CFI suggest that the alternative pathway is undergoing activation but also being regulated to prevent autoimmune insult. Recently, the link between coagulation and complement has been established through thrombin, plasmin and factor XIIa, these factors were not identified or significantly related to complement in this model. All activation pathways lead to the

promotion of cleavage of fluid phase complement component 3 (C3) in the bloodstream, with heavier β chains forming convertases downstream and lighter α chains such as C3a and complement component 5α (C5a) able to signal for potent local and systemic inflammation responses, with the most extreme being anaphylactic shock. The final stage of complement activation is the membrane attack complex (MAC), a large protein complex that is constructed on cells to disrupt the outer membrane and promotes active cytolysis. The profile of significant differences in MAC proteins such as C8G and C9 suggest plasma complement regulation by vitronectin (VTN) or clusterin (CLU).

Some of the regulatory proteins in the complement pathway (CLU & VTN) were not included in the pathway diagrams for KEGG and Reactome databases used in this analysis. Therefore, a correlation matrix with BH corrected *p*-values was created to explore the direction and strength of the regulatory effects between regulator and related proteins (Table 5). Fisher's transformation of Pearson's *r* to normally distributed *z* scores was computed to compare changes in regulatory effects between T2 and T1. While not overall presenting significant results, examination of the r_{diff}(T2-T1) complement regulators on their target proteins show that the middle stage complement regulators shifted their effects forward, resulting in a reduced inhibitory effect. Only the inhibitory effect of CLU on C9 was significantly increased from T1 to T2, with VTN also showing a large effect size for changes in C9 regulation.

Table 5

Pearson's Correlation matrix with Fisher *r* to *z* transformations for complement proteins and their regulators between T1 and T2

Stage	Regulator	Complement Protein	*r* (T1)	*r* (T2)	r_{diff}(T2-T1)	z_{diff}(T2-T1)	adj *p*
Early	SERPING1	C1QA	-0.01	0.06	0.07	0.37	0.71
	SERPING1	C1QB	0.13	0.16	0.03	0.15	0.88
	SERPING1	C1QC	0.25	-0.01	-0.26	-1.44	0.15
	SERPING1	C1R	-0.11	0.00	0.12	0.62	0.53
	SERPING1	C1S	-0.18	-0.15	0.02	0.12	0.90
Mid	C4BPA	C4A	-0.03	0.17	0.20	1.08	0.28
	C4BPA	C4B	-0.08	0.08	0.16	0.86	0.39
	C4BPA	C2	0.04	0.06	0.01	0.08	0.94
	C4BPB	C4A	-0.13	0.17	0.31	1.64	0.10
	C4BPB	C4B	-0.15	-0.08	0.07	0.40	0.69
	C4BPB	C2	-0.26*	-0.05	0.21	1.13	0.26
	CFH	C3	0.37*	0.54*	0.16	1.10	0.27
	CFI	C3	0.15	0.27*	0.12	0.67	0.50
Late	CLU	C6	0.13	0.05	-0.07	-0.40	0.69
	CLU	C7	0.01	0.04	0.03	0.14	0.89
	CLU	C8A	0.10	-0.03	-0.13	-0.68	0.50
	CLU	C8B	0.35*	0.14	-0.20	-1.15	0.25
	CLU	C8G	-0.24	-0.18	0.06	0.31	0.76
	CLU	C9	-0.02	-0.49*	-0.47**	-2.74	0.01
	VTN	C6	0.25	0.19	-0.06	-0.32	0.75
	VTN	C7	0.23	0.10	-0.12	-0.68	0.49
	VTN	C8A	0.42*	0.28*	-0.14	-0.85	0.40
	VTN	C8B	0.24	0.29*	0.05	0.28	0.78
	VTN	C8G	0.10	0.15	0.04	0.24	0.81
	VTN	C9	0.04	-0.30*	-0.34	-1.87	0.06

Note. The early stage of the complement pathway consists of activation units, the middle stage consists of protein activity related to cleaving C3 and C5, with the late stage consisting of proteins involved in membrane attack complex (MAC) formation. These correlations only show the effect of a single regulator on the related protein and does not take into consideration the downstream pathways.

p was BH adjusted * $p < .05$, ** $p < .01$

Discussion

Currently the interaction of the many biological mechanisms limiting post-stroke recovery remain poorly understood. To better understand the biochemical pathways impacting/affecting the post-stroke timeline of recovery, this study employed a discovery approach utilizing mass spectrometry to examine protein expression in patients at 3-7 days (T1), 3-months (T2), and 12-months (T3) post-stroke. Two sets of proteins were identified to be significantly different based on differential expression analysis from T1 to T2 and T1 to T3, but not between T2 and T3. Of these proteins, complement (C8A, C9, CFI), apolipoproteins (APOA4, APOA2, APOD) and membrane bound proteins (TF, ACTB) are highly abundant in human blood samples (von Zychlinski, Williams, McCormick, & Kleffmann, 2014). The lists of significantly differently expressed proteins were analysed using the GGEA bioinformatics algorithm based on the topological consistency of the observed data sets compared to database defined reference pathways. This revealed that the proteins identified in this longitudinal experiment significantly conformed to pathways central to the complement system in both KEGG and Reactome databases. To our knowledge, this will be first published use of the GGEA approach in clinical proteomics of human blood.

There is currently limited understanding of the changes in complement system in post-stroke recovery, especially relative to blood plasma concentrations in similar aged individuals (mean age 68 ± 14 years). Furthermore, most studies examining the complement system post-stroke have described its damage-exacerbating role in acute stroke human and animal models but have not examined changes in levels longitudinally over time post-stroke (Alawieh, Elvington, & Tomlinson, 2015; Alawieh, Langley, & Tomlinson, 2018; Di Napoli, 2001; Stokowska et al., 2017; Stokowska & Pekna, 2018). Many of the studies exploring complement in stroke have examined the system in the context of danger associated molecular pattern (DAMP) signalling and subsequent

neuronal repair in central nervous system (CNS) specific injury (Alawieh & Tomlinson, 2016), with plasma and serum-based studies examining levels in relation to clinical outcomes (Stokowska et al., 2011; Zhang, Yang, & Gao, 2015). The endothelial junctions forming the blood brain barrier (BBB) have been traditionally viewed as preventative to the entry of large peripherally derived molecules such as complement into the cerebral parenchyma (Jacob & Alexander, 2014; Veerhuis, Nielsen, & Tenner, 2011). The CNS has also been recognized as able to endogenously biosynthesize complement proteins in glial cells (Gasque, Fontaine, & Morgan, 1995). Although disruption to BBB permeability allows for the passage of complement proteins from periphery to CNS (Veerhuis et al., 2011), to our knowledge, no study has simultaneously examined the levels and functional activity of the complement system in both domains. Therefore, the results of this study have been interpreted as referring to the circulating fluid phase proteins that enhance the decay of complement proteins by cleaving active proteins through allosteric binding (Alawieh et al., 2015).

The complement pathway has three traditional modes of activation, the classical, lectin and alternative pathways. The classical and lectin pathways are activated by antigen and pathogen binding complexes respectively, whereas the alternative pathway is an internal biomechanical activation loop that regulates downstream pathway activity (Noris & Remuzzi, 2013). In the classical pathway of complement activation, upregulation of SERPING1 or the C1 inhibitor protein is responsible for regulation of the C1 complex. Specifically, SERPING1 strongly binds to the C1r and C1s proteases and also to the activation units of the lectin pathway (mannose binding lectin, MBL, and associated proteases MASP1 and MASP2 (Orsini et al., 2016)), effectively inhibiting the initial activation effects of the complement system in both antigen and pathogen related pathways. In addition to the effects on the complement system, high concentrations of this protease can further inhibit leukocyte-endothelial adhesion and vascular rolling via

interference of cell adhesion molecules (CAMs) or selectins, the first physiological requirements for trans-endothelial leukocyte infiltration (Cai et al., 2005). Although a fold change of .008 or .8% for SERPING1 was detected between T1 and T2 in this study, previous reports have indicated that a fold change can be observed during acute inflammation and that higher levels may be therapeutic in preventing autoimmune injury by providing a stop mechanism to complement system activation (Cicardi, Zingale, Zanichelli, Pappalardo, & Cicardi, 2005; Davis, 2004). These acute properties have been demonstrated in animal models of post-stroke middle cerebral artery occlusion (MCAO) and shown to attenuate ischemic reperfusion injury (Albert-Weissenberger et al., 2014; Heydenreich et al., 2012). Limited evidence for longitudinal changes in levels of SERPING1 in human cardiovascular disease have previously been suggested as a link to low-grade levels of chronic inflammation (Hertle et al., 2018). Furthermore, the incidence rates of post-stroke immunological trajectory of post-stroke recovery may also be explained by the immune inhibitory effects of SERPING1, following the initial phase of post-stroke immune flux and subsequent immunosuppression (Kamel & Iadecola, 2012).

The cleavage of C3 is the central amplification step of the whole complement pathway, with active C3b forming the C5 convertase and C3a inducing further phagocytotic chemotaxis (Stokowska & Pekna, 2018). Given that C3 is central in all complement activation pathways, especially with the alternative pathway requiring C3 hydrolysis, much of the theorized regulatory activity in this pathway is focused on inhibition of C3 convertase formation and decay rate (Noris & Remuzzi, 2013). Of the complement regulatory proteins, CFH and CFI were overexpressed, with only plasma CFI levels shown to be statistically different. Additionally, the changes in regulatory effect for middle stage complement proteins show that even though CFH and CFI are upregulated between T1 and T2, their inhibitory effects were reduced. Although CFI and CFH are cofactors of C3 convertase (not identified in this study) inhibition and not C3 itself, the

overall profile suggests a relative lack of effectiveness in the functional ability of these proteins to regulate alternative complement pathway activity in this older group. Furthermore, this profile may also be indicative of overexpression in levels of CD55 or decay accelerating factor (DAF), a protein that inhibits C3 and C5 convertase formation, albeit on cell surfaces (Carter, 2012). Complement C3a has previously been shown to be acutely elevated from 1 to 28 days post-stroke (Cojocaru et al., 2008; Mocco et al., 2006). Indeed, ongoing anaphylatoxin signalling from components such as C3a and C5a are damaging to host systems, especially in cerebral ischemia (Alawieh et al., 2015). The quantification of properdin, a protein that acts as a positive regulator of C3 and C5 convertases (Kouser et al., 2013), would further ameliorate understanding of alternative pathway activation in this sample.

Complement component C5 is the important functional unit of the pathway. Cleavage of soluble C5 protein into C5b is able to create the porous membrane attack complex (MAC) on targeted cells to induce cell death by cytolysis while C5a is able to produce local and system wide inflammatory cascades that are 20 times more potent than C3a (Mak & Saunders, 2006). The profile here shows that C5 is upregulated at 3-months, despite reduced C3 levels. This suggests that the reduction in C3 levels may be indicative of active cleaving to produce downstream increased C3b and C3a levels to produce or maintain constant immune system homeostasis. The detection of increased C5 may also imply involvement of the novel complement pathway or the extrinsic complement pathway, linking coagulation and complement systems. This pathway was originally thought to be activated based on thrombin functionally substituting for the C3 dependent C5 convertase (Huber-Lang et al., 2006). Recently, a study has also demonstrated that in thrombosis, interactions with plasmin in both surface bound and fluid phase complement is capable of upregulating C5a and C5b whether in venous or arterial thrombi (Foley et

al., 2016). In our analysis however, plasminogen was not found to be significantly differentially expressed or related to the complement pathway (Supplementary Figure 2).

Our methodology identified and quantified all the molecular components in MAC formation pathway. Our findings demonstrate that C6 and C8A were significantly upregulated and C8G and C9 were significantly downregulated between T1 and T2. Regulators of MAC formation exhibit function by preventing polymerisation of the C8 or C9 complexes, thereby inhibiting cytolytic function on cell surfaces (Hadders et al., 2012). Pore formation is heavily regulated by CD59 or MAC-inhibitory protein on host cell surfaces to prevent further immune insult resulting from early tissue induced hyperinflammation and gradual immune dysregulation (Ricklin, Reis, & Lambris, 2016). As this expression profile was obtained in plasma, the expression profile here is suggestive of fluid phase regulation of MAC formation, dependent on vitronectin (VTN or S-protein) and clusterin (CLU or apolipoprotein J). To create a successful MAC, from 12 to 16 molecules of C9 are needed as the final step to form a cyclical pore in addition to 1 of each previous MAC protein in the chain (Parsons et al., 2019). Both VTN and CLU inhibit the formation of the MAC insertion C5b-7 complex and also inhibit the ability of C9 to bind to inserted MAC complexes to create functioning lysis pores (Parsons et al., 2019). VTN and CLU were both found to be overexpressed, with the increases in VTN shown to be statistically significant (Noris & Remuzzi, 2013). Overall, the profile of the complement system here appears to be dysregulated, with an increase in alternative pathway activation by CFB, and reduced inhibitory effectiveness of all middle stage complement proteins but especially CFI and CFH. The apparent contradiction in expected abundances of MAC proteins suggests that the pathway is also being inhibited by CLU and VTN.

Although few studies have presented this specific profile of complement expression, and currently none in a complex proteomics pathway model, there are some

interpretations relevant to stroke survival and recovery that can be made here. From the upregulation of C5 and initial MAC proteins, it is possible that C5 is being continually cleaved to produce the pro-inflammatory chemokine C5a. This can be understood in theory of post-stroke immune profiles where both central and peripheral systems never fully recover to pre-stroke levels and establish new albeit dysfunctionally higher inflammation profiles (Kamel & Iadecola, 2012).

The profile of complement activation described in this study is likely to indicate that

1) the internal activation loop that readies the immune system for pathological threats and tissue damage (the alternative pathway) is activated,

2) middle stage complement regulators that are normally expected to be activated to prevent indiscriminatory autoimmune damage are overexpressed but their intended inhibitory effects (CFI -> C3) are reduced and

3) that although early chain MAC proteins (C6, C7, C8A, C8B) trend towards overexpression by time T2 are likely due to alternative pathway activation, CLU and VTN are also overexpressed, possibly as a compensatory mechanism for alternative pathway activation and thereby preventing autoimmune MAC formation by inhibiting C9 expression on host/friendly cells.

Our results did not find /demonstrate an association between C8G expression and CLU or VTN, although downregulation of C8G is theorized to be coincidental with immunosuppression (Morikis & Lambris, 2005). It is speculated that based on positional role of C8G as the protein that provides a binding site for the first C9 molecule in the structural formation of the MAC pore (see (Serna, Giles, Morgan, & Bubeck, 2016) for an overview of MAC protein structure), the significant downregulation of C8G between T1 and T2 that can likely be attributed to inhibitory CD55 or CD59 (that was not present in our proteomics analysis). Furthermore, there are additional molecules and mechanisms

that may regulate MAC formation through C8G expression such as anti-apoptotic processes that promote cell survival (Portt, Norman, Clapp, Greenwood, & Greenwood, 2011). This may also explain the immunosuppressed state of post-stroke patients and infections (Westendorp, Nederkoorn, Vermeij, Dijkgraaf, & de Beek, 2011), whereby the system is not able to react as efficiently to acquired pathogens such as viral pneumonia due to increased fluid phase MAC regulation (Armstrong & Mosher, 2011).

Limitations and Future Research

These results are limited to interpretation based on the sample characteristics of a cohort of mild stroke patients over a 12-month period post-stroke. During this period, it is recognised that other potential concurrent processes such as recurrent stroke, infections, venous thromboembolism could impact on the results: however, we were not able to control or adjust for such factors in the current analyses. The lack of an age matched control sample represents a challenge to external validity, and it is uncertain if these changes could also occur in people without stroke, as a process of senescence and aging or other comorbidities. Additionally, our results were not controlled for by age and baseline stroke as the analyses using the EnrichmentBrowser package did not have a streamlined function to address confounds. The profile of complement expression in this sample may also have been influenced by the myriad of post-stroke medications that currently include aspirin, warfarin, and clopidergel. Although these pharmacological agents target aspects of the coagulation system, it is possible that they also exhibit regulatory effects on the complement system (Arachchillage et al., 2016) via the intersection between hemostasis and innate immunity (Keragala, Draxler, McQuilten, & Medcalf, 2018). These confounds remain difficult to statistically control in the context of bioinformatics models, and should form part of the conceptualization in the design phase in future studies (Maes, Cho, & Baggerman, 2015). While the focus of the current study was to explore evidence of change in biological pathways over the first-year post-stroke,

it is recommended that future studies investigate the relationship between changes in these pathways and changes with age as well as in neurological and functional recovery profiles.

The sample preparation methods used in this study did not include fractionation or depletion of the plasma samples, leading to a low number of proteins identified relative to the blood proteome. Blood samples represent a very complex sample with high dynamic range for proteomics analysis, with a proportionally high abundance of proteins such as albumin, haptoglobin, fibrinogen and immunoglobulin G (Dayon & Kussmann, 2013). Future studies could increase the range of identified proteins by depleting abundant plasma proteins with various methods before MS analysis (Cao, Yende, Kellum, & Robinson, 2013; Tu et al., 2010); although some popular depletion columns also remove proteins identified here as differentially expressed such as those within the complement system (Smith et al., 2011). In addition to this, limitations to the discovery approach imply that there is a degree of uncertainty in the ability to identify a complete group of proteins (Doerr, 2009; Faria et al., 2017), especially where set and network-based bioinformatics are the intended analyses. Therefore, after pathways are identified by robust techniques such as GGEA, the integrity of these pathways can be enhanced by examining proteins that are missing in the pathway and directly related to pathway activation and function by targeted proteomics techniques such as single or mass reaction monitoring (SRM or MRM) in MS with immunoaffinity enrichment (Whiteaker et al., 2017).

In this study, the LC-MS methodology did not detect the specific formation of C3 or C5 chains or convertase protein complexes. Therefore, the pathway representations of relative abundance for C3 and C5 may not necessarily reflect the abundance of downstream protein chains of the active products. Specifically, C3a and C5a would be interesting proteins to target as they have potent effects on inflammation physiology and

would be of concern especially if shown to be elevated at 3 months post-stroke. Fluid phase identification of inactive pre-MAC complexes with protein S such as sC5b-7, 8 and 9 in addition to VTN and CLU may also provide further information on the regulatory activity of the MAC in stroke survivors. Furthermore, the longitudinal profile of complement in this study and others (Alawieh et al., 2015; Mocco et al., 2006) suggests that complement dysregulation begins in the first week post-stroke and continues for at least 3 months, and may set a post-stroke level of immune homeostasis. This finding needs to be further validated at other timepoints within this window, especially in relation to infection-based mortality (Shi, Wood, Shi, Wang, & Liu, 2018). To this extent, it is possible that the post-stroke immune profile is settled at 3 weeks to 1 month post-stroke when adaptive immunity is engaged in phagocytizing dead cells (Kamel & Iadecola, 2012) and resolution of cerebral microedema (Stokum, Gerzanich, & Simard, 2016).

Conclusion

The biology of post-stroke recovery is not well understood, with patients exhibiting varying profiles based on factors such as their age, location of lesion and degree of stroke damage (Pérez et al., 2016). This study aimed to characterize common peripheral biological systems involved in post-stroke recovery by examining the longitudinal proteomics profile of EDTA plasma in stroke survivors. Specifically, elements of the alternative pathway of complement system and MAC proteins, i.e. CFI, C6, C8A, C8G, C9 were found to be activated but also undergoing regulation at 3 months post-stroke. This increased turnover may lead to the upregulation of anaphylatoxins C3a and C5a that could explain the prolonged sterile inflammation profile of post-stroke survivors. These results suggest a biomechanism for post-stroke immunosuppression by complement system regulatory proteins. This knowledge may be useful to guide assessment in the clinical setting for post-stroke infections and immune recovery in the 3-month recovery window. For example, it may be used to inform future recommendations

for in-hospital and post-stroke infections by suggesting clinical utility of complement panel blood tests. Future investigations should examine the clinical implications of this biological profile and determine the temporal trajectory of immune homeostasis post-stroke.

Acknowledgements

The authors acknowledge the stroke patients and all research personnel associated with the START_EXTEND and START_PrePARE program of research.

Funding

We acknowledge financial support for conduct of the START_EXTEND and START_PrePARE studies from the Commonwealth Scientific and Industrial Research Organization (CSIRO) of Australia, Flagship Collaboration Fund through the Preventative Health Flagship; and support for analysis, write up and researchers from the James S. McDonnell Foundation 21st Century Science Initiative in Cognitive Rehabilitation - Collaborative Award (# 220020413); NHMRC Centre of Research Excellence in Stroke Rehabilitation and Brain Injury (#1077898); NHMRC program grant Saving brain and changing practice in stroke (#1113352); Victorian Government's Operational Infrastructure Support Program; an Australian Research Council Future Fellowship awarded to LMC [#FT0992299]; and a La Trobe University Post Graduate Scholarship supported by the Understanding Diseases research focus group awarded to VAN. The funding sources had no role in the conduct of this study or the writing of this report.

Competing Interests

The authors declare no competing interests.

References

Alawieh, A., Elvington, A., & Tomlinson, S. (2015). Complement in the homeostatic and ischemic brain. *Frontiers in Immunology, 6,* 417. doi:10.3389/fimmu.2015.00417

Alawieh, A., Langley, E. F., & Tomlinson, S. (2018). Targeted complement inhibition salvages stressed neurons and inhibits neuroinflammation after stroke in mice. *Science Translational Medicine, 10*(441).

Alawieh, A., & Tomlinson, S. (2016). Injury site-specific targeting of complement inhibitors for treating stroke. *Immunological Reviews, 274*(1), 270-280. doi:10.1111/imr.12470

Albert-Weissenberger, C., Mencl, S., Schuhmann, M. K., Salur, I., Göb, E., Langhauser, F., . . . Kleinschnitz, C. (2014). C1-inhibitor protects from focal brain trauma in a cortical cryolesion mice model by reducing thrombo-inflammation. *Frontiers in Cellular Neuroscience, 8,* 269. doi:10.3389/fncel.2014.00269

Arachchillage, D. R. J., Mackie, I. J., Efthymiou, M., Chitolie, A., Hunt, B. J., Isenberg, D. A., . . . Cohen, H. (2016). Rivaroxaban limits complement activation compared with warfarin in antiphospholipid syndrome patients with venous thromboembolism. *Journal of Thrombosis and Haemostasis, 14*(11), 2177-2186. doi:doi:10.1111/jth.13475

Armstrong, J. R., & Mosher, B. D. (2011). Aspiration pneumonia after stroke: Intervention and prevention. *The Neurohospitalist, 1*(2), 85-93. doi:10.1177/1941875210395775

Aslam, B., Basit, M., Nisar, M. A., Rasool, M. H., & Khurshid, M. (2017). Proteomics: Technologies and their applications. *Journal of Chromatographic Science, 55*(2), 182-196. doi:10.1093/chromsci/bmw167

Bang, O. Y. (2017). Advances in biomarker for stroke patients: From marker to regulator. *Precis Future Med, 1*(1), 32-42. doi:10.23838/pfm.2017.00052

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., . . . Muntner, P. (2017). Heart disease and stroke statistics—2017 update: A report from the american heart association. *Circulation, 135*(10), e146.

Cai, S., Dole, V. S., Bergmeier, W., Scafidi, J., Feng, H., Wagner, D. D., & Davis, A. E. (2005). A direct role for C1 inhibitor in regulation of leukocyte adhesion. *The Journal of Immunology*, *174*(10), 6462-6466.

Cao, Z., Yende, S., Kellum, J. A., & Robinson, R. A. (2013). Additions to the human plasma proteome via a tandem mars depletion itraq-based workflow. *International Journal of Proteomics, 2013*, 8. doi:10.1155/2013/654356

Carter, A. M. (2012). Complement activation: An emerging player in the pathogenesis of cardiovascular disease. *Scientifica, 2012*, 14. doi:10.6064/2012/402783

Chandramouli, K., & Qian, P. (2009). Proteomics: Challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics, 2009*, 239204. doi:10.4061/2009/239204

Cicardi, M., Zingale, L., Zanichelli, A., Pappalardo, E., & Cicardi, B. (2005). C1 inhibitor: Molecular and clinical aspects. *Springer Seminars in Immunopathology, 27*(3), 286-298. doi:10.1007/s00281-005-0001-4

Cojocaru, I. M., Cojocaru, M., Tanasescu, R., Burcin, C., Atanasiu, A. N., Petrescu, A.-M., . . . Dumitrescu, L. (2008). Changes in plasma levels of complement in patients with acute ischemic stroke. *Romanian Journal of Internal Medicine, 46*(1), 77-80.

Davis, A. (2004). Biological effects of c1 inhibitor. *Drug News Perspect*, *17*(7), 439-446.

Dayon, L., & Kussmann, M. (2013). Proteomics of human plasma: A critical comparison of analytical workflows in terms of effort, throughput and outcome. *EuPA Open Proteomics, 1*, 8-16. doi:https://doi.org/10.1016/j.euprot.2013.08.001

Di Napoli, M. (2001). Systemic complement activation in ischemic stroke. *Stroke, 32*(6), 1443.

Doerr, A. (2009). Targeted proteomics. *Nature Methods, 7*, 34. doi:10.1038/nmeth.f.284

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., . . . D'Eustachio, P. (2018). The Reactome pathway knowledgebase. *Nucleic Acids Research, 46*, 649-655. doi:10.1093/nar/gkx1132

Faria, S. S., Morris, C. F. M., Silva, A. R., Fonseca, M. P., Forget, P., Castro, M. S., & Fontes, W. (2017). A timely shift from shotgun to targeted proteomics and how it can be groundbreaking for cancer research. *Frontiers in Oncology, 7*, 13. doi:10.3389/fonc.2017.00013

Foley, J. H., Walton, B. L., Aleman, M. M., O'Byrne, A. M., Lei, V., Harrasser, M., . . . Conway, E. M. (2016). Complement activation in arterial and venous thrombosis is mediated by plasmin. *EBioMedicine, 5*, 175-182. doi:https://doi.org/10.1016/j.ebiom.2016.02.011

Gasque, P., Fontaine, M., & Morgan, B. P. (1995). Complement expression in human brain. Biosynthesis of terminal pathway components and regulators in human glial cells and cell lines. *Journal of Immunology, 154*(9), 4726-4733.

Geistlinger, L., Csaba, G., Küffner, R., Mulder, N., & Zimmer, R. (2011). From sets to graphs: Towards a realistic enrichment analysis of transcriptomic systems. *Bioinformatics, 27*(13), 366-373. doi:10.1093/bioinformatics/btr228

Gowda, G., & Djukovic, D. (2014). Overview of mass spectrometry-based metabolomics: Opportunities and challenges. *Methods in Molecular Biology, 1198*, 3-12. doi:10.1007/978-1-4939-1258-2_1

Hadders, Michael A., Bubeck, D., Roversi, P., Hakobyan, S., Forneris, F., Morgan, B. P., . . . Gros, P. (2012). Assembly and regulation of the membrane attack complex

based on structures of C5b6 and sC5b9. *Cell Reports, 1*(3), 200-207. doi:https://doi.org/10.1016/j.celrep.2012.02.003

Hatem, S. M., Saussez, G., della Faille, M., Prist, V., Zhang, X., Dispa, D., & Bleyenheuft, Y. (2016). Rehabilitation of motor function after stroke: A multiple systematic review focused on techniques to stimulate upper extremity recovery. *Frontiers in Human Neuroscience, 10*, 442. doi:10.3389/fnhum.2016.00442

Hertle, E., Arts, I. C. W., van der Kallen, C. J. H., Feskens, E. J. M., Schalkwijk, C. G., Stehouwer, C. D. A., & van Greevenbroek, M. M. J. (2018). Classical pathway of complement activation: Longitudinal associations of C1q and C1-inh with cardiovascular outcomes. *Arteriosclerosis, Thrombosis, and Vascular Biology, 38*(5), 1242-1244.

Heydenreich, N., Nolte, M. W., Göb, E., Langhauser, F., Hofmeister, M., Kraft, P., . . . Kleinschnitz, C. (2012). C1-inhibitor protects from brain ischemia-reperfusion injury by combined antiinflammatory and antithrombotic mechanisms. *Stroke, 43*(9), 2457-2467.

Huber-Lang, M., Sarma, J. V., Zetoune, F. S., Rittirsch, D., Neff, T. A., McGuire, S. R., . . . Ward, P. A. (2006). Generation of c5a in the absence of c3: A new complement activation pathway. *Nature Medicine, 12*, 682. doi:10.1038/nm1419

Jacob, A., & Alexander, J. J. (2014). Complement and blood–brain barrier integrity. *Molecular Immunology, 61*(2), 149-152. doi:https://doi.org/10.1016/j.molimm.2014.06.039

Kamel, H., & Iadecola, C. (2012). Brain-immune interactions and ischemic stroke: Clinical implications. *Archives of Neurology, 69*(5), 576-581. doi:10.1001/archneurol.2011.3590

Kanehisa, M., & Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27-30.

Keragala, C. B., Draxler, D. F., McQuilten, Z. K., & Medcalf, R. L. (2018). Haemostasis and innate immunity – a complementary relationship. *British Journal of Haematology, 180*(6), 782-798. doi:doi:10.1111/bjh.15062

Kouser, L., Abdul-Aziz, M., Nayak, A., Stover, C. M., Sim, R. B., & Kishore, U. (2013). Properdin and factor h: Opposing players on the alternative complement pathway "see-saw". *Frontiers in Immunology, 4*, 93. doi:10.3389/fimmu.2013.00093

Lyden, P., Raman, R., Liu, L., Emr, M., Warren, M., & Marler, J. (2009). NIHSS certification is reliable across multiple venues. *Stroke, 40*(7), 2507-2511. doi:10.1161/STROKEAHA.108.532069

Maes, E., Cho, W. C., & Baggerman, G. (2015). Translating clinical proteomics: The importance of study design. *Expert Review of Proteomics, 12*(3), 217-219. doi:10.1586/14789450.2015.1041512

Mak, T. W., & Saunders, M. E. (2006). Complement. In *The immune response* (pp. 553-581). Burlington: Academic Press.

Mocco, J., Wilson, D. A., Komotar, R. J., Sughrue, M. E., Coates, K., Sacco, R. L., . . . Connolly, E. S. (2006). Alterations in plasma complement levels following human ischemic stroke. *Neurosurgery, 59*(1), 1-6. doi:10.1227/01.neu.0000243280.75920.f4

Montaner, J., Mendioroz, M., Delgado, P., García-Berrocoso, T., Giralt, D., Merino, C., . . . Hernández-Guillamon, M. (2012). Differentiating ischemic from hemorrhagic stroke using plasma biomarkers: The s100b/rage pathway. *Journal of Proteomics, 75*(15), 4758-4765. doi:https://doi.org/10.1016/j.jprot.2012.01.033

Morikis, D., & Lambris, J. D. (2005). *Structural biology of the complement system*: CRC Press.

Nguyen, V. A., Carey, L. M., Giummarra, L., Faou, P., Cooke, I., Howells, D. W., . . . Crewther, S. G. (2016). A pathway proteomic profile of ischemic stroke survivors

reveals innate immune dysfunction in association with mild symptoms of depression – a pilot study. *Frontiers in Neurology, 7*, 85. doi:10.3389/fneur.2016.00085

Noris, M., & Remuzzi, G. (2013). Overview of complement activation and regulation. *Seminars in Nephrology, 33*(6), 479-492. doi:10.1016/j.semnephrol.2013.08.001

Orsini, F., Chrysanthou, E., Dudler, T., Cummings, W. J., Takahashi, M., Fujita, T., . . . Schwaeble, W. (2016). Mannan binding lectin-associated serine protease-2 (MASP-2) critically contributes to post-ischemic brain injury independent of MASP-1. *Journal of Neuroinflammation, 13*(1), 213. doi:10.1186/s12974-016-0684-6

Parsons, E. S., Stanley, G. J., Pyne, A. L. B., Hodel, A. W., Nievergelt, A. P., Menny, A., . . . Hoogenboom, B. W. (2019). Single-molecule kinetics of pore assembly by the membrane attack complex. *Nature Communications, 10*(1), 2066. doi:10.1038/s41467-019-10058-7

Penn, A. M., Bibok, M. B., Saly, V. K., Coutts, S. B., Lesperance, M. L., Balshaw, R. F., . . . Borchers, C. H. (2018). Validation of a proteomic biomarker panel to diagnose minor-stroke and transient ischaemic attack: Phase 2 of SPECTRA, a large scale translational study. *Biomarkers, 23*(8), 793-803. doi:10.1080/1354750X.2018.1499130

Penn, A. M., Saly, V., Trivedi, A., Lesperance, M. L., Votova, K., Jackson, A. M., . . . Borchers, C. H. (2018). Differential proteomics for distinguishing ischemic stroke from controls: A pilot study of the SPECTRA project. *Translational Stroke Research, 9*(6), 590-599. doi:10.1007/s12975-018-0609-z

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., . . . Vizcaíno, J. A. (2019). The PRIDE database and related tools and

resources in 2019: Improving support for quantification data. *Nucleic Acids Research, 47*(D1), D442-D450. doi:10.1093/nar/gky1106

Pérez, L. M., Inzitari, M., Quinn, T. J., Montaner, J., Gavaldà, R., Duarte, E., . . . Gallofré, M. (2016). Rehabilitation profiles of older adult stroke survivors admitted to intermediate care units: A multi-centre study. *PloS One, 11*(11), e0166304-e0166304. doi:10.1371/journal.pone.0166304

Portt, L., Norman, G., Clapp, C., Greenwood, M., & Greenwood, M. T. (2011). Anti-apoptosis and cell survival: A review. *Biochimica et Biophysica Acta, 1813*(1), 238-259. doi:10.1016/j.bbamcr.2010.10.010

Ricklin, D., Reis, E. S., & Lambris, J. D. (2016). Complement in disease: A defence system turning offensive. *Nature reviews Nephrology, 12*(7), 383-401. doi:10.1038/nrneph.2016.70

Robinson, R. G., & Jorge, R. E. (2016). Post-stroke depression: A review. *American Journal of Psychiatry, 173*(3), 221-231. doi:10.1176/appi.ajp.2015.15030363

Schmidt, A., Forne, I., & Imhof, A. (2014). Bioinformatic analysis of proteomics data. *BMC Systems Biology, 8*(2), S3. doi:10.1186/1752-0509-8-S2-S3

Serna, M., Giles, J. L., Morgan, B. P., & Bubeck, D. (2016). Structural basis of complement membrane attack complex formation. *Nature Communications, 7*(1), 10587. doi:10.1038/ncomms10587

Shi, K., Wood, K., Shi, F.-D., Wang, X., & Liu, Q. (2018). Stroke-induced immunosuppression and poststroke infection. *Stroke and Vascular Neurology, 3*(1), 34.

Shichita, T., Ito, M., & Yoshimura, A. (2014). Post-ischemic inflammation regulates neural damage and protection. *Frontiers in Cellular Neuroscience, 8*(319). doi:10.3389/fncel.2014.00319

Smith, M. P. W., Wood, S. L., Zougman, A., Ho, J. T. C., Peng, J., Jackson, D., . . . Banks, R. E. (2011). A systematic analysis of the effects of increasing degrees of serum immunodepletion in terms of depth of coverage and other key aspects in top-down and bottom-up proteomic analyses. *Proteomics, 11*(11), 2222-2235. doi:10.1002/pmic.201100005

Stokowska, A., Atkins, A. L., Morán, J., Pekny, T., Bulmer, L., Pascoe, M. C., . . . Pekna, M. (2017). Complement peptide c3a stimulates neural plasticity after experimental brain ischaemia. *Brain, 140*(2), 353-369. doi:10.1093/brain/aww314

Stokowska, A., Olsson, S., Holmegaard, L., Jood, K., Blomstrand, C., Jern, C., & Pekna, M. (2011). Plasma c3 and c3a levels in cryptogenic and large-vessel disease stroke: Associations with outcome. *Cerebrovascular Diseases, 32*(2), 114-122.

Stokowska, A., & Pekna, M. (2018). Complement C3a: Shaping the plasticity of the post-stroke brain. In P. A. Lapchak & J. H. Zhang (Eds.), *Cellular and molecular approaches to regeneration and repair* (pp. 521-541). Cham: Springer International Publishing.

Stokum, J. A., Gerzanich, V., & Simard, J. M. (2016). Molecular pathophysiology of cerebral edema. *Journal of Cerebral Blood Flow and Metabolism, 36*(3), 513-538. doi:10.1177/0271678X15617172

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences, 102*(43), 15545.

Sun, J.-H., Tan, L., & Yu, J.-T. (2014). Post-stroke cognitive impairment: Epidemiology, mechanisms and management. *Annals of Translational Medicine, 2*(8), 80. doi:10.3978/j.issn.2305-5839.2014.08.05

Tu, C., Rudnick, P. A., Martinez, M. Y., Cheek, K. L., Stein, S. E., Slebos, R. J. C., & Liebler, D. C. (2010). Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of Proteome Research, 9*(10), 4982-4991. doi:10.1021/pr100646w

Veerhuis, R., Nielsen, H. M., & Tenner, A. J. (2011). Complement in the brain. *Molecular Immunology, 48*(14), 1592-1603. doi:https://doi.org/10.1016/j.molimm.2011.04.003

von Zychlinski, A., Williams, M., McCormick, S., & Kleffmann, T. (2014). Absolute quantification of apolipoproteins and associated proteins on human plasma lipoproteins. *Journal of Proteomics, 106*, 181-190. doi:10.1016/j.jprot.2014.04.030

Westendorp, W. F., Nederkoorn, P. J., Vermeij, J.-D., Dijkgraaf, M. G., & de Beek, D. v. (2011). Post-stroke infection: A systematic review and meta-analysis. *BMC Neurology, 11*, 110-110. doi:10.1186/1471-2377-11-110

Whiteaker, J. R., Zhao, L., Schoenherr, R. M., Kennedy, J. J., Ivey, R. G., & Paulovich, A. G. (2017). Peptide immunoaffinity enrichment with targeted mass spectrometry: Application to quantification of ATM kinase phospho-signaling. *Methods in Molecular Biology, 1599*, 197-213. doi:10.1007/978-1-4939-6955-5_15

Xue, J., Huang, W., Chen, X., Li, Q., Cai, Z., Yu, T., & Shao, B. (2017). Neutrophil-to-lymphocyte ratio is a prognostic marker in acute ischemic stroke. *Journal of Stroke and Cerebrovascular Diseases, 26*(3), 650-657. doi:10.1016/j.jstrokecerebrovasdis.2016.11.010

Zhang, B., Yang, N., & Gao, C. (2015). Is plasma C3 and C4 levels useful in young cerebral ischemic stroke patients? Associations with prognosis at 3 months.

Journal of Thrombosis and Thrombolysis, 39(2), 209-214. doi:10.1007/s11239-014-1100-7

Zhou, F., Zhou, L., Guo, T., Wang, N., Hao, H., Zhou, Y., & Yu, D. (2017). Plasma proteomics reveals coagulation, inflammation, and metabolic shifts in H-type hypertension patients with and without acute ischemic stroke. *Oncotarget, 8*(59), 100384-100395. doi:10.18632/oncotarget.22233

Supplementary Table 1

Comparisons Between Current Subset for Proteomics Study (*n* = 60) and Total Cohort (*n* = 219)

	Current Subset (*n* = 60)		Total Cohort (*n* = 219)		
Baseline	*M*[†]	*SD*	*M*	*SD*	p
Age (years)	68.00	14.60	69.85	13.20	.13
NIHSS	4.70	4.76	7.99	6.74	.00*
Heart Rate (per minute)	74.52	11.06	76.06	13.83	.37
Systolic Blood Pressure (mm Hg)	141.47	23.27	144.87	23.04	.30
Diastolic Blood Pressure (mm Hg)	78.50	11.77	79.37	12.97	.77
3 Month					
NIHSS	1.18	2.31	2.54	4.90	.04*
mRS*	1.25	1.31	1.74	1.56	.02*
MoCA	25.93	4.57	24.61	5.92	.13
MADRS	8.72	8.81	7.40	7.56	.26

	Frequency[††]	Percentage	Frequency	Percentage	
Death	0	0.00%	21	9.59	.01*
Lost to Follow-up	0	0.00%	8	3.65	.13
Withdrawal	0	0.00%	2	0.91	.46
Past Atrial Fibrillation	4	6.7%	50	22.83	.01*
Hypertension	26	43.3%	124	56.62	.07
Lipid Disorder	24	40.0%	101	46.12	.40
Ischemic Heart Disease	11	18.3%	50	22.83	.46
Diabetes Mellitus	9	15.0%	42	19.18	.46

* $p < 0.05$

[†] T-tests were conducted between continuous variables.

[††] Chi squared tests were conducted for categorical variables.

Supplementary Methods

Blood collection and storage

All samples were collected in Benton-Dickson (BD) EDTA coated 4ml vacutainers and were mixed and left to stand in ambient room temperature for 30 minutes. The tubes were then centrifuged at 1100-1300 *g* at room temperature and the resulting plasma was aliquoted into cryotubes and immediately stored in a -80^{o-}C freezer. For transport from the central study freezer to the analysis site, temperature was kept at -70^{o-}C to -80^{o-}C on dry ice before transfer into a -80^{o-}C freezer.

Sample Processing and Trypsination

10µL of plasma from each patient was first stabilized in a prepared tube of 100µL 7M urea and 3M thiourea pH = 8.3. Samples were incubated overnight with 1µL tris-2-(carboxyethyl) phosphine (TCEP) at 22ºC. The following day, 4µL of 2-iodoacetamide was added to each sample and incubated for 45 minutes. An 800µL mixture of 1µg/µL trypsin and 50mM tris(hydroxymethyl)aminomethane was added to each sample and incubated overnight at 37ºC.

200µL was removed into a new tube and dried on a SpeedVac (Thermo Scientific, MA, USA) for 30 minutes. 10µL of 10% trifluoroacetic acid (TFA) was added to each sample to acidify. Samples were cleaned using stage tip pipette preparations of 3 plugs of Empore polystyrenedivinylbenzene (SBD-XC) copolymer disks (Sigma Aldrich, MO, USA) for solid phase extraction using washes of 50µL methanol, 50µL 5% acetonitrile (ACN)/0.5% TFA and elution with 50µL 85% ACN/0.5% TFA. Dry samples were resuspended in 20µL 5% ACN/0.5% TFA.

Chromatography and Mass Spectrometry

1µg of plasma peptides (no fractionation or protein depletion) were reconstituted in 0.1% TFA and 2% acetonitrile (ACN) and loaded at 45 ° C onto C_{18} PepMap 300 µm

ID × 2 cm trapping column (Thermo-Fisher Scientific) at 10 µl/min for 6 min, and washed for 6 minutes before switching the pre-column in line with the analytical column (BioSphere C_{18}, 1.9 µm, 120 Å and 75 µm ID × 40 cm, NanoSeparation). The separation of peptides was performed at 45 °C, 250 nl/min using a non-linear ACN gradient of buffer A (0.1% formic acid, 2% ACN) and buffer B (0.1% formic acid, 80% ACN), starting at 5% buffer B to 45% over 105 minutes, then 95% B for 5 min followed by an equilibration step of 15 min (0.1% formic acid, 2% ACN). Data were collected on a Q Exactive HF (Thermo-Fisher Scientific) in Data Dependent Acquisition mode using m/z 350–1500 as MS scan range at 120 000 resolution, HCD MS/MS spectra were collected for the 10 most intense ions per MS scan at 15 000 resolution. Dynamic exclusion parameters were set as follows: exclude isotope on, duration 30 s and peptide match preferred. Other instrument parameters for the Orbitrap were: MS maximum injection time 30 ms with AGC target 3×10^6, collision at 28% energy for a maximum injection time of 110 ms with AGT target of 1×10^5.

Protein Identification

Identification and label-free quantitation of proteins across all samples were performed using MaxQuant version version 1.6.1.0 and search against all reviewed and unreviewed human proteins in the Uniprot database (Jan 2017). Common contaminants and decoys were included automatically by Andromeda. MaxQuant settings included Carbamidomethyl C as a fixed modification, Oxidation of Methionine and Acetylation of protein N-terminus as variable modifications. Up to 2 missed cleavages were allowed, and peptides were required to be at least 7 amino acids in length. False discovery rate (FDR) cut-offs for both peptides and proteins in the database search were set to 1%. Both unique and razor peptides (peptides shared by different proteins of a group) were used for quantitation with a minimum of 2 peptides including at least 1 unique peptide required to calculate a protein quantitative value. The ‘match between runs’ setting in MaxQuant was

used to transfer peptide identifications from one run to another on the basis of matching retention time and mass to charge ratio.

Supplementary Data 1
Differential expression outputs from limma for T1 vs T2, T1 vs T3 and T2 vs T3

		T1 vs T2		
Gene.Symbol	**ENTREZID**	**FC**	**ADJ.PVAL**	**limma.STAT**
A1BG	1	-0.320972049	0.00044013	-4.59160606
A2M	2	-0.131701623	0.12196748	-2.21048661
ACTB	60	0.376597728	4.9685E-07	6.17281338
AFM	173	0.120497024	0.06893779	2.51772702
AGT	183	0.018343176	0.86348188	0.35159722
AHSG	197	-0.014796374	0.92279927	-0.19025546
AMBP	259	-0.108237084	0.43794967	-1.18980691
APCS	325	-0.092448156	0.4428891	-1.16262533
APOA1	335	0.046430282	0.83532034	0.40101632
APOA2	336	0.108729371	0.1583164	2.03811021
APOA4	337	0.31687119	4.9685E-07	6.17293768
APOB	338	-0.134745155	0.03888935	-2.80342222
APOC1	341	-0.169875768	0.02862888	-3.00490548
APOC3	345	0.058246347	0.67959423	0.73488093
APOC4-APOC2	100533990	0.027668675	0.74572884	0.60783821
APOD	347	0.158398281	0.01037244	3.50666589
APOE	348	0.108633204	0.59343023	0.88364583
APOF	319	-0.279339591	0.01094141	-3.46286275
APOH	350	0.032243321	0.72556329	0.64208351
APOL1	8542	-0.199126545	0.09411674	-2.38804382
APOM	55937	-0.037607173	0.8134155	-0.47770788
ATRN	8455	-0.041601928	0.71341135	-0.67126068
AZGP1	563	-0.007291854	0.92279927	-0.15470378
BCHE	590	0.102704847	0.35954571	1.37349835
BTD	686	-0.007455413	0.92279927	-0.14244677
C1QA	712	-0.036333853	0.81754353	-0.46607200
C1QB	713	-0.009347053	0.92279927	-0.14007650
C1QC	714	0.039943232	0.77388508	0.54905415
C1R	715	-0.004542277	0.9242303	-0.13114190
C1S	716	0.025392437	0.67959423	0.72952343
C2	717	0.089769045	0.38841432	1.28123675
C3	718	-0.082032732	0.25705494	-1.68464200
C4A	720	0.163469134	0.31144517	1.53241131
C4B	721	0.044525296	0.4428891	1.16706857
C4B_2	722	-0.005929642	0.92279927	-0.18693354
C4BPA	725	0.076493115	0.188562	1.89732981
C4BPB	100293534	-0.152410939	0.12314129	-2.19287246
C5	727	0.144858669	0.20878581	1.81666615
C6	729	0.098047526	0.04929592	2.66390676
C7	730	0.188405129	0.09878069	2.35741875
C8A	731	0.665669201	1.1091E-09	7.60057724
C8B	732	0.074324905	0.19367997	1.87654934

C8G	733	-0.151070477	0.03888935	-2.79889714
C9	735	-0.281430579	0.00270363	-4.02478303
CD14	929	-0.07840634	0.35804587	-1.40298360
CD5L	922	0.009016885	0.95970963	0.06543627
CFD	1675	0.164893072	0.10383755	2.31053141
CFH	3075	0.100785327	0.12196748	2.20711286
CFI	3426	0.241664973	0.00861516	3.62515549
CLEC3B	7123	0.181405499	0.00270363	4.06670342
CLU	1191	0.094602717	0.35315999	1.43507918
CNDP1	84735	0.061442742	0.78795616	0.52660495
CP	1356	0.029524543	0.67959423	0.76838934
CPB2	1361	0.003267318	0.96824604	0.03989022
CPN1	1369	-0.091933658	0.14806308	-2.08742619
CPN2	1370	0.059537466	0.27132024	1.62664572
CSPG4	1464	0.185948441	0.14806308	2.08609335
ECM1	1893	0.175217426	0.06001196	2.58078406
F10	2159	0.06967246	0.35804587	1.40358391
F12	2161	0.020356909	0.92279927	0.19175204
F13A1	2162	0.063200865	0.67959423	0.76340881
F13B	2165	0.029221985	0.83532034	0.39731727
F2	2147	-0.075218227	0.45328914	-1.14164394
F5	2153	-0.057487393	0.67959423	-0.74122267
F9	2158	-0.217981194	0.01328735	-3.34470914
FBLN1	2192	0.295280134	0.00444186	3.84854961
FCN3	8547	-0.09557323	0.36571861	-1.35008937
FETUB	26998	0.199544554	0.02862888	2.99389806
FGA	2243	0.11207122	0.10383755	2.30702077
FGB	2244	-0.234656514	0.03084161	-2.90590712
FGG	2266	0.040428175	0.50813827	1.03571592
FN1	2335	0.080614963	0.67959423	0.72646403
GC	2638	-0.037824074	0.79762463	-0.50899831
GPLD1	2822	0.168517936	0.15120419	2.06741216
GPX3	2878	0.036312395	0.7561752	0.58938884
GSN	2934	0.101128213	0.26618056	1.65938711
HABP2	3026	0.011993535	0.92279927	0.24492401
HBA1	3039	-0.214717869	0.37093102	-1.33512712
HBA2	3040	-0.227520198	0.35954571	-1.38218758
HBB	3043	0.212789626	0.35804587	1.39770539
HP	3240	0.009357265	0.92279927	0.19458812
HPR	3250	-0.167294289	0.02950888	-2.96576563
HPX	3263	0.091768188	0.04331365	2.72207539
HRG	3273	-0.035887613	0.59343023	-0.88876584
IGFALS	3483	0.197474947	0.01306815	3.38393876
IGHA1	3493	0.143923642	0.14654049	2.11053843
IGHG1	3500	0.122291974	0.18255081	1.92064202
IGHG2	3501	0.121674254	0.4428891	1.17439916

IGHG3	3502	0.122041704	0.52608641	1.00623995
IGHG4	3503	0.039744815	0.92279927	0.18695641
IGHM	3507	0.171884624	0.10383755	2.30375204
IGHV1-24	28467	0.022133275	0.92279927	0.18710595
IGHV1-46	28465	0.144767753	0.35315999	1.43927302
IGHV2-26	28455	0.017854694	0.92279927	0.16361137
IGHV2-5	28457	0.098777379	0.63605358	0.82722750
IGHV3-15	28448	0.05724629	0.8134155	0.47763726
IGHV3-49	28423	0.08387574	0.69389505	0.69994038
IGHV3-7	28452	0.101831687	0.25645357	1.69398550
IGHV3OR16-9	28307	0.117481305	0.37564652	1.32101548
IGHV4-38-2	28389	-0.018328184	0.92279927	-0.18691986
IGHV5-51	28388	0.175660604	0.18134977	1.95408663
IGHV6-1	28385	0.200760603	0.18255081	1.92444082
IGKC	3514	0.050957532	0.77388508	0.55401146
IGKV1-17	28937	0.040268363	0.84834723	0.37566059
IGKV1-27	28935	0.152608333	0.20687507	1.82929386
IGKV1-33	28933	0.020174989	0.92279927	0.21623206
IGKV1-5	28299	-0.049319202	0.35954571	-1.36664070
IGKV1D-39	28893	0.103570028	0.50536638	1.05320665
IGKV2-24	28923	0.063167801	0.72556329	0.63710369
IGKV2-30	28919	-0.01378048	0.92279927	-0.14937777
IGKV2-40	28916	-0.046109127	0.8191584	-0.43633433
IGKV2D-29	28882	0.113795903	0.35804587	1.41647915
IGKV3-11	28914	0.096601737	0.2934494	1.57082743
IGKV3-15	28913	0.080097346	0.55526318	0.94987535
IGKV3-20	28912	0.055731292	0.27132024	1.62603138
IGKV4-1	28908	0.064805124	0.41867295	1.22361144
IGLC3	3539	0.063627348	0.50536638	1.04599077
IGLL5	100423062	0.128907368	0.29214403	1.58089220
IGLV1-47	28822	0.06075647	0.67959423	0.72289847
IGLV1-51	28820	0.127159119	0.50536638	1.04946359
IGLV3-10	28803	-0.014139631	0.92279927	-0.15475061
IGLV3-21	28796	0.043927592	0.86885261	0.33846929
IGLV3-25	28793	0.048904043	0.83532034	0.41103149
IGLV7-46	28775	0.037521756	0.8191584	0.43719369
ITIH1	3697	0.074747152	0.18255081	1.92671776
ITIH2	3693	-0.005385507	0.93602574	-0.10941225
ITIH3	3699	-0.047390465	0.77388508	-0.55382884
ITIH4	3700	-0.002966908	0.95833562	-0.07454706
JCHAIN	3512	-0.139182134	0.48580825	-1.09157065
KLKB1	3818	-0.021069805	0.8191584	-0.44049371
KNG1	3827	0.050983844	0.35315999	1.44316244
LBP	3929	-0.161216029	0.18134977	-1.95060437
LCAT	3931	0.098682462	0.17981244	1.97316003
LGALS3BP	3959	0.005940929	0.96688193	0.04906226

LRG1	116844	-0.150668519	0.12196748	-2.22574107
LUM	4060	0.00928301	0.92279927	0.15901181
ORM1	5004	-0.089875156	0.56038677	-0.93688966
ORM2	5005	-0.068838196	0.38366255	-1.29594816
PGLYRP2	114770	0.163406685	0.01037244	3.52931638
PLG	5340	-0.036153745	0.87346771	-0.32616626
PLTP	5360	0.332219006	0.02950888	2.93548753
PON1	5444	-0.007784141	0.92279927	-0.16390495
PROS1	5627	0.027136356	0.67959423	0.73324809
RBP4	5950	0.183622671	0.01461113	3.26234055
SAA1	6288	-0.924774864	0.02950888	-2.94384305
SAA2-SAA4	100528017	-0.246208476	0.01409697	-3.29320205
SELENOP	6414	0.093096068	0.38366255	1.29516518
SERPINA1	5265	0.100166794	0.35954571	1.37467875
SERPINA10	51156	-0.214340595	0.04195581	-2.74594424
SERPINA3	12	0.133240883	0.01328735	3.33231107
SERPINA4	5267	0.0647216	0.53197312	0.98543982
SERPINA6	866	-0.030125919	0.72361999	-0.65303049
SERPINA7	6906	-0.028966443	0.8191584	-0.43739972
SERPINC1	462	0.008444189	0.92279927	0.21393791
SERPIND1	3053	0.047941848	0.52727442	0.99802864
SERPINF1	5176	-0.113281234	0.34045392	-1.47785879
SERPINF2	5345	-0.074369647	0.39114473	-1.27038740
SERPING1	710	0.139084239	0.25258005	1.70967669
SHBG	6462	0.184362041	0.19569814	1.86329255
TF	7018	0.202556876	0.01463906	3.24300643
TTR	7276	0.086077331	0.21279931	1.79952372
VTN	7448	0.072308044	0.0412557	2.76493074
VWF	7450	-0.349118801	0.26687454	-1.65005831

T1 vs T3				
Gene.Symbol	**ENTREZID**	**FC**	**ADJ.PVAL**	**limma.STAT**
A1BG	1	-0.235261746	0.007318	-3.447054547
A2M	2	-0.185984309	0.009541893	-3.324682752
ACTB	60	0.37116762	1.47E-07	6.423751715
AFM	173	0.148684169	0.016041302	3.133591495
AGT	183	-0.050819707	0.48259229	-1.069303065
AHSG	197	-0.207356632	0.090390277	-2.291896079
AMBP	259	-0.197441467	0.119074531	-2.158287987
APCS	325	-0.277207259	0.019804038	-3.024753131
APOA1	335	-0.030213122	0.899864358	-0.253742912
APOA2	336	0.234901236	0.00030232	4.629218562
APOA4	337	0.358527559	2.97E-10	7.720958022
APOB	338	-0.175380091	0.004902842	-3.751634146
APOC1	341	-0.037862919	0.703907685	-0.652242555
APOC3	345	0.177912497	0.086270428	2.325579434

APOC4-APOC2	1.01E+08	0.032308387	0.629720667	0.827615688
APOD	347	0.139916903	0.006848688	3.490703449
APOE	348	0.1493346	0.405654481	1.245856619
APOF	319	-0.088412223	0.532618543	-0.991336944
APOH	350	-0.001589512	0.988965296	-0.037827448
APOL1	8542	-0.144014488	0.153779235	-1.955405367
APOM	55937	-0.085051851	0.48259229	-1.070631882
ATRN	8455	-0.078869591	0.359950311	-1.345162788
AZGP1	563	0.006648636	0.960762772	0.146032935
BCHE	590	0.141596376	0.129478224	2.114241362
BTD	686	0.049406965	0.542413997	0.955377911
C1QA	712	-0.223109766	0.065472761	-2.481254509
C1QB	713	0.119305922	0.150775899	1.995869148
C1QC	714	0.006472757	0.988965296	0.091458566
C1R	715	-0.016442951	0.788900698	-0.47535174
C1S	716	-0.046937811	0.370093074	-1.317103429
C2	717	0.089758797	0.37608023	1.299846604
C3	718	-0.043391315	0.539585165	-0.96925557
C4A	720	0.017127467	0.960762772	0.150898456
C4B	721	0.008542444	0.899864358	0.238966842
C4B_2	722	-0.054391947	0.234024827	-1.648002278
C4BPA	725	0.120572879	0.018738026	3.055944445
C4BPB	1E+08	-0.20110305	0.016041302	-3.14795198
C5	727	0.121164273	0.266342079	1.568862491
C6	729	0.078633647	0.110005178	2.20006708
C7	730	0.19695625	0.065472761	2.482294782
C8A	731	0.739414378	1.19E-11	8.45444335
C8B	732	0.000547193	0.988965296	0.013858899
C8G	733	-0.134872116	0.067464658	-2.459677329
C9	735	-0.323119274	0.000346003	-4.551461162
CD14	929	-0.076643888	0.289464748	-1.511774115
CD5L	922	-0.079577787	0.75582347	-0.568160763
CFD	1675	0.245920399	0.006518495	3.588932974
CFH	3075	0.080352366	0.156488577	1.935321075
CFI	3426	0.334603466	3.07E-06	5.727298878
CLEC3B	7123	0.199242861	0.000619039	4.333927487
CLU	1191	0.072371147	0.48259229	1.069007018
CNDP1	84735	0.235530352	0.150775899	1.989247305
CP	1356	0.026914865	0.662553357	0.742337828
CPB2	1361	0.140881143	0.175892364	1.863936314
CPN1	1369	-0.143521893	0.006518495	-3.610661077
CPN2	1370	0.052240729	0.306758122	1.474151957
CSPG4	1464	0.238568967	0.016047021	3.119341884
ECM1	1893	0.146408017	0.090390277	2.288362431
F10	2159	0.068014619	0.331393687	1.418586178
F12	2161	0.047855845	0.789359776	0.463159749

F13A1	2162	0.130400269	0.218214175	1.71818655
F13B	2165	0.001039253	0.988965296	0.015981084
F2	2147	-0.021374011	0.855073122	-0.354034567
F5	2153	0.005592428	0.988965296	0.079052024
F9	2158	-0.110075084	0.227265325	-1.676543393
FBLN1	2192	0.211689347	0.029565162	2.830450935
FCN3	8547	-0.090887969	0.359950311	-1.349952823
FETUB	26998	0.106726645	0.256110404	1.595818373
FGA	2243	0.150383019	0.006833387	3.51031349
FGB	2244	-0.147732647	0.156488577	-1.932046069
FGG	2266	0.031209869	0.633775541	0.813666787
FN1	2335	0.141772817	0.406103014	1.221590516
GC	2638	-0.15324368	0.150565628	-2.014316466
GPLD1	2822	0.233452204	0.027416831	2.891698768
GPX3	2878	-0.004123277	0.988965296	-0.071653773
GSN	2934	0.04170521	0.662553357	0.747591085
HABP2	3026	-0.024355105	0.772349838	-0.502844231
HBA1	3039	-0.3374191	0.148513168	-2.028997131
HBA2	3040	-0.35530315	0.133567556	-2.092092264
HBB	3043	-0.081344427	0.703907685	-0.647990702
HP	3240	-0.02919195	0.703907685	-0.653355826
HPR	3250	-0.089205822	0.288517832	-1.520505157
HPX	3263	0.089819711	0.043964065	2.65872653
HRG	3273	-0.049843813	0.405654481	-1.240461354
IGFALS	3483	0.190639391	0.007318	3.436245967
IGHA1	3493	0.136478663	0.150775899	1.980128196
IGHG1	3500	0.12049311	0.169987989	1.887135503
IGHG2	3501	0.079346082	0.652453142	0.771379374
IGHG3	3502	0.201494243	0.234024827	1.654764443
IGHG4	3503	-0.116597906	0.767640611	-0.521728284
IGHM	3507	0.143720227	0.178167908	1.850468941
IGHV1-24	28467	-0.055584757	0.788900698	-0.470468299
IGHV1-46	28465	0.083480606	0.542413997	0.95246649
IGHV2-26	28455	-0.058957127	0.75582347	-0.550613653
IGHV2-5	28457	0.236655352	0.150775899	1.987914571
IGHV3-15	28448	0.054096628	0.772349838	0.505965444
IGHV3-49	28423	0.105165345	0.584105045	0.888060125
IGHV3-7	28452	0.104076891	0.226688058	1.684935608
IGHV3OR16-9	28307	0.023533071	0.899864358	0.260532367
IGHV4-38-2	28389	0.001587962	0.988965296	0.016086963
IGHV5-51	28388	-0.026674185	0.899864358	-0.25800976
IGHV6-1	28385	0.257871299	0.065472761	2.490805862
IGKC	3514	0.027730247	0.87454891	0.31784779
IGKV1-17	28937	0.131232387	0.406103014	1.22603356
IGKV1-27	28935	0.192257207	0.07953957	2.376701206
IGKV1-33	28933	0.037697572	0.828609701	0.404888439

IGKV1-5	28299	-0.077557084	0.086270428	-2.332378726
IGKV1D-39	28893	0.127424088	0.370093074	1.315790882
IGKV2-24	28923	0.060386454	0.75582347	0.567324417
IGKV2-30	28919	0.001341903	0.988965296	0.014193318
IGKV2-40	28916	0.056315865	0.767640611	0.526644129
IGKV2D-29	28882	0.065041937	0.633775541	0.809300965
IGKV3-11	28914	0.085965255	0.326865051	1.433065252
IGKV3-15	28913	0.057649836	0.682486514	0.706050794
IGKV3-20	28912	0.056778952	0.214337185	1.748859771
IGKV4-1	28908	0.002230497	0.988965296	0.047890533
IGLC3	3539	0.019841907	0.87454891	0.321932874
IGLL5	1E+08	0.03783961	0.813263943	0.429152257
IGLV1-47	28822	0.046819189	0.75582347	0.548645294
IGLV1-51	28820	0.100521694	0.550996511	0.935171039
IGLV3-10	28803	-0.050498089	0.75582347	-0.558024059
IGLV3-21	28796	0.077376306	0.700780891	0.671756762
IGLV3-25	28793	-0.028047447	0.899864358	-0.245879797
IGLV7-46	28775	-0.065164323	0.652453142	-0.767441597
ITIH1	3697	0.077806114	0.137047794	2.072259013
ITIH2	3693	-0.028862139	0.75582347	-0.576880409
ITIH3	3699	-0.063503575	0.652453142	-0.767933508
ITIH4	3700	0.000719994	0.988965296	0.018176468
JCHAIN	3512	-0.129541909	0.465946167	-1.117828953
KLKB1	3818	-0.042951432	0.532618543	-0.992121008
KNG1	3827	0.000866796	0.988965296	0.025678443
LBP	3929	-0.249732745	0.028175055	-2.858366083
LCAT	3931	0.032942777	0.679696456	0.715992498
LGALS3BP	3959	-0.029655492	0.899864358	-0.246456188
LRG1	116844	-0.228463616	0.006600169	-3.562167952
LUM	4060	-0.023798822	0.834668832	-0.391097084
ORM1	5004	-0.34968034	0.02741787	-2.879440898
ORM2	5005	-0.086736991	0.218214175	-1.725532562
PGLYRP2	114770	0.186261074	0.000566237	4.390301838
PLG	5340	-0.040089983	0.846383684	-0.37087381
PLTP	5360	0.181434462	0.221569954	1.703273882
PON1	5444	0.009846742	0.914090253	0.215201121
PROS1	5627	-0.008683228	0.899864358	-0.244485205
RBP4	5950	0.215326111	0.006681216	3.537065065
SAA1	6288	-0.938034156	0.033440032	-2.776852253
SAA2-SAA4	1.01E+08	-0.272777303	0.008147999	-3.387960675
SELENOP	6414	0.164633927	0.076013737	2.404157115
SERPINA1	5265	0.121420789	0.205561609	1.776291668
SERPINA10	51156	-0.04943912	0.695138407	-0.684919842
SERPINA3	12	0.100017734	0.037422183	2.726604716
SERPINA4	5267	0.09056407	0.359950311	1.363223028
SERPINA6	866	-0.043614697	0.48259229	-1.08106702

SERPINA7	6906	0.072457342	0.405654481	1.235569173
SERPINC1	462	0.003355845	0.988965296	0.085767507
SERPIND1	3053	-0.053261881	0.534381606	-0.982429651
SERPINF1	5176	-0.085391918	0.449652215	-1.153347404
SERPINF2	5345	0.006323648	0.980168974	0.115791184
SERPING1	710	0.13742734	0.218214175	1.71793811
SHBG	6462	0.131227357	0.359950311	1.345174463
TF	7018	0.224290653	0.006518495	3.637917158
TTR	7276	0.057196989	0.452786119	1.142339398
VTN	7448	0.073673874	0.024423125	2.942936507
VWF	7450	-0.430586346	0.152330385	-1.967546069

T2 vs T3

Gene.Symbol	ENTREZID	FC	ADJ.PVAL	limma.STAT
A1BG	1	0.08571	0.989162	1.26207
A2M	2	-0.05428	0.989162	-0.99164
ACTB	60	-0.00543	0.989162	-0.10057
AFM	173	0.028187	0.989162	0.558297
AGT	183	-0.06916	0.989162	-1.39849
AHSG	197	-0.19256	0.700271	-2.36197
AMBP	259	-0.0892	0.989162	-1.05936
APCS	325	-0.18476	0.700271	-2.29229
APOA1	335	-0.07664	0.989162	-0.70891
APOA2	336	0.126172	0.700271	2.505999
APOA4	337	0.041656	0.989162	0.840316
APOB	338	-0.04063	0.989162	-0.72721
APOC1	341	0.132013	0.700271	2.476721
APOC3	345	0.119666	0.989162	1.448291
APOC4-APOC2	1.01E+08	0.00464	0.989162	0.106051
APOD	347	-0.01848	0.989162	-0.40209
APOE	348	0.040701	0.989162	0.338167
APOF	319	0.190927	0.700271	2.19481
APOH	350	-0.03383	0.989162	-0.70583
APOL1	8542	0.055112	0.989162	0.65153
APOM	55937	-0.04744	0.989162	-0.58475
ATRN	8455	-0.03727	0.989162	-0.60911
AZGP1	563	0.01394	0.989162	0.281215
BCHE	590	0.038892	0.989162	0.586021
BTD	686	0.056862	0.989162	0.963225
C1QA	712	-0.18678	0.700271	-2.28547
C1QB	713	0.128653	0.768374	1.867032
C1QC	714	-0.03347	0.989162	-0.44461
C1R	715	-0.0119	0.989162	-0.32864
C1S	716	-0.07233	0.768374	-1.8739
C2	717	-1.02E-05	0.999887	-0.00014
C3	718	0.038641	0.989162	0.828456

C4A	720	-0.14634	0.989162	-1.27059
C4B	721	-0.03598	0.989162	-1.00244
C4B_2	722	-0.04846	0.989162	-1.44519
C4BPA	725	0.04408	0.989162	1.145277
C4BPB	1E+08	-0.04869	0.989162	-0.66359
C5	727	-0.02369	0.989162	-0.31241
C6	729	-0.01941	0.989162	-0.52865
C7	730	0.008551	0.989162	0.111064
C8A	731	0.073745	0.989162	0.958989
C8B	732	-0.07378	0.768374	-1.8765
C8G	733	0.016198	0.989162	0.241502
C9	735	-0.04169	0.989162	-0.52994
CD14	929	0.001762	0.989162	0.034386
CD5L	922	-0.08859	0.989162	-0.6945
CFD	1675	0.081027	0.989162	1.142014
CFH	3075	-0.02043	0.989162	-0.4629
CFI	3426	0.092938	0.989162	1.519736
CLEC3B	7123	0.017837	0.989162	0.386061
CLU	1191	-0.02223	0.989162	-0.33824
CNDP1	84735	0.174088	0.989162	1.44989
CP	1356	-0.00261	0.989162	-0.06776
CPB2	1361	0.137614	0.768374	1.823301
CPN1	1369	-0.05159	0.989162	-1.18708
CPN2	1370	-0.0073	0.989162	-0.19192
CSPG4	1464	0.052621	0.989162	0.643604
ECM1	1893	-0.02881	0.989162	-0.40355
F10	2159	-0.00166	0.989162	-0.03264
F12	2161	0.027499	0.989162	0.266464
F13A1	2162	0.067199	0.989162	0.846544
F13B	2165	-0.02818	0.989162	-0.39147
F2	2147	0.053844	0.989162	0.786037
F5	2153	0.06308	0.989162	0.836821
F9	2158	0.107906	0.989162	1.430346
FBLN1	2192	-0.08359	0.989162	-1.23783
FCN3	8547	0.004685	0.989162	0.073651
FETUB	26998	-0.09282	0.989162	-1.45331
FGA	2243	0.038312	0.989162	0.824699
FGB	2244	0.086924	0.989162	1.124622
FGG	2266	-0.00922	0.989162	-0.22844
FN1	2335	0.061158	0.989162	0.556994
GC	2638	-0.11542	0.989162	-1.56792
GPLD1	2822	0.064934	0.989162	0.791159
GPX3	2878	-0.04044	0.989162	-0.67516
GSN	2934	-0.05942	0.989162	-0.99644
HABP2	3026	-0.03635	0.989162	-0.83574
HBA1	3039	-0.1227	0.989162	-0.9472

HBA2	3040	-0.12778	0.989162	-0.97905
HBB	3043	-0.29413	0.756719	-2.01201
HP	3240	-0.03855	0.989162	-0.81266
HPR	3250	0.078088	0.989162	1.140278
HPX	3263	-0.00195	0.989162	-0.05221
HRG	3273	-0.01396	0.989162	-0.3514
IGFALS	3483	-0.00684	0.989162	-0.12379
IGHA1	3493	-0.00744	0.989162	-0.10945
IGHG1	3500	-0.0018	0.989162	-0.02968
IGHG2	3501	-0.04233	0.989162	-0.43268
IGHG3	3502	0.079453	0.989162	0.663729
IGHG4	3503	-0.15634	0.989162	-0.6917
IGHM	3507	-0.02816	0.989162	-0.36156
IGHV1-24	28467	-0.07772	0.989162	-0.59811
IGHV1-46	28465	-0.06129	0.989162	-0.68635
IGHV2-26	28455	-0.07681	0.989162	-0.71172
IGHV2-5	28457	0.137878	0.989162	1.166776
IGHV3-15	28448	-0.00315	0.989162	-0.02928
IGHV3-49	28423	0.02129	0.989162	0.172036
IGHV3-7	28452	0.002245	0.989162	0.036331
IGHV3OR16-9	28307	-0.09395	0.989162	-1.02428
IGHV4-38-2	28389	0.019916	0.989162	0.201942
IGHV5-51	28388	-0.20233	0.756719	-2.06797
IGHV6-1	28385	0.057111	0.989162	0.611895
IGKC	3514	-0.02323	0.989162	-0.23343
IGKV1-17	28937	0.090964	0.989162	0.809456
IGKV1-27	28935	0.039649	0.989162	0.504801
IGKV1-33	28933	0.017523	0.989162	0.181568
IGKV1-5	28299	-0.02824	0.989162	-0.7765
IGKV1D-39	28893	0.023854	0.989162	0.243346
IGKV2-24	28923	-0.00278	0.989162	-0.02738
IGKV2-30	28919	0.015122	0.989162	0.168592
IGKV2-40	28916	0.102425	0.989162	1.025819
IGKV2D-29	28882	-0.04875	0.989162	-0.63828
IGKV3-11	28914	-0.01064	0.989162	-0.17229
IGKV3-15	28913	-0.02245	0.989162	-0.27259
IGKV3-20	28912	0.001048	0.989162	0.030878
IGKV4-1	28908	-0.06257	0.989162	-1.24219
IGLC3	3539	-0.04379	0.989162	-0.68548
IGLL5	1E+08	-0.09107	0.989162	-1.14282
IGLV1-47	28822	-0.01394	0.989162	-0.16066
IGLV1-51	28820	-0.02664	0.989162	-0.22126
IGLV3-10	28803	-0.03636	0.989162	-0.38109
IGLV3-21	28796	0.033449	0.989162	0.273868
IGLV3-25	28793	-0.07695	0.989162	-0.67296
IGLV7-46	28775	-0.10269	0.989162	-1.14759

ITIH1	3697	0.003059	0.989162	0.075566
ITIH2	3693	-0.02348	0.989162	-0.47044
ITIH3	3699	-0.01611	0.989162	-0.1876
ITIH4	3700	0.003687	0.989162	0.090896
JCHAIN	3512	0.00964	0.989162	0.076817
KLKB1	3818	-0.02188	0.989162	-0.46951
KNG1	3827	-0.05012	0.989162	-1.3599
LBP	3929	-0.08852	0.989162	-1.08519
LCAT	3931	-0.06574	0.989162	-1.32989
LGALS3BP	3959	-0.0356	0.989162	-0.29636
LRG1	116844	-0.0778	0.989162	-1.2141
LUM	4060	-0.03308	0.989162	-0.54281
ORM1	5004	-0.25981	0.756719	-2.03967
ORM2	5005	-0.0179	0.989162	-0.31046
PGLYRP2	114770	0.022854	0.989162	0.471991
PLG	5340	-0.00394	0.989162	-0.03425
PLTP	5360	-0.15078	0.989162	-1.409
PON1	5444	0.017631	0.989162	0.356769
PROS1	5627	-0.03582	0.989162	-0.91816
RBP4	5950	0.031703	0.989162	0.559515
SAA1	6288	-0.01326	0.989162	-0.05086
SAA2-SAA4	1.01E+08	-0.02657	0.989162	-0.38677
SELENOP	6414	0.071538	0.989162	1.063039
SERPINA1	5265	0.021254	0.989162	0.289322
SERPINA10	51156	0.164901	0.700271	2.21512
SERPINA3	12	-0.03322	0.989162	-0.84407
SERPINA4	5267	0.025842	0.989162	0.364929
SERPINA6	866	-0.01349	0.989162	-0.30517
SERPINA7	6906	0.101424	0.989162	1.559608
SERPINC1	462	-0.00509	0.989162	-0.13296
SERPIND1	3053	-0.1012	0.768374	-1.82339
SERPINF1	5176	0.027889	0.989162	0.35977
SERPINF2	5345	0.080693	0.989162	1.443485
SERPING1	710	-0.00166	0.989162	-0.02123
SHBG	6462	-0.05313	0.989162	-0.53488
TF	7018	0.021734	0.989162	0.366482
TTR	7276	-0.02888	0.989162	-0.55873
VTN	7448	0.001366	0.989162	0.050969
VWF	7450	-0.08147	0.989162	-0.38744

Supplementary Data 2

Complete Outputs from Gene Graph Enrichment Analysis on Enrichment Browser For T1 vs T2 using KEGG and Reactome

Kyoto Encyclopedia of Genes and Genomes (KEGG)				
GENE.SET	NR.RELS	RAW.SCORE	NORM.SCORE	ADJ.P.VALUE
hsa04610_Complement_and_coagulation_cascades	6	1.91	0.319	0.003996
hsa05150_Staphylococcus_aureus_infection	5	1.6	0.32	0.39
hsa05322_Systemic_lupus_erythematosus	3	0.819	0.273	0.8
hsa05133_Pertussis	3	0.819	0.273	0.8
Reactome				
GENE.SET	NR.RELS	RAW.SCORE	NORM.SCORE	ADJ.P.VALUE
REACTOME_INNATE_IMMUNE_SYSTEM	10	4.97	0.497	0.011
REACTOME_IMMUNE_SYSTEM	10	4.97	0.497	0.011
REACTOME_COMPLEMENT_CASCADE	10	4.97	0.497	0.011
REACTOME_LIPOPROTEIN_METABOLISM	3	0.925	0.308	1
REACTOME_LIPID_DIGESTION_MOBILIZATION_AND_TRANSPORT	3	0.925	0.308	1
REACTOME_METABOLISM_OF_LIPIDS_AND_LIPOPROTEINS	3	0.925	0.308	1
REACTOME_RESPONSE_TO_ELEVATED_PLATELET_CYTOSOLIC_CA2_	2	0.5	0.25	1
REACTOME_PLATELET_ACTIVATION_SIGNALING_AND_AGGREGATION	3	0.802	0.267	1
REACTOME_COMMON_PATHWAY	4	1.12	0.281	1
REACTOME_FORMATION_OF_FIBRIN_CLOT_CLOTTING_CASCADE	5	1.33	0.266	1
REACTOME_HEMOSTASIS	6	1.57	0.261	1

Chapter 8

General Discussion

Chapter Overview

The overall aim of this thesis was to advance current understanding of the biological systems involved in the post-stroke timeline, extending from acute hospital periods to the recovery phase over the first-year post-stroke. To achieve these aims, an exploration was conducted of acute stroke hospital biological data against clinical recovery measures. Building upon these analyses, this thesis employed several 'discovery-based' technologies and analyses to review currently published studies in stroke systems biology as well as conduct proteomics analyses of patients from a longitudinal stroke cohort over time. Data were examined for single molecules within biosystems, using additional bioinformatics methods that seek to understand molecules in the context of biological pathways. This thesis has established a methodology that will shift the traditional 'candidate biomarker' approach to a more comprehensive 'candidate biosystems' approach, allowing qualitative and quantitative explanation of complete pathways. These findings have immediate implications for clinical adoption of immune related medical tests and therapies that modulate immune system functioning.

A Biosystems Approach

This thesis began with presentation of available evidence for the involvement of biosystems in the ischemic stroke (IS) timeline, from pre-stroke risk to the longitudinal recovery phase. As noted in Chapter 3, while evidence for biological systems such as lipid metabolism (Stoll & Bendszus, 2006), coagulation (Sfredel et al., 2018), inflammation (Barrington, Lemarchand, & Allan, 2017; Guruswamy & El Ali, 2017) and oxidative stress (Allen & Bayraktutan, 2009) have been established in acute IS, evidence for these systems extending to the recovery phase is not as well developed (Kamel & Iadecola, 2012). Understood as a collective thesis, the general findings of Chapters 3, 5, 6 and 7 suggest a temporal association with immune related pathways and the trajectory of the first-year post-stroke, along with the involvement of relative suppression of the

immune related pathways over the year post-stroke. This final discussion chapter will address the immune related findings considering the temporal context.

The Peripheral Immune System in Stroke

Acute phase

Understandably, the majority of stroke human and animal studies on immunity have been conducted in the central nervous system (CNS), addressing themes such as response to damage (Samary, Pelosi, Leme Silva, & Rieken Macedo Rocco, 2016), autophagy (Sun et al., 2019) neuroprotection and restoration (Cruz, Cantú-Saldaña, & Ibarra, 2016). Studies of peripheral immune responses have been limited to immunosuppression and infection (Shi, Wood, Shi, Wang, & Liu, 2018). The most well studied interaction between central and peripheral immune systems in stroke has been leukocyte infiltration, or the early movement of blood-borne peripheral immune cells into the cerebral parenchyma. Indeed, leukocyte infiltration peaks within 24-72 hours in both human and animal models (Edwards & Bix, 2019), corresponding with similarly timed CNS microglia and macrophage responses (Figure 1a-b) (Grønberg, Johansen, Kristiansen, & Hasseldam, 2013). Furthermore, diffusion weighted imaging studies in humans have shown that 24-48 hours is the time of greatest blood brain barrier permeability in acute IS (Figure 2) (Merali, Huang, Mikulis, Silver, & Kassner, 2017). Consistent with established literature, the findings in Chapter 5 show a small significant increase in blood neutrophil count from <12 hours post-stroke to 24-48 hours post-stroke (Nguyen et al., 2020a). While the protocol utilized in Chapter 5 did not have measures to corroborate infiltration of leukocytes into the CNS, these findings may represent a temporal signature of the peripheral immune response to IS damage. Indeed, a review of leukocyte responses in stroke has noted that infiltration is accompanied by increased white blood cell count in the periphery, with further correlation to infarct size (Audebert, Rott, Eck, & Haberl, 2004).

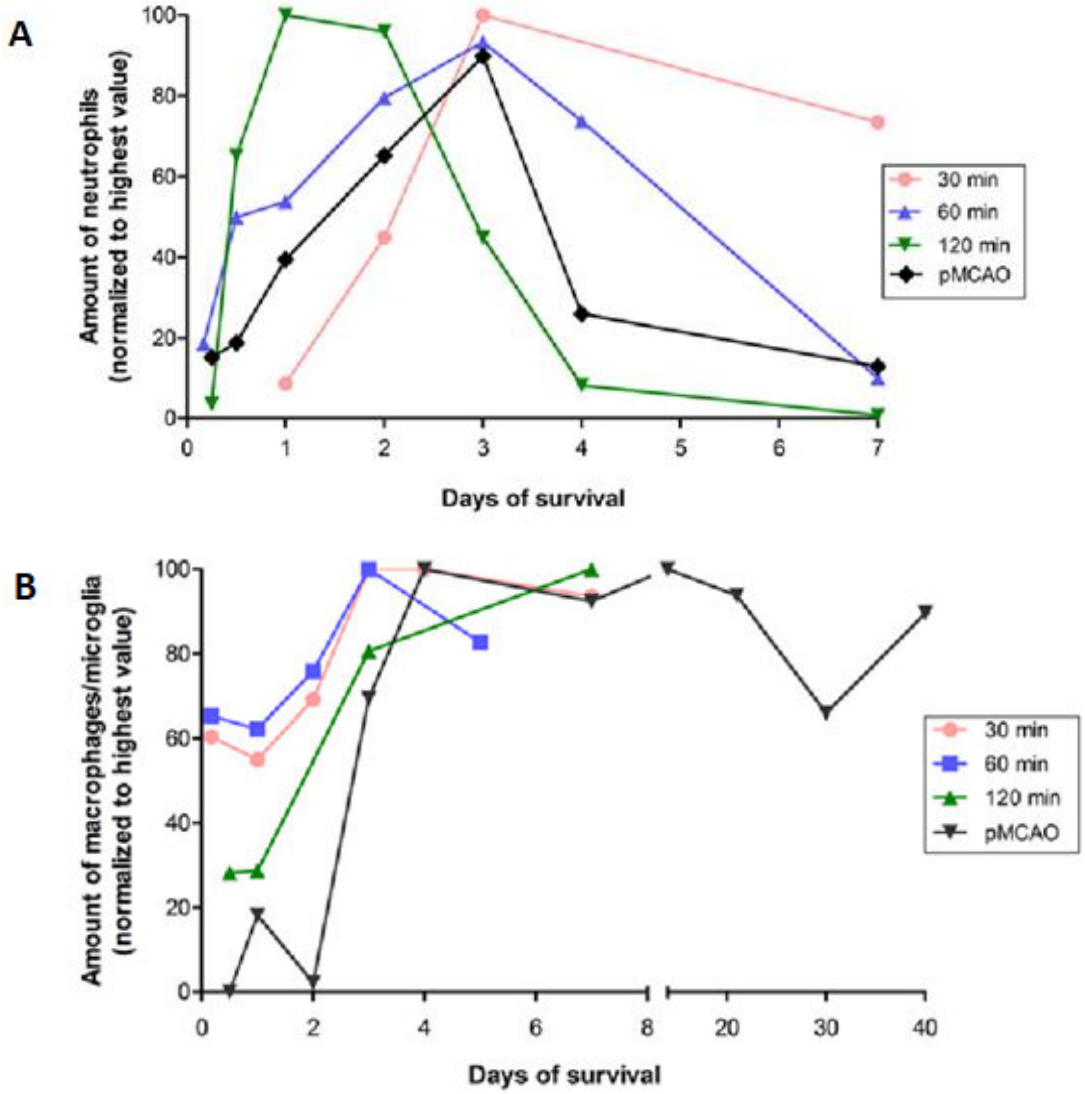

Figure 1. A review conducted of stroke rodent studies exploring the temporal association of peripheral neutrophil (A) and central macrophage or microglia (B) accumulation in transient occlusion models at 30 min, 60 min, 120 min or permanent middle cerebral artery occlusion. Values of neutrophils and macrophages given are normalized to the highest value reported by the reviewed studies then aggregated to produce these plots. Reprinted by permission from the Licensor: Springer Nature, Leukocyte infiltration in experimental stroke, Grønberg, N. V., Johansen, F. F., & Hasseldam, H. COPYRIGHT, 2013.

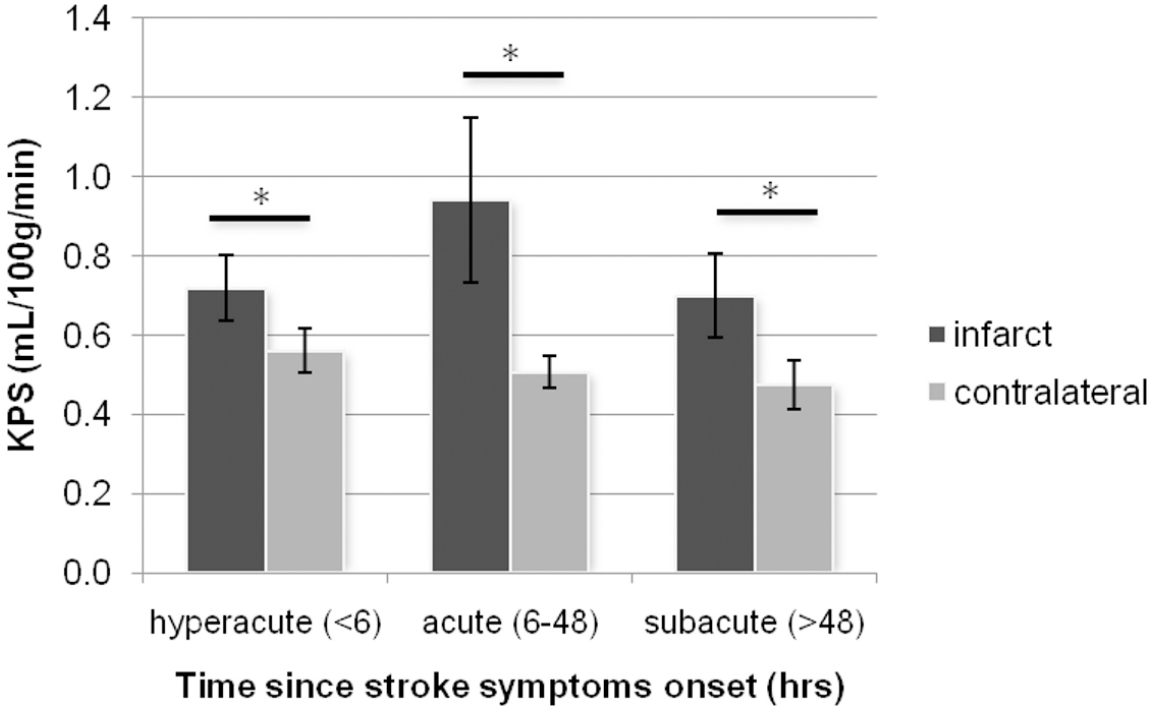

Figure 2. A comparison of blood brain barrier permeability at hyperacute, acute and subacute times between infarct and contralateral areas using diffusion weighted imaging and permeability-surface area product. Reprinted from Merali, Z., Huang, K., Mikulis, D., Silver, F., & Kassner A. Evolution of blood-brain-barrier permeability after acute ischemic stroke, *PLoS ONE*, 2013, 12(2), e0171558 under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Immunosuppression is a common sequalae following stroke and is associated with 20% of post-stroke mortality by predisposing patients to further complications such as pneumonia (Wilson, 2012) and urinary tract infections (Shi et al., 2018). The timing of communicable diseases in stroke attributes the highest rate of infection to 7 days post-stroke (Westendorp, Nederkoorn, Vermeij, Dijkgraaf, & de Beek, 2011), considering both immunological response to stroke and infectious hospital environments (Grabska, Gromadzka, & Członkowska, 2011). The functional adaptation of immunosuppression after stroke has been suggested to be reduced risk of autoimmune damage (Becker, 2012; Javidi & Magnus, 2019) and may be indicative of a system responding to sterile inflammation as opposed to external pathogens. However, the signalling mechanisms for immunosuppression have not been well established. Findings from the study reported in Chapter 6 suggest that the most likely biological pathway to contribute to immunosuppression is the complement system, a system that contributes to the activation

and reactivity status of the peripheral immune system. Proteins that regulate the complement system to effectively attack cells by forming lysis pores such as vitronectin (VTN) and clusterin (CLU) (Chauhan & Moore, 2006) were shown to greatly increase their regulatory effects between 3-5 days and 90 days post-stroke (Nguyen et. al, 2020b). In relation to this finding, the systematic review conducted in Chapter 3 identified platelet degranulation as the most likely biological system to overlap between transcriptomics and proteomics studies within 48 hours post-stroke, of which clusterin was an interacting messenger ribonucleic acid (mRNA) and protein. Although clusterin is readily available in human blood, it is one of the molecules present in the alpha granules of platelets including proteins involved in coagulation and thromboembolism such as von Willebrand factor and fibrinogen (Rijkers et al., 2017). Given these converging points of evidence for coagulation and immunity, it is likely that activated platelets in the acute phase of stroke contribute greatly to peripheral immune signalling systems (Zeller, Lenz, Eschenfelder, Zunker, & Deuschl, 2005) such as complement and therefore induce immunosuppression.

Recovery phase

Following from acute ischemic damage, the immune system is likely to be the central biosystem involved in stroke recovery processes. Understanding of a dynamic system such as immunity over a temporal period should adopt a model centred around immune homeostasis. In this paradigm, homeostasis refers to the stability of the whole immune system considering both the timing and amplitude of a variety of immune functions, such as the pathogen response, endogenous or sterile inflammation and maintenance of natural cell death (Sirisinha, 2011). In this context, protective repair mechanisms can only begin once destructive cell death signalling is resolved (Tobin et al., 2014), describing immune homeostasis as the interplay between pro- and anti-inflammatory processes (Rayasam et al., 2018). The findings of Chapter 4 suggest that an increased number of leukocytes and neutrophils in blood at 24-48 hours correlates with

increased cognitive outcomes at 90 days post-stroke. In this context, it is possible that an active immune response accompanied by a necessary and complementary regulatory response early in stroke influences the speed and trajectory of individual stroke recovery. As discussed in the acute phase, it is speculated that system-wide autoimmune disruption occurs as a result of stroke (Becker et al., 2011) and likely in other examples of CNS insults such as traumatic brain injury (Yang et al., 2017). The findings of Chapter 5 that show downregulated immune and complement system systems are correlated with increasing depression symptomology at 90 days (Nguyen et al., 2016). In addition, Chapter 6 further suggests that increasing complement regulation at 90 days may explain the mechanism for innate immune immunosuppression (Nguyen et. al, 2020b). Indeed, post-stroke infections are known to occur even at 76 days post-stroke (Learoyd et al., 2017) and CNS inflammation likely persists for over a year after stroke (Kamel & Iadecola, 2012). Given this longitudinal timeline, peripheral immune homeostasis may reflect ongoing chronic CNS inflammation and it is possible that assessment of the immune system by a combination of leukocyte counts and blood complement testing may be useful in determining if an individual has entered into a positive state of recovery, if issues such as immunosuppression and inflammation are resolved in the longitudinal phase, relative to acute stroke. Therefore, explicit understanding of the timing and amplitude of an individual's immune responses as the biological system that is likely to be most involved with stroke recovery could provide clinicians with the ability to personalize pharmacological treatments and/or rehabilitation to the appropriate programmes and dosages.

Clinical Implications

Currently, there is a need for biomarkers both as a quantitative tool for assessing individual recovery patterns closely linked to the biology of recovery to understand the effects of current and future stroke therapies. As discussed previously in Chapter 2,

current clinical neurological stroke practice has not yet adopted blood biomarkers for the purpose of predicting recovery. The complement system is comprised of over 50 distinct proteins including complement components C1q, C2b, C3a, C4b, C5a, C5b and sC5b-9 that may be targeted in the clinical setting by commercial assays utilizing neo-epitope monoclonal antibodies or enzyme linked immunosorbent assays (Ekdahl et al., 2018). Indeed, utilizing neo-epitopes also reflects an endogenous immune response to novel foreign antigens that may be useful to consider in the context of stroke immunosuppression. Although few studies have examined peripheral complement system functioning at months post-stroke onset, current findings have suggested that increased complement C3 and C4 from acute IS (24 hours) to discharge (M = 16 days) was associated with worse neurological disability (Cojocaru et al., 2008). Furthermore, increases in admission levels of serum mannose-binding lectin (MBL) were associated with increased mortality and reduced functional recovery at 90 days (Zhang et al., 2015). A preliminary analysis of urine clusterin (CLU) in acute stroke has also been associated with cortical lesions but not with inflammatory markers including apolipoprotein A1 (APOA1), B1 (APOB1) and serum amyloid A (SAA). Given the understanding of the complement system and interacting inflammatory processes, the anaphylactic proteins such as C3a, C4a and C5a may be strong indicators of immune induced inflammation in post-stroke patients and should be investigated within the first week of stroke.

Assessing immunological proteins and immune system functioning in acute stroke may be useful to determine the course and dosage of treatments. Pharmacological treatments such as intravenous immunoglobulin (IVIG) and subcutaneous immunoglobulin (SCIG) have been utilized in animal studies (Widiapradja et al., 2014) and stroke human (Basta, 2014). As a powerful immunological intervention for primary immunodeficiency disorders, IVIG has also been shown to address stroke damage by attenuating leukocyte infiltration and maintaining blood brain barrier permeability

(Widiapradja et al., 2014). In a case study of a patient with post-stroke pneumonia, administration of SCIG successfully recovered immunoglobulin levels and prevented further complications (Moss & Keith, 2013). However, use of expensive immunoglobulin therapies ($2000-$3000 a dose) is often cautioned, as use of IVIG in stroke patients has been suggested to cause vascular complications such as thromboembolisms and death (Basta, 2014). Indeed, IVIG treatments interacts with the later stages of the complement system and reduces inflammatory signalling from C3a and C5a (Galeotti, Kaveri, & Bayry, 2017). Considering the vascular age of stroke patients, development of subcutaneous delivery devices has largely reduced the thromboembolism concerns of IVIG treatments (Kobrynski, 2012). Therefore, although stroke pharmacotherapeutics are relatively unused, there is a clinical need to be able to identify if patients can safely receive advanced immunotherapies in stroke that may be accomplished by directly assessing the biomarkers and functioning of the target system and molecules of these therapies. Although these new immunomodulatory therapies have been shown to have little to no effect on stroke outcomes in clinical trials for natalizumab at both acute (Elkind 2020) and 90-days (Elkins 2017), systems biology techniques can be further used to investigate the effects that these drugs have on patient suitability and long-term immune system functioning.

Limitations

The proteomics approaches of the studies reported in this thesis have examined protein concentrations in plasma. As previously described in Chapter 4, mass spectrometric analyses of plasma samples will often be limited by the presence of relative abundance of large molecules in blood such as human serum albumin. This issue can be seen when comparing the length of lists of identified proteins between Chapter 5 where depletion was achieved with spin columns (475 proteins) and Chapter 6 where there was no abundant protein depletion (163 proteins). Steps to deplete high abundance proteins

are essential to consider in proteomics analyses of plasma as they remove proteins such as albumin from the sample by utilizing affinity columns (Cao, Yende, Kellum, & Robinson, 2013), increased column processing time, with different column parameters such as temperature or using spin columns in the sample preparation phase (Polaskova, Kapur, Khan, Molloy, & Baker, 2010). In the methodology for Chapter 5, each complete mass spectrometric analysis per sample with depletion methods occurred over an average of 9 hours whilst runs with no depletion in Chapter 6 occurred over 3 hours. These considerations add additional costs and run time on the machine, an issue that may be prohibitive to large datasets and the fast turnaround times required in a clinical environment.

The bioinformatics methods employed in the experimental series of this thesis assume that protein and transcriptomics are comparatively similar at the identifier level. Although separate nomenclature and annotations exist for protein species (UniProt, 2008), many of the databases used throughout this thesis that examine biochemical reactions and pathways such as the Kyoto Encyclopedia of Genes and Genomes (Kanehisa & Goto, 2000) and Reactome (Fabregat et al., 2018) are based on information from genomics and transcriptomics studies. Compared to DNA and RNA sequences, proteins or amino acid sequences may be able to be expressed in a variety of isoforms that may not be a complete copy of the original genomic sequence, whereby structure dictates function (Stastna & Van Eyk, 2012). Overall, this limitation may be addressed by designing studies that address protein post-translational modifications (Perchey et al., 2019) and protein-protein interactions (Smits & Vermeulen, 2016), although noting that more complicated analytical chemistry and bioinformatics methodologies are needed than those presented in the current thesis.

Future Directions

Investigating the involvement of the human immune system will lead to better understanding of the post-stroke recovery timeline. Although this thesis suggested that the innate immune dysregulation follows ischemic stroke, the biological circumstances leading to stroke complement-induced immunosuppression at 3 months are relatively unknown. Further proteomics studies between the first week of stroke and 3 months should provide further evidence of the temporal dynamics of complement expression. There are a variety of molecular signalling pathways, especially those that are initiated by platelet driven coagulation responses (Sharp et al., 2011) that should be considered as the link between acute stroke and peripheral immunity during the first-year recovery phase. Adopting a framework similar to that described in Chapter 3, i.e. a multi-omics framework adopting genomics, proteomics and metabolomics, will help guide examination of a complete range of blood products including red blood cells, platelets, peripheral blood mononuclear cells from the hours to years post-stroke. This comprehensive approach would lead to a much more nuanced understanding of the biological systems involved in stroke recovery. Furthermore, understanding of the immune system in stroke patients should consider that metabolism, oxidative stress and vascular fragility are components of an aging system (Mattson & Arumugam, 2018). As suggested in Chapter 5, it is likely that the linking of immune systems to a bioenergetics framework affected by cellular senescence may be the main determinant of how well a system will react to stroke and other CNS trauma and thus explain heterogeneity in recovery profiles. Indeed, prior studies have successfully explained how proteins and metabolites are linked to functional measures of mitochondrial metabolism in diabetes (Haythorne et al., 2019) and kidney cancer (Wettersten et al., 2015), using high resolution respirometers such as the Seahorse XF or the OxyGraph-2k (Horan, Pichaud, & Ballard, 2012). Utilizing a combination of systems biology techniques and cellular respirometry within the first months post-stroke will be able to provide clinicians and rehabilitation

specialists with a comprehensive approach to understanding and predicting clinical outcomes such as post-stroke cognition and depression.

As gender is often viewed as a statistical confound in the context of stroke research, a better understanding of gender differences in biological systems responses to stroke is needed. When compared to males, females have been shown to have worsened outcomes, including mortality, physical impairments and activities of daily living (Reeves et al., 2008). Studies have explored the biological underpinnings of sexual dimorphism in stroke aetiology as products of aging (Appelros, Stegmayr, & Terént, 2009) and hormonal differences (Girijala, Sohrabji, & Bush, 2017). Although the evidence is limited, both transcriptomic (Stamova, et al., 2019) and proteomic (Cao, Cobbs, Simon, & Zhou, 2019) studies suggest that females have a stronger initial inflammatory response to stroke than males, including upregulation in complement, interleukins, B cell and lymphocyte pathways. Furthermore, females have a more robust treatment response to acute treatments such as tPA but may exhibit a worsened response to nitric oxide (Sohrabji, Park, & Mahnke, 2017). Therefore, there appears to be clear gender differences to initial stroke damage that are mediated by immune response. However, the systems driving gender differences towards long-term post-stroke recovery are less well understood and have great potential to be investigated utilizing further systems biology approaches.

Discovery approaches are often the beginning phase of research programs to understand disease. Often, it is intended that the understanding gained from discovery approaches will be used to develop further treatments. In the field of stroke, there are no well-established and commonly used pharmacological interventions for positive recovery. Given that this thesis has shown that genomics, transcriptomics and proteomics with the conjunction of bioinformatics methods offer simultaneous understanding and quantification for both biomarkers and biosystems, these paradigms could provide the formulative basis for drug discovery programs. Pharmacogenetics and

pharmacoproteomics are relatively new fields of research that seek to identify and tailor novel drugs and monitor their responses to patients with complex diseases (Chambliss & Chan, 2016). These approaches have led to the development of gene regulated dosage considerations in clinical practice for colon cancer patients (Innocenti et al., 2014). However, it is unlikely that simple candidate-gene associations will lead to advances in stroke pharmacology as stroke is not a monogenic disorder and does not share strong associations with any one genotype (Malik et al., 2018). Instead, an encompassing discovery approach examining patients that recover well or recover worse utilizing a variety of physiological and clinical measures could provide the most comprehensive understanding of how biological systems are involved in stroke recovery. This pharma-omics approach may provide the biomarkers and biosystems level understanding necessary to recommend existing drugs or generate novel pharmacotherapeutics.

Conclusions

The premise of this thesis was to advance understanding of the biological systems involved in stroke recovery by using new and powerful approaches in genomics, transcriptomics and proteomics. Stroke recovery outcomes of cognition and depression were the major focus given the relative lack of quantitative biology-based predictors of post-stroke cognitive impairment and post-stroke depression. The experimental series suggest that immune system suppression follows from the first week to 3 months post-stroke and is related to the degree of depression symptoms. Blood based measures of immune cells are also elevated in early acute stroke and peripheral immune response at 48 hours may also be linked to leukocyte infiltration and BBB permeability. It is likely that variability and in immune response will be the most important biological system to explain heterogenous stroke recovery patterns, whereby early acute immune responses may indicate the ability of the immune system to re-establish homeostasis and therefore initiate efficient neural repair in the months to years post-stroke. As a collection of

discovery studies, these findings provide direction for future studies to better understand the temporal trajectory of post-stroke immunological changes to develop candidate biomarkers and candidate biosystems to improve post-stroke care.

References

Allen, C. L., & Bayraktutan, U. (2009). Oxidative stress and its role in the pathogenesis of ischaemic stroke. *International Journal of Stroke, 4*(6), 461-470. doi:10.1111/j.1747-4949.2009.00387.x

Audebert, H. J., Rott, M. M., Eck, T., & Haberl, R. L. (2004). Systemic inflammatory response depends on initial stroke severity but is attenuated by successful thrombolysis. *Stroke, 35*(9), 2128-2133. doi:10.1161/01.STR.0000137607.61697.77

Barrington, J., Lemarchand, E., & Allan, S. M. (2017). A brain in flame; do inflammasomes and pyroptosis influence stroke pathology? *Brain Pathology, 27*(2), 205-212. doi:10.1111/bpa.12476

Basta, M. (2014). Intravenous immunoglobulin-related thromboembolic events - an accusation that proves the opposite. *Clinical and Experimental Immunology, 178 Suppl 1*(Suppl 1), 153-155. doi:10.1111/cei.12551

Becker, K. (2012). Autoimmune responses to brain following stroke. *Translational Stroke Research, 3*(3), 310-317. doi:10.1007/s12975-012-0154-0

Becker, K. J., Kalil, A. J., Tanzi, P., Zierath, D. K., Savos, A. V., Gee, J. M., . . . Cain, K. C. (2011). Autoimmune responses to brain following stroke are associated with worse outcome. *Stroke, 42*(10), 2763-2769. doi:10.1161/STROKEAHA.111.619593

Cao, L., Cobbs, A., Simon, R. P., & Zhou, A. (2019). Distinct plasma proteomic changes in male and female african american stroke patients. *International Journal of Physiology, Pathophysiology and Pharmacology, 11*(2), 12-20.

Cao, Z., Yende, S., Kellum, J. A., & Robinson, R. A. (2013). Additions to the human plasma proteome via a tandem MARS depletion ITRAQ-based workflow. *International Journal of Proteomics, 2013*, 8. doi:10.1155/2013/654356

Chambliss, A. B., & Chan, D. W. (2016). Precision medicine: From pharmacogenomics to pharmacoproteomics. *Clinical Proteomics, 13*, 25-25. doi:10.1186/s12014-016-9127-8

Chauhan, A. K., & Moore, T. L. (2006). Presence of plasma complement regulatory proteins clusterin (APOJ) and vitronectin (S40) on circulating immune complexes (CIC). *Clinical and Experimental Immunology, 145*(3), 398-406. doi:10.1111/j.1365-2249.2006.03135.x

Cojocaru, I. M., Cojocaru, M., Tanasescu, R., Burcin, C., Atanasiu, A. N., Petrescu, A.-M., . . . Dumitrescu, L. (2008). Changes in plasma levels of complement in patients with acute ischemic stroke. *Romanian Journal of Internal Medicine, 46*(1), 77-80.

Cruz, Y., Cantú-Saldaña, K., & Ibarra, A. (2016). Immune system involvement in the degeneration, neuroprotection, and restoration after stroke. *Ischemic Stroke: Updates*, 107.

Edwards, D. N., & Bix, G. J. (2019). The inflammatory response after ischemic stroke: Targeting β2 and β1 integrins. *Frontiers in Neuroscience, 13*(540). doi:10.3389/fnins.2019.00540

Ekdahl, K. N., Persson, B., Mohlin, C., Sandholm, K., Skattum, L., & Nilsson, B. (2018). Interpretation of serological complement biomarkers in disease. *Frontiers in Immunology, 9*, 2237-2237. doi:10.3389/fimmu.2018.02237

Elkins, J., Veltkamp, R., Montaner, J., Johnston, S. C., Singhal, A. B., . . . Elkind, M. (2017). Safety and efficacy of natalizumab in patients with acute ischaemic stroke (ACTION): a randomised, placebo-controlled, double-blind phase 2 trial. The Lancet. Neurology, 16(3), 217–226. https://doi.org/10.1016/S1474-4422(16)30357-X

Elkind, M., Veltkamp, R., Montaner, J., Johnston, S. C., Singhal, A. B., Becker, K., . . . Elkins, J. (2020). Natalizumab in acute ischemic stroke (ACTION II): A randomized, placebo-controlled trial. Neurology, 95(8), e1091–e1104.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., . . . D'Eustachio, P. (2018). The reactome pathway knowledgebase. *Nucleic Acids Research, 46*, 649-655. doi:10.1093/nar/gkx1132

Galeotti, C., Kaveri, S. V., & Bayry, J. (2017). IVIG-mediated effector functions in autoimmune and inflammatory diseases. *International Immunology, 29*(11), 491-498. doi:10.1093/intimm/dxx039

Grabska, K., Gromadzka, G., & Członkowska, A. (2011). Infections and ischemic stroke outcome. *Neurology Research International, 2011*.

Grønberg, N. V., Johansen, F. F., Kristiansen, U., & Hasseldam, H. (2013). Leukocyte infiltration in experimental stroke. *Journal of Neuroinflammation, 10*(1), 892.

Guruswamy, R., & El Ali, A. (2017). Complex roles of microglial cells in ischemic stroke pathobiology: New insights and future directions. *International Journal of Molecular Sciences, 18*(3), 496. doi:10.3390/ijms18030496

Haythorne, E., Rohm, M., van de Bunt, M., Brereton, M. F., Tarasov, A. I., Blacker, T. S., . . . Ashcroft, F. M. (2019). Diabetes causes marked inhibition of mitochondrial metabolism in pancreatic β-cells. *Nature Communications, 10*(1), 2474. doi:10.1038/s41467-019-10189-x

Horan, M. P., Pichaud, N., & Ballard, J. W. O. (2012). Review: Quantifying mitochondrial dysfunction in complex diseases of aging. *The Journals of Gerontology: Series A, 67*(10), 1022-1035. doi:10.1093/gerona/glr263

Innocenti, F., Schilsky, R. L., Ramírez, J., Janisch, L., Undevia, S., House, L. K., . . . Ratain, M. J. (2014). Dose-finding and pharmacokinetic study to optimize the

dosing of irinotecan according to the ugt1a1 genotype of patients with cancer. *Journal of Clinical Oncology, 32*(22), 2328-2334. doi:10.1200/JCO.2014.55.2307

Javidi, E., & Magnus, T. (2019). Autoimmunity after ischemic stroke and brain injury. *Frontiers in Immunology, 10*(686). doi:10.3389/fimmu.2019.00686

Kamel, H., & Iadecola, C. (2012). Brain-immune interactions and ischemic stroke: Clinical implications. *Archives of Neurology, 69*(5), 576-581. doi:10.1001/archneurol.2011.3590

Kanehisa, M., & Goto, S. (2000). Kegg: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research, 28*(1), 27-30.

Kobrynski, L. (2012). Subcutaneous immunoglobulin therapy: A new option for patients with primary immunodeficiency diseases. *Biologics : targets & therapy, 6*, 277-287. doi:10.2147/BTT.S25188

Learoyd, A. E., Woodhouse, L., Shaw, L., Sprigg, N., Bereczki, D., Berge, E., . . . investigators, o. b. o. t. E. T. (2017). Infections up to 76 days after stroke increase disability and death. *Translational stroke research, 8*(6), 541-548. doi:10.1007/s12975-017-0553-3

Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., . . . Attia, J. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics, 50*(4), 524-537. doi:10.1038/s41588-018-0058-3

Mattson, M. P., & Arumugam, T. V. (2018). Hallmarks of brain aging: Adaptive and pathological modification by metabolic states. *Cell Metabolism, 27*(6), 1176-1199. doi:https://doi.org/10.1016/j.cmet.2018.05.011

Merali, Z., Huang, K., Mikulis, D., Silver, F., & Kassner, A. (2017). Evolution of blood-brain-barrier permeability after acute ischemic stroke. *PloS One, 12*(2), e0171558. doi:10.1371/journal.pone.0171558

Moss, L. A., & Keith, P. (2013). Basic and clinical immunology – 3030. Subcutaneous immunoglobulin therapy: An option for patients who have experienced thrombotic complications with intravenous therapy. *The World Allergy Organization Journal, 6*(Suppl 1), P205-P205. doi:10.1186/1939-4551-6-S1-P205

Nguyen, V. A., Carey, L. M., Giummarra, L., Faou, P., Cooke, I., Howells, D. W., . . . Crewther, S. G. (2016). A pathway proteomic profile of ischemic stroke survivors reveals innate immune dysfunction in association with mild symptoms of depression – a pilot study. *Frontiers in Neurology, 7*, 85. doi:10.3389/fneur.2016.00085

Perchey, R. T., Tonini, L., Tosolini, M., Fournié, J.-J., Lopez, F., Besson, A., & Pont, F. (2019). Ptmselect: Optimization of protein modifications discovery by mass spectrometry. *Scientific Reports, 9*(1), 4181. doi:10.1038/s41598-019-40873-3

Polaskova, V., Kapur, A., Khan, A., Molloy, M. P., & Baker, M. S. (2010). High-abundance protein depletion: Comparison of methods for human plasma biomarker discovery. *Electrophoresis, 31*(3), 471-482. doi:10.1002/elps.200900286

Rayasam, A., Hsu, M., Kijak, J. A., Kissel, L., Hernandez, G., Sandor, M., & Fabry, Z. (2018). Immune responses in stroke: How the immune system contributes to damage and healing after stroke and how this knowledge could be translated to better cures? *Immunology, 154*(3), 363-376. doi:10.1111/imm.12918

Reeves, M. J., Bushnell, C. D., Howard, G., Gargano, J. W., Duncan, P. W., Lynch, G., . . . Lisabeth, L. (2008). Sex differences in stroke: Epidemiology, clinical presentation, medical care, and outcomes. *The Lancet. Neurology, 7*(10), 915-926. doi:10.1016/S1474-4422(08)70193-5

Rijkers, M., van den Eshof, B. L., van der Meer, P. F., van Alphen, F. P. J., de Korte, D., Leebeek, F. W. G., . . . Jansen, A. J. G. (2017). Monitoring storage induced

changes in the platelet proteome employing label free quantitative mass spectrometry. *Scientific Reports, 7*(1), 11045. doi:10.1038/s41598-017-11643-w

Samary, C., Pelosi, P., Leme Silva, P., & Rieken Macedo Rocco, P. (2016). Immunomodulation after ischemic stroke: Potential mechanisms and implications for therapy. *Critical Care, 20*(1), 391. doi:10.1186/s13054-016-1573-1

Sfredel, M. D., Burada, E., Cătălin, B., Dinescu, V., Târtea, G., Iancău, M., & Osiac, E. (2018). Blood coagulation following an acute ischemic stroke. *Current Health Sciences Journal, 44*(2), 118-121. doi:10.12865/CHSJ.44.02.04

Sharp, F. R., Jickling, G. C., Stamova, B., Tian, Y., Zhan, X., Liu, D., . . . Ander, B. P. (2011). Molecular markers and mechanisms of stroke: Rna studies of blood in animals and humans. *Journal of Cerebral Blood Flow and Metabolism, 31*(7), 1513-1531. doi:10.1038/jcbfm.2011.45

Shi, K., Wood, K., Shi, F.-D., Wang, X., & Liu, Q. (2018). Stroke-induced immunosuppression and poststroke infection. *Stroke and Vascular Neurology, 3*(1), 34.

Sirisinha, S. (2011). Insight into the mechanisms regulating immune homeostasis in health and disease. *Asian Pacific Journal of Allergy and Immunology, 29*(1), 1-14.

Smits, A. H., & Vermeulen, M. (2016). Characterizing protein-protein interactions using mass spectrometry: Challenges and opportunities. *Trends in Biotechnology, 34*(10), 825-834. doi:10.1016/j.tibtech.2016.02.014

Sohrabji, F., Park, M. J., & Mahnke, A. H. (2017). Sex differences in stroke therapies. *Journal of Neuroscience Research, 95*(1-2), 681-691. doi:10.1002/jnr.23855

Stamova, B., Jickling, G. C., Ander, B. P., Zhan, X., Liu, D., Turner, R., . . . Sharp, F. R. (2014). Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. *PLOS ONE, 9*(7), e102550. doi:10.1371/journal.pone.0102550

Stastna, M., & Van Eyk, J. E. (2012). Analysis of protein isoforms: Can we do it better? *Proteomics, 12*(19-20), 2937-2948. doi:10.1002/pmic.201200161

Stoll, G., & Bendszus, M. (2006). Inflammation and atherosclerosis. *Stroke, 37*(7), 1923-1932. doi:doi:10.1161/01.STR.0000226901.34927.10

Sun, Y., Zhu, Y., Zhong, X., Chen, X., Wang, J., & Ying, G. (2019). Crosstalk between autophagy and cerebral ischemia. *Frontiers in Neuroscience, 12*(1022). doi:10.3389/fnins.2018.01022

Tobin, M. K., Bonds, J. A., Minshall, R. D., Pelligrino, D. A., Testai, F. D., & Lazarov, O. (2014). Neurogenesis and inflammation after ischemic stroke: What is known and where we go from here. *Journal of Cerebral Blood Flow and Metabolism, 34*(10), 1573-1584. doi:10.1038/jcbfm.2014.130

UniProt, C. (2008). The universal protein resource (UNIPROT). *Nucleic Acids Research, 36*(Database issue), D190-D195. doi:10.1093/nar/gkm895

Westendorp, W. F., Nederkoorn, P. J., Vermeij, J.-D., Dijkgraaf, M. G., & de Beek, D. v. (2011). Post-stroke infection: A systematic review and meta-analysis. *BMC Neurology, 11*, 110-110. doi:10.1186/1471-2377-11-110

Wettersten, H. I., Hakimi, A. A., Morin, D., Bianchi, C., Johnstone, M. E., Donohoe, D. R., . . . Weiss, R. H. (2015). Grade-dependent metabolic reprogramming in kidney cancer revealed by combined proteomics and metabolomics analysis. *Cancer Research, 75*(12), 2541-2552. doi:10.1158/0008-5472.CAN-14-1703

Widiapradja, A., Santro, T., Basta, M., Sobey, C. G., Manzanero, S., & Arumugam, T. V. (2014). Intravenous immunoglobulin (IVIG) provides protection against endothelial cell dysfunction and death in ischemic stroke. *Experimental & Translational Stroke Medicine, 6*(1), 7. doi:10.1186/2040-7378-6-7

Wilson, R. D. (2012). Mortality and cost of pneumonia after stroke for different risk groups. *Journal of Stroke and Cerebrovascular Diseases, 21*(1), 61-67. doi:10.1016/j.jstrokecerebrovasdis.2010.05.002

Yang, Z., Zhu, T., Weissman, A. S., Jaalouk, E., Rathore, D. S., Romo, P., . . . Wang, K. K. W. (2017). Autoimmunity and traumatic brain injury. *Current Physical Medicine and Rehabilitation Reports, 5*(1), 22-29. doi:10.1007/s40141-017-0146-9

Zeller, J. A., Lenz, A., Eschenfelder, C. C., Zunker, P., & Deuschl, G. (2005). Platelet-leukocyte interaction and platelet activation in acute stroke with and without preceding infection. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *25*(7), 1519-1523. doi:doi:10.1161/01.ATV.0000167524.69092.16

Zhang, Z.-G., Wang, C., Wang, J., Zhang, Z., Yang, Y.-L., Gao, L., . . . Li, L.-H. (2015). Prognostic value of mannose-binding lectin: 90-day outcome in patients with acute ischemic stroke. *Molecular Neurobiology*, *51*(1), 230-239. doi:10.1007/s12035-014-8682-0

Appendices

Appendix 1 – List of Copyrighted Figures

Figure 1.1a
Original Image:
Hemorrhagic and Ischemic Strokes Compared.
Anderson, K. K., Olsen, T. S., Dehlendorff, C., & Kammersgaard, L. P.
Stroke, 40(6), 2009, pp. 2070, Figure 2 (as it appears in publication)
https://www.ahajournals.org/doi/full/10.1161/strokeaha.108.540112
Original Caption:
Hazard Ratio for patients with hemorrhagic stroke compared to ischemic stroke obtained from the Poisson survival model. It is seen that the HR is not constant but has a sharp decrease as a function of time since stroke.
Permission obtained with Wolters Kluwer Health, Inc. 26/02/2020
RightsLink License: 4776831509908

Figure 1.1b
Original Image:
Differences in Outcome and Predictors Between Ischemic and Intracerebral Hemorrhage
Bhalla, A., Wang, Y., Rudd, A., Wolfe, C.
Stroke, 44(8), 2013, pp. 2197, Figure 2 (as it appears in publication)
https://www.ahajournals.org/doi/10.1161/STROKEAHA.113.001263
Original Caption:
Figure 2. Time course recovery by subtype for survivors.
Permission obtained with Wolters Kluwer Health, Inc. 16/03/2020
RightsLink License: 4790910663589

Figure 2.1
Original Image:
Metabolic Pathways map01100
https://www.genome.jp/dbget-bin/www_bget?pathway+map01100
Permission obtained via correspondence with Miwako Matsumoto, Kanehisa Laboratories. 14/02/2020

Figure 2.2
Original Image:
The complexity of neurobiological processes in acute ischemic stroke
Brouns, R., De Deyn, P. P.
Clinical Neurology and Neurosurgery, 111(6), 2009, pp. 485, Figure 2 (as it appears in publication)
https://www.sciencedirect.com/science/article/abs/pii/S0303846709000821?via%3Dihub
Original Caption:
Fig. 2. Graph representing the temporal profile of the main pathophysiological mechanisms underlying acute focal cerebral ischemia and their impact on the final ischemicdamage. In absence of early reperfusion, cells in the ischemic penumbra subside due to ongoing ischemic injury, resulting in expansion of the infarcted core.
Permission obtained with Wolters Kluwer Health, Inc. 09/02/2020
RightsLink License: 4764901435391

Figure 2.3
Original Image:

Ionic Regulation of Cell Volume Changes and Cell Death after Ischemic Stroke
Song, M., Yu, S. P.
Translational Stroke Research, 5(1), 2014, pp. 20, Figure 1 (as it appears in publication)
Original Caption:
Fig. 1 Ionic mechanism of ischemia-induced cell death. The graphic diagram illustrates a simplified model of ischemia-induced neuronal cell death. Excessive activation of Ca2+ and Na+ permeable channels and receptors leads to intracellular Ca2+ and Na+ accumulation. The resulted cell swelling and destructive consequences are characteristics of necrosis. On the other hand, ischemia can cause over-activation of K+ permeable channels and receptors that mediate pro-apoptotic K+ efflux. The resulted intracellular K+ reduction consequently induces caspase activation and apoptotic cell death. Concurrent activation of these cell death events likely occurs after an ischemic insult and may result in hybrid cell death features in the same cells. The dysfunction of Na+ /K+ pump plays a critical role in the development of hybrid cell death. Recent data suggest that excessive autophage also contributes to ischemic cellular damage and is an integrated component of hybrid cell death.
Permission obtained with Springer Nature, 10/02/2020.
https://doi-org.ez.library.latrobe.edu.au/10.1007/s12975-013-0314-x
RightsLink License: 4765331004071

Figure 2.4
Original Image:
Post-ischemic inflammation regulates neural damage and protection
Shichita, T., Ito, M., & Yoshimura, A.
Frontiers in Cellular Neuroscience, 8, 2014, Article 319, pp. 2.
Creative Commons Attribution 4.0 International (CC BY 4.0)
Original Caption:
Figure 1. | Mechanisms of post-ischemic inflammation. DAMPs are released into extracellular compartment and activate infiltrating immune cells by two ways: Signal 1 (via the activation of pattern recognition receptor) and Signal 2 (via the activation of inflammasome). Various inflammatory cytokines promote neuronal injury, and induce further inflammation mediated by T cells in subacute phase. After days and week after stroke onset, the resolution of post-ischemic inflammation is brought by the clearance of debris including DAMPs or inflammatory mediators, and the production of anti-inflammatory molecules or neurotrophic factors. In this recover phase, inflammatory immune cells turn into neuroprotective cells.
Used under the Creative Commons International Attribution License 4.0
https://www.frontiersin.org/articles/10.3389/fncel.2014.00319/full

Figure 2.5
Original Image:
Blood-brain barrier tight junction permeability and ischemic stroke
Sandoval, K. E., Witt, K. A.
Neurobiology of Disease, 32(2), 2008, pp. 206, Figure 2 (as it appears in publication)
Original Caption:
Fig. 2. Schematic of blood-brain barrier phasic events associated with cerebral ischemia and reperfusion time-course, as defined in the text. Variability occurs in time-frame of mediators identified, dependent upon tissue distance from ischemic core and duration of

ischemic insult.
Permission obtained with Elseiver, 09/02/2020.
https://www.sciencedirect.com/science/article/abs/pii/S0969996108001927?via%3Dihub
Rightslink License: 4764921286660

Figure 4.2
Original Image:
State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research
Carette, O., Burkhard, P. R., Sanchez, J., & Hochstrasser
Nature Protocols, 1(2), 2006, pp. 819. Figure 2 (as it appears in publication)
Original Caption:
Figure 2. 2D-PAGE gel of human plasma stained with sensitive silver. 45 micrograms of protein was loaded. As indicated by the arrows, the protein migration in the first dimension occurs from right to left (for cathodic cup loading); the second-dimension separation occurs from top to bottom. This applies to all gels shown except Figure 4a.
Permission obtained with Springer Nature, 10/02/2020.
https://www.nature.com/articles/nprot.2006.104
RightsLink License: 4765241443203

Figure 4.4
Original Image:
Analytical strategies for characterization of oxysterol lipidomes:Liver X receptor ligands in plasma
Griffiths, W. J., Crick, P. J., Wang, Y., Ogundare, M., Tuschl, K., Morris, A. A., Bigger, B. W., Clayton, P. T. & Wang, Y.
Free Radical Biology and Medicine, 59, 2013, pp.81. Figure **6B** (as it appears in publication)
Original Caption:
Fig. 6. Negative-ion ESI-MS spectra of plasma samples from (A) a control infant and **(B) a boy with CTX not on treatment.** (C) ESI-MS2spectrum of the ionm/z367 from thecontrol sample. (D) ESI-MS2spectrum of the ionm/z611 from the boy with CTX. Peaks labeled with an asterisk correspond to C19and C21steroid sulfates, those with afilled square to oxysterol glucuronides, and those with a filled circle to fatty acids. Spectra were recorded on the LTQ-Orbitrap XL. Spectra in (A) and (B) were recorded athigh resolution, (C) and (D) utilized the ion-trap detector.
Used under the Creative Commons 3.0 License (Unported)
https://www.sciencedirect.com/science/article/pii/S0891584912004224?via%3Dihub

Figure 7.1a & b
Original Image:
Leukocyte infiltration in experimental stroke
Grønberg, N. V., Johansen, F. F., & Hasseldam, H.
Journal of Neuroinflammation, 10, Article 892, 2013, pp. 4. Figure 1 (as it appears in publication) and pp. 5. Figure 2 (as it appears in publication)
Original Captions:
Figure 1 Mean of normalized amount of infiltrating neutrophils according to survival time. Duration of occlusion has an influence on the temporal profile of neutrophilic

infiltration, suggesting that 120 minutes of MCAO results in an early massive influx of neutrophil granulocytes compared to the shorter occlusion times. Data are also presented in Table 2. MCAO, middle cerebral artery occlusion; pMCAO, permanent middle cerebral artery occlusion.
Figure 2 Mean of normalized amount of infiltrating macrophages/microglia according to survival time. The amount of macrophages/microglia in the infarcted area increases until day 4 to 7 after experimental stroke, whereafter a plateau is reached. Data are also presented in Table 3. pMCAO, permanent middle cerebral artery occlusion.
https://jneuroinflammation.biomedcentral.com/articles/10.1186/1742-2094-10-115
RightsLink License: 4792740089500

Figure 7.2
Original Image:
Evolution of blood-brain-barrier permeability after acute ischemic stroke
Merali, Z., Huang, K., Mikulis, D., Silver, F., Kassner, A.
PloS One, 2014,12(2), Article: e0171558 , 2017, pp. 6, Figure 2 (as it appears in publication)
Creative Commons Attribution 4.0 International (CC BY 4.0)
Original Caption:
Fig 2. Blood-brain-barrier permeability within the infarct and a homologous region in the contralateral hemisphere stratified by time since stroke symptom onset.
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171558

Appendix 2 – Chapter 3, Supplementary Data 1
List of genes from the MEGASTROKE study based on VEGAS $p < 0.05*10^{-4}$

ENTREZID	Gene Symbol
64218	SEMA4A
102467003	MIR7851
79957	PAQR6
100527963	PMF1-BGLAP
11243	PMF1
632	BGLAP
9673	SLC25A44
23381	SMG5
94234	FOXQ1
102466720	MIR6720
2295	FOXF2
84283	TMEM79
23143	LRCH1
6311	ATXN2
8315	BRAP
4000	LMNA
1030	CDKN2B
10019	SH2B3
80724	ACAD10
100048912	CDKN2B-AS1
4321	MMP12
4314	MMP3
100288077	WTAPP1
9644	SH3PXD2A
102465456	MIR6761
217	ALDH2
4312	MMP1
144717	PHETA1
23316	CUX2
79991	STN1
3949	LDLR
6827	SUPT4H1
9256	TSPOAP1
6597	SMARCA4
6128	RPL6
100506779	TSPOAP1-AS1
100616220	MIR4736
102465534	MIR6886
5308	PITX2
406934	MIR142
147727	ILF3-DT
8550	MAPKAPK5
54894	RNF43

ENTREZID	Gene Symbol
3609	ILF3
57153	SLC44A2
10961	ERP29
8759	ADAM1A
89894	TMEM116
4319	MMP10
2244	FGB
81890	QTRT1
59338	PLEKHA1
2243	FGA
2266	FGG
728066	FAM133DP
257415	FAM133B
10906	TRAFD1
4025	LPO
4353	MPO
693223	MIR638
100500866	MIR3941
124535	HSF5
24149	ZNF318
5781	PTPN11
1785	DNM2
84466	MEGF10
102465455	MIR6760
1021	CDK6
727910	TLCD2
10864	SLC22A7
10053	AP1M2
23584	VSIG2
401262	CRIP3
4900	NRGN
1012	CDH13
10594	PRPF8
3664	IRF6
1282	COL4A1
112770	GLMP
5654	HTRA1
5356	PLRG1
90952	ESAM
283450	HECTD4
83547	RILP
102465519	MIR6861
27042	UTP25

80018	NAA25
51275	MAPKAPK5-AS1
387715	ARMS2
100874039	THOC7-AS1
1026	CDKN1A
100422825	MIR3160-2
84333	PCGF5
100422827	MIR3160-1
124935	SLC43A2
56931	DUS3L
163154	PRR22
100422976	MIR4256
9734	HDAC9
63976	PRDM16
613210	DEFB136
55585	UBE2Q1
2780	GNAT2
1141	CHRNB2
23258	DENND5A
53340	SPA17
103	ADAR
134957	STXBP5
81688	C6orf62
285973	ATG9B
153770	PLAC8L1
8497	PPFIA4
55856	ACOT13
6175	RPLP0
130872	AHSA2P
51567	TDP2
51520	LARS
440026	TMEM41B
284083	C17orf47
2773	GNAI3
283254	HARBI1
100506649	PXN-AS1
473	RERE
101154753	PANDAR
126669	SHE
406974	MIR197

157807	CLVS1
10985	GCN1
8578	SCARF1
653659	TMEM183B
79684	MSANTD2
407004	MIR22
101928417	LOC101928417
84981	MIR22HG
100616179	MIR4498
287	ANK2
119369	NUDT9P1
9736	USP34
283902	HCCAT5
92703	TMEM183A
221687	RNF182
80342	TRAF3IP3
133923	ZNF474
148304	C1orf74
4846	NOS3
22853	LMTK2
126668	TDRD10
3757	KCNH2
83873	GPR61
9856	KIAA0319
11021	RAB35
124997	WDR81
4656	MYOG
9110	MTMR4
133619	PRRC1
55626	AMBRA1
54439	RBM27
84630	TTBK1

Appendix 3 – Chapter 3, Supplementary Data 2
List of mRNA transcripts from untargeted stroke transcriptomics studies

ENTREZID	SYMBOL	GENENAME	PROBEID
383	ARG1	arginase 1	NA
762	CA4	carbonic anhydrase 4	NA
1236	CCR7	C-C motif chemokine receptor 7	NA
8826	IQGAP1	IQ motif containing GTPase activating protein 1	NA
23643	LY96	lymphocyte antigen 96	NA
4318	MMP9	matrix metallopeptidase 9	NA
5004	ORM1	orosomucoid 1	NA
6283	S100A12	S100 calcium binding protein A12	NA
5144	PDE4D	phosphodiesterase 4D	204491_at
54741	LEPROT	leptin receptor overlapping transcript	202377_at
8607	RUVBL1	RuvB like AAA ATPase 1	201614_s_at
599	BCL2L2	BCL2 like 2	209311_at
79188	TMEM43	transmembrane protein 43	217795_s_at
7791	ZYX	zyxin	200808_s_at
9414	TJP2	tight junction protein 2	202085_at
100272147	CMC4	C-X9-C motif containing 4	210212_x_at
22878	TRAPPC8	trafficking protein particle complex 8	207305_s_at
54987	CZIB	CXXC motif containing zinc binding protein	203197_s_at
50848	F11R	F11 receptor	221664_s_at
10726	NUDC	nuclear distribution C, dynein complex regulator	201173_x_at
8945	BTRC	beta-transducin repeat containing E3 ubiquitin protein ligase	216091_s_at
5872	RAB13	RAB13, member RAS oncogene family	202252_at
6185	RPN2	ribophorin II	213399_x_at
4650	MYO9B	myosin IXB	217297_s_at
83442	SH3BGRL3	SH3 domain binding glutamate rich protein like 3	221269_s_at
4190	MDH1	malate dehydrogenase 1	200978_at
3735	KARS	lysyl-tRNA synthetase	200840_at
6233	RPS27A	ribosomal protein S27a	200017_at
10982	MAPRE2	microtubule associated protein RP/EB family member 2	213489_at
140885	SIRPA	signal regulatory protein alpha	202896_s_at
114885	OSBPL11	oxysterol binding protein like 11	218304_s_at
9368	SLC9A3R1	SLC9A3 regulator 1	201349_at
51108	METTL9	methyltransferase like 9	217868_s_at
79882	ZC3H14	zinc finger CCCH-type containing 14	213064_at
22796	COG2	component of oligomeric golgi complex 2	203073_at
80896	NPL	N-acetylneuraminate pyruvate lyase	221210_s_at
64747	MFSD1	major facilitator superfamily domain containing 1	218109_s_at
79887	PLBD1	phospholipase B domain containing 1	218454_at
2983	GUCY1B1	guanylate cyclase 1 soluble subunit beta 1	211555_s_at
10632	ATP5MG	ATP synthase membrane subunit g	210453_x_at
334	APLP2	amyloid beta precursor like protein 2	208248_x_at
6132	RPL8	ribosomal protein L8	200936_at

83989	FAM172A	family with sequence similarity 172 member A	212936_at
79090	TRAPPC6A	trafficking protein particle complex 6A	204985_s_at
4068	SH2D1A	SH2 domain containing 1A	211210_x_at
58497	PRUNE1	prune exopolyphosphatase 1	209599_s_at
829	CAPZA1	capping actin protein of muscle Z-line subunit alpha 1	208374_s_at
833	CARS	cysteinyl-tRNA synthetase	212971_at
9197	SLC33A1	solute carrier family 33 member 1	203164_at
54069	MIS18A	MIS18 kinetochore protein A	219004_s_at
1178	CLC	Charcot-Leyden crystal galectin	206207_at
977	CD151	CD151 molecule (Raph blood group)	204306_s_at
51465	UBE2J1	ubiquitin conjugating enzyme E2 J1	222435_s_at
23564	DDAH2	dimethylarginine dimethylaminohydrolase 2	202262_x_at
114769	CARD16	caspase recruitment domain family member 16	1552701_a_at
644	BLVRA	biliverdin reductase A	203773_x_at
283149	BCL9L	BCL9 like	227616_at
1995	ELAVL3	ELAV like RNA binding protein 3	227612_at
100507507	LOC100507507	uncharacterized LOC100507507	238016_s_at
6583	SLC22A4	solute carrier family 22 member 4	205896_at
10020	GNE	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	205042_at
5110	PCMT1	protein-L-isoaspartate (D-aspartate) O-methyltransferase	210156_s_at
2213	FCGR2B	Fc fragment of IgG receptor IIb	210889_s_at
23181	DIP2A	disco interacting protein 2 homolog A	1552677_a_at
857	CAV1	caveolin 1	203065_s_at
166378	SPATA5	spermatogenesis associated 5	242251_at
57507	ZNF608	zinc finger protein 608	232303_at
773	CACNA1A	calcium voltage-gated channel subunit alpha1 A	214933_at
4709	NDUFB3	NADH:ubiquinone oxidoreductase subunit B3	203371_s_at
834	CASP1	caspase 1	211366_x_at
23210	JMJD6	jumonji domain containing 6, arginine demethylase and lysine hydroxylase	212722_s_at
26289	AK5	adenylate kinase 5	219308_s_at
81606	LBH	LBH regulator of WNT signaling pathway	221011_s_at
51651	PTRH2	peptidyl-tRNA hydrolase 2	218732_at
2113	ETS1	ETS proto-oncogene 1, transcription factor	214447_at
896	CCND3	cyclin D3	1562028_at
25797	QPCT	glutaminyl-peptide cyclotransferase	NA
1052	CEBPD	CCAAT enhancer binding protein delta	NA
9846	GAB2	GRB2 associated binding protein 2	NA
55024	BANK1	B cell scaffold protein with ankyrin repeats 1	NA
6503	SLA	Src like adaptor	NA
11027	LILRA2	leukocyte immunoglobulin like receptor A2	NA
64386	MMP25	matrix metallopeptidase 25	NA
10409	BASP1	brain abundant membrane attached signal protein 1	NA
5836	PYGL	glycogen phosphorylase L	NA
4783	NFIL3	nuclear factor, interleukin 3 regulated	NA

4689	NCF4	neutrophil cytosolic factor 4	NA
10135	NAMPT	nicotinamide phosphoribosyltransferase	NA
894	CCND2	cyclin D2	NA
23569	PADI4	peptidyl arginine deiminase 4	NA
8837	CFLAR	CASP8 and FADD like apoptosis regulator	NA
3820	KLRB1	killer cell lectin like receptor B1	NA
9885	OSBPL2	oxysterol binding protein like 2	NA
8972	MGAM	maltase-glucoamylase	NA
2180	ACSL1	acyl-CoA synthetase long chain family member 1	NA
1933	EEF1B2	eukaryotic translation elongation factor 1 beta 2	NA
6222	RPS18	ribosomal protein S18	NA
79746	ECHDC3	enoyl-CoA hydratase domain containing 3	NA
57655	GRAMD1A	GRAM domain containing 1A	NA
9054	NFS1	NFS1 cysteine desulfurase	NA
706	TSPO	translocator protein	NA
4688	NCF2	neutrophil cytosolic factor 2	NA
54543	TOMM7	translocase of outer mitochondrial membrane 7	NA
3148	HMGB2	high mobility group box 2	NA
1831	TSC22D3	TSC22 domain family member 3	NA
9489	PGS1	phosphatidylglycerophosphate synthase 1	NA
2289	FKBP5	FKBP prolyl isomerase 5	NA
3921	RPSA	ribosomal protein SA	NA
11224	RPL35	ribosomal protein L35	NA
83862	TMEM120A	transmembrane protein 120A	NA
6688	SPI1	Spi-1 proto-oncogene	NA
4835	NQO2	N-ribosyldihydronicotinamide:quinone reductase 2	NA
26020	LRP10	LDL receptor related protein 10	NA
7850	IL1R2	interleukin 1 receptor type 2	NA
8291	DYSF	dysferlin	NA
6194	RPS6	ribosomal protein S6	NA
3094	HINT1	histidine triad nucleotide binding protein 1	NA
6253	RTN2	reticulon 2	NA
11213	IRAK3	interleukin 1 receptor associated kinase 3	NA
8673	VAMP8	vesicle associated membrane protein 8	NA
6155	RPL27	ribosomal protein L27	NA
2992	GYG1	glycogenin 1	NA
6232	RPS27	ribosomal protein S27	NA
53917	RAB24	RAB24, member RAS oncogene family	NA
6139	RPL17	ribosomal protein L17	NA
197259	MLKL	mixed lineage kinase domain like pseudokinase	NA
23531	MMD	monocyte to macrophage differentiation associated	NA
1043	CD52	CD52 molecule	NA
8809	IL18R1	interleukin 18 receptor 1	NA
10460	TACC3	transforming acidic coiled-coil containing protein 3	NA
160364	CLEC12A	C-type lectin domain family 12 member A	NA
6189	RPS3A	ribosomal protein S3A	NA

84525	HOPX	HOP homeobox	NA
6133	RPL9	ribosomal protein L9	NA
26118	WSB1	WD repeat and SOCS box containing 1	NA
170622	COMMD6	COMM domain containing 6	NA
147372	CCBE1	collagen and calcium binding EGF domains 1	NA
6208	RPS14	ribosomal protein S14	NA
10857	PGRMC1	progesterone receptor membrane component 1	NA
445329	SULT1A4	sulfotransferase family 1A member 4	NA
8460	TPST1	tyrosylprotein sulfotransferase 1	NA
6352	CCL5	C-C motif chemokine ligand 5	NA
963	CD53	CD53 molecule	NA
51065	RPS27L	ribosomal protein S27 like	NA
23406	COTL1	coactosin like F-actin binding protein 1	NA
51187	RSL24D1	ribosomal L24 domain containing 1	NA
2185	PTK2B	protein tyrosine kinase 2 beta	NA
5204	PFDN5	prefoldin subunit 5	NA
8527	DGKD	diacylglycerol kinase delta	NA
56729	RETN	resistin	NA
1350	COX7C	cytochrome c oxidase subunit 7C	NA
6218	RPS17	ribosomal protein S17	NA
3071	NCKAP1L	NCK associated protein 1 like	NA
2791	GNG11	G protein subunit gamma 11	NA
7133	TNFRSF1B	TNF receptor superfamily member 1B	NA
8807	IL18RAP	interleukin 18 receptor accessory protein	NA
6125	RPL5	ribosomal protein L5	NA
6210	RPS15A	ribosomal protein S15a	NA
387066	SNHG5	small nucleolar RNA host gene 5	NA
23764	MAFF	MAF bZIP transcription factor F	NA
4303	FOXO4	forkhead box O4	NA
7178	TPT1	tumor protein, translationally-controlled 1	NA
10628	TXNIP	thioredoxin interacting protein	NA
6230	RPS25	ribosomal protein S25	NA
1345	COX6C	cytochrome c oxidase subunit 6C	NA
10578	GNLY	granulysin	NA
2215	FCGR3B	Fc fragment of IgG receptor IIIb	NA
3133	HLA-E	major histocompatibility complex, class I, E	NA
8993	PGLYRP1	peptidoglycan recognition protein 1	NA
57107	PDSS2	decaprenyl diphosphate synthase subunit 2	NA
84901	NFATC2IP	nuclear factor of activated T cells 2 interacting protein	NA
6170	RPL39	ribosomal protein L39	NA
2352	FOLR3	folate receptor 3	NA
6165	RPL35A	ribosomal protein L35a	NA
93058	COQ10A	coenzyme Q10A	NA
1455	CSNK1G2	casein kinase 1 gamma 2	NA
6152	RPL24	ribosomal protein L24	NA
25793	FBXO7	F-box protein 7	NA

7388	UQCRH	ubiquinol-cytochrome c reductase hinge protein	NA
6286	S100P	S100 calcium binding protein P	NA
6160	RPL31	ribosomal protein L31	NA
253018	HCG27	HLA complex group 27	NA
339231	ARL16	ADP ribosylation factor like GTPase 16	NA
126364	LRRC25	leucine rich repeat containing 25	NA
4694	NDUFA1	NADH:ubiquinone oxidoreductase subunit A1	NA
8970	HIST1H2BJ	histone cluster 1 H2B family member j	NA
10410	IFITM3	interferon induced transmembrane protein 3	NA
6130	RPL7A	ribosomal protein L7a	NA
55454	CSGALNACT2	chondroitin sulfate N-acetylgalactosaminyltransferase 2	NA
283450	HECTD4	HECT domain E3 ubiquitin protein ligase 4	NA
2768	GNA12	G protein subunit alpha 12	NA
115207	KCTD12	potassium channel tetramerization domain containing 12	NA
28962	OSTM1	osteoclastogenesis associated transmembrane protein 1	NA
26781	SNORA67	small nucleolar RNA, H/ACA box 67	NA
57396	CLK4	CDC like kinase 4	NA
3039	HBA1	hemoglobin subunit alpha 1	NA
677777	SCARNA12	small Cajal body-specific RNA 12	NA
7097	TLR2	toll like receptor 2	NA
10766	TOB2	transducer of ERBB2, 2	NA
1460	CSNK2B	casein kinase 2 beta	NA
4205	MEF2A	myocyte enhancer factor 2A	NA
3717	JAK2	Janus kinase 2	NA
84164	ASCC2	activating signal cointegrator 1 complex subunit 2	NA
6374	CXCL5	C-X-C motif chemokine ligand 5	NA
79364	ZXDC	ZXD family zinc finger C	NA
26511	CHIC2	cysteine rich hydrophobic domain 2	NA
5763	PTMS	parathymosin	NA
5742	PTGS1	prostaglandin-endoperoxide synthase 1	NA
3267	AGFG1	ArfGAP with FG repeats 1	NA
285533	RNF175	ring finger protein 175	NA
619383	SCARNA9	small Cajal body-specific RNA 9	NA
7038	TG	thyroglobulin	NA
527	ATP6V0C	ATPase H+ transporting V0 subunit c	NA
57602	USP36	ubiquitin specific peptidase 36	NA
2171	FABP5	fatty acid binding protein 5	NA
400058	MKRN9P	makorin ring finger protein 9, pseudogene	NA
407037	MIR320A	microRNA 320a	NA
3050	HBZ	hemoglobin subunit zeta	NA
203228	C9orf72	C9orf72-SMCR8 complex subunit	NA
54997	TESC	tescalcin	NA
643036	SLED1	proteoglycan 3, pro eosinophil major basic protein 2 pseudogene	NA
28987	NOB1	NIN1 (RPN12) binding protein 1 homolog	NA

407008	MIR223	microRNA 223	NA
57690	TNRC6C	trinucleotide repeat containing adaptor 6C	NA
285172	FAM126B	family with sequence similarity 126 member B	NA
57610	RANBP10	RAN binding protein 10	NA
54880	BCOR	BCL6 corepressor	NA
23503	ZFYVE26	zinc finger FYVE-type containing 26	NA
55824	PAG1	phosphoprotein membrane anchor with glycosphingolipid microdomains 1	NA
79595	SAP130	Sin3A associated protein 130	NA
598	BCL2L1	BCL2 like 1	NA
342510	CD300E	CD300e molecule	NA
339487	ZBTB8OS	zinc finger and BTB domain containing 8 opposite strand	NA
5079	PAX5	paired box 5	NA
28996	HIPK2	homeodomain interacting protein kinase 2	NA
167227	DCP2	decapping mRNA 2	NA
51629	SLC25A39	solute carrier family 25 member 39	NA
171017	ZNF384	zinc finger protein 384	NA
6711	SPTBN1	spectrin beta, non-erythrocytic 1	NA
57605	PITPNM2	phosphatidylinositol transfer protein membrane associated 2	NA
55665	URGCP	upregulator of cell proliferation	NA
7917	BAG6	BCL2 associated athanogene 6	NA
3300	DNAJB2	DnaJ heat shock protein family (Hsp40) member B2	NA
6497	SKI	SKI proto-oncogene	NA
57532	NUFIP2	nuclear FMR1 interacting protein 2	NA
55858	TMEM165	transmembrane protein 165	NA
7779	SLC30A1	solute carrier family 30 member 1	NA
84232	MAF1	MAF1 homolog, negative regulator of RNA polymerase III	NA
9682	KDM4A	lysine demethylase 4A	NA
80306	MED28	mediator complex subunit 28	NA
55362	TMEM63B	transmembrane protein 63B	NA
25942	SIN3A	SIN3 transcription regulator family member A	NA
663	BNIP2	BCL2 interacting protein 2	NA
10644	IGF2BP2	insulin like growth factor 2 mRNA binding protein 2	NA
51351	ZNF117	zinc finger protein 117	NA
10333	TLR6	toll like receptor 6	NA
8874	ARHGEF7	Rho guanine nucleotide exchange factor 7	NA
2995	GYPC	glycophorin C (Gerbich blood group)	NA
6117	RPA1	replication protein A1	NA
9332	CD163	CD163 molecule	215049_x_at
945	CD33	CD33 molecule	206120_at
4048	LTA4H	leukotriene A4 hydrolase	208771_s_at
8522	GAS7	growth arrest specific 7	210872_x_at
2114	ETS2	ETS proto-oncogene 2, transcription factor	201328_at
55357	TBC1D2	TBC1 domain family member 2	222173_s_at

3597	IL13RA1	interleukin 13 receptor subunit alpha 1	211612_s_at
29992	PILRA	paired immunoglobin like type 2 receptor alpha	219788_at
948	CD36	CD36 molecule	206488_s_at
54504	CPVL	carboxypeptidase vitellogenic like	208146_s_at
23191	CYFIP1	cytoplasmic FMR1 interacting protein 1	208923_at
8321	FZD1	frizzled class receptor 1	204452_s_at
683	BST1	bone marrow stromal cell antigen 1	205715_at
965	CD58	CD58 molecule	216942_s_at
59342	SCPEP1	serine carboxypeptidase 1	218217_at
55905	RNF114	ring finger protein 114	200868_s_at
133	ADM	adrenomedullin	202912_at
953	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1	207691_x_at
4615	MYD88	MYD88 innate immune signal transduction adaptor	209124_at
1462	VCAN	versican	204620_s_at
6280	S100A9	S100 calcium binding protein A9	203535_at
22918	CD93	CD93 molecule	202878_s_at
4005	LMO2	LIM domain only 2	204249_s_at
7905	REEP5	receptor accessory protein 5	208872_s_at
1730	DIAPH2	diaphanous related formin 2	205603_s_at
1312	COMT	catechol-O-methyltransferase	208818_s_at
6038	RNASE4	ribonuclease A family member 4	205158_at
1495	CTNNA1	catenin alpha 1	200765_x_at
55711	FAR2	fatty acyl-CoA reductase 2	220615_s_at
11010	GLIPR1	GLI pathogenesis related 1	204222_s_at
929	CD14	CD14 molecule	201743_at
9555	H2AFY	H2A histone family member Y	207168_s_at
6604	SMARCD3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3	204099_at
2799	GNS	glucosamine (N-acetyl)-6-sulfatase	212335_at
11025	LILRB3	leukocyte immunoglobulin like receptor B3	211135_x_at
9517	SPTLC2	serine palmitoyltransferase long chain base subunit 2	203127_s_at
1843	DUSP1	dual specificity phosphatase 1	201041_s_at
1536	CYBB	cytochrome b-245 beta chain	203922_s_at
1508	CTSB	cathepsin B	200838_at
5223	PGAM1	phosphoglycerate mutase 1	200886_s_at
3958	LGALS3	galectin 3	208949_s_at
2896	GRN	granulin precursor	211284_s_at
9103	FCGR2C	Fc fragment of IgG receptor IIc (gene/pseudogene)	210992_x_at
4671	NAIP	NLR family apoptosis inhibitory protein	204860_s_at
2512	FTL	ferritin light chain	212788_x_at
23136	EPB41L3	erythrocyte membrane protein band 4.1 like 3	211776_s_at
29887	SNX10	sorting nexin 10	218404_at
2210	FCGR1B	Fc fragment of IgG receptor Ib	214511_x_at
8804	CREG1	cellular repressor of E1A stimulated genes 1	201200_at
2353	FOS	Fos proto-oncogene, AP-1 transcription factor subunit	209189_at

4668	NAGA	alpha-N-acetylgalactosaminidase	202943_s_at
5728	PTEN	phosphatase and tensin homolog	204053_x_at
1102	RCBTB2	RCC1 and BTB domain containing protein 2	204759_at
53826	FXYD6	FXYD domain containing ion transport regulator 6	217897_at
5873	RAB27A	RAB27A, member RAS oncogene family	210951_x_at
10404	CPQ	carboxypeptidase Q	208454_s_at
1806	DPYD	dihydropyrimidine dehydrogenase	204646_at
55625	ZDHHC7	zinc finger DHHC-type containing 7	218606_at
55640	FLVCR2	feline leukemia virus subgroup C cellular receptor family member 2	219316_s_at
4616	GADD45B	growth arrest and DNA damage inducible beta	207574 s at
6272	SORT1	sortilin 1	212807_s_at
5359	PLSCR1	phospholipid scramblase 1	202446 s at
2212	FCGR2A	Fc fragment of IgG receptor IIa	203561_at
11006	LILRB4	leukocyte immunoglobulin like receptor B4	210152 at
302	ANXA2	annexin A2	210427_x_at
1955	MEGF9	multiple EGF like domains 9	212830 at
3614	IMPDH1	inosine monophosphate dehydrogenase 1	204169_at
8741	TNFSF13	TNF superfamily member 13	209500 x at
407977	TNFSF12-TNFSF13	TNFSF12-TNFSF13 readthrough	209500_x_at
847	CAT	catalase	201432 at
10437	IFI30	IFI30 lysosomal thiol reductase	201422_at
5296	PIK3R2	phosphoinositide-3-kinase regulatory subunit 2	201422 at
3176	HNMT	histamine N-methyltransferase	204112_s_at
10129	FRY	FRY microtubule binding protein	214318 s at
9056	SLC7A7	solute carrier family 7 member 7	204588_s_at
55819	RNF130	ring finger protein 130	217865 at
51100	SH3GLB1	SH3 domain containing GRB2 like, endophilin B1	209091_s_at
7431	VIM	vimentin	201426 s at
3021	H3F3B	H3 histone family member 3B	209069_s_at
3020	H3F3A	H3 histone family member 3A	209069 s at
100616282	MIR4738	microRNA 4738	209069_s_at
6916	TBXAS1	thromboxane A synthase 1	208130 s at
81671	VMP1	vacuole membrane protein 1	220990_s_at
406991	MIR21	microRNA 21	220990 s at
604	BCL6	BCL6 transcription repressor	203140_at
433	ASGR2	asialoglycoprotein receptor 2	206130 s at
3034	HAL	histidine ammonia-lyase	217521_at
1051	CEBPB	CCAAT enhancer binding protein beta	212501 at
51164	DCTN4	dynactin subunit 4	218013_x_at
1810	DR1	down-regulator of transcription 1	209188 x at
5604	MAP2K1	mitogen-activated protein kinase kinase 1	202670_at
11191	PTENP1	phosphatase and tensin homolog pseudogene 1	217492 s at
9121	SLC16A5	solute carrier family 16 member 5	206600_s_at
23071	ERP44	endoplasmic reticulum protein 44	208959 s at
8650	NUMB	NUMB endocytic adaptor protein	209073_s_at
101928143	LOC101928143	uncharacterized LOC101928143	209073 s at

3084	NRG1	neuregulin 1	206237_s_at
8660	IRS2	insulin receptor substrate 2	209185_s_at
84669	USP32	ubiquitin specific peptidase 32	211702_s_at
1200	TPP1	tripeptidyl peptidase 1	200742_s_at
23433	RHOQ	ras homolog family member Q	214449_s_at
10875	FGL2	fibrinogen like 2	204834_at
4758	NEU1	neuraminidase 1	208926_at
3074	HEXB	hexosaminidase subunit beta	201944_at
3459	IFNGR1	interferon gamma receptor 1	202727_s_at
637	BID	BH3 interacting domain death agonist	204493_at
79868	ALG13	ALG13 UDP-N-acetylglucosaminyltransferase subunit	219015_s_at
4200	ME2	malic enzyme 2	209397_at
7763	ZFAND5	zinc finger AN1-type containing 5	217741_s_at
54491	OTULINL	OTU deubiquitinase with linear linkage specificity like	219694_at
47	ACLY	ATP citrate lyase	201127_s_at
3684	ITGAM	integrin subunit alpha M	205786_s_at
1992	SERPINB1	serpin family B member 1	212268_at
7867	MAPKAPK3	MAPK activated protein kinase 3	202787_s_at
7056	THBD	thrombomodulin	203888_at
9314	KLF4	Kruppel like factor 4	221841_s_at
4035	LRP1	LDL receptor related protein 1	200785_s_at
7077	TIMP2	TIMP metallopeptidase inhibitor 2	203167_at
3417	IDH1	isocitrate dehydrogenase (NADP(+)) 1, cytosolic	201193_at
3055	HCK	HCK proto-oncogene, Src family tyrosine kinase	208018_s_at
23312	DMXL2	Dmx like 2	212820_at
3052	HCCS	holocytochrome c synthase	203746_s_at
10370	CITED2	Cbp/p300 interacting transactivator with Glu/Asp rich carboxy-terminal domain 2	207980_s_at
6279	S100A8	S100 calcium binding protein A8	202917_s_at
5861	RAB1A	RAB1A, member RAS oncogene family	207791_s_at
55288	RHOT1	ras homolog family member T1	222148_s_at
3689	ITGB2	integrin subunit beta 2	202803_s_at
412	STS	steroid sulfatase	203767_s_at
23001	WDFY3	WD repeat and FYVE domain containing 3	212606_at
2153	F5	coagulation factor V	204714_s_at
7099	TLR4	toll like receptor 4	221060_s_at
3557	IL1RN	interleukin 1 receptor antagonist	216243_s_at
3613	IMPA2	inositol monophosphatase 2	203126_at
9473	THEMIS2	thymocyte selection associated family member 2	210785_s_at
2207	FCER1G	Fc fragment of IgE receptor Ig	204232_at
2752	GLUL	glutamate-ammonia ligase	200648_s_at
55332	DRAM1	DNA damage regulated autophagy modulator 1	218627_at
5271	SERPINB8	serpin family B member 8	206034_at
7462	LAT2	linker for activation of T cells family member 2	221581_s_at
9936	CD302	CD302 molecule	203799_at
4065	LY75	lymphocyte antigen 75	203799_at

100526664	LY75-CD302	LY75-CD302 readthrough	203799_at
3920	LAMP2	lysosomal associated membrane protein 2	203041_s_at
26234	FBXL5	F-box and leucine rich repeat protein 5	209004_s_at
58472	SQOR	sulfide quinone oxidoreductase	217995_at
55701	ARHGEF40	Rho guanine nucleotide exchange factor 40	220326_s_at
10859	LILRB1	leukocyte immunoglobulin like receptor B1	207104_x_at
79901	CYBRD1	cytochrome b reductase 1	217889_s_at
25840	METTL7A	methyltransferase like 7A	207761_s_at
2322	FLT3	fms related tyrosine kinase 3	206674_at
79627	OGFRL1	opioid growth factor receptor like 1	219582_at
11024	LILRA1	leukocyte immunoglobulin like receptor A1	207872_s_at
308	ANXA5	annexin A5	200782_at
307	ANXA4	annexin A4	201301_s_at
960	CD44	CD44 molecule (Indian blood group)	209835_x_at
8723	SNX4	sorting nexin 4	205329_s_at
1520	CTSS	cathepsin S	202902_s_at
221	ALDH3B1	aldehyde dehydrogenase 3 family member B1	205640_at
8819	SAP30	Sin3A associated protein 30	204900_x_at
831	CAST	calpastatin	208908_s_at
26292	MYCBP	MYC binding protein	203360_s_at
81025	GJA9	gap junction protein alpha 9	203360_s_at
100527950	GJA9-MYCBP	GJA9-MYCBP readthrough	203360_s_at
6036	RNASE2	ribonuclease A family member 2	206111_at
55379	LRRC59	leucine rich repeat containing 59	222231_s_at
1845	DUSP3	dual specificity phosphatase 3	201537_s_at
645	BLVRB	biliverdin reductase B	202201_at
1441	CSF3R	colony stimulating factor 3 receptor	203591_s_at
240	ALOX5	arachidonate 5-lipoxygenase	214366_s_at
51734	MSRB1	methionine sulfoxide reductase B1	217977_at
27351	DESI1	desumoylating isopeptidase 1	212527_at
1438	CSF2RA	colony stimulating factor 2 receptor alpha subunit	211286_x_at
54788	DNAJB12	DnaJ heat shock protein family (Hsp40) member B12	202867_s_at
9935	MAFB	MAF bZIP transcription factor B	218559_s_at
2209	FCGR1A	Fc fragment of IgG receptor Ia	216950_s_at
100132417	FCGR1CP	Fc fragment of IgG receptor Ic, pseudogene	216950_s_at
654816	NCF1B	neutrophil cytosolic factor 1B pseudogene	214084_x_at
653361	NCF1	neutrophil cytosolic factor 1	214084_x_at
55233	MOB1A	MOB kinase activator 1A	201298_s_at
1362	CPD	carboxypeptidase D	201940_at
8514	KCNAB2	potassium voltage-gated channel subfamily A regulatory beta subunit 2	211791_s_at
23307	FKBP15	FKBP prolyl isomerase 15	31826_at
7084	TK2	thymidine kinase 2	204227_s_at
2204	FCAR	Fc fragment of IgA receptor	207674_at
27180	SIGLEC9	sialic acid binding Ig like lectin 9	210569_s_at
6745	SSR1	signal sequence receptor subunit 1	200889_s_at

10288	LILRB2	leukocyte immunoglobulin like receptor B2	207697_x_at
5281	PIGF	phosphatidylinositol glycan anchor biosynthesis class F	212117_at
1124	CHN2	chimerin 2	213385_at
23673	STX12	syntaxin 12	212112_s_at
8500	PPFIA1	PTPRF interacting protein alpha 1	210235_s_at
57142	RTN4	reticulon 4	211509_s_at
1861	TOR1A	torsin family 1 member A	202349_at
8677	STX10	syntaxin 10	212625_at
11031	RAB31	RAB31, member RAS oncogene family	217764_s_at
728	C5AR1	complement C5a receptor 1	220088_at
51573	GDE1	glycerophosphodiester phosphodiesterase 1	202593_s_at
7832	BTG2	BTG anti-proliferation factor 2	201235_s_at
6556	SLC11A1	solute carrier family 11 member 1	217473_x_at
5594	MAPK1	mitogen-activated protein kinase 1	212271_at
25939	SAMHD1	SAM and HD domain containing deoxynucleoside triphosphate triphosphohydrolase 1	204502_at
1512	CTSH	cathepsin H	202295_s_at
55862	ECHDC1	ethylmalonyl-CoA decarboxylase 1	219974_x_at
10159	ATP6AP2	ATPase H+ transporting accessory protein 2	201444_s_at
6948	TCN2	transcobalamin 2	204043_at
7518	XRCC4	X-ray repair cross complementing 4	205071_x_at
163	AP2B1	adaptor related protein complex 2 subunit beta 1	200615_s_at
912	CD1D	CD1d molecule	205789_at
57178	ZMIZ1	zinc finger MIZ-type containing 1	212124_at
1545	CYP1B1	cytochrome P450 family 1 subfamily B member 1	202436_s_at
1317	SLC31A1	solute carrier family 31 member 1	203971_at
53346	TM6SF1	transmembrane 6 superfamily member 1	219892_at
79168	LILRA6	leukocyte immunoglobulin like receptor A6	208594_x_at
217	ALDH2	aldehyde dehydrogenase 2 family member	201425_at
51363	CHST15	carbohydrate sulfotransferase 15	203066_at
8932	MBD2	methyl-CpG binding domain protein 2	202484_s_at
7316	UBC	ubiquitin C	211296_x_at
8763	CD164	CD164 molecule	208653_s_at
5862	RAB2A	RAB2A, member RAS oncogene family	208734_x_at
51099	ABHD5	abhydrolase domain containing 5	218739_at
3964	LGALS8	galectin 8	208934_s_at
196	AHR	aryl hydrocarbon receptor	202820_at
6451	SH3BGRL	SH3 domain binding glutamate rich protein like	201311_s_at
3281	HSBP1	heat shock factor binding protein 1	200942_s_at
2773	GNAI3	G protein subunit alpha i3	201179_s_at
2242	FES	FES proto-oncogene, tyrosine kinase	205418_at
6453	ITSN1	intersectin 1	209297_at
10326	SIRPB1	signal regulatory protein beta 1	206934_at
10023	FRAT1	FRAT regulator of WNT signaling pathway 1	219889_at
7805	LAPTM5	lysosomal protein transmembrane 5	201720_s_at
5899	RALB	RAS like proto-oncogene B	202100_at

1066	CES1	carboxylesterase 1	209616_s_at
9770	RASSF2	Ras association domain family member 2	203185_at
56994	CHPT1	choline phosphotransferase 1	221675_s_at
4332	MNDA	myeloid cell nuclear differentiation antigen	204959_at
55	ACPP	acid phosphatase, prostate	204393_s_at
8935	SKAP2	src kinase associated phosphoprotein 2	216899_s_at
978	CDA	cytidine deaminase	205627_at
26509	MYOF	myoferlin	211864_s_at
10154	PLXNC1	plexin C1	213241_at
5558	PRIM2	DNA primase subunit 2	215708_s_at
366	AQP9	aquaporin 9	205568_at
10327	AKR1A1	aldo-keto reductase family 1 member A1	201900_s_at
114548	NLRP3	NLR family pyrin domain containing 3	216015_s_at
602	BCL3	BCL3 transcription coactivator	204908_s_at
102465879	MIR8085	microRNA 8085	204908_s_at
10261	IGSF6	immunoglobulin superfamily member 6	206420_at
9021	SOCS3	suppressor of cytokine signaling 3	206359_at
6768	ST14	suppression of tumorigenicity 14	216905_s_at
51397	COMMD10	COMM domain containing 10	218439_s_at
10287	RGS19	regulator of G protein signaling 19	204336_s_at
950	SCARB2	scavenger receptor class B member 2	201647_s_at
51313	GASK1B	golgi associated kinase 1B	219872_at
7422	VEGFA	vascular endothelial growth factor A	211527_x_at
9341	VAMP3	vesicle associated membrane protein 3	211749_s_at
64231	MS4A6A	membrane spanning 4-domains A6A	219666_at
23232	TBC1D12	TBC1 domain family member 12	221858_at
54502	RBM47	RNA binding motif protein 47	218035_s_at
2745	GLRX	glutaredoxin	209276_s_at
6515	SLC2A3	solute carrier family 2 member 3	202497_x_at
6303	SAT1	spermidine/spermine N1-acetyltransferase 1	213988_s_at
8754	ADAM9	ADAM metallopeptidase domain 9	202381_at
3101	HK3	hexokinase 3	205936_s_at
10602	CDC42EP3	CDC42 effector protein 3	209287_s_at
51136	RNFT1	ring finger protein, transmembrane 1	221194_s_at
653645	TBC1D3P1-DHX40P1	TBC1D3P1-DHX40P1 readthrough, transcribed pseudogene	221194_s_at
8724	SNX3	sorting nexin 3	210648_x_at
2219	FCN1	ficolin 1	205237_at
5937	RBMS1	RNA binding motif single stranded interacting protein 1	207266_x_at
64422	ATG3	autophagy related 3	221492_s_at
2710	GK	glycerol kinase	207387_s_at
7305	TYROBP	TYRO protein tyrosine kinase binding protein	204122_at
7439	BEST1	bestrophin 1	207671_s_at
822	CAPG	capping actin protein, gelsolin like	201850_at
3099	HK2	hexokinase 2	202934_at
2588	GALNS	galactosamine (N-acetyl)-6-sulfatase	206335_at

55704	CCDC88A	coiled-coil domain containing 88A	221078_s_at
4055	LTBR	lymphotoxin beta receptor	203005_at
8540	AGPS	alkylglycerone phosphate synthase	205401_at
64757	MAR1	mitochondrial amidoxime reducing component 1	218865_at
3726	JUNB	JunB proto-oncogene, AP-1 transcription factor subunit	201473_at
8778	SIGLEC5	sialic acid binding Ig like lectin 5	220000_at
5175	PECAM1	platelet and endothelial cell adhesion molecule 1	208983_s_at
55240	STEAP3	STEAP3 metalloreductase	218424_s_at
4043	LRPAP1	LDL receptor related protein associated protein 1	201186_at
6715	SRD5A1	steroid 5 alpha-reductase 1	210959_s_at
1794	DOCK2	dedicator of cytokinesis 2	213160_at
6888	TALDO1	transaldolase 1	201463_s_at
533	ATP6V0B	ATPase H+ transporting V0 subunit b	200078_s_at
7045	TGFBI	transforming growth factor beta induced	201506_at
23166	STAB1	stabilin 1	38487_at
51324	SPG21	SPG21 abhydrolase domain containing, maspardin	217827_s_at
3460	IFNGR2	interferon gamma receptor 2	201642_at
5925	RB1	RB transcriptional corepressor 1	211540_s_at
83464	APH1B	aph-1 homolog B, gamma-secretase subunit	221036_s_at
81542	TMX1	thioredoxin related transmembrane protein 1	208097_s_at
8933	RTL8C	retrotransposon Gag like 8C	201828_x_at
64759	TNS3	tensin 3	217853_at
10871	CD300C	CD300c molecule	207270_x_at
27036	SIGLEC7	sialic acid binding Ig like lectin 7	217159_x_at
199	AIF1	allograft inflammatory factor 1	209901_x_at
144195	SLC2A14	solute carrier family 2 member 14	216236_s_at
654817	NCF1C	neutrophil cytosolic factor 1C pseudogene	204961_s_at
4647	MYO7A	myosin VIIA	208189_s_at
51816	ADA2	adenosine deaminase 2	219505_at
4853	NOTCH2	notch receptor 2	202445_s_at
3903	LAIR1	leukocyte associated immunoglobulin like receptor 1	208071_s_at
51311	TLR8	toll like receptor 8	220832_at
219654	ZCCHC24	zinc finger CCHC-type containing 24	212419_at
10954	PDIA5	protein disulfide isomerase family A member 5	203857_s_at
102465667	MIR7110	microRNA 7110	203857_s_at
10226	PLIN3	perilipin 3	202122_s_at
56935	SMCO4	single-pass membrane protein with coiled-coil domains 4	219806_s_at
8875	VNN2	vanin 2	205922_at
266629	SEC14L3	SEC14 like lipid binding 3	216346_at
4084	MXD1	MAX dimerization protein 1	206877_at
643246	MAP1LC3B2	microtubule associated protein 1 light chain 3 beta 2	208785_s_at
81631	MAP1LC3B	microtubule associated protein 1 light chain 3 beta	208785_s_at
5184	PEPD	peptidase D	202108_at
1604	CD55	CD55 molecule (Cromer blood group)	201926_s_at

3295	HSD17B4	hydroxysteroid 17-beta dehydrogenase 4	201413_at
8676	STX11	syntaxin 11	210190_at
23250	ATP11A	ATPase phospholipid transporting 11A	215842_s_at
6692	SPINT1	serine peptidase inhibitor, Kunitz type 1	202826_at
4170	MCL1	MCL1 apoptosis regulator, BCL2 family member	200798_x_at
5341	PLEK	pleckstrin	203471_s_at
6868	ADAM17	ADAM metallopeptidase domain 17	213532_at
91860	CALML4	calmodulin like 4	221879_at
754	PTTG1IP	PTTG1 interacting protein	200677_at
5226	PGD	phosphogluconate dehydrogenase	201118_at
5781	PTPN11	protein tyrosine phosphatase non-receptor type 11	205868_s_at
10645	CAMKK2	calcium/calmodulin dependent protein kinase kinase 2	212252_at
378	ARF4	ADP ribosylation factor 4	201096_s_at
28231	SLCO4A1	solute carrier organic anion transporter family member 4A1	219911 s at
23601	CLEC5A	C-type lectin domain containing 5A	219890_at
10116	FEM1B	fem-1 homolog B	212374 at
50862	RNF141	ring finger protein 141	219104_at
2358	FPR2	formyl peptide receptor 2	210773 s at
51338	MS4A4A	membrane spanning 4-domains A4A	219607_s_at
5165	PDK3	pyruvate dehydrogenase kinase 3	206348 s at
284266	SIGLEC15	sialic acid binding Ig like lectin 15	215856_at
5230	PGK1	phosphoglycerate kinase 1	200737 at
2217	FCGRT	Fc fragment of IgG receptor and transporter	218831_s_at
26253	CLEC4E	C-type lectin domain family 4 member E	219859 at
23061	TBC1D9B	TBC1 domain family member 9B	212054_x_at
206358	SLC36A1	solute carrier family 36 member 1	213119 at
2357	FPR1	formyl peptide receptor 1	205119_s_at
2720	GLB1	galactosidase beta 1	201576 s at
643853	TMPPE	transmembrane protein with metallophosphoesterase domain	201576_s_at
10325	RRAGB	Ras related GTP binding B	205540_s_at
50856	CLEC4A	C-type lectin domain family 4 member A	221724 s at
23185	LARP4B	La ribonucleoprotein domain family member 4B	208952_s_at
5199	CFP	complement factor properdin	206380_s_at
9819	TSC22D2	TSC22 domain family member 2	210953_at
3936	LCP1	lymphocyte cytosolic protein 1	208885_at
8566	PDXK	pyridoxal kinase	202671_s_at
10237	SLC35B1	solute carrier family 35 member B1	202433 at
55293	UEVLD	UEV and lactate/malate dehyrogenase domains	220775_s_at
10295	BCKDK	branched chain keto acid dehydrogenase kinase	202030 at
313	AOAH	acyloxyacyl hydrolase	205639_at
10221	TRIB1	tribbles pseudokinase 1	202241_at
6386	SDCBP	syndecan binding protein	200958_s_at
7113	TMPRSS2	transmembrane serine protease 2	211689 s at
2787	GNG5	G protein subunit gamma 5	207157_s_at

2280	FKBP1A	FKBP prolyl isomerase 1A	210186_s_at
10197	PSME3	proteasome activator subunit 3	200987_x_at
821	CANX	calnexin	208853_s_at
23265	EXOC7	exocyst complex component 7	212026_s_at
7319	UBE2A	ubiquitin conjugating enzyme E2 A	201898_s_at
5058	PAK1	p21 (RAC1) activated kinase 1	209615_s_at
1378	CR1	complement C3b/C4b receptor 1 (Knops blood group)	208488_s_at
8773	SNAP23	synaptosome associated protein 23	209131_s_at
1432	MAPK14	mitogen-activated protein kinase 14	211087_x_at
6533	SLC6A6	solute carrier family 6 member 6	205920_at
51167	CYB5R4	cytochrome b5 reductase 4	219079_at
10935	PRDX3	peroxiredoxin 3	201619_at
3142	HLX	H2.0 like homeobox	214438_at
8897	MTMR3	myotubularin related protein 3	211507_s_at
55969	RAB5IF	RAB5 interacting factor	217835_x_at
100527943	TGIF2-RAB5IF	TGIF2-RAB5IF readthrough	217835_x_at
597	BCL2A1	BCL2 related protein A1	205681_at
5351	PLOD1	procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	200827_at
4802	NFYC	nuclear transcription factor Y subunit gamma	211797_s_at
571	BACH1	BTB domain and CNC homolog 1	204194_at
9375	TM9SF2	transmembrane 9 superfamily member 2	201078_at
3663	IRF5	interferon regulatory factor 5	205468_s_at
10972	TMED10	transmembrane p24 trafficking protein 10	200929_at
11026	LILRA3	leukocyte immunoglobulin like receptor A3	206881_s_at
51014	TMED7	transmembrane p24 trafficking protein 7	209404_s_at
100302736	TMED7-TICAM2	TMED7-TICAM2 readthrough	209404_s_at
4179	CD46	CD46 molecule	207549_x_at
54849	DEF8	differentially expressed in FDCP 8 homolog	219646_at
56606	SLC2A9	solute carrier family 2 member 9	219991_at
5147	PDE6D	phosphodiesterase 6D	216883_x_at
5265	SERPINA1	serpin family A member 1	202833_s_at
23288	IQCE	IQ motif containing E	217124_at
8202	NCOA3	nuclear receptor coactivator 3	209062_x_at
11167	FSTL1	follistatin like 1	208310_s_at
51622	CCZ1	CCZ1 homolog, vacuolar protein trafficking and biogenesis associated	208310_s_at
221960	CCZ1B	CCZ1 homolog B, vacuolar protein trafficking and biogenesis associated	208310_s_at
9114	ATP6V0D1	ATPase H+ transporting V0 subunit d1	212041_at
10015	PDCD6IP	programmed cell death 6 interacting protein	217746_s_at
79707	NOL9	nucleolar protein 9	218754_at
54470	ARMCX6	armadillo repeat containing X-linked 6	214749_s_at
9587	MAD2L1BP	MAD2L1 binding protein	203094_at
51742	ARID4B	AT-rich interaction domain 4B	221230_s_at
6196	RPS6KA2	ribosomal protein S6 kinase A2	212912_at
55922	NKRF	NFKB repressing factor	205004_at

26240	FAM50B	family with sequence similarity 50 member B	205775_at
348995	NUP43	nucleoporin 43	219007_at
10866	HCP5	HLA complex P5	206082_at
8519	IFITM1	interferon induced transmembrane protein 1	214022_s_at
8703	B4GALT3	beta-1,4-galactosyltransferase 3	210243_s_at
10465	PPIH	peptidylprolyl isomerase H	204228_at
23046	KIF21B	kinesin family member 21B	204411_at
9659	PDE4DIP	phosphodiesterase 4D interacting protein	213388_at
100996724	LOC100996724	phosphodiesterase 4D interacting protein-like	213388_at
9093	DNAJA3	DnaJ heat shock protein family (Hsp40) member A3	205963_s_at
55341	LSG1	large 60S subunit nuclear export GTPase 1	221535_at
5437	POLR2H	RNA polymerase II subunit H	209302_at
9683	N4BP1	NEDD4 binding protein 1	221867_at
4798	NFRKB	nuclear factor related to kappaB binding protein	213028_at
321	APBA2	amyloid beta precursor protein binding family A member 2	209870_s_at
8718	TNFRSF25	TNF receptor superfamily member 25	210847_x_at
55290	BRF2	BRF2 RNA polymerase III transcription initiation factor subunit	218955_at
11276	SYNRG	synergin gamma	64418_at
23354	HAUS5	HAUS augmin like complex subunit 5	36888_at
50615	IL21R	interleukin 21 receptor	219971_at
100293516	ZNF587B	zinc finger protein 587B	221963_x_at
23	ABCF1	ATP binding cassette subfamily F member 1	200045_at
28990	ASTE1	asteroid homolog 1	221135_s_at
27079	RPUSD2	RNA pseudouridine synthase domain containing 2	221940_at
9882	TBC1D4	TBC1 domain family member 4	203386_at
23338	JADE2	jade family PHD finger 2	212660_at
7695	ZNF136	zinc finger protein 136	206240_s_at
5810	RAD1	RAD1 checkpoint DNA exonuclease	204461_x_at
84861	KLHL22	kelch like family member 22	49329_at
1616	DAXX	death domain associated protein	201763_s_at
8629	JRK	Jrk helix-turn-helix protein	216309_x_at
9295	SRSF11	serine and arginine rich splicing factor 11	213742_at
6932	TCF7	transcription factor 7	205255_x_at
7775	ZNF232	zinc finger protein 232	219123_at
360	AQP3	aquaporin 3 (Gill blood group)	39248_at
6137	RPL13	ribosomal protein L13	214351_x_at
606500	SNORD68	small nucleolar RNA, C/D box 68	214351_x_at
100101267	POM121C	POM121 transmembrane nucleoporin C	213360_s_at
9883	POM121	POM121 transmembrane nucleoporin	213360_s_at
919	CD247	CD247 molecule	210031_at
5287	PIK3C2B	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta	204484_at
4848	CNOT2	CCR4-NOT transcription complex subunit 2	217798_at
6749	SSRP1	structure specific recognition protein 1	200957_s_at
9831	ZNF623	zinc finger protein 623	206188_at

55031	USP47	ubiquitin specific peptidase 47	221518_s_at
3134	HLA-F	major histocompatibility complex, class I, F	221978_at
51337	THEM6	thioesterase superfamily member 6	218500_at
79673	ZNF329	zinc finger protein 329	219765_at
4050	LTB	lymphotoxin beta	207339_s_at
246243	RNASEH1	ribonuclease H1	218496_at
3932	LCK	LCK proto-oncogene, Src family tyrosine kinase	204891_s_at
7014	TERF2	telomeric repeat binding factor 2	203611_at
388650	DIPK1A	divergent protein kinase domain 1A	213689_x_at
8567	MADD	MAP kinase activating death domain	38398_at
90864	SPSB3	splA/ryanodine receptor domain and SOCS box containing 3	46256_at
5252	PHF1	PHD finger protein 1	40446 at
64710	NUCKS1	nuclear casein kinase and cyclin dependent kinase substrate 1	217802_s_at
28986	MAGEH1	MAGE family member H1	218573_at
83480	PUS3	pseudouridine synthase 3	221277_s_at
23099	ZBTB43	zinc finger and BTB domain containing 43	204182_s_at
23301	EHBP1	EH domain binding protein 1	212653_s_at
740	MRPL49	mitochondrial ribosomal protein L49	201717_at
8934	RAB29	RAB29, member RAS oncogene family	218700_s_at
51070	NOSIP	nitric oxide synthase interacting protein	217950_at
471	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	208758 at
4204	MECP2	methyl-CpG binding protein 2	202617_s_at
23263	MCF2L	MCF.2 cell line derived transforming sequence like	212935 at
29127	RACGAP1	Rac GTPase activating protein 1	222077_s_at
80833	APOL3	apolipoprotein L3	221087 s at
7374	UNG	uracil DNA glycosylase	202330_s_at
940	CD28	CD28 molecule	206545 at
54820	NDE1	nudE neurodevelopment protein 1	218414_s_at
5631	PRPS1	phosphoribosyl pyrophosphate synthetase 1	209440 at
54971	BANP	BTG3 associated nuclear protein	219966_x_at
51710	ZNF44	zinc finger protein 44	215359 x at
26036	ZNF451	zinc finger protein 451	215012_at
9020	MAP3K14	mitogen-activated protein kinase kinase kinase 14	205192 at
6775	STAT4	signal transducer and activator of transcription 4	206118_at
3837	KPNB1	karyopherin subunit beta 1	213574 s at
65108	MARCKSL1	MARCKS like 1	200644_at
55139	ANKZF1	ankyrin repeat and zinc finger domain containing 1	218274 s at
5411	PNN	pinin, desmosome associated protein	212037_at
3707	ITPKB	inositol-trisphosphate 3-kinase B	203723 at
8445	DYRK2	dual specificity tyrosine phosphorylation regulated kinase 2	202970 at
51106	TFB1M	transcription factor B1, mitochondrial	219169_s_at
11161	ERG28	ergosterol biosynthesis 28 homolog	202562_s_at
23016	EXOSC7	exosome component 7	213648_at

9848	MFAP3L	microfibril associated protein 3 like	205442_at
79810	PTCD2	pentatricopeptide repeat domain 2	219658_at
126231	ZNF573	zinc finger protein 573	217627_at
10126	DNAL4	dynein axonemal light chain 4	204008_at
25776	CBY1	chibby family member 1, beta catenin antagonist	203450_at
79037	PVRIG	PVR related immunoglobulin domain containing	219812_at
79582	SPAG16	sperm associated antigen 16	219109_at
8315	BRAP	BRCA1 associated protein	213473_at
375449	MAST4	microtubule associated serine/threonine kinase family member 4	40016_g_at
22882	ZHX2	zinc fingers and homeoboxes 2	203556_at
4863	NPAT	nuclear protein, coactivator of histone transcription	209798_at
80095	ZNF606	zinc finger protein 606	219635_at
22800	RRAS2	RAS related 2	212589_at
7753	ZNF202	zinc finger protein 202	204327_s_at
60436	TGIF2	TGFB induced factor homeobox 2	216262_s_at
23225	NUP210	nucleoporin 210	220035_at
23370	ARHGEF18	Rho/Rac guanine nucleotide exchange factor 18	213039_at
23344	ESYT1	extended synaptotagmin 1	208858_s_at
55340	GIMAP5	GTPase, IMAP family member 5	218805_at
100527949	GIMAP1-GIMAP5	GIMAP1-GIMAP5 readthrough	218805_at
9026	HIP1R	huntingtin interacting protein 1 related	209558_s_at
7696	ZNF137P	zinc finger protein 137, pseudogene	207394_at
53347	UBASH3A	ubiquitin associated and SH3 domain containing A	220418_at
26207	PITPNC1	phosphatidylinositol transfer protein cytoplasmic 1	219155_at
84859	LRCH3	leucine rich repeats and calponin homology domain containing 3	214739_at
29078	NDUFAF4	NADH:ubiquinone oxidoreductase complex assembly factor 4	219006_at
3298	HSF2	heat shock transcription factor 2	209657_s_at
79088	ZNF426	zinc finger protein 426	205964_at
9252	RPS6KA5	ribosomal protein S6 kinase A5	204635_at
203068	TUBB	tubulin beta class I	212320_at
81576	CCDC130	coiled-coil domain containing 130	208094_s_at
112869	SGF29	SAGA complex associated factor 29	48117_at
80764	THAP7	THAP domain containing 7	218492_s_at
54509	RHOF	ras homolog family member F, filopodia associated	219045_at
2744	GLS	glutaminase	203159_at
57125	PLXDC1	plexin domain containing 1	219700_at
923	CD6	CD6 molecule	213958_at
259197	NCR3	natural cytotoxicity triggering receptor 3	210763_x_at
274	BIN1	bridging integrator 1	202931_x_at
5567	PRKACB	protein kinase cAMP-activated catalytic subunit beta	202741_at
57496	MRTFB	myocardin related transcription factor B	218259_at
2308	FOXO1	forkhead box O1	202724_s_at
64118	DUS1L	dihydrouridine synthase 1 like	217912_at
79652	TMEM204	transmembrane protein 204	219315_s_at

54463	RETREG1	reticulophagy regulator 1	218510_x_at
7767	ZNF224	zinc finger protein 224	216983_s_at
27300	ZNF544	zinc finger protein 544	218735_s_at
9125	CNOT9	CCR4-NOT transcription complex subunit 9	213179_at
5813	PURA	purine rich element binding protein A	204020_at
9527	GOSR1	golgi SNAP receptor complex member 1	204630_s_at
994	CDC25B	cell division cycle 25B	201853_s_at
23164	MPRIP	myosin phosphatase Rho interacting protein	214771_x_at
915	CD3D	CD3d molecule	213539_at
9263	STK17A	serine/threonine kinase 17a	202693_s_at
10125	RASGRP1	RAS guanyl releasing protein 1	205590_at
28639	TRBC1	T cell receptor beta constant 1	213193_x_at
80224	NUBPL	nucleotide binding protein like	220176_at
10061	ABCF2	ATP binding cassette subfamily F member 2	209246_at
50717	DCAF8	DDB1 and CUL4 associated factor 8	202250_s_at
330	BIRC3	baculoviral IAP repeat containing 3	210538_s_at
56616	DIABLO	diablo IAP-binding mitochondrial protein	219350_s_at
9500	MAGED1	MAGE family member D1	209014_at
1901	S1PR1	sphingosine-1-phosphate receptor 1	204642_at
959	CD40LG	CD40 ligand	207892_at
29105	CFAP20	cilia and flagella associated protein 20	217957_at
25940	FAM98A	family with sequence similarity 98 member A	212333_at
5590	PRKCZ	protein kinase C zeta	202178_at
2841	GPR18	G protein-coupled receptor 18	210279_at
3978	LIG1	DNA ligase 1	202726_at
23157	SEPTIN6	septin 6	214298_x_at
7292	TNFSF4	TNF superfamily member 4	207426_s_at
23468	CBX5	chromobox 5	212126_at
939	CD27	CD27 molecule	206150_at
25865	PRKD2	protein kinase D2	209282_at
91782	CHMP7	charged multivesicular body protein 7	212313_at
874	CBR3	carbonyl reductase 3	205379_at
54977	SLC25A38	solute carrier family 25 member 38	217961_at
3652	IPP	intracisternal A particle-promoted polypeptide	219843_at
79744	ZNF419	zinc finger protein 419	219826_at
868	CBLB	Cbl proto-oncogene B	209682_at
26119	LDLRAP1	low density lipoprotein receptor adaptor protein 1	221790_s_at
6304	SATB1	SATB homeobox 1	203408_s_at
51174	TUBD1	tubulin delta 1	210389_x_at
9214	FCMR	Fc fragment of IgM receptor	221601_s_at
28951	TRIB2	tribbles pseudokinase 2	202478_at
9814	SFI1	SFI1 centrin binding protein	36545_s_at
26018	LRIG1	leucine rich repeats and immunoglobulin like domains 1	211596_s_at
155066	ATP6V0E2	ATPase H+ transporting V0 subunit e2	213587_s_at
1803	DPP4	dipeptidyl peptidase 4	203717_at
9797	TATDN2	TatD DNase domain containing 2	203648_at

28984	RGCC	regulator of cell cycle	218723_s_at
4171	MCM2	minichromosome maintenance complex component 2	202107_s_at
2145	EZH1	enhancer of zeste 1 polycomb repressive complex 2 subunit	32259_at
56911	MAP3K7CL	MAP3K7 C-terminal like	221211_s_at
2026	ENO2	enolase 2	201313_at
60468	BACH2	BTB domain and CNC homolog 2	221234_s_at
54910	SEMA4C	semaphorin 4C	46665_at
51611	DPH5	diphthamide biosynthesis 5	219590_x_at
10868	USP20	ubiquitin specific peptidase 20	203965_at
5051	PAFAH2	platelet activating factor acetylhydrolase 2	205233_s_at
27040	LAT	linker for activation of T cells	209881_s_at
4176	MCM7	minichromosome maintenance complex component 7	208795_s_at
80264	ZNF430	zinc finger protein 430	206829_x_at
26999	CYFIP2	cytoplasmic FMR1 interacting protein 2	215785_s_at
57326	PBXIP1	PBX homeobox interacting protein 1	214177_s_at
5883	RAD9A	RAD9 checkpoint clamp component A	204828_at
135	ADORA2A	adenosine A2a receptor	205013_s_at
101730217	SPECC1L-ADORA2A	SPECC1L-ADORA2A readthrough (NMD candidate)	205013_s_at
2189	FANCG	FA complementation group G	203564_at
9249	DHRS3	dehydrogenase/reductase 3	202481_at
23403	FBXO46	F-box protein 46	205310_at
6709	SPTAN1	spectrin alpha, non-erythrocytic 1	215235_at
22880	MORC2	MORC family CW-type zinc finger 2	203956_at
9725	TMEM63A	transmembrane protein 63A	214833_at
5001	ORC5	origin recognition complex subunit 5	204957_at
84656	GLYR1	glyoxylate reductase 1 homolog	212414_s_at
6526	SLC5A3	solute carrier family 5 member 3	213164_at
28755	TRAC	T cell receptor alpha constant	209670_at
22954	TRIM32	tripartite motif containing 32	203846_at
3983	ABLIM1	actin binding LIM protein 1	200965_s_at
9363	RAB33A	RAB33A, member RAS oncogene family	206039_at
56941	HMCES	5-hydroxymethylcytosine binding, ES cell specific	201677_at
9656	MDC1	mediator of DNA damage checkpoint 1	203062_s_at
9666	DZIP3	DAZ interacting zinc finger protein 3	207231_at
54900	LAX1	lymphocyte transmembrane adaptor 1	207734_at
7994	KAT6A	lysine acetyltransferase 6A	202423_at
4175	MCM6	minichromosome maintenance complex component 6	201930_at
92340	PRR29	proline rich 29	213620_s_at
3384	ICAM2	intercellular adhesion molecule 2	213620_s_at
2625	GATA3	GATA binding protein 3	209603_at
56257	MEPCE	methylphosphate capping enzyme	219798_s_at
5588	PRKCQ	protein kinase C theta	210038_at
3429	IFI27	interferon alpha inducible protein 27	NA

249	ALPL	alkaline phosphatase, biomineralization associated	NA
5551	PRF1	perforin 1	NA
925	CD8A	CD8a molecule	NA
53637	S1PR5	sphingosine-1-phosphate receptor 5	NA
3001	GZMA	granzyme A	NA
3560	IL2RB	interleukin 2 receptor subunit beta	NA
2999	GZMH	granzyme H	NA
4818	NKG7	natural killer cell granule protein 7	NA
83888	FGFBP2	fibroblast growth factor binding protein 2	NA
8320	EOMES	eomesodermin	NA
1524	CX3CR1	C-X3-C motif chemokine receptor 1	NA
3008	HIST1H1E	histone cluster 1 H1 family member e	NA
858	CAV2	caveolin 2	NA
100130253	RCBTB2P1	RCC1 and BTB domain containing protein 2 pseudogene 1	NA
147495	APCDD1	APC down-regulated 1	NA
2119	ETV5	ETS variant 5	NA
728921	HAUS1P1	HAUS augmin like complex subunit 1 pseudogene 1	NA
11144	DMC1	DNA meiotic recombinase 1	NA
170261	ZCCHC12	zinc finger CCHC-type containing 12	NA
247	ALOX15B	arachidonate 15-lipoxygenase, type B	NA
4661	MYT1	myelin transcription factor 1	NA
6422	SFRP1	secreted frizzled related protein 1	NA
406992	MIR210	microRNA 210	NA
164714	TTLL8	tubulin tyrosine ligase like 8	NA
10457	GPNMB	glycoprotein nmb	NA
83999	KREMEN1	kringle containing transmembrane protein 1	NA
3242	HPD	4-hydroxyphenylpyruvate dioxygenase	NA
6383	SDC2	syndecan 2	NA
7053	TGM3	transglutaminase 3	NA
147719	LYPD4	LY6/PLAUR domain containing 4	NA
4233	MET	MET proto-oncogene, receptor tyrosine kinase	NA
64359	NXN	nucleoredoxin	NA
106480770	RNA5SP498	RNA, 5S ribosomal pseudogene 498	NA
6853	SYN1	synapsin I	NA
29122	PRSS50	serine protease 50	NA
440073	IQSEC3	IQ motif and Sec7 domain 3	NA
374907	B3GNT8	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 8	NA
387747	OR52B3P	olfactory receptor family 52 subfamily B member 3 pseudogene	NA
89886	SLAMF9	SLAM family member 9	NA
9052	GPRC5A	G protein-coupled receptor class C group 5 member A	NA
220108	FAM124A	family with sequence similarity 124 member A	NA
342035	GLDN	gliomedin	NA
152831	KLB	klotho beta	NA

129881	CCDC173	coiled-coil domain containing 173	NA
439996	IFIT1B	interferon induced protein with tetratricopeptide repeats 1B	NA
8367	HIST1H4E	histone cluster 1 H4 family member e	NA
57619	SHROOM3	shroom family member 3	NA
3036	HAS1	hyaluronan synthase 1	NA
106480276	UBE2Q2P9	ubiquitin conjugating enzyme E2 Q2 pseudogene 9	NA
51334	PRR16	proline rich 16	NA
390940	PINLYP	phospholipase A2 inhibitor and LY6/PLAUR domain containing	NA
387700	SLC16A12	solute carrier family 16 member 12	NA
128676	RPS18P1	ribosomal protein S18 pseudogene 1	NA
221438	TREML5P	triggering receptor expressed on myeloid cells like 5, pseudogene	NA
79411	GLB1L	galactosidase beta 1 like	NA
100287495	GCSHP4	glycine cleavage system protein H pseudogene 4	NA
438	ASMT	acetylserotonin O-methyltransferase	NA
2865	FFAR3	free fatty acid receptor 3	NA
79930	DOK3	docking protein 3	NA
126003	TRAPPC5	trafficking protein particle complex 5	NA
7057	THBS1	thrombospondin 1	NA
6405	SEMA3F	semaphorin 3F	NA
2501	FTH1P8	ferritin heavy chain 1 pseudogene 8	NA
2866	GPR42	G protein-coupled receptor 42 (gene/pseudogene)	NA
273	AMPH	amphiphysin	NA
54550	NECAB2	N-terminal EF-hand calcium binding protein 2	NA
340146	SLC35D3	solute carrier family 35 member D3	NA
910	CD1B	CD1b molecule	NA
343702	XKR7	XK related 7	NA
100129827	MRVI1-AS1	MRVI1 antisense RNA 1	NA
1588	CYP19A1	cytochrome P450 family 19 subfamily A member 1	NA
26873	OPLAH	5-oxoprolinase, ATP-hydrolysing	NA
3397	ID1	inhibitor of DNA binding 1, HLH protein	NA
51330	TNFRSF12A	TNF receptor superfamily member 12A	NA
284958	NT5DC4	5'-nucleotidase domain containing 4	NA
26526	TSPAN16	tetraspanin 16	NA
9536	PTGES	prostaglandin E synthase	NA
1258	CNGB1	cyclic nucleotide gated channel beta 1	NA
80270	HSD3B7	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7	NA
100422691	ITGB1P1	integrin subunit beta 1 pseudogene 1	NA
1593	CYP27A1	cytochrome P450 family 27 subfamily A member 1	NA
4327	MMP19	matrix metallopeptidase 19	NA
55350	VNN3	vanin 3	NA
5909	RAP1GAP	RAP1 GTPase activating protein	NA
57817	HAMP	hepcidin antimicrobial peptide	NA
348013	TMEM255B	transmembrane protein 255B	NA

943	TNFRSF8	TNF receptor superfamily member 8	NA
55365	TMEM176A	transmembrane protein 176A	NA
1669	DEFA4	defensin alpha 4	NA
5328	PLAU	plasminogen activator, urokinase	NA
65124	SOWAHC	sosondowah ankyrin repeat domain family member C	NA
389170	LEKR1	leucine, glutamate and lysine rich 1	NA
54567	DLL4	delta like canonical Notch ligand 4	NA
3625	INHBB	inhibin subunit beta B	NA
106479075	RNU7-133P	RNA, U7 small nuclear 133 pseudogene	NA
4254	KITLG	KIT ligand	NA
6354	CCL7	C-C motif chemokine ligand 7	NA
106479034	LARP1BP2	La ribonucleoprotein domain family member 1B pseudogene 2	NA
374897	SBSN	suprabasin	NA
2810	SFN	stratifin	NA
153643	FAM81B	family with sequence similarity 81 member B	NA
100873694	RNA5SP449	RNA, 5S ribosomal pseudogene 449	NA
144423	GLT1D1	glycosyltransferase 1 domain containing 1	NA
151579	BZW1P2	basic leucine zipper and W2 domains 1 pseudogene 2	NA
91947	ARRDC4	arrestin domain containing 4	NA
6590	SLPI	secretory leukocyte peptidase inhibitor	NA
25822	DNAJB5	DnaJ heat shock protein family (Hsp40) member B5	NA
1903	S1PR3	sphingosine-1-phosphate receptor 3	NA
171177	RHOV	ras homolog family member V	NA
693201	MIR616	microRNA 616	NA
2258	FGF13	fibroblast growth factor 13	NA
5003	SLC22A18AS	solute carrier family 22 member 18 antisense	NA
80243	PREX2	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 2	NA
4608	MYBPH	myosin binding protein H	NA
3898	LAD1	ladinin 1	NA
53632	PRKAG3	protein kinase AMP-activated non-catalytic subunit gamma 3	NA
340206	TREML3P	triggering receptor expressed on myeloid cells like 3, pseudogene	NA
106479820	RNU6-574P	RNA, U6 small nuclear 574, pseudogene	NA
4923	NTSR1	neurotensin receptor 1	NA
3162	HMOX1	heme oxygenase 1	NA
55509	BATF3	basic leucine zipper ATF-like transcription factor 3	NA
7869	SEMA3B	semaphorin 3B	NA
6262	RYR2	ryanodine receptor 2	NA
1299	COL9A3	collagen type IX alpha 3 chain	NA
9970	NR1I3	nuclear receptor subfamily 1 group I member 3	NA
875	CBS	cystathionine-beta-synthase	NA
256355	RPS2P32	ribosomal protein S2 pseudogene 32	NA
6406	SEMG1	semenogelin 1	NA

83716	CRISPLD2	cysteine rich secretory protein LCCL domain containing 2	NA
3127	HLA-DRB5	major histocompatibility complex, class II, DR beta 5	NA
3240	HP	haptoglobin	NA
260429	PRSS33	serine protease 33	NA
100500894	MIR3690	microRNA 3690	NA
150142	ZNF295-AS1	ZNF295 antisense RNA 1	NA
79729	SH3D21	SH3 domain containing 21	NA
9609	RAB36	RAB36, member RAS oncogene family	NA
22895	RPH3A	rabphilin 3A	NA
57761	TRIB3	tribbles pseudokinase 3	NA
10272	FSTL3	follistatin like 3	NA
3309	HSPA5	heat shock protein family A (Hsp70) member 5	NA
1308	COL17A1	collagen type XVII alpha 1 chain	NA
9732	DOCK4	dedicator of cytokinesis 4	NA
10008	KCNE3	potassium voltage-gated channel subfamily E regulatory subunit 3	NA
1668	DEFA3	defensin alpha 3	NA
28959	TMEM176B	transmembrane protein 176B	NA
7136	TNNI2	troponin I2, fast skeletal type	NA
11326	VSIG4	V-set and immunoglobulin domain containing 4	NA
338442	HCAR2	hydroxycarboxylic acid receptor 2	NA
100506681	JARID2-AS1	JARID2 antisense RNA 1	NA
729359	PLIN4	perilipin 4	NA
54757	FAM20A	FAM20A, golgi associated secretory pathway pseudokinase	NA
8061	FOSL1	FOS like 1, AP-1 transcription factor subunit	NA
201191	SAMD14	sterile alpha motif domain containing 14	NA
100847046	MIR5010	microRNA 5010	NA
84057	MND1	meiotic nuclear divisions 1	NA
3816	KLK1	kallikrein 1	NA
91050	CCDC149	coiled-coil domain containing 149	NA
728358	DEFA1B	defensin alpha 1B	NA
375686	SPATC1	spermatogenesis and centriole associated 1	NA
29948	OSGIN1	oxidative stress induced growth inhibitor 1	NA
106481850	RN7SL566P	RNA, 7SL, cytoplasmic 566, pseudogene	NA
58191	CXCL16	C-X-C motif chemokine ligand 16	NA
64855	FAM129B	family with sequence similarity 129 member B	NA
100506413	DOCK4-AS1	DOCK4 antisense RNA 1	NA
6690	SPINK1	serine peptidase inhibitor, Kazal type 1	NA
7288	TULP2	tubby like protein 2	NA
57617	VPS18	VPS18, CORVET/HOPS core subunit	NA
51266	CLEC1B	C-type lectin domain family 1 member B	NA
196383	RILPL2	Rab interacting lysosomal protein like 2	NA
290	ANPEP	alanyl aminopeptidase, membrane	NA
3586	IL10	interleukin 10	NA
5029	P2RY2	purinergic receptor P2Y2	NA

84649	DGAT2	diacylglycerol O-acyltransferase 2	NA
9454	HOMER3	homer scaffold protein 3	NA
1809	DPYSL3	dihydropyrimidinase like 3	NA
9322	TRIP10	thyroid hormone receptor interactor 10	NA
64109	CRLF2	cytokine receptor like factor 2	NA
8338	HIST2H2AC	histone cluster 2 H2A family member c	NA
92737	DNER	delta/notch like EGF repeat containing	NA
7076	TIMP1	TIMP metallopeptidase inhibitor 1	NA
2354	FOSB	FosB proto-oncogene, AP-1 transcription factor subunit	NA
10268	RAMP3	receptor activity modifying protein 3	NA
3202	HOXA5	homeobox A5	NA
23753	SDF2L1	stromal cell derived factor 2 like 1	NA
2331	FMOD	fibromodulin	NA
7718	ZNF165	zinc finger protein 165	NA
54512	EXOSC4	exosome component 4	NA
9771	RAPGEF5	Rap guanine nucleotide exchange factor 5	NA
5768	QSOX1	quiescin sulfhydryl oxidase 1	NA
339855	KY	kyphoscoliosis peptidase	NA
127550	A3GALT2	alpha 1,3-galactosyltransferase 2	NA
780851	SNORD3A	small nucleolar RNA, C/D box 3A	NA
1428	CRYM	crystallin mu	NA
2907	GRINA	glutamate ionotropic receptor NMDA type subunit associated protein 1	NA
3291	HSD11B2	hydroxysteroid 11-beta dehydrogenase 2	NA
241	ALOX5AP	arachidonate 5-lipoxygenase activating protein	NA
65268	WNK2	WNK lysine deficient protein kinase 2	NA
9468	PCYT1B	phosphate cytidylyltransferase 1, choline, beta	NA
440503	PLIN5	perilipin 5	NA
5210	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	NA
27242	TNFRSF21	TNF receptor superfamily member 21	NA
6580	SLC22A1	solute carrier family 22 member 1	NA
56666	PANX2	pannexin 2	NA
84134	TOMM40L	translocase of outer mitochondrial membrane 40 like	NA
89795	NAV3	neuron navigator 3	NA
84617	TUBB6	tubulin beta 6 class V	NA
389015	SLC9A4	solute carrier family 9 member A4	NA
154664	ABCA13	ATP binding cassette subfamily A member 13	NA
171389	NLRP6	NLR family pyrin domain containing 6	NA
10804	GJB6	gap junction protein beta 6	NA
64090	GAL3ST2	galactose-3-O-sulfotransferase 2	NA
122953	JDP2	Jun dimerization protein 2	NA
130951	M1AP	meiosis 1 associated protein	NA
84225	ZMYND15	zinc finger MYND-type containing 15	NA
3732	CD82	CD82 molecule	NA
84689	MS4A14	membrane spanning 4-domains A14	NA

1649	DDIT3	DNA damage inducible transcript 3	NA
51458	RHCG	Rh family C glycoprotein	NA
167555	FAM151B	family with sequence similarity 151 member B	NA
151516	ASPRV1	aspartic peptidase retroviral like 1	NA
100271229	RPL9P33	ribosomal protein L9 pseudogene 33	NA
22874	PLEKHA6	pleckstrin homology domain containing A6	NA
206	AK4P1	adenylate kinase 4 pseudogene 1	NA
10398	MYL9	myosin light chain 9	NA
3914	LAMB3	laminin subunit beta 3	NA
4856	NOV	nephroblastoma overexpressed	NA
7784	ZP3	zona pellucida glycoprotein 3	NA
6483	ST3GAL2	ST3 beta-galactoside alpha-2,3-sialyltransferase 2	NA
9094	UNC119	unc-119 lipid binding chaperone	NA
253559	CADM2	cell adhesion molecule 2	NA
2509	FTH1P5	ferritin heavy chain 1 pseudogene 5	NA
201181	ZNF385C	zinc finger protein 385C	NA
646467	CNN2P1	calponin 2 pseudogene 1	NA
390211	KRT8P26	keratin 8 pseudogene 26	NA
29950	SERTAD1	SERTA domain containing 1	NA
3554	IL1R1	interleukin 1 receptor type 1	NA
27287	VENTX	VENT homeobox	NA
10769	PLK2	polo like kinase 2	NA
10970	CKAP4	cytoskeleton associated protein 4	NA
8348	HIST1H2BO	histone cluster 1 H2B family member o	NA
4323	MMP14	matrix metallopeptidase 14	NA
440829	SHISA8	shisa family member 8	NA
2161	F12	coagulation factor XII	NA
967	CD63	CD63 molecule	NA
8876	VNN1	vanin 1	NA
10402	ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	NA
80301	PLEKHO2	pleckstrin homology domain containing O2	NA
1440	CSF3	colony stimulating factor 3	NA
25907	TMEM158	transmembrane protein 158 (gene/pseudogene)	NA
51733	UPB1	beta-ureidopropionase 1	NA
9123	SLC16A3	solute carrier family 16 member 3	NA
10656	KHDRBS3	KH RNA binding domain containing, signal transduction associated 3	NA
283651	HMGN2P46	high mobility group nucleosomal binding domain 2 pseudogene 46	NA
79908	BTNL8	butyrophilin like 8	NA
8740	TNFSF14	TNF superfamily member 14	NA
10053	AP1M2	adaptor related protein complex 1 subunit mu 2	NA
5443	POMC	proopiomelanocortin	NA
80256	FAM214B	family with sequence similarity 214 member B	NA
10732	TCFL5	transcription factor like 5	NA
5055	SERPINB2	serpin family B member 2	NA
9306	SOCS6	suppressor of cytokine signaling 6	NA

54210	TREM1	triggering receptor expressed on myeloid cells 1	NA
222235	FBXL13	F-box and leucine rich repeat protein 13	NA
2239	GPC4	glypican 4	NA
55647	RAB20	RAB20, member RAS oncogene family	NA
64866	CDCP1	CUB domain containing protein 1	NA
124599	CD300LB	CD300 molecule like family member b	NA
402778	IFITM10	interferon induced transmembrane protein 10	NA
124222	PAQR4	progestin and adipoQ receptor family member 4	NA
9704	DHX34	DExH-box helicase 34	NA
222389	BEND7	BEN domain containing 7	NA
8858	PROZ	protein Z, vitamin K dependent plasma glycoprotein	NA
8974	P4HA2	prolyl 4-hydroxylase subunit alpha 2	NA
9260	PDLIM7	PDZ and LIM domain 7	NA
171427	CYP2F2P	cytochrome P450 family 2 subfamily F member 2, pseudogene	NA
1116	CHI3L1	chitinase 3 like 1	NA
84557	MAP1LC3A	microtubule associated protein 1 light chain 3 alpha	NA
5957	RCVRN	recoverin	NA
2769	GNA15	G protein subunit alpha 15	NA
116173	CMTM5	CKLF like MARVEL transmembrane domain containing 5	NA
9856	KIAA0319	KIAA0319	NA
5266	PI3	peptidase inhibitor 3	NA
55561	CDC42BPG	CDC42 binding protein kinase gamma	NA
115727	RASGRP4	RAS guanyl releasing protein 4	NA
10215	OLIG2	oligodendrocyte transcription factor 2	NA
22822	PHLDA1	pleckstrin homology like domain family A member 1	NA
3128	HLA-DRB6	major histocompatibility complex, class II, DR beta 6 (pseudogene)	NA
58160	NFE4	nuclear factor, erythroid 4	NA
117144	CATSPER1	cation channel sperm associated 1	NA
148170	CDC42EP5	CDC42 effector protein 5	NA
8808	IL1RL2	interleukin 1 receptor like 2	NA
10610	ST6GALNAC2	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	NA
388931	MFSD2B	major facilitator superfamily domain containing 2B	NA
106479346	RN7SL336P	RNA, 7SL, cytoplasmic 336, pseudogene	NA
28943	IGKV1-6	immunoglobulin kappa variable 1-6	NA
285489	DOK7	docking protein 7	NA
100113378	SNORD119	small nucleolar RNA, C/D box 119	NA
9966	TNFSF15	TNF superfamily member 15	NA
91156	IGFN1	immunoglobulin-like and fibronectin type III domain containing 1	NA
1145	CHRNE	cholinergic receptor nicotinic epsilon subunit	NA
7504	XK	X-linked Kx blood group	NA
340205	TREML1	triggering receptor expressed on myeloid cells like 1	NA

3040	HBA2	hemoglobin subunit alpha 2	NA
57110	HRASLS	HRAS like suppressor	NA
10105	PPIF	peptidylprolyl isomerase F	NA
5034	P4HB	prolyl 4-hydroxylase subunit beta	NA
3934	LCN2	lipocalin 2	NA
221692	PHACTR1	phosphatase and actin regulator 1	NA
100144748	KLLN	killin, p53 regulated DNA replication inhibitor	NA
388323	GLTPD2	glycolipid transfer protein domain containing 2	NA
2500	FTH1P7	ferritin heavy chain 1 pseudogene 7	NA
353514	LILRA5	leukocyte immunoglobulin like receptor A5	NA
3674	ITGA2B	integrin subunit alpha 2b	NA
54682	MANSC1	MANSC domain containing 1	NA
440926	H3F3AP4	H3 histone, family 3A, pseudogene 4	NA
23657	SLC7A11	solute carrier family 7 member 11	NA
55194	EVA1B	eva-1 homolog B	NA
6854	SYN2	synapsin II	NA
1514	CTSL	cathepsin L	NA
387590	TPTEP1	TPTE pseudogene 1	NA
56927	GPR108	G protein-coupled receptor 108	NA
79865	TREML2	triggering receptor expressed on myeloid cells like 2	NA
122402	TDRD9	tudor domain containing 9	NA
9914	ATP2C2	ATPase secretory pathway Ca2+ transporting 2	NA
3043	HBB	hemoglobin subunit beta	NA
116362	RBP7	retinol binding protein 7	NA
171220	DSTNP2	destrin, actin depolymerizing factor pseudogene 2	NA
1475	CSTA	cystatin A	NA
1088	CEACAM8	carcinoembryonic antigen related cell adhesion molecule 8	NA
8578	SCARF1	scavenger receptor class F member 1	NA
2646	GCKR	glucokinase regulator	NA
84981	MIR22HG	MIR22 host gene	NA
81832	NETO1	neuropilin and tolloid like 1	NA
200942	KLHDC8B	kelch domain containing 8B	NA
92162	TMEM88	transmembrane protein 88	NA
8843	HCAR3	hydroxycarboxylic acid receptor 3	NA
121256	TMEM132D	transmembrane protein 132D	NA
131540	ZDHHC19	zinc finger DHHC-type containing 19	NA
79689	STEAP4	STEAP4 metalloreductase	NA
5627	PROS1	protein S	NA
51296	SLC15A3	solute carrier family 15 member 3	NA
5913	RAPSN	receptor associated protein of the synapse	NA
54863	TOR4A	torsin family 4 member A	NA
85364	ZCCHC3	zinc finger CCHC-type containing 3	NA
23401	FRAT2	FRAT2, WNT signaling pathway regulator	NA
2497	FTH1P2	ferritin heavy chain 1 pseudogene 2	NA
147906	DACT3	dishevelled binding antagonist of beta catenin 3	NA

10990	LILRB5	leukocyte immunoglobulin like receptor B5	NA
83742	MARVELD1	MARVEL domain containing 1	NA
147798	TMC4	transmembrane channel like 4	NA
128854	TSPY26P	testis specific protein Y-linked 26, pseudogene	NA
665	BNIP3L	BCL2 interacting protein 3 like	NA
10184	LHFPL2	LHFPL tetraspan subfamily member 2	NA
11343	MGLL	monoglyceride lipase	NA
202309	GAPT	GRB2 binding adaptor protein, transmembrane	NA
4239	MFAP4	microfibril associated protein 4	NA
56975	FAM20C	FAM20C, golgi associated secretory pathway kinase	NA
4326	MMP17	matrix metallopeptidase 17	NA
7125	TNNC2	troponin C2, fast skeletal type	NA
5365	PLXNB3	plexin B3	NA
671	BPI	bactericidal permeability increasing protein	NA
7165	TPD52L2	TPD52 like 2	NA
5900	RALGDS	ral guanine nucleotide dissociation stimulator	NA
4097	MAFG	MAF bZIP transcription factor G	NA
6817	SULT1A1	sulfotransferase family 1A member 1	NA
10924	SMPDL3A	sphingomyelin phosphodiesterase acid like 3A	NA
10043	TOM1	target of myb1 membrane trafficking protein	NA
79167	LILRP1	leukocyte immunoglobulin-like receptor pseudogene 1	NA
286006	LSMEM1	leucine rich single-pass membrane protein 1	NA
5618	PRLR	prolactin receptor	NA
55002	TMCO3	transmembrane and coiled-coil domains 3	NA
4317	MMP8	matrix metallopeptidase 8	NA
85301	COL27A1	collagen type XXVII alpha 1 chain	NA
26978	PABPC1P3	poly(A) binding protein cytoplasmic 1 pseudogene 3	NA
10011	SRA1	steroid receptor RNA activator 1	NA
9775	EIF4A3	eukaryotic translation initiation factor 4A3	NA
134549	SHROOM1	shroom family member 1	NA
5724	PTAFR	platelet activating factor receptor	NA
2495	FTH1	ferritin heavy chain 1	NA
23223	RRP12	ribosomal RNA processing 12 homolog	NA
820	CAMP	cathelicidin antimicrobial peptide	NA
5553	PRG2	proteoglycan 2, pro eosinophil major basic protein	NA
91752	ZNF804A	zinc finger protein 804A	NA
100874272	KLF7-IT1	KLF7 intronic transcript 1	NA
537	ATP6AP1	ATPase H+ transporting accessory protein 1	NA
5777	PTPN6	protein tyrosine phosphatase, non-receptor type 6	NA
1950	EGF	epidermal growth factor	NA
8854	ALDH1A2	aldehyde dehydrogenase 1 family member A2	NA
55911	APOBR	apolipoprotein B receptor	NA
9365	KL	klotho	NA
10981	RAB32	RAB32, member RAS oncogene family	NA
55022	PID1	phosphotyrosine interaction domain containing 1	NA

9853	RUSC2	RUN and SH3 domain containing 2	NA
168455	CCDC71L	coiled-coil domain containing 71 like	NA
5598	MAPK7	mitogen-activated protein kinase 7	NA
91663	MYADM	myeloid associated differentiation marker	NA
80045	GPR157	G protein-coupled receptor 157	NA
7450	VWF	von Willebrand factor	NA
6364	CCL20	C-C motif chemokine ligand 20	NA
89872	AQP10	aquaporin 10	NA
84879	MFSD2A	major facilitator superfamily domain containing 2A	NA
9088	PKMYT1	protein kinase, membrane associated tyrosine/threonine 1	NA
272	AMPD3	adenosine monophosphate deaminase 3	NA
148252	DIRAS1	DIRAS family GTPase 1	NA
168544	ZNF467	zinc finger protein 467	NA
2268	FGR	FGR proto-oncogene, Src family tyrosine kinase	NA
126014	OSCAR	osteoclast associated, immunoglobulin-like receptor	NA
402665	IGLON5	IgLON family member 5	NA
7167	TPI1	triosephosphate isomerase 1	NA
1084	CEACAM3	carcinoembryonic antigen related cell adhesion molecule 3	NA
2152	F3	coagulation factor III, tissue factor	NA
55505	NOP10	NOP10 ribonucleoprotein	NA
90874	ZNF697	zinc finger protein 697	NA
51312	SLC25A37	solute carrier family 25 member 37	NA
2681	GGTA1P	glycoprotein, alpha-galactosyltransferase 1 pseudogene	NA
374	AREG	amphiregulin	NA
25798	BRI3	brain protein I3	NA
81029	WNT5B	Wnt family member 5B	NA
124935	SLC43A2	solute carrier family 43 member 2	NA
944	TNFSF8	TNF superfamily member 8	NA
441478	NRARP	NOTCH regulated ankyrin repeat protein	NA
254122	SNX32	sorting nexin 32	NA
79888	LPCAT1	lysophosphatidylcholine acyltransferase 1	NA
205	AK4	adenylate kinase 4	NA
11015	KDELR3	KDEL endoplasmic reticulum protein retention receptor 3	NA
692075	SNORD6	small nucleolar RNA, C/D box 6	NA
8125	ANP32A	acidic nuclear phosphoprotein 32 family member A	NA
10321	CRISP3	cysteine rich secretory protein 3	NA
2760	GM2A	GM2 ganglioside activator	NA
91056	AP5B1	adaptor related protein complex 5 subunit beta 1	NA
387837	CLEC12B	C-type lectin domain family 12 member B	NA
11148	HHLA2	HERV-H LTR-associating 2	NA
148113	CILP2	cartilage intermediate layer protein 2	NA
25805	BAMBI	BMP and activin membrane bound inhibitor	NA
4947	OAZ2	ornithine decarboxylase antizyme 2	NA
2876	GPX1	glutathione peroxidase 1	NA

10842	PPP1R17	protein phosphatase 1 regulatory subunit 17	NA
5580	PRKCD	protein kinase C delta	NA
1230	CCR1	C-C motif chemokine receptor 1	NA
440695	ETV3L	ETS variant 3 like	NA
6324	SCN1B	sodium voltage-gated channel beta subunit 1	NA
66002	CYP4F12	cytochrome P450 family 4 subfamily F member 12	NA
3269	HRH1	histamine receptor H1	NA
10758	TRAF3IP2	TRAF3 interacting protein 2	NA
84418	CYSTM1	cysteine rich transmembrane module containing 1	NA
11337	GABARAP	GABA type A receptor-associated protein	NA
1889	ECE1	endothelin converting enzyme 1	NA
83547	RILP	Rab interacting lysosomal protein	NA
348094	ANKDD1A	ankyrin repeat and death domain containing 1A	NA
60490	PPCDC	phosphopantothenoylcysteine decarboxylase	NA
29946	SERTAD3	SERTA domain containing 3	NA
3696	ITGB8	integrin subunit beta 8	NA
266675	BEST4	bestrophin 4	NA
56953	NT5M	5',3'-nucleotidase, mitochondrial	NA
11156	PTP4A3	protein tyrosine phosphatase type IVA, member 3	NA
7409	VAV1	vav guanine nucleotide exchange factor 1	NA
4607	MYBPC3	myosin binding protein C, cardiac	NA
3915	LAMC1	laminin subunit gamma 1	NA
3656	IRAK2	interleukin 1 receptor associated kinase 2	NA
10318	TNIP1	TNFAIP3 interacting protein 1	NA
8497	PPFIA4	PTPRF interacting protein alpha 4	NA
7873	MANF	mesencephalic astrocyte derived neurotrophic factor	NA
4665	NAB2	NGFI-A binding protein 2	NA
162466	PHOSPHO1	phosphoethanolamine/phosphocholine phosphatase	NA
80820	EEPD1	endonuclease/exonuclease/phosphatase family domain containing 1	NA
1263	PLK3	polo like kinase 3	NA
9060	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	NA
2919	CXCL1	C-X-C motif chemokine ligand 1	NA
100270974	RPL29P14	ribosomal protein L29 pseudogene 14	NA
57801	HES4	hes family bHLH transcription factor 4	NA
58528	RRAGD	Ras related GTP binding D	NA
106480955	RN7SL124P	RNA, 7SL, cytoplasmic 124, pseudogene	NA
3017	HIST1H2BD	histone cluster 1 H2B family member d	NA
143689	PIWIL4	piwi like RNA-mediated gene silencing 4	NA
226	ALDOA	aldolase, fructose-bisphosphate A	NA
10942	PRSS21	serine protease 21	NA
1152	CKB	creatine kinase B	NA
100113407	TMEM170B	transmembrane protein 170B	NA
55048	VPS37C	VPS37C, ESCRT-I subunit	NA
23533	PIK3R5	phosphoinositide-3-kinase regulatory subunit 5	NA
10625	IVNS1ABP	influenza virus NS1A binding protein	NA

100631383	FAM47E-STBD1	FAM47E-STBD1 readthrough	NA
8828	NRP2	neuropilin 2	NA
2077	ERF	ETS2 repressor factor	NA
6813	STXBP2	syntaxin binding protein 2	NA
6236	RRAD	RRAD, Ras related glycolysis inhibitor and calcium channel regulator	NA
10603	SH2B2	SH2B adaptor protein 2	NA
115761	ARL11	ADP ribosylation factor like GTPase 11	NA
8632	DNAH17	dynein axonemal heavy chain 17	NA
90952	ESAM	endothelial cell adhesion molecule	NA
28665	TRAV18	T cell receptor alpha variable 18	NA
5329	PLAUR	plasminogen activator, urokinase receptor	NA
6678	SPARC	secreted protein acidic and cysteine rich	NA
116844	LRG1	leucine rich alpha-2-glycoprotein 1	NA
56978	PRDM8	PR/SET domain 8	NA
57053	CHRNA10	cholinergic receptor nicotinic alpha 10 subunit	NA
125336	LOXHD1	lipoxygenase homology domains 1	NA
129642	MBOAT2	membrane bound O-acyltransferase domain containing 2	NA
10211	FLOT1	flotillin 1	NA
3012	HIST1H2AE	histone cluster 1 H2A family member e	NA
7408	VASP	vasodilator stimulated phosphoprotein	NA
8685	MARCO	macrophage receptor with collagenous structure	NA
3588	IL10RB	interleukin 10 receptor subunit beta	NA
116448	OLIG1	oligodendrocyte transcription factor 1	NA
728262	FAM157A	family with sequence similarity 157 member A	NA
27189	IL17C	interleukin 17C	NA
338339	CLEC4D	C-type lectin domain family 4 member D	NA
1476	CSTB	cystatin B	NA
100418750	KRT8P42	keratin 8 pseudogene 42	NA
9244	CRLF1	cytokine receptor like factor 1	NA
8477	GPR65	G protein-coupled receptor 65	NA
4493	MT1E	metallothionein 1E	NA
4257	MGST1	microsomal glutathione S-transferase 1	NA
10501	SEMA6B	semaphorin 6B	NA
221806	VWDE	von Willebrand factor D and EGF domains	NA
26996	GPR160	G protein-coupled receptor 160	NA
3006	HIST1H1C	histone cluster 1 H1 family member c	NA
10938	EHD1	EH domain containing 1	NA
2597	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	NA
29780	PARVB	parvin beta	NA
3687	ITGAX	integrin subunit alpha X	NA
79628	SH3TC2	SH3 domain and tetratricopeptide repeats 2	NA
8337	HIST2H2AA3	histone cluster 2 H2A family member a3	NA
117286	CIB3	calcium and integrin binding family member 3	NA
266747	RGL4	ral guanine nucleotide dissociation stimulator like 4	NA
2649	NR6A1	nuclear receptor subfamily 6 group A member 1	NA

83706	FERMT3	fermitin family member 3	NA
10383	TUBB4B	tubulin beta 4B class IVb	NA
4354	MPP1	membrane palmitoylated protein 1	NA
3753	KCNE1	potassium voltage-gated channel subfamily E regulatory subunit 1	NA
5552	SRGN	serglycin	NA
55287	TMEM40	transmembrane protein 40	NA
444	ASPH	aspartate beta-hydroxylase	NA
106480740	GK-IT1	GK intronic transcript 1	NA
124044	SPATA2L	spermatogenesis associated 2 like	NA
140701	ABHD16B	abhydrolase domain containing 16B	NA
146225	CMTM2	CKLF like MARVEL transmembrane domain containing 2	NA
84519	ACRBP	acrosin binding protein	NA
6648	SOD2	superoxide dismutase 2	NA
79850	FAM57A	family with sequence similarity 57 member A	NA
8178	ELL	elongation factor for RNA polymerase II	NA
4082	MARCKS	myristoylated alanine rich protein kinase C substrate	NA
51162	EGFL7	EGF like domain multiple 7	NA
128866	CHMP4B	charged multivesicular body protein 4B	NA
10588	MTHFS	methenyltetrahydrofolate synthetase	NA
94059	LENG9	leukocyte receptor cluster member 9	NA
5473	PPBP	pro-platelet basic protein	NA
5460	POU5F1	POU class 5 homeobox 1	NA
83478	ARHGAP24	Rho GTPase activating protein 24	NA
50486	G0S2	G0/G1 switch 2	NA
10634	GAS2L1	growth arrest specific 2 like 1	NA
162989	DEDD2	death effector domain containing 2	NA
29015	SLC43A3	solute carrier family 43 member 3	NA
1912	PHC2	polyhomeotic homolog 2	NA
5791	PTPRE	protein tyrosine phosphatase, receptor type E	NA
10089	KCNK7	potassium two pore domain channel subfamily K member 7	NA
100616164	MIR4420	microRNA 4420	NA
112574	SNX18	sorting nexin 18	NA
3340	NDST1	N-deacetylase and N-sulfotransferase 1	NA
100422964	MIR3150A	microRNA 3150a	NA
2887	GRB10	growth factor receptor bound protein 10	NA
9625	AATK	apoptosis associated tyrosine kinase	NA
9654	TTLL4	tubulin tyrosine ligase like 4	NA
4051	CYP4F3	cytochrome P450 family 4 subfamily F member 3	NA
177	AGER	advanced glycosylation end-product specific receptor	NA
6282	S100A11	S100 calcium binding protein A11	NA
91662	NLRP12	NLR family pyrin domain containing 12	NA
535	ATP6V0A1	ATPase H+ transporting V0 subunit a1	NA
80758	PRR7	proline rich 7, synaptic	NA

3654	IRAK1	interleukin 1 receptor associated kinase 1	NA
5315	PKM	pyruvate kinase M1/2	NA
29935	RPA4	replication protein A4	NA
7846	TUBA1A	tubulin alpha 1a	NA
5025	P2RX4	purinergic receptor P2X 4	NA
642559	POU5F1P3	POU class 5 homeobox 1 pseudogene 3	NA
23498	HAAO	3-hydroxyanthranilate 3,4-dioxygenase	NA
664	BNIP3	BCL2 interacting protein 3	NA
7132	TNFRSF1A	TNF receptor superfamily member 1A	NA
7421	VDR	vitamin D receptor	NA
10148	EBI3	Epstein-Barr virus induced 3	NA
80149	ZC3H12A	zinc finger CCCH-type containing 12A	NA
79143	MBOAT7	membrane bound O-acyltransferase domain containing 7	NA
199731	CADM4	cell adhesion molecule 4	NA
85450	ITPRIP	inositol 1,4,5-trisphosphate receptor interacting protein	NA
2852	GPER1	G protein-coupled estrogen receptor 1	NA
2023	ENO1	enolase 1	NA
15	AANAT	aralkylamine N-acetyltransferase	NA
2321	FLT1	fms related tyrosine kinase 1	NA
171558	PTCRA	pre T cell antigen receptor alpha	NA
5997	RGS2	regulator of G protein signaling 2	NA
1191	CLU	clusterin	NA
400569	MED11	mediator complex subunit 11	NA
64127	NOD2	nucleotide binding oligomerization domain containing 2	NA
4000	LMNA	lamin A/C	NA
23308	ICOSLG	inducible T cell costimulator ligand	NA
64092	SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	NA
25946	ZNF385A	zinc finger protein 385A	NA
692212	SNORD99	small nucleolar RNA, C/D box 99	NA
64778	FNDC3B	fibronectin type III domain containing 3B	NA
7140	TNNT3	troponin T3, fast skeletal type	NA
83931	STK40	serine/threonine kinase 40	NA
121260	SLC15A4	solute carrier family 15 member 4	NA
11182	SLC2A6	solute carrier family 2 member 6	NA
7262	PHLDA2	pleckstrin homology like domain family A member 2	NA
8971	H1FX	H1 histone family member X	NA
407013	MIR24-2	microRNA 24-2	NA
3987	LIMS1	LIM zinc finger domain containing 1	NA
29982	NRBF2	nuclear receptor binding factor 2	NA
54498	SMOX	spermine oxidase	NA
83693	HSDL1	hydroxysteroid dehydrogenase like 1	NA
23130	ATG2A	autophagy related 2A	NA
4067	LYN	LYN proto-oncogene, Src family tyrosine kinase	NA

1509	CTSD	cathepsin D	NA
6403	SELP	selectin P	NA
23327	NEDD4L	neural precursor cell expressed, developmentally down-regulated 4-like, E3 ubiquitin protein ligase	NA
22885	ABLIM3	actin binding LIM protein family member 3	NA
3310	HSPA6	heat shock protein family A (Hsp70) member 6	NA
29940	DSE	dermatan sulfate epimerase	NA
116369	SLC26A8	solute carrier family 26 member 8	NA
6617	SNAPC1	small nuclear RNA activating complex polypeptide 1	NA
6947	TCN1	transcobalamin 1	NA
10581	IFITM2	interferon induced transmembrane protein 2	NA
5326	PLAGL2	PLAG1 like zinc finger 2	NA
4086	SMAD1	SMAD family member 1	NA
1820	ARID3A	AT-rich interaction domain 3A	NA
3679	ITGA7	integrin subunit alpha 7	NA
3815	KIT	KIT proto-oncogene receptor tyrosine kinase	NA
85236	HIST1H2BK	histone cluster 1 H2B family member k	NA
90427	BMF	Bcl2 modifying factor	NA
427	ASAH1	N-acylsphingosine amidohydrolase 1	NA
409	ARRB2	arrestin beta 2	NA
1277	COL1A1	collagen type I alpha 1 chain	NA
4189	DNAJB9	DnaJ heat shock protein family (Hsp40) member B9	NA
50626	CYHR1	cysteine and histidine rich 1	NA
643418	LIPN	lipase family member N	NA
5033	P4HA1	prolyl 4-hydroxylase subunit alpha 1	NA
79759	ZNF668	zinc finger protein 668	NA
2997	GYS1	glycogen synthase 1	NA
57172	CAMK1G	calcium/calmodulin dependent protein kinase IG	NA
1869	E2F1	E2F transcription factor 1	NA
3693	ITGB5	integrin subunit beta 5	NA
9446	GSTO1	glutathione S-transferase omega 1	NA
586	BCAT1	branched chain amino acid transaminase 1	NA
5739	PTGIR	prostaglandin I2 receptor	NA
391	RHOG	ras homolog family member G	NA
1141	CHRNB2	cholinergic receptor nicotinic beta 2 subunit	NA
123872	DNAAF1	dynein axonemal assembly factor 1	NA
51726	DNAJB11	DnaJ heat shock protein family (Hsp40) member B11	NA
65220	NADK	NAD kinase	NA
339929	LPP-AS2	LPP antisense RNA 2	NA
5209	PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	NA
9173	IL1RL1	interleukin 1 receptor like 1	NA
11153	FICD	FIC domain containing	NA
4520	MTF1	metal regulatory transcription factor 1	NA
2921	CXCL3	C-X-C motif chemokine ligand 3	NA

51738	GHRL	ghrelin and obestatin prepropeptide	NA
407006	MIR221	microRNA 221	NA
148022	TICAM1	toll like receptor adaptor molecule 1	NA
6611	SMS	spermine synthase	NA
106479357	RN7SL368P	RNA, 7SL, cytoplasmic 368, pseudogene	NA
4860	PNP	purine nucleoside phosphorylase	NA
51700	CYB5R2	cytochrome b5 reductase 2	NA
23254	KAZN	kazrin, periplakin interacting protein	NA
80741	LY6G5C	lymphocyte antigen 6 family member G5C	NA
11345	GABARAPL2	GABA type A receptor associated protein like 2	NA
27243	CHMP2A	charged multivesicular body protein 2A	NA
9034	CCRL2	C-C motif chemokine receptor like 2	NA
7980	TFPI2	tissue factor pathway inhibitor 2	NA
9050	PSTPIP2	proline-serine-threonine phosphatase interacting protein 2	NA
6804	STX1A	syntaxin 1A	NA
65249	ZSWIM4	zinc finger SWIM-type containing 4	NA
64218	SEMA4A	semaphorin 4A	NA
100996541	FAM157C	family with sequence similarity 157 member C	NA
7378	UPP1	uridine phosphorylase 1	NA
84002	B3GNT5	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5	NA
93129	ORAI3	ORAI calcium release-activated calcium modulator 3	NA
200186	CRTC2	CREB regulated transcription coactivator 2	NA
7130	TNFAIP6	TNF alpha induced protein 6	NA
92799	SHKBP1	SH3KBP1 binding protein 1	NA
255488	RNF144B	ring finger protein 144B	NA
4291	MLF1	myeloid leukemia factor 1	NA
51114	ZDHHC9	zinc finger DHHC-type containing 9	NA
221749	PXDC1	PX domain containing 1	NA
54883	CWC25	CWC25 spliceosome associated protein homolog	NA
7226	TRPM2	transient receptor potential cation channel subfamily M member 2	NA
56675	NRIP3	nuclear receptor interacting protein 3	NA
8870	IER3	immediate early response 3	NA
5306	PITPNA	phosphatidylinositol transfer protein alpha	NA
2010	EMD	emerin	NA
55122	AKIRIN2	akirin 2	NA
1192	CLIC1	chloride intracellular channel 1	NA
9903	KLHL21	kelch like family member 21	NA
58513	EPS15L1	epidermal growth factor receptor pathway substrate 15 like 1	NA
9589	WTAP	WT1 associated protein	NA
10663	CXCR6	C-X-C motif chemokine receptor 6	NA
2766	GMPR	guanosine monophosphate reductase	NA
157695	TDRP	testis development related protein	NA
131566	DCBLD2	discoidin, CUB and LCCL domain containing 2	NA

9997	SCO2	SCO2, cytochrome c oxidase assembly protein	NA
2107	ETF1	eukaryotic translation termination factor 1	NA
64411	ARAP3	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 3	NA
84034	EMILIN2	elastin microfibril interfacer 2	NA
3720	JARID2	jumonji and AT-rich interaction domain containing 2	NA
58484	NLRC4	NLR family CARD domain containing 4	NA
2101	ESRRA	estrogen related receptor alpha	NA
5990	RFX2	regulatory factor X2	NA
8878	SQSTM1	sequestosome 1	NA
400410	ST20	suppressor of tumorigenicity 20	NA
6277	S100A6	S100 calcium binding protein A6	NA
8407	TAGLN2	transgelin 2	NA
6809	STX3	syntaxin 3	NA
441481	GPX1P1	glutathione peroxidase pseudogene 1	NA
400709	SIGLEC16	sialic acid binding Ig like lectin 16 (gene/pseudogene)	NA
306	ANXA3	annexin A3	NA
54823	SWT1	SWT1, RNA endoribonuclease homolog	NA
7280	TUBB2A	tubulin beta 2A class IIa	NA
51206	GP6	glycoprotein VI platelet	NA
85403	EAF1	ELL associated factor 1	NA
51043	ZBTB7B	zinc finger and BTB domain containing 7B	NA
254263	CNIH2	cornichon family AMPA receptor auxiliary protein 2	NA
84067	FAM160A2	family with sequence similarity 160 member A2	NA
60370	AVPI1	arginine vasopressin induced 1	NA
27161	AGO2	argonaute 2, RISC catalytic component	NA
5606	MAP2K3	mitogen-activated protein kinase kinase 3	NA
25953	PNKD	PNKD, MBL domain containing	NA
6622	SNCA	synuclein alpha	NA
57085	AGTRAP	angiotensin II receptor associated protein	NA
79174	CRELD2	cysteine rich with EGF like domains 2	NA
4066	LYL1	LYL1, basic helix-loop-helix family member	NA
5799	PTPRN2	protein tyrosine phosphatase, receptor type N2	NA
51582	AZIN1	antizyme inhibitor 1	NA
5023	P2RX1	purinergic receptor P2X 1	NA
65059	RAPH1	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1	NA
284723	SLC25A34	solute carrier family 25 member 34	NA
10253	SPRY2	sprouty RTK signaling antagonist 2	NA
161253	REM2	RRAD and GEM like GTPase 2	NA
346653	FAM71F2	family with sequence similarity 71 member F2	NA
85329	LGALS12	galectin 12	NA
57567	ZNF319	zinc finger protein 319	NA
26037	SIPA1L1	signal induced proliferation associated 1 like 1	NA
113878	DTX2	deltex E3 ubiquitin ligase 2	NA

9628	RGS6	regulator of G protein signaling 6	NA
10045	SH2D3A	SH2 domain containing 3A	NA
55281	TMEM140	transmembrane protein 140	NA
1398	CRK	CRK proto-oncogene, adaptor protein	NA
117289	TAGAP	T cell activation RhoGTPase activating protein	NA
84444	DOT1L	DOT1 like histone lysine methyltransferase	NA
51205	ACP6	acid phosphatase 6, lysophosphatidic	NA
27348	TOR1B	torsin family 1 member B	NA
6320	CLEC11A	C-type lectin domain containing 11A	NA
80331	DNAJC5	DnaJ heat shock protein family (Hsp40) member C5	NA
84925	DIRC2	disrupted in renal carcinoma 2	NA
400617	KCNJ2-AS1	KCNJ2 antisense RNA 1	NA
64581	CLEC7A	C-type lectin domain containing 7A	NA
131601	TPRA1	transmembrane protein adipocyte associated 1	NA
6884	TAF13	TATA-box binding protein associated factor 13	NA
8303	SNN	stannin	NA
127829	ARL8A	ADP ribosylation factor like GTPase 8A	NA
1030	CDKN2B	cyclin dependent kinase inhibitor 2B	NA
51177	PLEKHO1	pleckstrin homology domain containing O1	NA
5196	PF4	platelet factor 4	NA
26024	PTCD1	pentatricopeptide repeat domain 1	NA
84243	ZDHHC18	zinc finger DHHC-type containing 18	NA
51725	FBXO40	F-box protein 40	NA
90204	ZSWIM1	zinc finger SWIM-type containing 1	NA
6520	SLC3A2	solute carrier family 3 member 2	NA
10046	MAMLD1	mastermind like domain containing 1	NA
7090	TLE3	transducin like enhancer of split 3	NA
396	ARHGDIA	Rho GDP dissociation inhibitor alpha	NA
5447	POR	cytochrome p450 oxidoreductase	NA
100505881	MAGI2-AS3	MAGI2 antisense RNA 3	NA
4311	MME	membrane metalloendopeptidase	NA
4582	MUC1	mucin 1, cell surface associated	NA
55766	H2AFJ	H2A histone family member J	NA
10555	AGPAT2	1-acylglycerol-3-phosphate O-acyltransferase 2	NA
811	CALR	calreticulin	NA
23480	SEC61G	Sec61 translocon gamma subunit	NA
55257	MRGBP	MRG domain binding protein	NA
54541	DDIT4	DNA damage inducible transcript 4	NA
10826	FAXDC2	fatty acid hydroxylase domain containing 2	NA
5187	PER1	period circadian regulator 1	NA
9479	MAPK8IP1	mitogen-activated protein kinase 8 interacting protein 1	NA
1471	CST3	cystatin C	NA
468	ATF4	activating transcription factor 4	NA
7171	TPM4	tropomyosin 4	NA
10867	TSPAN9	tetraspanin 9	NA

11142	PKIG	cAMP-dependent protein kinase inhibitor gamma	NA
85012	TCEAL3	transcription elongation factor A like 3	NA
1844	DUSP2	dual specificity phosphatase 2	NA
57471	ERMN	ermin	NA
23416	KCNH3	potassium voltage-gated channel subfamily H member 3	NA
23603	CORO1C	coronin 1C	NA
101	ADAM8	ADAM metallopeptidase domain 8	NA
5531	PPP4C	protein phosphatase 4 catalytic subunit	NA
4482	MSRA	methionine sulfoxide reductase A	NA
6455	SH3GL1	SH3 domain containing GRB2 like 1, endophilin A2	NA
58526	MID1IP1	MID1 interacting protein 1	NA
4046	LSP1	lymphocyte specific protein 1	NA
2623	GATA1	GATA binding protein 1	NA
490	ATP2B1	ATPase plasma membrane Ca2+ transporting 1	NA
282969	FUOM	fucose mutarotase	NA
7494	XBP1	X-box binding protein 1	NA
3005	H1F0	H1 histone family member 0	NA
150209	AIFM3	apoptosis inducing factor, mitochondria associated 3	NA
283050	ZMIZ1-AS1	ZMIZ1 antisense RNA 1	NA
4597	MVD	mevalonate diphosphate decarboxylase	NA
5795	PTPRJ	protein tyrosine phosphatase, receptor type J	NA
26088	GGA1	golgi associated, gamma adaptin ear containing, ARF binding protein 1	NA
55565	ZNF821	zinc finger protein 821	NA
9709	HERPUD1	homocysteine inducible ER protein with ubiquitin like domain 1	NA
9150	CTDP1	CTD phosphatase subunit 1	NA
56270	WDR45B	WD repeat domain 45B	NA
85365	ALG2	ALG2, alpha-1,3/1,6-mannosyltransferase	NA
10952	SEC61B	Sec61 translocon beta subunit	NA
91300	R3HDM4	R3H domain containing 4	NA
10533	ATG7	autophagy related 7	NA
4824	NKX3-1	NK3 homeobox 1	NA
5468	PPARG	peroxisome proliferator activated receptor gamma	NA
57333	RCN3	reticulocalbin 3	NA
28991	COMMD5	COMM domain containing 5	NA
2057	EPOR	erythropoietin receptor	NA
4163	MCC	MCC, WNT signaling pathway regulator	NA
9455	HOMER2	homer scaffold protein 2	NA
23584	VSIG2	V-set and immunoglobulin domain containing 2	NA
25911	DPCD	deleted in primary ciliary dyskinesia homolog (mouse)	NA
9647	PPM1F	protein phosphatase, Mg2+/Mn2+ dependent 1F	NA
3779	KCNMB1	potassium calcium-activated channel subfamily M regulatory beta subunit 1	NA
9040	UBE2M	ubiquitin conjugating enzyme E2 M	NA

9572	NR1D1	nuclear receptor subfamily 1 group D member 1	NA
54440	SASH3	SAM and SH3 domain containing 3	NA
10124	ARL4A	ADP ribosylation factor like GTPase 4A	NA
64319	FBRS	fibrosin	NA
90203	SNX21	sorting nexin family member 21	NA
3638	INSIG1	insulin induced gene 1	NA
5611	DNAJC3	DnaJ heat shock protein family (Hsp40) member C3	NA
2923	PDIA3	protein disulfide isomerase family A member 3	NA
23580	CDC42EP4	CDC42 effector protein 4	NA
6016	RIT1	Ras like without CAAX 1	NA
2014	EMP3	epithelial membrane protein 3	NA
5355	PLP2	proteolipid protein 2	NA
401588	ZNF674-AS1	ZNF674 antisense RNA 1 (head to head)	NA
100506311	HOTAIRM1	HOXA transcript antisense RNA, myeloid-specific 1	NA
79650	USB1	U6 snRNA biogenesis phosphodiesterase 1	NA
54896	PQLC2	PQ loop repeat containing 2	NA
9185	REPS2	RALBP1 associated Eps domain containing 2	NA
431705	ASTL	astacin like metalloendopeptidase	NA
9371	KIF3B	kinesin family member 3B	NA
119391	GSTO2	glutathione S-transferase omega 2	NA
8915	BCL10	B cell CLL/lymphoma 10	NA
2274	FHL2	four and a half LIM domains 2	NA
253430	IPMK	inositol polyphosphate multikinase	NA
10397	NDRG1	N-myc downstream regulated 1	NA
284021	MILR1	mast cell immunoglobulin like receptor 1	NA
4974	OMG	oligodendrocyte myelin glycoprotein	NA
8877	SPHK1	sphingosine kinase 1	NA
58489	ABHD17C	abhydrolase domain containing 17C	NA
6029	RN7SL1	RNA, 7SL, cytoplasmic 1	NA
4778	NFE2	nuclear factor, erythroid 2	NA
79583	TMEM231	transmembrane protein 231	NA
6018	RLF	rearranged L-myc fusion	NA
5476	CTSA	cathepsin A	NA
2162	F13A1	coagulation factor XIII A chain	NA
51271	UBAP1	ubiquitin associated protein 1	NA
118788	PIK3AP1	phosphoinositide-3-kinase adaptor protein 1	NA
388	RHOB	ras homolog family member B	NA
149478	BTBD19	BTB domain containing 19	NA
90637	ZFAND2A	zinc finger AN1-type containing 2A	NA
3383	ICAM1	intercellular adhesion molecule 1	NA
27128	CYTH4	cytohesin 4	NA
23258	DENND5A	DENN domain containing 5A	NA
255631	COL24A1	collagen type XXIV alpha 1 chain	NA
387496	RASL11A	RAS like family 11 member A	NA
7086	TKT	transketolase	NA

7940	LST1	leukocyte specific transcript 1	NA
3164	NR4A1	nuclear receptor subfamily 4 group A member 1	NA
1984	EIF5A	eukaryotic translation initiation factor 5A	NA
56548	CHST7	carbohydrate sulfotransferase 7	NA
4794	NFKBIE	NFKB inhibitor epsilon	NA
407021	MIR29A	microRNA 29a	NA
6810	STX4	syntaxin 4	NA
79774	GRTP1	growth hormone regulated TBC protein 1	NA
80273	GRPEL1	GrpE like 1, mitochondrial	NA
2987	GUK1	guanylate kinase 1	NA
150372	NFAM1	NFAT activating protein with ITAM motif 1	NA
644962	TNRC18P1	trinucleotide repeat containing 18 pseudogene 1	NA
10189	ALYREF	Aly/REF export factor	NA
8476	CDC42BPA	CDC42 binding protein kinase alpha	NA
375	ARF1	ADP ribosylation factor 1	NA
8535	CBX4	chromobox 4	NA
8996	NOL3	nucleolar protein 3	NA
22903	BTBD3	BTB domain containing 3	NA
84957	RELT	RELT, TNF receptor	NA
92815	HIST3H2A	histone cluster 3 H2A	NA
1902	LPAR1	lysophosphatidic acid receptor 1	NA
1435	CSF1	colony stimulating factor 1	NA
221184	CPNE2	copine 2	NA
54732	TMED9	transmembrane p24 trafficking protein 9	NA
3068	HDGF	heparin binding growth factor	NA
3500	IGHG1	immunoglobulin heavy constant gamma 1 (G1m marker)	NA
5603	MAPK13	mitogen-activated protein kinase 13	NA
440400	RNASEK	ribonuclease K	NA
100616259	MIR378I	microRNA 378i	NA
10577	NPC2	NPC intracellular cholesterol transporter 2	NA
9545	RAB3D	RAB3D, member RAS oncogene family	NA
55201	MAP1S	microtubule associated protein 1S	NA
100873902	GK-AS1	GK antisense RNA 1	NA
81622	UNC93B1	unc-93 homolog B1, TLR signaling regulator	NA
64859	NABP1	nucleic acid binding protein 1	NA
50640	PNPLA8	patatin like phospholipase domain containing 8	NA
2039	DMTN	dematin actin binding protein	NA
57622	LRFN1	leucine rich repeat and fibronectin type III domain containing 1	NA
9895	TECPR2	tectonin beta-propeller repeat containing 2	NA
9985	REC8	REC8 meiotic recombination protein	NA
2664	GDI1	GDP dissociation inhibitor 1	NA
8829	NRP1	neuropilin 1	NA
284207	METRNL	meteorin like, glial cell differentiation regulator	NA
60685	ZFAND3	zinc finger AN1-type containing 3	NA
3183	HNRNPC	heterogeneous nuclear ribonucleoprotein C (C1/C2)	NA

55063	ZCWPW1	zinc finger CW-type and PWWP domain containing 1	NA
60	ACTB	actin beta	NA
6532	SLC6A4	solute carrier family 6 member 4	NA
63940	GPSM3	G protein signaling modulator 3	NA
7184	HSP90B1	heat shock protein 90 beta family member 1	NA
4210	MEFV	MEFV, pyrin innate immunity regulator	NA
83699	SH3BGRL2	SH3 domain binding glutamate rich protein like 2	NA
27146	FAM184B	family with sequence similarity 184 member B	NA
7531	YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon	NA
6478	SIAH2	siah E3 ubiquitin protein ligase 2	NA
79939	SLC35E1	solute carrier family 35 member E1	NA
51567	TDP2	tyrosyl-DNA phosphodiesterase 2	NA
970	CD70	CD70 molecule	NA
4245	MGAT1	mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	NA
112398	EGLN2	egl-9 family hypoxia inducible factor 2	NA
54936	ADPRHL2	ADP-ribosylhydrolase like 2	NA
64121	RRAGC	Ras related GTP binding C	NA
9948	WDR1	WD repeat domain 1	NA
55508	SLC35E3	solute carrier family 35 member E3	NA
9261	MAPKAPK2	mitogen-activated protein kinase-activated protein kinase 2	NA
7376	NR1H2	nuclear receptor subfamily 1 group H member 2	NA
2683	B4GALT1	beta-1,4-galactosyltransferase 1	NA
5479	PPIB	peptidylprolyl isomerase B	NA
79180	EFHD2	EF-hand domain family member D2	NA
116092	DNTTIP1	deoxynucleotidyltransferase terminal interacting protein 1	NA
80271	ITPKC	inositol-trisphosphate 3-kinase C	NA
5360	PLTP	phospholipid transfer protein	NA
6275	S100A4	S100 calcium binding protein A4	NA
9754	STARD8	StAR related lipid transfer domain containing 8	NA
10130	PDIA6	protein disulfide isomerase family A member 6	NA
5660	PSAP	prosaposin	NA
10252	SPRY1	sprouty RTK signaling antagonist 1	NA
9124	PDLIM1	PDZ and LIM domain 1	NA
8775	NAPA	NSF attachment protein alpha	NA
51278	IER5	immediate early response 5	NA
26301	GBGT1	globoside alpha-1,3-N-acetylgalactosaminyltransferase 1 (FORS blood group)	NA
64080	RBKS	ribokinase	NA
9120	SLC16A6	solute carrier family 16 member 6	NA
9450	LY86	lymphocyte antigen 86	NA
84777	DLGAP1-AS2	DLGAP1 antisense RNA 2	NA
6722	SRF	serum response factor	NA
55361	PI4K2A	phosphatidylinositol 4-kinase type 2 alpha	NA

146223	CMTM4	CKLF like MARVEL transmembrane domain containing 4	NA
91	ACVR1B	activin A receptor type 1B	NA
51195	RAPGEFL1	Rap guanine nucleotide exchange factor like 1	NA
9435	CHST2	carbohydrate sulfotransferase 2	NA
301	ANXA1	annexin A1	NA
50807	ASAP1	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	NA
23597	ACOT9	acyl-CoA thioesterase 9	NA
4791	NFKB2	nuclear factor kappa B subunit 2	NA
441024	MTHFD2L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2 like	NA
94056	SYAP1	synapse associated protein 1	NA
214	ALCAM	activated leukocyte cell adhesion molecule	NA
51526	OSER1	oxidative stress responsive serine rich 1	NA
27033	ZBTB32	zinc finger and BTB domain containing 32	NA
53831	GPR84	G protein-coupled receptor 84	NA
81928	CABLES2	Cdk5 and Abl enzyme substrate 2	NA
83871	RAB34	RAB34, member RAS oncogene family	NA
3784	KCNQ1	potassium voltage-gated channel subfamily Q member 1	NA
7295	TXN	thioredoxin	NA
10525	HYOU1	hypoxia up-regulated 1	NA
5879	RAC1	Rac family small GTPase 1	NA
642475	MROH6	maestro heat like repeat family member 6	NA
10312	TCIRG1	T cell immune regulator 1, ATPase H+ transporting V0 subunit a3	NA
7739	ZNF185	zinc finger protein 185 with LIM domain	NA
79651	RHBDF2	rhomboid 5 homolog 2	NA
5133	PDCD1	programmed cell death 1	NA
3241	HPCAL1	hippocalcin like 1	NA
84727	SPSB2	splA/ryanodine receptor domain and SOCS box containing 2	NA
201305	SPNS3	sphingolipid transporter 3 (putative)	NA
3139	HLA-L	major histocompatibility complex, class I, L (pseudogene)	NA
1326	MAP3K8	mitogen-activated protein kinase kinase kinase 8	NA
57592	ZNF687	zinc finger protein 687	NA
9643	MORF4L2	mortality factor 4 like 2	NA
29115	SAP30BP	SAP30 binding protein	NA
9540	TP53I3	tumor protein p53 inducible protein 3	NA
5608	MAP2K6	mitogen-activated protein kinase kinase 6	NA
1522	CTSZ	cathepsin Z	NA
23765	IL17RA	interleukin 17 receptor A	NA
375387	NRROS	negative regulator of reactive oxygen species	NA
8495	PPFIBP2	PPFIA binding protein 2	NA
6714	SRC	SRC proto-oncogene, non-receptor tyrosine kinase	NA
3956	LGALS1	galectin 1	NA
1627	DBN1	drebrin 1	NA

222229	LRWD1	leucine rich repeats and WD repeat domain containing 1	NA
23558	WBP2	WW domain binding protein 2	NA
65996	CENPBD1P1	CENPB DNA-binding domains containing 1 pseudogene 1	NA
3606	IL18	interleukin 18	NA
347734	SLC35B2	solute carrier family 35 member B2	NA
4001	LMNB1	lamin B1	NA
5331	PLCB3	phospholipase C beta 3	NA
11076	TPPP	tubulin polymerization promoting protein	NA
53918	PELO	pelota mRNA surveillance and ribosome rescue factor	NA
81553	FAM49A	family with sequence similarity 49 member A	NA
641649	TMEM91	transmembrane protein 91	NA
5005	ORM2	orosomucoid 2	NA
116985	ARAP1	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	NA
26267	FBXO10	F-box protein 10	NA
2632	GBE1	1,4-alpha-glucan branching enzyme 1	NA
57673	BEND3	BEN domain containing 3	NA
89790	SIGLEC10	sialic acid binding Ig like lectin 10	NA
692205	SNORD89	small nucleolar RNA, C/D box 89	NA
23315	SLC9A8	solute carrier family 9 member A8	NA
55521	TRIM36	tripartite motif containing 36	NA
11261	CHP1	calcineurin like EF-hand protein 1	NA
638	BIK	BCL2 interacting killer	NA
54331	GNG2	G protein subunit gamma 2	NA
7263	TST	thiosulfate sulfurtransferase	NA
9175	MAP3K13	mitogen-activated protein kinase kinase kinase 13	NA
2309	FOXO3	forkhead box O3	NA
2783	GNB2	G protein subunit beta 2	NA
7071	KLF10	Kruppel like factor 10	NA
2920	CXCL2	C-X-C motif chemokine ligand 2	NA
1611	DAP	death associated protein	NA
284131	ENDOV	endonuclease V	NA
1022	CDK7	cyclin dependent kinase 7	NA
55909	BIN3	bridging integrator 3	NA
151056	PLB1	phospholipase B1	NA
10903	MTMR11	myotubularin related protein 11	NA
10957	PNRC1	proline rich nuclear receptor coactivator 1	NA
5538	PPT1	palmitoyl-protein thioesterase 1	NA
4637	MYL6	myosin light chain 6	NA
7358	UGDH	UDP-glucose 6-dehydrogenase	NA
10765	KDM5B	lysine demethylase 5B	NA
285958	SNHG15	small nucleolar RNA host gene 15	NA
415116	PIM3	Pim-3 proto-oncogene, serine/threonine kinase	NA
58500	ZNF250	zinc finger protein 250	NA
8347	HIST1H2BC	histone cluster 1 H2B family member c	NA

5973	RENBP	renin binding protein	NA
153830	RNF145	ring finger protein 145	NA
3937	LCP2	lymphocyte cytosolic protein 2	NA
741	ZNHIT2	zinc finger HIT-type containing 2	NA
11040	PIM2	Pim-2 proto-oncogene, serine/threonine kinase	NA
9308	CD83	CD83 molecule	NA
112495	GTF3C6	general transcription factor IIIC subunit 6	NA
8334	HIST1H2AC	histone cluster 1 H2A family member c	NA
79693	YRDC	yrdC N6-threonylcarbamoyltransferase domain containing	NA
1265	CNN2	calponin 2	NA
8649	LAMTOR3	late endosomal/lysosomal adaptor, MAPK and MTOR activator 3	NA
3556	IL1RAP	interleukin 1 receptor accessory protein	NA
7325	UBE2E2	ubiquitin conjugating enzyme E2 E2	NA
2526	FUT4	fucosyltransferase 4	NA
401258	RAB44	RAB44, member RAS oncogene family	NA
29110	TBK1	TANK binding kinase 1	NA
200232	FAM209A	family with sequence similarity 209 member A	NA
81620	CDT1	chromatin licensing and DNA replication factor 1	NA
928	CD9	CD9 molecule	NA
10365	KLF2	Kruppel like factor 2	NA
7903	ST8SIA4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	NA
285550	FAM200B	family with sequence similarity 200 member B	NA
39	ACAT2	acetyl-CoA acetyltransferase 2	NA
8938	BAIAP3	BAI1 associated protein 3	NA
10044	SH2D3C	SH2 domain containing 3C	NA
10307	APBB3	amyloid beta precursor protein binding family B member 3	NA
55095	SAMD4B	sterile alpha motif domain containing 4B	NA
51439	FAM8A1	family with sequence similarity 8 member A1	NA
5325	PLAGL1	PLAG1 like zinc finger 1	NA
80714	PBX4	PBX homeobox 4	NA
90007	MIDN	midnolin	NA
100	ADA	adenosine deaminase	NA
57561	ARRDC3	arrestin domain containing 3	NA
4946	OAZ1	ornithine decarboxylase antizyme 1	NA
4830	NME1	NME/NM23 nucleoside diphosphate kinase 1	NA
63971	KIF13A	kinesin family member 13A	NA
89932	PAPLN	papilin, proteoglycan like sulfated glycoprotein	NA
51429	SNX9	sorting nexin 9	NA
7386	UQCRFS1	ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1	NA
80148	PQLC1	PQ loop repeat containing 1	NA
8682	PEA15	proliferation and apoptosis adaptor protein 15	NA
5515	PPP2CA	protein phosphatase 2 catalytic subunit alpha	NA
2678	GGT1	gamma-glutamyltransferase 1	NA

7050	TGIF1	TGFB induced factor homeobox 1	NA
79572	ATP13A3	ATPase 13A3	NA
51307	FAM53C	family with sequence similarity 53 member C	NA
3091	HIF1A	hypoxia inducible factor 1 subunit alpha	NA
56658	TRIM39	tripartite motif containing 39	NA
1535	CYBA	cytochrome b-245 alpha chain	NA
8942	KYNU	kynureninase	NA
5878	RAB5C	RAB5C, member RAS oncogene family	NA
6009	RHEB	Ras homolog, mTORC1 binding	NA
5093	PCBP1	poly(rC) binding protein 1	NA
126298	IRGQ	immunity related GTPase Q	NA
54926	UBE2R2	ubiquitin conjugating enzyme E2 R2	NA
81831	NETO2	neuropilin and tolloid like 2	NA
84975	MFSD5	major facilitator superfamily domain containing 5	NA
91445	RNF185	ring finger protein 185	NA
9516	LITAF	lipopolysaccharide induced TNF factor	NA
9559	VPS26A	VPS26, retromer complex component A	NA
54413	NLGN3	neuroligin 3	NA
29927	SEC61A1	Sec61 translocon alpha 1 subunit	NA
5127	CDK16	cyclin dependent kinase 16	NA
5091	PC	pyruvate carboxylase	NA
9541	CIR1	corepressor interacting with RBPJ, 1	NA
10209	EIF1	eukaryotic translation initiation factor 1	NA
5687	PSMA6	proteasome subunit alpha 6	NA
150465	TTL	tubulin tyrosine ligase	NA
2879	GPX4	glutathione peroxidase 4	NA
1647	GADD45A	growth arrest and DNA damage inducible alpha	NA
10488	CREB3	cAMP responsive element binding protein 3	NA
389537	OR7E38P	olfactory receptor family 7 subfamily E member 38 pseudogene	NA
114990	VASN	vasorin	NA
9361	LONP1	lon peptidase 1, mitochondrial	NA
147015	DHRS13	dehydrogenase/reductase 13	NA
6452	SH3BP2	SH3 domain binding protein 2	NA
10095	ARPC1B	actin related protein 2/3 complex subunit 1B	NA
3490	IGFBP7	insulin like growth factor binding protein 7	NA
91404	SESTD1	SEC14 and spectrin domain containing 1	NA
201163	FLCN	folliculin	NA
407975	MIR17HG	miR-17-92a-1 cluster host gene	NA
80727	TTYH3	tweety family member 3	NA
79789	CLMN	calmin	NA
5817	PVR	poliovirus receptor	NA
8834	TMEM11	transmembrane protein 11	NA
7709	ZBTB17	zinc finger and BTB domain containing 17	NA
51548	SIRT6	sirtuin 6	NA
55654	TMEM127	transmembrane protein 127	NA
27433	TOR2A	torsin family 2 member A	NA

79753	SNIP1	Smad nuclear interacting protein 1	NA
6840	SVIL	supervillin	NA
4839	NOP2	NOP2 nucleolar protein	NA
100874322	BACH1-IT2	BACH1 intronic transcript 2	NA
23207	PLEKHM2	pleckstrin homology and RUN domain containing M2	NA
7277	TUBA4A	tubulin alpha 4a	NA
11014	KDELR2	KDEL endoplasmic reticulum protein retention receptor 2	NA
753	LDLRAD4	low density lipoprotein receptor class A domain containing 4	NA
55893	ZNF395	zinc finger protein 395	NA
30845	EHD3	EH domain containing 3	NA
7030	TFE3	transcription factor binding to IGHM enhancer 3	NA
51371	POMP	proteasome maturation protein	NA
9700	ESPL1	extra spindle pole bodies like 1, separase	NA
10435	CDC42EP2	CDC42 effector protein 2	NA
27309	ZNF330	zinc finger protein 330	NA
26286	ARFGAP3	ADP ribosylation factor GTPase activating protein 3	NA
28514	DLL1	delta like canonical Notch ligand 1	NA
5272	SERPINB9	serpin family B member 9	NA
7251	TSG101	tumor susceptibility 101	NA
1119	CHKA	choline kinase alpha	NA
51005	AMDHD2	amidohydrolase domain containing 2	NA
2548	GAA	glucosidase alpha, acid	NA
5597	MAPK6	mitogen-activated protein kinase 6	NA
6039	RNASE6	ribonuclease A family member k6	NA
71	ACTG1	actin gamma 1	NA
1368	CPM	carboxypeptidase M	NA
22898	DENND3	DENN domain containing 3	NA
84299	MIEN1	migration and invasion enhancer 1	NA
2230	FDX1	ferredoxin 1	NA
4117	MAK	male germ cell associated kinase	NA
2634	GBP2	guanylate binding protein 2	NA
1890	TYMP	thymidine phosphorylase	NA
55851	PSENEN	presenilin enhancer, gamma-secretase subunit	NA
1612	DAPK1	death associated protein kinase 1	NA
22850	ADNP2	ADNP homeobox 2	NA
842	CASP9	caspase 9	NA
6535	SLC6A8	solute carrier family 6 member 8	NA
112483	SAT2	spermidine/spermine N1-acetyltransferase family member 2	NA
57182	ANKRD50	ankyrin repeat domain 50	NA
9907	AP5Z1	adaptor related protein complex 5 subunit zeta 1	NA
6281	S100A10	S100 calcium binding protein A10	NA
9170	LPAR2	lysophosphatidic acid receptor 2	NA
57007	ACKR3	atypical chemokine receptor 3	NA

3632	INPP5A	inositol polyphosphate-5-phosphatase A	NA
734	OSGIN2	oxidative stress induced growth inhibitor family member 2	NA
5954	RCN1	reticulocalbin 1	NA
4047	LSS	lanosterol synthase	NA
9592	IER2	immediate early response 2	NA
148823	GCSAML	germinal center associated signaling and motility like	NA
64342	HS1BP3	HCLS1 binding protein 3	NA
25801	GCA	grancalcin	NA
8099	CDK2AP1	cyclin dependent kinase 2 associated protein 1	NA
126208	ZNF787	zinc finger protein 787	NA
3122	HLA-DRA	major histocompatibility complex, class II, DR alpha	NA
941	CD80	CD80 molecule	NA
92	ACVR2A	activin A receptor type 2A	NA
3067	HDC	histidine decarboxylase	NA
79042	TSEN34	tRNA splicing endonuclease subunit 34	NA
6573	SLC19A1	solute carrier family 19 member 1	NA
122416	ANKRD9	ankyrin repeat domain 9	NA
10554	AGPAT1	1-acylglycerol-3-phosphate O-acyltransferase 1	NA
2201	FBN2	fibrillin 2	NA
23092	ARHGAP26	Rho GTPase activating protein 26	NA
932	MS4A3	membrane spanning 4-domains A3	NA
6184	RPN1	ribophorin I	NA
81502	HM13	histocompatibility minor 13	NA
340591	CA5BP1	carbonic anhydrase 5B pseudogene 1	NA
55669	MFN1	mitofusin 1	NA
4891	SLC11A2	solute carrier family 11 member 2	NA
27166	PRELID1	PRELI domain containing 1	NA
4580	MTX1	metaxin 1	NA
23761	PISD	phosphatidylserine decarboxylase	NA
64763	ZNF574	zinc finger protein 574	NA
51009	DERL2	derlin 2	NA
51094	ADIPOR1	adiponectin receptor 1	NA
5007	OSBP	oxysterol binding protein	NA
1453	CSNK1D	casein kinase 1 delta	NA
6525	SMTN	smoothelin	NA
136	ADORA2B	adenosine A2b receptor	NA
84868	HAVCR2	hepatitis A virus cellular receptor 2	NA
3311	HSPA7	heat shock protein family A (Hsp70) member 7	NA
93323	HAUS8	HAUS augmin like complex subunit 8	NA
91373	UAP1L1	UDP-N-acetylglucosamine pyrophosphorylase 1 like 1	NA
10296	MAEA	macrophage erythroblast attacher	NA
55356	SLC22A15	solute carrier family 22 member 15	NA
10094	ARPC3	actin related protein 2/3 complex subunit 3	NA
3385	ICAM3	intercellular adhesion molecule 3	NA

10169	SERF2	small EDRK-rich factor 2	NA
9130	FAM50A	family with sequence similarity 50 member A	NA
10491	CRTAP	cartilage associated protein	NA
146850	PIK3R6	phosphoinositide-3-kinase regulatory subunit 6	NA
7355	SLC35A2	solute carrier family 35 member A2	NA
55959	SULF2	sulfatase 2	NA
79176	FBXL15	F-box and leucine rich repeat protein 15	NA
54850	FBXL12	F-box and leucine rich repeat protein 12	NA
51133	KCTD3	potassium channel tetramerization domain containing 3	NA
3149	HMGB3	high mobility group box 3	NA
10165	SLC25A13	solute carrier family 25 member 13	NA
51125	GOLGA7	golgin A7	NA
7083	TK1	thymidine kinase 1	NA
55041	PLEKHB2	pleckstrin homology domain containing B2	NA
5192	PEX10	peroxisomal biogenesis factor 10	NA
1513	CTSK	cathepsin K	NA
4928	NUP98	nucleoporin 98	NA
2273	FHL1	four and a half LIM domains 1	NA
9938	ARHGAP25	Rho GTPase activating protein 25	NA
5718	PSMD12	proteasome 26S subunit, non-ATPase 12	NA
113402	SFT2D1	SFT2 domain containing 1	NA
374882	TMEM205	transmembrane protein 205	NA
91544	UBXN11	UBX domain protein 11	NA
5903	RANBP2	RAN binding protein 2	NA
81552	VOPP1	VOPP1, WBP1/VOPP1 family member	NA
5880	RAC2	Rac family small GTPase 2	NA
729230	CCR2	C-C motif chemokine receptor 2	NA
7328	UBE2H	ubiquitin conjugating enzyme E2 H	NA
79142	PHF23	PHD finger protein 23	NA
6732	SRPK1	SRSF protein kinase 1	NA
25824	PRDX5	peroxiredoxin 5	NA
387	RHOA	ras homolog family member A	NA
10809	STARD10	StAR related lipid transfer domain containing 10	NA
6628	SNRPB	small nuclear ribonucleoprotein polypeptides B and B1	NA
340348	TSPAN33	tetraspanin 33	NA
2000	ELF4	E74 like ETS transcription factor 4	NA
11282	MGAT4B	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase B	NA
1026	CDKN1A	cyclin dependent kinase inhibitor 1A	NA
7852	CXCR4	C-X-C motif chemokine receptor 4	NA
677	ZFP36L1	ZFP36 ring finger protein like 1	NA
7416	VDAC1	voltage dependent anion channel 1	NA
27154	BRPF3	bromodomain and PHD finger containing 3	NA
84817	TXNDC17	thioredoxin domain containing 17	NA
100216545	KMT2E-AS1	KMT2E antisense RNA 1	NA
54995	OXSM	3-oxoacyl-ACP synthase, mitochondrial	NA

2184	FAH	fumarylacetoacetate hydrolase	NA
11344	TWF2	twinfilin actin binding protein 2	NA
3628	INPP1	inositol polyphosphate-1-phosphatase	NA
860	RUNX2	runt related transcription factor 2	NA
6449	SGTA	small glutamine rich tetratricopeptide repeat containing alpha	NA
26608	TBL2	transducin beta like 2	NA
90441	ZNF622	zinc finger protein 622	NA
3570	IL6R	interleukin 6 receptor	NA
23616	SH3BP1	SH3 domain binding protein 1	NA
7114	TMSB4X	thymosin beta 4 X-linked	NA
5770	PTPN1	protein tyrosine phosphatase, non-receptor type 1	NA
1072	CFL1	cofilin 1	NA
7941	PLA2G7	phospholipase A2 group VII	NA
83658	DYNLRB1	dynein light chain roadblock-type 1	NA
2203	FBP1	fructose-bisphosphatase 1	NA
5894	RAF1	Raf-1 proto-oncogene, serine/threonine kinase	NA
54469	ZFAND6	zinc finger AN1-type containing 6	NA
10093	ARPC4	actin related protein 2/3 complex subunit 4	NA
80005	DOCK5	dedicator of cytokinesis 5	NA
83667	SESN2	sestrin 2	NA
388796	SNHG17	small nucleolar RNA host gene 17	NA
84919	PPP1R15B	protein phosphatase 1 regulatory subunit 15B	NA
54968	TMEM70	transmembrane protein 70	NA
7922	SLC39A7	solute carrier family 39 member 7	NA
23770	FKBP8	FK506 binding protein 8	NA
131408	FAM131A	family with sequence similarity 131 member A	NA
7936	NELFE	negative elongation factor complex member E	NA
220323	OAF	out at first homolog	NA
55004	LAMTOR1	late endosomal/lysosomal adaptor, MAPK and MTOR activator 1	NA
55113	XKR8	XK related 8	NA
7127	TNFAIP2	TNF alpha induced protein 2	NA
2934	GSN	gelsolin	NA
2647	BLOC1S1	biogenesis of lysosomal organelles complex 1 subunit 1	NA
5514	PPP1R10	protein phosphatase 1 regulatory subunit 10	NA
6782	HSPA13	heat shock protein family A (Hsp70) member 13	NA
56681	SAR1A	secretion associated Ras related GTPase 1A	NA
81565	NDEL1	nudE neurodevelopment protein 1 like 1	NA
84266	ALKBH7	alkB homolog 7	NA
79626	TNFAIP8L2	TNF alpha induced protein 8 like 2	NA
151987	PPP4R2	protein phosphatase 4 regulatory subunit 2	NA
3159	HMGA1	high mobility group AT-hook 1	NA
4130	MAP1A	microtubule associated protein 1A	NA
7320	UBE2B	ubiquitin conjugating enzyme E2 B	NA
6484	ST3GAL4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	NA

145567	TTC7B	tetratricopeptide repeat domain 7B	NA
11054	OGFR	opioid growth factor receptor	NA
23352	UBR4	ubiquitin protein ligase E3 component n-recognin 4	NA
28511	NKIRAS2	NFKB inhibitor interacting Ras like 2	NA
441212	RP9P	RP9 pseudogene	NA
9443	MED7	mediator complex subunit 7	NA
4780	NFE2L2	nuclear factor, erythroid 2 like 2	NA
147808	ZNF784	zinc finger protein 784	NA
4792	NFKBIA	NFKB inhibitor alpha	NA
4234	METTL1	methyltransferase like 1	NA
26355	FAM162A	family with sequence similarity 162 member A	NA
84769	MPV17L2	MPV17 mitochondrial inner membrane protein like 2	NA
5467	PPARD	peroxisome proliferator activated receptor delta	NA
64116	SLC39A8	solute carrier family 39 member 8	NA
9441	MED26	mediator complex subunit 26	NA
55422	ZNF331	zinc finger protein 331	NA
25915	NDUFAF3	NADH:ubiquinone oxidoreductase complex assembly factor 3	NA
5291	PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta	NA
2124	EVI2B	ecotropic viral integration site 2B	NA
6836	SURF4	surfeit 4	NA
10567	RABAC1	Rab acceptor 1	NA
29937	NENF	neudesin neurotrophic factor	NA
3475	IFRD1	interferon related developmental regulator 1	NA
56063	TMEM234	transmembrane protein 234	NA
381	ARF5	ADP ribosylation factor 5	NA
8896	BUD31	BUD31 homolog	NA
1031	CDKN2C	cyclin dependent kinase inhibitor 2C	NA
54386	TERF2IP	TERF2 interacting protein	NA
100128731	OST4	oligosaccharyltransferase complex subunit 4, non-catalytic	NA
10670	RRAGA	Ras related GTP binding A	NA
239	ALOX12	arachidonate 12-lipoxygenase, 12S type	NA
2885	GRB2	growth factor receptor bound protein 2	NA
6046	BRD2	bromodomain containing 2	NA
4682	NUBP1	nucleotide binding protein 1	NA
2286	FKBP2	FK506 binding protein 2	NA
51545	ZNF581	zinc finger protein 581	NA
51571	FAM49B	family with sequence similarity 49 member B	NA
200734	SPRED2	sprouty related EVH1 domain containing 2	NA
54458	PRR13	proline rich 13	NA
25932	CLIC4	chloride intracellular channel 4	NA
7321	UBE2D1	ubiquitin conjugating enzyme E2 D1	NA
11316	COPE	coatomer protein complex subunit epsilon	NA
59286	UBL5	ubiquitin like 5	NA

55845	BRK1	BRICK1, SCAR/WAVE actin nucleating complex subunit	NA
5089	PBX2	PBX homeobox 2	NA
100131017	ZNF316	zinc finger protein 316	NA
902	CCNH	cyclin H	NA
64114	TMBIM1	transmembrane BAX inhibitor motif containing 1	NA
10440	TIMM17A	translocase of inner mitochondrial membrane 17A	NA
79171	RBM42	RNA binding motif protein 42	NA
7782	SLC30A4	solute carrier family 30 member 4	NA
6103	RPGR	retinitis pigmentosa GTPase regulator	NA
79862	ZNF669	zinc finger protein 669	NA
9451	EIF2AK3	eukaryotic translation initiation factor 2 alpha kinase 3	NA
55577	NAGK	N-acetylglucosamine kinase	NA
11047	ADRM1	adhesion regulating molecule 1	NA
2002	ELK1	ELK1, ETS transcription factor	NA
84269	CHCHD5	coiled-coil-helix-coiled-coil-helix domain containing 5	NA
255919	CNEP1R1	CTD nuclear envelope phosphatase 1 regulatory subunit 1	NA
5710	PSMD4	proteasome 26S subunit, non-ATPase 4	NA
1652	DDT	D-dopachrome tautomerase	NA
80228	ORAI2	ORAI calcium release-activated calcium modulator 2	NA
23646	PLD3	phospholipase D family member 3	NA
23011	RAB21	RAB21, member RAS oncogene family	NA
137994	LETM2	leucine zipper and EF-hand containing transmembrane protein 2	NA
23158	TBC1D9	TBC1 domain family member 9	NA
84106	PRAM1	PML-RARA regulated adaptor molecule 1	NA
285381	DPH3	diphthamide biosynthesis 3	NA
5989	RFX1	regulatory factor X1	NA
2355	FOSL2	FOS like 2, AP-1 transcription factor subunit	NA
10950	BTG3	BTG anti-proliferation factor 3	NA
27338	UBE2S	ubiquitin conjugating enzyme E2 S	NA
79803	HPS6	HPS6, biogenesis of lysosomal organelles complex 2 subunit 3	NA
79594	MUL1	mitochondrial E3 ubiquitin protein ligase 1	NA
9871	SEC24D	SEC24 homolog D, COPII coat complex component	NA
9990	SLC12A6	solute carrier family 12 member 6	NA
142891	SAMD8	sterile alpha motif domain containing 8	NA
10652	YKT6	YKT6 v-SNARE homolog	NA
6388	SDF2	stromal cell derived factor 2	NA
23145	SSPO	SCO-spondin	NA
7803	PTP4A1	protein tyrosine phosphatase type IVA, member 1	NA
8263	F8A1	coagulation factor VIII associated 1	NA
7538	ZFP36	ZFP36 ring finger protein	NA
23787	MTCH1	mitochondrial carrier 1	NA
9839	ZEB2	zinc finger E-box binding homeobox 2	NA

81628	TSC22D4	TSC22 domain family member 4	NA
1871	E2F3	E2F transcription factor 3	NA
23214	XPO6	exportin 6	NA
8576	STK16	serine/threonine kinase 16	NA
100861453	DNAJC19P5	DnaJ heat shock protein family (Hsp40) member C19 pseudogene 5	NA
9741	LAPTM4A	lysosomal protein transmembrane 4 alpha	NA
10653	SPINT2	serine peptidase inhibitor, Kunitz type 2	NA
89853	MVB12B	multivesicular body subunit 12B	NA
100873638	RNA5SP383	RNA, 5S ribosomal pseudogene 383	NA
1106	CHD2	chromodomain helicase DNA binding protein 2	NA
64419	MTMR14	myotubularin related protein 14	NA
8744	TNFSF9	TNF superfamily member 9	NA
9218	VAPA	VAMP associated protein A	NA
28232	SLCO3A1	solute carrier organic anion transporter family member 3A1	NA
3014	H2AFX	H2A histone family member X	NA
83606	GUCD1	guanylyl cyclase domain containing 1	NA
6482	ST3GAL1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1	NA
90850	ZNF598	zinc finger protein 598	NA
84447	SYVN1	synoviolin 1	NA
51646	YPEL5	yippee like 5	NA
10978	CLP1	cleavage and polyadenylation factor I subunit 1	NA
7553	ZNF7	zinc finger protein 7	NA
9802	DAZAP2	DAZ associated protein 2	NA
8764	TNFRSF14	TNF receptor superfamily member 14	NA
124402	UBALD1	UBA like domain containing 1	NA
100129195	ZSCAN16-AS1	ZSCAN16 antisense RNA 1	NA
83982	IFI27L2	interferon alpha inducible protein 27 like 2	NA
140823	ROMO1	reactive oxygen species modulator 1	NA
22845	DOLK	dolichol kinase	NA
1201	CLN3	CLN3, battenin	NA
221830	TWISTNB	TWIST neighbor	NA
9535	GMFG	glia maturation factor gamma	NA
4236	MFAP1	microfibril associated protein 1	NA
64925	CCDC71	coiled-coil domain containing 71	NA
81873	ARPC5L	actin related protein 2/3 complex subunit 5 like	NA
79666	PLEKHF2	pleckstrin homology and FYVE domain containing 2	NA
84254	CAMKK1	calcium/calmodulin dependent protein kinase kinase 1	NA
784	CACNB3	calcium voltage-gated channel auxiliary subunit beta 3	NA
26093	CCDC9	coiled-coil domain containing 9	NA
57186	RALGAPA2	Ral GTPase activating protein catalytic alpha subunit 2	NA
26521	TIMM8B	translocase of inner mitochondrial membrane 8 homolog B	NA
7391	USF1	upstream transcription factor 1	NA

3608	ILF2	interleukin enhancer binding factor 2	NA
10313	RTN3	reticulon 3	NA
147007	TMEM199	transmembrane protein 199	NA
10524	KAT5	lysine acetyltransferase 5	NA
55957	LIN37	lin-37 DREAM MuvB core complex component	NA
27304	MOCS3	molybdenum cofactor synthesis 3	NA
5452	POU2F2	POU class 2 homeobox 2	NA
57410	SCYL1	SCY1 like pseudokinase 1	NA
2931	GSK3A	glycogen synthase kinase 3 alpha	NA
7392	USF2	upstream transcription factor 2, c-fos interacting	NA
2878	GPX3	glutathione peroxidase 3	NA
8569	MKNK1	MAP kinase interacting serine/threonine kinase 1	NA
1198	CLK3	CDC like kinase 3	NA
2771	GNAI2	G protein subunit alpha i2	NA
9618	TRAF4	TNF receptor associated factor 4	NA
51720	UIMC1	ubiquitin interaction motif containing 1	NA
10557	RPP38	ribonuclease P/MRP subunit p38	NA
9922	IQSEC1	IQ motif and Sec7 domain 1	NA
8408	ULK1	unc-51 like autophagy activating kinase 1	NA
1351	COX8A	cytochrome c oxidase subunit 8A	NA
6396	SEC13	SEC13 homolog, nuclear pore and COPII coat complex component	NA
222166	MTURN	maturin, neural progenitor differentiation regulator homolog	NA
5863	RGL2	ral guanine nucleotide dissociation stimulator like 2	NA
51605	TRMT6	tRNA methyltransferase 6	NA
7920	ABHD16A	abhydrolase domain containing 16A	NA
51141	INSIG2	insulin induced gene 2	NA
23135	KDM6B	lysine demethylase 6B	NA
406934	MIR142	microRNA 142	NA
9444	QKI	QKI, KH domain containing RNA binding	NA
5782	PTPN12	protein tyrosine phosphatase, non-receptor type 12	NA
260425	MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3	NA
26959	HBP1	HMG-box transcription factor 1	NA
7760	ZNF213	zinc finger protein 213	NA
10695	CNPY3	canopy FGF signaling regulator 3	NA
23474	ETHE1	ETHE1, persulfide dioxygenase	NA
25844	YIPF3	Yip1 domain family member 3	NA
11140	CDC37	cell division cycle 37	NA
9689	BZW1	basic leucine zipper and W2 domains 1	NA
5338	PLD2	phospholipase D2	NA
9798	IST1	IST1, ESCRT-III associated factor	NA
55700	MAP7D1	MAP7 domain containing 1	NA
5732	PTGER2	prostaglandin E receptor 2	NA
29886	SNX8	sorting nexin 8	NA
10362	HMG20B	high mobility group 20B	NA

808	CALM3	calmodulin 3	NA
11237	RNF24	ring finger protein 24	NA
51550	CINP	cyclin dependent kinase 2 interacting protein	NA
63874	ABHD4	abhydrolase domain containing 4	NA
1396	CRIP1	cysteine rich protein 1	NA
9367	RAB9A	RAB9A, member RAS oncogene family	NA
7185	TRAF1	TNF receptor associated factor 1	NA
8140	SLC7A5	solute carrier family 7 member 5	NA
57154	SMURF1	SMAD specific E3 ubiquitin protein ligase 1	NA
64005	MYO1G	myosin IG	NA
8073	PTP4A2	protein tyrosine phosphatase type IVA, member 2	NA
10376	TUBA1B	tubulin alpha 1b	NA
55352	COPRS	coordinator of PRMT5 and differentiation stimulator	NA
6166	RPL36AL	ribosomal protein L36a like	NA
10459	MAD2L2	mitotic arrest deficient 2 like 2	NA
65080	MRPL44	mitochondrial ribosomal protein L44	NA
23030	KDM4B	lysine demethylase 4B	NA
5595	MAPK3	mitogen-activated protein kinase 3	NA
7533	YWHAH	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein eta	NA
705	BYSL	bystin like	NA
1983	EIF5	eukaryotic translation initiation factor 5	NA
5208	PFKFB2	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2	NA
116988	AGAP3	ArfGAP with GTPase domain, ankyrin repeat and PH domain 3	NA
826	CAPNS1	calpain small subunit 1	NA
79134	TMEM185B	transmembrane protein 185B	NA
91582	RPS19BP1	ribosomal protein S19 binding protein 1	NA
29097	CNIH4	cornichon family AMPA receptor auxiliary protein 4	NA
60626	RIC8A	RIC8 guanine nucleotide exchange factor A	NA
727736	SH3GL1P3	SH3 domain containing GRB2 like 1, endophilin A2 pseudogene 3	NA
9043	SPAG9	sperm associated antigen 9	NA
10844	TUBGCP2	tubulin gamma complex associated protein 2	NA
2166	FAAH	fatty acid amide hydrolase	NA
56006	SMG9	SMG9, nonsense mediated mRNA decay factor	NA
3727	JUND	JunD proto-oncogene, AP-1 transcription factor subunit	NA
7879	RAB7A	RAB7A, member RAS oncogene family	NA
5049	PAFAH1B2	platelet activating factor acetylhydrolase 1b catalytic subunit 2	NA
4700	NDUFA6	NADH:ubiquinone oxidoreductase subunit A6	NA
84674	CARD6	caspase recruitment domain family member 6	NA
22906	TRAK1	trafficking kinesin protein 1	NA
4643	MYO1E	myosin IE	NA
5566	PRKACA	protein kinase cAMP-activated catalytic subunit alpha	NA

23529	CLCF1	cardiotrophin like cytokine factor 1	NA
4267	CD99	CD99 molecule (Xg blood group)	NA
2120	ETV6	ETS variant 6	NA
50999	TMED5	transmembrane p24 trafficking protein 5	NA
79798	ARMC5	armadillo repeat containing 5	NA
64412	GZF1	GDNF inducible zinc finger protein 1	NA
6397	SEC14L1	SEC14 like lipid binding 1	NA
6548	SLC9A1	solute carrier family 9 member A1	NA
9191	DEDD	death effector domain containing	NA
11094	CACFD1	calcium channel flower domain containing 1	NA
1175	AP2S1	adaptor related protein complex 2 subunit sigma 1	NA
5694	PSMB6	proteasome subunit beta 6	NA
10573	MRPL28	mitochondrial ribosomal protein L28	NA
8273	SLC10A3	solute carrier family 10 member 3	NA
359948	IRF2BP2	interferon regulatory factor 2 binding protein 2	NA
51121	RPL26L1	ribosomal protein L26 like 1	NA
83443	SF3B5	splicing factor 3b subunit 5	NA
56478	EIF4ENIF1	eukaryotic translation initiation factor 4E nuclear import factor 1	NA
91694	LONRF1	LON peptidase N-terminal domain and ring finger 1	NA
2869	GRK5	G protein-coupled receptor kinase 5	NA
10342	TFG	TRK-fused gene	NA
2801	GOLGA2	golgin A2	NA
91012	CERS5	ceramide synthase 5	NA
6050	RNH1	ribonuclease/angiogenin inhibitor 1	NA
400960	PCBP1-AS1	PCBP1 antisense RNA 1	NA
51142	CHCHD2	coiled-coil-helix-coiled-coil-helix domain containing 2	NA
8864	PER2	period circadian regulator 2	NA
63905	MANBAL	mannosidase beta like	NA
1195	CLK1	CDC like kinase 1	NA
81603	TRIM8	tripartite motif containing 8	NA
54918	CMTM6	CKLF like MARVEL transmembrane domain containing 6	NA
10870	HCST	hematopoietic cell signal transducer	NA
57461	ISY1	ISY1 splicing factor homolog	NA
1069	CETN2	centrin 2	NA
9448	MAP4K4	mitogen-activated protein kinase kinase kinase kinase 4	NA
351	APP	amyloid beta precursor protein	NA
22919	MAPRE1	microtubule associated protein RP/EB family member 1	NA
11179	ZNF277	zinc finger protein 277	NA
199746	U2AF1L4	U2 small nuclear RNA auxiliary factor 1 like 4	NA
1727	CYB5R3	cytochrome b5 reductase 3	NA
254102	EHBP1L1	EH domain binding protein 1 like 1	NA
55370	PPP4R1L	protein phosphatase 4 regulatory subunit 1 like (pseudogene)	NA

984	CDK11B	cyclin dependent kinase 11B	NA
25978	CHMP2B	charged multivesicular body protein 2B	NA
4833	NME4	NME/NM23 nucleoside diphosphate kinase 4	NA
310	ANXA7	annexin A7	NA
3108	HLA-DMA	major histocompatibility complex, class II, DM alpha	NA
3106	HLA-B	major histocompatibility complex, class I, B	NA
6717	SRI	sorcin	NA
382	ARF6	ADP ribosylation factor 6	NA
2950	GSTP1	glutathione S-transferase pi 1	NA
51228	GLTP	glycolipid transfer protein	NA
6793	STK10	serine/threonine kinase 10	NA
164	AP1G1	adaptor related protein complex 1 subunit gamma 1	NA
1376	CPT2	carnitine palmitoyltransferase 2	NA
84967	LSM10	LSM10, U7 small nuclear RNA associated	NA
7417	VDAC2	voltage dependent anion channel 2	NA
3398	ID2	inhibitor of DNA binding 2	NA
9777	TM9SF4	transmembrane 9 superfamily member 4	NA
1176	AP3S1	adaptor related protein complex 3 subunit sigma 1	NA
10241	CALCOCO2	calcium binding and coiled-coil domain 2	NA
3622	ING2	inhibitor of growth family member 2	NA
64857	PLEKHG2	pleckstrin homology and RhoGEF domain containing G2	NA
79571	GCC1	GRIP and coiled-coil domain containing 1	NA
8859	STK19	serine/threonine kinase 19	NA
8568	RRP1	ribosomal RNA processing 1	NA
154467	CCDC167	coiled-coil domain containing 167	NA
997	CDC34	cell division cycle 34	NA
201294	UNC13D	unc-13 homolog D	NA
219988	PATL1	PAT1 homolog 1, processing body mRNA decay factor	NA
9554	SEC22B	SEC22 homolog B, vesicle trafficking protein (gene/pseudogene)	NA
4953	ODC1	ornithine decarboxylase 1	NA
11267	SNF8	SNF8, ESCRT-II complex subunit	NA
5814	PURB	purine rich element binding protein B	NA
118487	CHCHD1	coiled-coil-helix-coiled-coil-helix domain containing 1	NA
23152	CIC	capicua transcriptional repressor	NA
7673	ZNF222	zinc finger protein 222	NA
130617	ZFAND2B	zinc finger AN1-type containing 2B	NA
116068	LYSMD3	LysM domain containing 3	NA
7067	THRA	thyroid hormone receptor alpha	NA
1774	DNASE1L1	deoxyribonuclease 1 like 1	NA
37	ACADVL	acyl-CoA dehydrogenase very long chain	NA
64072	CDH23	cadherin related 23	NA
6001	RGS10	regulator of G protein signaling 10	NA
29109	FHOD1	formin homology 2 domain containing 1	NA

160622	GRASP	general receptor for phosphoinositides 1 associated scaffold protein	NA
1340	COX6B1	cytochrome c oxidase subunit 6B1	NA
6881	TAF10	TATA-box binding protein associated factor 10	NA
6748	SSR4	signal sequence receptor subunit 4	NA
55937	APOM	apolipoprotein M	NA
9852	EPM2AIP1	EPM2A interacting protein 1	NA
5250	SLC25A3	solute carrier family 25 member 3	NA
1193	CLIC2	chloride intracellular channel 2	NA
51529	ANAPC11	anaphase promoting complex subunit 11	NA
80345	ZSCAN16	zinc finger and SCAN domain containing 16	NA
23212	RRS1	ribosome biogenesis regulator homolog	NA
8992	ATP6V0E1	ATPase H+ transporting V0 subunit e1	NA
23604	DAPK2	death associated protein kinase 2	NA
10010	TANK	TRAF family member associated NFKB activator	NA
84298	LLPH	LLP homolog, long-term synaptic facilitation factor	NA
9989	PPP4R1	protein phosphatase 4 regulatory subunit 1	NA
1211	CLTA	clathrin light chain A	NA
11171	STRAP	serine/threonine kinase receptor associated protein	NA
8560	DEGS1	delta 4-desaturase, sphingolipid 1	NA
4125	MAN2B1	mannosidase alpha class 2B member 1	NA
80207	OPA3	OPA3, outer mitochondrial membrane lipid metabolism regulator	NA
5216	PFN1	profilin 1	NA
170954	PPP1R18	protein phosphatase 1 regulatory subunit 18	NA
1349	COX7B	cytochrome c oxidase subunit 7B	NA
5868	RAB5A	RAB5A, member RAS oncogene family	NA
51343	FZR1	fizzy and cell division cycle 20 related 1	NA
548645	DNAJC25	DnaJ heat shock protein family (Hsp40) member C25	NA
29095	ORMDL2	ORMDL sphingolipid biosynthesis regulator 2	NA
9789	SPCS2	signal peptidase complex subunit 2	NA
7307	U2AF1	U2 small nuclear RNA auxiliary factor 1	NA
6903	TBCC	tubulin folding cofactor C	NA
64061	TSPYL2	TSPY like 2	NA
23399	CTDNEP1	CTD nuclear envelope phosphatase 1	NA
51398	WDR83OS	WD repeat domain 83 opposite strand	NA
10956	OS9	OS9, endoplasmic reticulum lectin	NA
131965	METTL6	methyltransferase like 6	NA
78994	PRR14	proline rich 14	NA
204	AK2	adenylate kinase 2	NA
10134	BCAP31	B cell receptor associated protein 31	NA
114785	MBD6	methyl-CpG binding domain protein 6	NA
26000	TBC1D10B	TBC1 domain family member 10B	NA
5692	PSMB4	proteasome subunit beta 4	NA
6195	RPS6KA1	ribosomal protein S6 kinase A1	NA
3801	KIFC3	kinesin family member C3	NA

27076	LYPD3	LY6/PLAUR domain containing 3	NA
10001	MED6	mediator complex subunit 6	NA
64320	RNF25	ring finger protein 25	NA
83855	KLF16	Kruppel like factor 16	NA
51291	GMIP	GEM interacting protein	NA
23220	DTX4	deltex E3 ubiquitin ligase 4	NA
64781	CERK	ceramide kinase	NA
3188	HNRNPH2	heterogeneous nuclear ribonucleoprotein H2	NA
353322	ANKRD37	ankyrin repeat domain 37	NA
54784	ALKBH4	alkB homolog 4, lysine demethylase	NA
4209	MEF2D	myocyte enhancer factor 2D	NA
1973	EIF4A1	eukaryotic translation initiation factor 4A1	NA
54985	HCFC1R1	host cell factor C1 regulator 1	NA
9616	RNF7	ring finger protein 7	NA
27013	CNPPD1	cyclin Pas1/PHO80 domain containing 1	NA
10542	LAMTOR5	late endosomal/lysosomal adaptor, MAPK and MTOR activator 5	NA
85025	TMEM60	transmembrane protein 60	NA
79797	ZNF408	zinc finger protein 408	NA
3104	ZBTB48	zinc finger and BTB domain containing 48	NA
1445	CSK	C-terminal Src kinase	NA
81037	CLPTM1L	CLPTM1 like	NA
23295	MGRN1	mahogunin ring finger 1	NA
4357	MPST	mercaptopyruvate sulfurtransferase	NA
83641	FAM107B	family with sequence similarity 107 member B	NA
4126	MANBA	mannosidase beta	NA
79897	RPP21	ribonuclease P/MRP subunit p21	NA
63935	PCIF1	PDX1 C-terminal inhibiting factor 1	NA
80854	SETD7	SET domain containing lysine methyltransferase 7	NA
6776	STAT5A	signal transducer and activator of transcription 5A	NA
100131801	PET100	PET100 homolog	NA
9612	NCOR2	nuclear receptor corepressor 2	NA
3156	HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase	NA
3516	RBPJ	recombination signal binding protein for immunoglobulin kappa J region	NA
529	ATP6V1E1	ATPase H+ transporting V1 subunit E1	NA
79165	LENG1	leukocyte receptor cluster member 1	NA
7170	TPM3	tropomyosin 3	NA
54998	AURKAIP1	aurora kinase A interacting protein 1	NA
541578	CXorf40B	chromosome X open reading frame 40B	NA
91283	MSANTD3	Myb/SANT DNA binding domain containing 3	NA
7324	UBE2E1	ubiquitin conjugating enzyme E2 E1	NA
998	CDC42	cell division cycle 42	NA
84187	TMEM164	transmembrane protein 164	NA
79585	CORO7	coronin 7	NA
100507297	TMEM44-AS1	TMEM44 antisense RNA 1	NA
54867	TMEM214	transmembrane protein 214	NA

1337	COX6A1	cytochrome c oxidase subunit 6A1	NA
5663	PSEN1	presenilin 1	NA
56900	TMEM167B	transmembrane protein 167B	NA
6238	RRBP1	ribosome binding protein 1	NA
56261	GPCPD1	glycerophosphocholine phosphodiesterase 1	NA
27020	NPTN	neuroplastin	NA
1603	DAD1	defender against cell death 1	NA
84282	RNF135	ring finger protein 135	NA
7532	YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma	NA
51316	PLAC8	placenta specific 8	NA
5908	RAP1B	RAP1B, member of RAS oncogene family	NA
4793	NFKBIB	NFKB inhibitor beta	NA
11070	TMEM115	transmembrane protein 115	NA
219541	MED19	mediator complex subunit 19	NA
4296	MAP3K11	mitogen-activated protein kinase kinase kinase 11	NA
10063	COX17	cytochrome c oxidase copper chaperone COX17	NA
51053	GMNN	geminin, DNA replication inhibitor	NA
483	ATP1B3	ATPase Na+/K+ transporting subunit beta 3	NA
51282	SCAND1	SCAN domain containing 1	NA
8705	B3GALT4	beta-1,3-galactosyltransferase 4	NA
9878	TOX4	TOX high mobility group box family member 4	NA
83695	RHNO1	RAD9-HUS1-RAD1 interacting nuclear orphan 1	NA
7458	EIF4H	eukaryotic translation initiation factor 4H	NA
4942	OAT	ornithine aminotransferase	NA
6746	SSR2	signal sequence receptor subunit 2	NA
4790	NFKB1	nuclear factor kappa B subunit 1	NA
55974	SLC50A1	solute carrier family 50 member 1	NA
220929	ZNF438	zinc finger protein 438	NA
6626	SNRPA	small nuclear ribonucleoprotein polypeptide A	NA
90522	YIF1B	Yip1 interacting factor homolog B, membrane trafficking protein	NA
7918	GPANK1	G-patch domain and ankyrin repeats 1	NA
284129	SLC26A11	solute carrier family 26 member 11	NA
6624	FSCN1	fascin actin-bundling protein 1	NA
6675	UAP1	UDP-N-acetylglucosamine pyrophosphorylase 1	NA
80347	COASY	Coenzyme A synthase	NA
23389	MED13L	mediator complex subunit 13 like	NA
124446	TMEM219	transmembrane protein 219	NA
9146	HGS	hepatocyte growth factor-regulated tyrosine kinase substrate	NA
54930	HAUS4	HAUS augmin like complex subunit 4	NA
5691	PSMB3	proteasome subunit beta 3	NA
90326	THAP3	THAP domain containing 3	NA
10113	PREB	prolactin regulatory element binding	NA
10569	SLU7	SLU7 homolog, splicing factor	NA
7415	VCP	valosin containing protein	NA

10109	ARPC2	actin related protein 2/3 complex subunit 2	NA
10745	PHTF1	putative homeodomain transcription factor 1	NA
79155	TNIP2	TNFAIP3 interacting protein 2	NA
6879	TAF7	TATA-box binding protein associated factor 7	NA
64743	WDR13	WD repeat domain 13	NA
8079	MLF2	myeloid leukemia factor 2	NA
2539	G6PD	glucose-6-phosphate dehydrogenase	NA
196883	ADCY4	adenylate cyclase 4	NA
81876	RAB1B	RAB1B, member RAS oncogene family	NA
80325	ABTB1	ankyrin repeat and BTB domain containing 1	NA
112950	MED8	mediator complex subunit 8	NA
27230	SERP1	stress associated endoplasmic reticulum protein 1	NA
11021	RAB35	RAB35, member RAS oncogene family	NA
7916	PRRC2A	proline rich coiled-coil 2A	NA
7511	XPNPEP1	X-prolyl aminopeptidase 1	NA
58496	LY6G5B	lymphocyte antigen 6 family member G5B	NA
9275	BCL7B	BCL tumor suppressor 7B	NA
29888	STRN4	striatin 4	NA
11078	TRIOBP	TRIO and F-actin binding protein	NA
10078	TSSC4	tumor suppressing subtransferable candidate 4	NA
8420	SNHG3	small nucleolar RNA host gene 3	NA
55207	ARL8B	ADP ribosylation factor like GTPase 8B	NA
6448	SGSH	N-sulfoglucosamine sulfohydrolase	NA
54776	PPP1R12C	protein phosphatase 1 regulatory subunit 12C	NA
6609	SMPD1	sphingomyelin phosphodiesterase 1	NA
5571	PRKAG1	protein kinase AMP-activated non-catalytic subunit gamma 1	NA
5362	PLXNA2	plexin A2	NA
5901	RAN	RAN, member RAS oncogene family	NA
54432	YIPF1	Yip1 domain family member 1	NA
6844	VAMP2	vesicle associated membrane protein 2	NA
56970	ATXN7L3	ataxin 7 like 3	NA
79778	MICALL2	MICAL like 2	NA
27173	SLC39A1	solute carrier family 39 member 1	NA
402055	SRRD	SRR1 domain containing	NA
79089	TMUB2	transmembrane and ubiquitin like domain containing 2	NA
51341	ZBTB7A	zinc finger and BTB domain containing 7A	NA
5688	PSMA7	proteasome subunit alpha 7	NA
126321	MFSD12	major facilitator superfamily domain containing 12	NA
10552	ARPC1A	actin related protein 2/3 complex subunit 1A	NA
85028	SNHG12	small nucleolar RNA host gene 12	NA
84330	ZNF414	zinc finger protein 414	NA
23466	CBX6	chromobox 6	NA
138716	RPP25L	ribonuclease P/MRP subunit p25 like	NA
1212	CLTB	clathrin light chain B	NA
24144	TFIP11	tuftelin interacting protein 11	NA

81875	ISG20L2	interferon stimulated exonuclease gene 20 like 2	NA
80700	UBXN6	UBX domain protein 6	NA
2197	FAU	FAU, ubiquitin like and ribosomal protein S30 fusion	NA
5202	PFDN2	prefoldin subunit 2	NA
113444	SMIM12	small integral membrane protein 12	NA
106480027	RNU6-1048P	RNA, U6 small nuclear 1048, pseudogene	NA
148641	SLC35F3	solute carrier family 35 member F3	NA
28948	IGKJ3	immunoglobulin kappa joining 3	NA
1823	DSC1	desmocollin 1	NA
22813	TDGF1P6	teratocarcinoma-derived growth factor 1 pseudogene 6	NA
6969	TRGJ2	T cell receptor gamma joining 2	NA
5342	PLGLB2	plasminogen-like B2	NA
164832	LONRF2	LON peptidase N-terminal domain and ring finger 2	NA
645160	RPL5P30	ribosomal protein L5 pseudogene 30	NA
8322	FZD4	frizzled class receptor 4	NA
9127	P2RX6	purinergic receptor P2X 6	NA
344787	ZNF860	zinc finger protein 860	NA
11181	TREH	trehalase	NA
392752	OR7E7P	olfactory receptor family 7 subfamily E member 7 pseudogene	NA
51421	AMOTL2	angiomotin like 2	NA
57615	ZNF492	zinc finger protein 492	NA
100271325	RPL36AP15	ribosomal protein L36a pseudogene 15	NA
220594	USP32P2	ubiquitin specific peptidase 32 pseudogene 2	NA
100873206	MTND5P1	MT-ND5 pseudogene 1	NA
642446	TRIM64B	tripartite motif containing 64B	NA
100506070	RBFADN	RBFA downstream neighbor	NA
406976	MIR199A1	microRNA 199a-1	NA
28751	TRAJ4	T cell receptor alpha joining 4	NA
283417	DPY19L2	dpy-19 like 2	NA
2900	GRIK4	glutamate ionotropic receptor kainate type subunit 4	NA
140766	ADAMTS14	ADAM metallopeptidase with thrombospondin type 1 motif 14	NA
1740	DLG2	discs large MAGUK scaffold protein 2	NA
100526835	FPGT-TNNI3K	FPGT-TNNI3K readthrough	NA
376693	RPS10P7	ribosomal protein S10 pseudogene 7	NA
653073	GOLGA8J	golgin A8 family member J	NA
28	ABO	ABO, alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase	NA
170541	PSMD10P1	proteasome 26S subunit, non-ATPase, 10 pseudogene 1	NA
400696	LGALS17A	galectin 14 pseudogene	NA
100873263	MTND5P25	MT-ND5 pseudogene 25	NA
219736	STOX1	storkhead box 1	NA
255877	BCL6B	B cell CLL/lymphoma 6B	NA
107057644	RN7SKP70	RNA, 7SK small nuclear pseudogene 70	NA

4897	NRCAM	neuronal cell adhesion molecule	NA
26534	OR10G2	olfactory receptor family 10 subfamily G member 2	NA
6343	SCT	secretin	NA
153478	PLEKHG4B	pleckstrin homology and RhoGEF domain containing G4B	NA
29850	TRPM5	transient receptor potential cation channel subfamily M member 5	NA
126069	ZNF491	zinc finger protein 491	NA
6493	SIM2	SIM bHLH transcription factor 2	NA
28517	TRDV2	T cell receptor delta variable 2	NA
100381270	ZBED6	zinc finger BED-type containing 6	NA
56103	PCDHGB2	protocadherin gamma subfamily B, 2	NA
143630	UBQLNL	ubiquilin like	NA
1109	AKR1C4	aldo-keto reductase family 1 member C4	NA
8973	CHRNA6	cholinergic receptor nicotinic alpha 6 subunit	NA
28455	IGHV2-26	immunoglobulin heavy variable 2-26	NA
28692	TRAV1-2	T cell receptor alpha variable 1-2	NA
143872	ARHGAP42	Rho GTPase activating protein 42	NA
653316	FAM153C	family with sequence similarity 153 member C	NA
693206	MIR621	microRNA 621	NA
401427	OR2A7	olfactory receptor family 2 subfamily A member 7	NA
727909	GOLGA8Q	golgin A8 family member Q	NA
100873245	MTND4P23	MT-ND4 pseudogene 23	NA
100312828	VN1R83P	vomeronasal 1 receptor 83 pseudogene	NA
648947	AKR1C7P	aldo-keto reductase family 1 member C7, pseudogene	NA
142680	SLC34A3	solute carrier family 34 member 3	NA
131890	GRK7	G protein-coupled receptor kinase 7	NA
2099	ESR1	estrogen receptor 1	NA
128229	TSACC	TSSK6 activating cochaperone	NA
93034	NT5C1B	5'-nucleotidase, cytosolic IB	NA
60494	CCDC81	coiled-coil domain containing 81	NA
85379	KIAA1671	KIAA1671	NA
401409	RAB19	RAB19, member RAS oncogene family	NA
100874001	NCBP2-AS1	NCBP2 antisense RNA 1	NA
100312826	VN1R81P	vomeronasal 1 receptor 81 pseudogene	NA
387036	GUSBP2	glucuronidase, beta pseudogene 2	NA
10643	IGF2BP3	insulin like growth factor 2 mRNA binding protein 3	NA
100420920	ZDHHC4P1	zinc finger DHHC-type containing 4 pseudogene 1	NA
2619	GAS1	growth arrest specific 1	NA
23285	KIAA1107	KIAA1107	NA
134728	IRAK1BP1	interleukin 1 receptor associated kinase 1 binding protein 1	NA
23086	EXPH5	exophilin 5	NA
23322	RPGRIP1L	RPGRIP1 like	NA
1768	DNAH6	dynein axonemal heavy chain 6	NA

81248	OR52P1P	olfactory receptor family 52 subfamily P member 1 pseudogene	NA
440078	FAM66C	family with sequence similarity 66 member C	NA
692234	SNORD103A	small nucleolar RNA, C/D box 103A	NA
11162	NUDT6	nudix hydrolase 6	NA
55304	SPTLC3	serine palmitoyltransferase long chain base subunit 3	NA
1638	DCT	dopachrome tautomerase	NA
100507300	ALKBH3-AS1	ALKBH3 antisense RNA 1	NA
6972	TRGJP2	T cell receptor gamma joining P2	NA
55296	TBC1D19	TBC1 domain family member 19	NA
645700	ZNF890P	zinc finger protein 890, pseudogene	NA
100506686	IQCH-AS1	IQCH antisense RNA 1	NA
101927682	CD81-AS1	CD81 antisense RNA 1	NA
57664	PLEKHA4	pleckstrin homology domain containing A4	NA
765	CA6	carbonic anhydrase 6	NA
28941	IGKV1-9	immunoglobulin kappa variable 1-9	NA
342909	ZNF284	zinc finger protein 284	NA
284161	GDPD1	glycerophosphodiester phosphodiesterase domain containing 1	NA
100505678	STARD4-AS1	STARD4 antisense RNA 1	NA
105377105	FLNB-AS1	FLNB antisense RNA 1	NA
196968	DNM1P46	dynamin 1 pseudogene 46	NA
553112	FERP1	FER tyrosine kinase pseudogene 1	NA
100874214	TM4SF19-AS1	TM4SF19 antisense RNA 1	NA
100874110	GLYCTK-AS1	GLYCTK antisense RNA 1	NA
10913	EDAR	ectodysplasin A receptor	NA
85417	CCNB3	cyclin B3	NA
28457	IGHV2-5	immunoglobulin heavy variable 2-5	NA
146822	CDRT15	CMT1A duplicated region transcript 15	NA
50937	CDON	cell adhesion associated, oncogene regulated	NA
64283	ARHGEF28	Rho guanine nucleotide exchange factor 28	NA
643699	GOLGA8N	golgin A8 family member N	NA
114827	FHAD1	forkhead associated phosphopeptide binding domain 1	NA
81854	ZNF205-AS1	ZNF205 antisense RNA 1	NA
25937	WWTR1	WW domain containing transcription regulator 1	NA
51209	RAB9B	RAB9B, member RAS oncogene family	NA
54674	LRRN3	leucine rich repeat neuronal 3	NA
203238	CCDC171	coiled-coil domain containing 171	NA
727826	RPS11P5	ribosomal protein S11 pseudogene 5	NA
163183	SYNE4	spectrin repeat containing nuclear envelope family member 4	NA
7694	ZNF135	zinc finger protein 135	NA
1573	CYP2J2	cytochrome P450 family 2 subfamily J member 2	NA
138639	PTPDC1	protein tyrosine phosphatase domain containing 1	NA
22806	IKZF3	IKAROS family zinc finger 3	NA
3818	KLKB1	kallikrein B1	NA

100874007	VIPR1-AS1	VIPR1 antisense RNA 1	NA
100132116	ACTA2-AS1	ACTA2 antisense RNA 1	NA
100130742	LRRC69	leucine rich repeat containing 69	NA
79413	ZBED2	zinc finger BED-type containing 2	NA
7617	ZNF66	zinc finger protein 66	NA
100507582	BHLHE40-AS1	BHLHE40 antisense RNA 1	NA
163351	GBP6	guanylate binding protein family member 6	NA
7621	ZNF70	zinc finger protein 70	NA
2315	MLANA	melan-A	NA
1268	CNR1	cannabinoid receptor 1	NA
27085	MTBP	MDM2 binding protein	NA
387104	SOGA3	SOGA family member 3	NA
92565	FANK1	fibronectin type III and ankyrin repeat domains 1	NA
3812	KIR3DL2	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 2	NA
5357	PLS1	plastin 1	NA
84911	ZNF382	zinc finger protein 382	NA
162963	ZNF610	zinc finger protein 610	NA
9456	HOMER1	homer scaffold protein 1	NA
2491	CENPI	centromere protein I	NA
650655	ABCA17P	ATP binding cassette subfamily A member 17, pseudogene	NA
1894	ECT2	epithelial cell transforming 2	NA
124783	SPATA32	spermatogenesis associated 32	NA
285440	CYP4V2	cytochrome P450 family 4 subfamily V member 2	NA
1092	CEACAMP3	carcinoembryonic antigen related cell adhesion molecule pseudogene 3	NA
440288	UBL7-AS1	UBL7 antisense RNA 1 (head to head)	NA
84239	ATP13A4	ATPase 13A4	NA
390816	THA1P	threonine aldolase 1, pseudogene	NA
80333	KCNIP4	potassium voltage-gated channel interacting protein 4	NA
649946	RPL23AP64	ribosomal protein L23a pseudogene 64	NA
26974	ZNF285	zinc finger protein 285	NA
6334	SCN8A	sodium voltage-gated channel alpha subunit 8	NA
388182	SPATA41	spermatogenesis associated 41	NA
124535	HSF5	heat shock transcription factor 5	NA
286205	SCAI	suppressor of cancer cell invasion	NA
641518	LEF1-AS1	LEF1 antisense RNA 1	NA
158431	ZNF782	zinc finger protein 782	NA
63898	SH2D4A	SH2 domain containing 4A	NA
1952	CELSR2	cadherin EGF LAG seven-pass G-type receptor 2	NA
100499483	CCDC180	coiled-coil domain containing 180	NA
654231	OCM	oncomodulin	NA
84100	ARL6	ADP ribosylation factor like GTPase 6	NA
6444	SGCD	sarcoglycan delta	NA
441733	PRKXP1	PRKX pseudogene 1	NA
1621	DBH	dopamine beta-hydroxylase	NA

140807	KRT72	keratin 72	NA
30832	ZNF354C	zinc finger protein 354C	NA
4306	NR3C2	nuclear receptor subfamily 3 group C member 2	NA
100507266	STX18-AS1	STX18 antisense RNA 1 (head to head)	NA
139322	APOOL	apolipoprotein O like	NA
3627	CXCL10	C-X-C motif chemokine ligand 10	NA
6956	TRAV6	T cell receptor alpha variable 6	NA
128553	TSHZ2	teashirt zinc finger homeobox 2	NA
79598	CEP97	centrosomal protein 97	NA
79839	CCDC102B	coiled-coil domain containing 102B	NA
8411	EEA1	early endosome antigen 1	NA
28690	TRAV3	T cell receptor alpha variable 3 (gene/pseudogene)	NA
319101	KRT73	keratin 73	NA
59277	NTN4	netrin 4	NA
100118954	NRBF2P5	nuclear receptor binding factor 2 pseudogene 5	NA
10896	OCLM	oculomedin	NA
284349	ZNF283	zinc finger protein 283	NA
100131089	SRP14-AS1	SRP14 antisense RNA1 (head to head)	NA
657	BMPR1A	bone morphogenetic protein receptor type 1A	NA
80310	PDGFD	platelet derived growth factor D	NA
10223	GPA33	glycoprotein A33	NA
28653	TRAV29DV5	T cell receptor alpha variable 29/delta variable 5 (gene/pseudogene)	NA
56164	STK31	serine/threonine kinase 31	NA
345778	MTX3	metaxin 3	NA
100049076	GUSBP9	glucuronidase, beta pseudogene 9	NA
100887750	MRPS31P5	mitochondrial ribosomal protein S31 pseudogene 5	NA
22807	IKZF2	IKAROS family zinc finger 2	NA
672	BRCA1	BRCA1, DNA repair associated	NA
284194	LGALS9B	galectin 9B	NA
2838	GPR15	G protein-coupled receptor 15	NA
100528032	KLRC4-KLRK1	KLRC4-KLRK1 readthrough	NA
163131	ZNF780B	zinc finger protein 780B	NA
387751	GVINP1	GTPase, very large interferon inducible pseudogene 1	NA
84073	MYCBPAP	MYCBP associated protein	NA
26047	CNTNAP2	contactin associated protein like 2	NA
387492	DSTNP1	destrin, actin depolymerizing factor pseudogene 1	NA
1310	COL19A1	collagen type XIX alpha 1 chain	NA
644297	FAM215B	family with sequence similarity 215 member B	NA
283571	PROX2	prospero homeobox 2	NA
285498	RNF212	ring finger protein 212	NA
401474	SAMD12	sterile alpha motif domain containing 12	NA
8496	PPFIBP1	PPFIA binding protein 1	NA
28601	TRBV6-6	T cell receptor beta variable 6-6	NA
284252	KCTD1	potassium channel tetramerization domain containing 1	NA
133522	PPARGC1B	PPARG coactivator 1 beta	NA

151556	GPR155	G protein-coupled receptor 155	NA
55175	KLHL11	kelch like family member 11	NA
4168	MCF2	MCF.2 cell line derived transforming sequence	NA
8857	FCGBP	Fc fragment of IgG binding protein	NA
22852	ANKRD26	ankyrin repeat domain 26	NA
89884	LHX4	LIM homeobox 4	NA
4862	NPAS2	neuronal PAS domain protein 2	NA
5152	PDE9A	phosphodiesterase 9A	NA
54532	USP53	ubiquitin specific peptidase 53	NA
9994	CASP8AP2	caspase 8 associated protein 2	NA
92797	HELB	DNA helicase B	NA
2310	FOXO3B	forkhead box O3B pseudogene	NA
11174	ADAMTS6	ADAM metallopeptidase with thrombospondin type 1 motif 6	NA
56172	ANKH	ANKH inorganic pyrophosphate transport regulator	NA
606495	CYB5RL	cytochrome b5 reductase like	NA
285598	ARL10	ADP ribosylation factor like GTPase 10	NA
4646	MYO6	myosin VI	NA
9866	TRIM66	tripartite motif containing 66	NA
7915	ALDH5A1	aldehyde dehydrogenase 5 family member A1	NA
199704	ZNF585A	zinc finger protein 585A	NA
127428	TCEANC2	transcription elongation factor A N-terminal and central domain containing 2	NA
55691	FRMD4A	FERM domain containing 4A	NA
10936	GPR75	G protein-coupled receptor 75	NA
725	C4BPB	complement component 4 binding protein beta	NA
84443	FRMPD3	FERM and PDZ domain containing 3	NA
79841	AGBL2	ATP/GTP binding protein like 2	NA
56624	ASAH2	N-acylsphingosine amidohydrolase 2	NA
1773	DNASE1	deoxyribonuclease 1	NA
100507218	RNF157-AS1	RNF157 antisense RNA 1	NA
126626	GABPB2	GA binding protein transcription factor subunit beta 2	NA
30837	SOCS7	suppressor of cytokine signaling 7	NA
342618	SLFN14	schlafen family member 14	NA
22986	SORCS3	sortilin related VPS10 domain containing receptor 3	NA
659	BMPR2	bone morphogenetic protein receptor type 2	NA
654346	LGALS9C	galectin 9C	NA
647121	EMBP1	embigin pseudogene 1	NA
221322	TBC1D32	TBC1 domain family member 32	NA
83481	EPPK1	epiplakin 1	NA
9832	JAKMIP2	janus kinase and microtubule interacting protein 2	NA
59084	ENPP5	ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative)	NA
100506915	CHRM3-AS2	CHRM3 antisense RNA 2	NA
79899	PRR5L	proline rich 5 like	NA
100289635	ZNF605	zinc finger protein 605	NA

54830	NUP62CL	nucleoporin 62 C-terminal like	NA
389114	ZNF662	zinc finger protein 662	NA
472	ATM	ATM serine/threonine kinase	NA
1380	CR2	complement C3d receptor 2	NA
79632	FAM184A	family with sequence similarity 184 member A	NA
101927983	EDRF1-AS1	EDRF1 antisense RNA 1	NA
257068	PLCXD2	phosphatidylinositol specific phospholipase C X domain containing 2	NA
23467	NPTXR	neuronal pentraxin receptor	NA
55252	ASXL2	ASXL transcriptional regulator 2	NA
50649	ARHGEF4	Rho guanine nucleotide exchange factor 4	NA
6296	ACSM3	acyl-CoA synthetase medium chain family member 3	NA
64097	EPB41L4A	erythrocyte membrane protein band 4.1 like 4A	NA
84976	DISP1	dispatched RND transporter family member 1	NA
130399	ACVR1C	activin A receptor type 1C	NA
28434	IGHV3-33	immunoglobulin heavy variable 3-33	NA
339327	ZNF546	zinc finger protein 546	NA
152877	FAM53A	family with sequence similarity 53 member A	NA
64105	CENPK	centromere protein K	NA
7275	TUB	tubby bipartite transcription factor	NA
342926	ZNF677	zinc finger protein 677	NA
112849	L3HYPDH	trans-L-3-hydroxyproline dehydratase	NA
80206	FHOD3	formin homology 2 domain containing 3	NA
28616	TRBV4-2	T cell receptor beta variable 4-2	NA
28685	TRAV8-1	T cell receptor alpha variable 8-1	NA
494470	RNF165	ring finger protein 165	NA
100310846	ANKRD61	ankyrin repeat domain 61	NA
340252	ZNF680	zinc finger protein 680	NA
79800	CARF	calcium responsive transcription factor	NA
100507549	HMGB3P32	high mobility group box 3 pseudogene 32	NA
57683	ZDBF2	zinc finger DBF-type containing 2	NA
28682	TRAV8-4	T cell receptor alpha variable 8-4	NA
440482	ANKRD20A5P	ankyrin repeat domain 20 family member A5, pseudogene	NA
93627	TBCK	TBC1 domain containing kinase	NA
124923	SGK494	uncharacterized serine/threonine-protein kinase SgK494	NA
5796	PTPRK	protein tyrosine phosphatase, receptor type K	NA
5382	PMS2P4	PMS1 homolog 2, mismatch repair system component pseudogene 4	NA
129049	SGSM1	small G protein signaling modulator 1	NA
122773	KLHDC1	kelch domain containing 1	NA
440253	WHAMMP2	WAS protein homolog associated with actin, golgi membranes and microtubules pseudogene 2	NA
55285	RBM41	RNA binding motif protein 41	NA
83448	PUS7L	pseudouridine synthase 7 like	NA
10495	ENOX2	ecto-NOX disulfide-thiol exchanger 2	NA
125206	SLC5A10	solute carrier family 5 member 10	NA

55073	LRRC37A4P	leucine rich repeat containing 37 member A4, pseudogene	NA
23382	AHCYL2	adenosylhomocysteinase like 2	NA
6490	PMEL	premelanosome protein	NA
4012	LNPEP	leucyl and cystinyl aminopeptidase	NA
100506736	SLFN12L	schlafen family member 12 like	NA
2121	EVC	EvC ciliary complex subunit 1	NA
158135	TTLL11	tubulin tyrosine ligase like 11	NA
596	BCL2	BCL2, apoptosis regulator	NA
100996567	ANTXRLP1	ANTXR like pseudogene 1	NA
8671	SLC4A4	solute carrier family 4 member 4	NA
5378	PMS1	PMS1 homolog 1, mismatch repair system component	NA
57209	ZNF248	zinc finger protein 248	NA
57669	EPB41L5	erythrocyte membrane protein band 4.1 like 5	NA
107105253	RNU6-322P	RNA, U6 small nuclear 322, pseudogene	NA
140862	ISM1	isthmin 1	NA
64167	ERAP2	endoplasmic reticulum aminopeptidase 2	NA
55328	RNLS	renalase, FAD dependent amine oxidase	NA
79613	TANGO6	transport and golgi organization 6 homolog	NA
554225	STRCP1	stereocilin pseudogene 1	NA
1062	CENPE	centromere protein E	NA
644397	LRRC37A17P	leucine rich repeat containing 37 member A17, pseudogene	NA
26249	KLHL3	kelch like family member 3	NA
642517	AGAP9	ArfGAP with GTPase domain, ankyrin repeat and PH domain 9	NA
10160	FARP1	FERM, ARH/RhoGEF and pleckstrin domain protein 1	NA
646300	COL6A4P2	collagen type VI alpha 4 pseudogene 2	NA
9310	ZNF235	zinc finger protein 235	NA
152185	SPICE1	spindle and centriole associated protein 1	NA
8821	INPP4B	inositol polyphosphate-4-phosphatase type II B	NA
100129792	CCDC152	coiled-coil domain containing 152	NA
6003	RGS13	regulator of G protein signaling 13	NA
2786	GNG4	G protein subunit gamma 4	NA
51348	KLRF1	killer cell lectin like receptor F1	NA
4610	MYCL	MYCL proto-oncogene, bHLH transcription factor	NA
51463	GPR89B	G protein-coupled receptor 89B	NA
4649	MYO9A	myosin IXA	NA
340351	AGBL3	ATP/GTP binding protein like 3	NA
22997	IGSF9B	immunoglobulin superfamily member 9B	NA
374946	DRAXIN	dorsal inhibitory axon guidance protein	NA
51540	SCLY	selenocysteine lyase	NA
4750	NEK1	NIMA related kinase 1	NA
80055	PGAP1	post-GPI attachment to proteins 1	NA
283314	C1RL-AS1	C1RL antisense RNA 1	NA
100418744	KRT8P46	keratin 8 pseudogene 46	NA

91975	ZNF300	zinc finger protein 300	NA
57511	COG6	component of oligomeric golgi complex 6	NA
9881	TRANK1	tetratricopeptide repeat and ankyrin repeat containing 1	NA
128178	EDARADD	EDAR associated death domain	NA
6983	TRGV9	T cell receptor gamma variable 9	NA
23331	TTC28	tetratricopeptide repeat domain 28	NA
84515	MCM8	minichromosome maintenance 8 homologous recombination repair factor	NA
79675	FASTKD1	FAST kinase domains 1	NA
3077	HFE	homeostatic iron regulator	NA
100129460	DPY19L1P1	DPY19L1 pseudogene 1	NA
11278	KLF12	Kruppel like factor 12	NA
546	ATRX	ATRX, chromatin remodeler	NA
729375	FAM86HP	family with sequence similarity 86, member A pseudogene	NA
28557	TRBV30	T cell receptor beta variable 30 (gene/pseudogene)	NA
57181	SLC39A10	solute carrier family 39 member 10	NA
55796	MBNL3	muscleblind like splicing regulator 3	NA
100287898	TTC34	tetratricopeptide repeat domain 34	NA
10592	SMC2	structural maintenance of chromosomes 2	NA
55333	SYNJ2BP	synaptojanin 2 binding protein	NA
728121	CSPG4P12	chondroitin sulfate proteoglycan 4 pseudogene 12	NA
64864	RFX7	regulatory factor X7	NA
57474	ZNF490	zinc finger protein 490	NA
10157	AASS	aminoadipate-semialdehyde synthase	NA
9194	SLC16A7	solute carrier family 16 member 7	NA
55251	PCMTD2	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 2	NA
23475	QPRT	quinolinate phosphoribosyltransferase	NA
6489	ST8SIA1	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	NA
342892	ZNF850	zinc finger protein 850	NA
28655	TRAV27	T cell receptor alpha variable 27	NA
112487	DTD2	D-tyrosyl-tRNA deacylase 2 (putative)	NA
375593	TRIM73	tripartite motif containing 73	NA
4329	ALDH6A1	aldehyde dehydrogenase 6 family member A1	NA
83394	PITPNM3	PITPNM family member 3	NA
9475	ROCK2	Rho associated coiled-coil containing protein kinase 2	NA
399671	HEATR4	HEAT repeat containing 4	NA
200894	ARL13B	ADP ribosylation factor like GTPase 13B	NA
84612	PARD6B	par-6 family cell polarity regulator beta	NA
100131827	ZNF717	zinc finger protein 717	NA
55653	BCAS4	breast carcinoma amplified sequence 4	NA
26049	FAM169A	family with sequence similarity 169 member A	NA
259283	MDS2	myelodysplastic syndrome 2 translocation associated	NA
147166	TRIM16L	tripartite motif containing 16 like	NA

23060	ZNF609	zinc finger protein 609	NA
90624	LYRM7	LYR motif containing 7	NA
150709	ANKAR	ankyrin and armadillo repeat containing	NA
84144	SYDE2	synapse defective Rho GTPase homolog 2	NA
729597	SPDYE6	speedy/RINGO cell cycle regulator family member E6	NA
114804	RNF157	ring finger protein 157	NA
54765	TRIM44	tripartite motif containing 44	NA
57185	NIPAL3	NIPA like domain containing 3	NA
728953	RPS19P3	ribosomal protein S19 pseudogene 3	NA
157869	SBSPON	somatomedin B and thrombospondin type 1 domain containing	NA
222256	CDHR3	cadherin related family member 3	NA
5959	RDH5	retinol dehydrogenase 5	NA
79414	LRFN3	leucine rich repeat and fibronectin type III domain containing 3	NA
693220	MIR635	microRNA 635	NA
3655	ITGA6	integrin subunit alpha 6	NA
10771	ZMYND11	zinc finger MYND-type containing 11	NA
23348	DOCK9	dedicator of cytokinesis 9	NA
54903	MKS1	Meckel syndrome, type 1	NA
100885789	IFNG-AS1	IFNG antisense RNA 1	NA
6920	TCEA3	transcription elongation factor A3	NA
51474	LIMA1	LIM domain and actin binding 1	NA
22858	ICK	intestinal cell kinase	NA
6322	SCML1	Scm polycomb group protein like 1	NA
160418	TMTC3	transmembrane and tetratricopeptide repeat containing 3	NA
23335	WDR7	WD repeat domain 7	NA
257218	SHPRH	SNF2 histone linker PHD RING helicase	NA
89978	DPH6	diphthamine biosynthesis 6	NA
6935	ZEB1	zinc finger E-box binding homeobox 1	NA
100129482	ZNF37BP	zinc finger protein 37B, pseudogene	NA
6297	SALL2	spalt like transcription factor 2	NA
899	CCNF	cyclin F	NA
92255	LMBRD2	LMBR1 domain containing 2	NA
26259	FBXW8	F-box and WD repeat domain containing 8	NA
202134	FAM153B	family with sequence similarity 153 member B	NA
126375	ZNF792	zinc finger protein 792	NA
57579	FAM135A	family with sequence similarity 135 member A	NA
360023	ZBTB41	zinc finger and BTB domain containing 41	NA
118980	SFXN2	sideroflexin 2	NA
440248	HERC2P9	hect domain and RLD 2 pseudogene 9	NA
646080	RAB11FIP1P1	RAB11 family interacting protein 1 pseudogene 1	NA
92521	SPECC1	sperm antigen with calponin homology and coiled-coil domains 1	NA
84224	NBPF3	NBPF member 3	NA
5934	RBL2	RB transcriptional corepressor like 2	NA

253558	LCLAT1	lysocardiolipin acyltransferase 1	NA
23547	LILRA4	leukocyte immunoglobulin like receptor A4	NA
8115	TCL1A	T cell leukemia/lymphoma 1A	NA
3	A2MP1	alpha-2-macroglobulin pseudogene 1	NA
55888	ZKSCAN7	zinc finger with KRAB and SCAN domains 7	NA
9057	SLC7A6	solute carrier family 7 member 6	NA
8722	CTSF	cathepsin F	NA
5727	PTCH1	patched 1	NA
79723	SUV39H2	suppressor of variegation 3-9 homolog 2	NA
493	ATP2B4	ATPase plasma membrane Ca2+ transporting 4	NA
594	BCKDHB	branched chain keto acid dehydrogenase E1 subunit beta	NA
3213	HOXB3	homeobox B3	NA
93	ACVR2B	activin A receptor type 2B	NA
28693	TRAV1-1	T cell receptor alpha variable 1-1	NA
101928492	OR2A1-AS1	OR2A1 antisense RNA 1	NA
339829	CCDC39	coiled-coil domain containing 39	NA
57506	MAVS	mitochondrial antiviral signaling protein	NA
114814	GNRHR2	gonadotropin releasing hormone receptor 2 (pseudogene)	NA
4436	MSH2	mutS homolog 2	NA
200558	APLF	aprataxin and PNKP like factor	NA
83990	BRIP1	BRCA1 interacting protein C-terminal helicase 1	NA
55614	KIF16B	kinesin family member 16B	NA
4280	MICE	MHC class I polypeptide-related sequence E (pseudogene)	NA
643641	ZNF862	zinc finger protein 862	NA
57560	IFT80	intraflagellar transport 80	NA
10290	SPEG	SPEG complex locus	NA
2241	FER	FER tyrosine kinase	NA
93474	ZNF670	zinc finger protein 670	NA
28677	TRAV9-2	T cell receptor alpha variable 9-2	NA
94015	TTYH2	tweety family member 2	NA
151742	PPM1L	protein phosphatase, Mg2+/Mn2+ dependent 1L	NA
84295	PHF6	PHD finger protein 6	NA
6815	STYX	serine/threonine/tyrosine interacting protein	NA
92285	ZNF585B	zinc finger protein 585B	NA
160897	GPR180	G protein-coupled receptor 180	NA
26053	AUTS2	AUTS2, activator of transcription and developmental regulator	NA
79694	MANEA	mannosidase endo-alpha	NA
285346	ZNF852	zinc finger protein 852	NA
34	ACADM	acyl-CoA dehydrogenase medium chain	NA
152485	ZNF827	zinc finger protein 827	NA
55166	CENPQ	centromere protein Q	NA
647174	SERPINE3	serpin family E member 3	NA
340481	ZDHHC21	zinc finger DHHC-type containing 21	NA

8029	CUBN	cubilin	NA
10466	COG5	component of oligomeric golgi complex 5	NA
57509	MTUS1	microtubule associated scaffold protein 1	NA
55601	DDX60	DExD/H-box helicase 60	NA
9786	KIAA0586	KIAA0586	NA
390213	DOC2GP	double C2 domain gamma pseudogene	NA
26040	SETBP1	SET binding protein 1	NA
79690	GAL3ST4	galactose-3-O-sulfotransferase 4	NA
22909	FAN1	FANCD2 and FANCI associated nuclease 1	NA
286128	ZFP41	ZFP41 zinc finger protein	NA
353497	POLN	DNA polymerase nu	NA
29760	BLNK	B cell linker	NA
126017	ZNF813	zinc finger protein 813	NA
90102	PHLDB2	pleckstrin homology like domain family B member 2	NA
79659	DYNC2H1	dynein cytoplasmic 2 heavy chain 1	NA
117145	THEM4	thioesterase superfamily member 4	NA
84961	FBXL20	F-box and leucine rich repeat protein 20	NA
28672	TRAV12-3	T cell receptor alpha variable 12-3	NA
79068	FTO	FTO, alpha-ketoglutarate dependent dioxygenase	NA
9650	MTFR1	mitochondrial fission regulator 1	NA
7571	ZNF23	zinc finger protein 23	NA
5991	RFX3	regulatory factor X3	NA
814	CAMK4	calcium/calmodulin dependent protein kinase IV	NA
2135	EXTL2	exostosin like glycosyltransferase 2	NA
11169	WDHD1	WD repeat and HMG-box DNA binding protein 1	NA
148213	ZNF681	zinc finger protein 681	NA
7840	ALMS1	ALMS1, centrosome and basal body associated protein	NA
100130600	CICP14	capicua transcriptional repressor pseudogene 14	NA
389677	RBM12B	RNA binding motif protein 12B	NA
26146	TRAF3IP1	TRAF3 interacting protein 1	NA
29104	N6AMT1	N-6 adenine-specific DNA methyltransferase 1	NA
10780	ZNF234	zinc finger protein 234	NA
825	CAPN3	calpain 3	NA
147686	ZNF418	zinc finger protein 418	NA
85315	PAQR8	progestin and adipoQ receptor family member 8	NA
285025	CCDC141	coiled-coil domain containing 141	NA
80012	PHC3	polyhomeotic homolog 3	NA
285596	FAM153A	family with sequence similarity 153 member A	NA
225	ABCD2	ATP binding cassette subfamily D member 2	NA
285268	ZNF621	zinc finger protein 621	NA
100131031	ANKRD18EP	ankyrin repeat domain 18E, pseudogene	NA
4040	LRP6	LDL receptor related protein 6	NA
158801	NKAPP1	NFKB activating protein pseudogene 1	NA
100272205	TATDN2P2	TatD DNase domain containing 2 pseudogene 2	NA
90529	STPG1	sperm tail PG-rich repeat containing 1	NA

26512	INTS6	integrator complex subunit 6	NA
10152	ABI2	abl interactor 2	NA
5140	PDE3B	phosphodiesterase 3B	NA
23098	SARM1	sterile alpha and TIR motif containing 1	NA
56479	KCNQ5	potassium voltage-gated channel subfamily Q member 5	NA
221914	GPC2	glypican 2	NA
84874	ZNF514	zinc finger protein 514	NA
2875	GPT	glutamic--pyruvic transaminase	NA
374900	ZNF568	zinc finger protein 568	NA
55814	BDP1	B double prime 1, subunit of RNA polymerase III transcription initiation factor IIIB	NA
32	ACACB	acetyl-CoA carboxylase beta	NA
54916	TMEM260	transmembrane protein 260	NA
128486	FITM2	fat storage inducing transmembrane protein 2	NA
11016	ATF7	activating transcription factor 7	NA
196951	FAM227B	family with sequence similarity 227 member B	NA
9849	ZNF518A	zinc finger protein 518A	NA
80318	GKAP1	G kinase anchoring protein 1	NA
22866	CNKSR2	connector enhancer of kinase suppressor of Ras 2	NA
84460	ZMAT1	zinc finger matrin-type 1	NA
100131193	CCDC183-AS1	CCDC183 antisense RNA 1	NA
55227	LRRC1	leucine rich repeat containing 1	NA
4603	MYBL1	MYB proto-oncogene like 1	NA
100129842	ZNF737	zinc finger protein 737	NA
84636	GPR174	G protein-coupled receptor 174	NA
57380	MRS2	magnesium transporter MRS2	NA
117608	ZNF354B	zinc finger protein 354B	NA
353189	SLCO4C1	solute carrier organic anion transporter family member 4C1	NA
375190	FAM228B	family with sequence similarity 228 member B	NA
9580	SOX13	SRY-box 13	NA
3676	ITGA4	integrin subunit alpha 4	NA
80014	WWC2	WW and C2 domain containing 2	NA
148103	ZNF599	zinc finger protein 599	NA
162967	ZNF320	zinc finger protein 320	NA
55435	AP1AR	adaptor related protein complex 1 associated regulatory protein	NA
26035	GLCE	glucuronic acid epimerase	NA
55609	ZNF280C	zinc finger protein 280C	NA
25913	POT1	protection of telomeres 1	NA
28609	TRBV5-6	T cell receptor beta variable 5-6	NA
55605	KIF21A	kinesin family member 21A	NA
85442	KNDC1	kinase non-catalytic C-lobe domain containing 1	NA
387921	NHLRC3	NHL repeat containing 3	NA
11215	AKAP11	A-kinase anchoring protein 11	NA
390927	ZNF793	zinc finger protein 793	NA
728411	GUSBP1	glucuronidase, beta pseudogene 1	NA

6619	SNAPC3	small nuclear RNA activating complex polypeptide 3	NA
123036	TC2N	tandem C2 domains, nuclear	NA
23272	FAM208A	family with sequence similarity 208 member A	NA
25938	HEATR5A	HEAT repeat containing 5A	NA
23413	NCS1	neuronal calcium sensor 1	NA
55320	MIS18BP1	MIS18 binding protein 1	NA
441273	SPDYE2	speedy/RINGO cell cycle regulator family member E2	NA
23229	ARHGEF9	Cdc42 guanine nucleotide exchange factor 9	NA
80267	EDEM3	ER degradation enhancing alpha-mannosidase like protein 3	NA
4753	NELL2	neural EGFL like 2	NA
27065	NSG1	neuronal vesicle trafficking associated 1	NA
401647	GOLGA7B	golgin A7 family member B	NA
256987	SERINC5	serine incorporator 5	NA
100419031	PRDX3P1	peroxiredoxin 3 pseudogene 1	NA
26610	ELP4	elongator acetyltransferase complex subunit 4	NA
10384	BTN3A3	butyrophilin subfamily 3 member A3	NA
23269	MGA	MGA, MAX dimerization protein	NA
4285	MIPEP	mitochondrial intermediate peptidase	NA
5924	RASGRF2	Ras protein specific guanine nucleotide releasing factor 2	NA
8642	DCHS1	dachsous cadherin-related 1	NA
51520	LARS	leucyl-tRNA synthetase	NA
349152	DPY19L2P2	DPY19L2 pseudogene 2	NA
6584	SLC22A5	solute carrier family 22 member 5	NA
91687	CENPL	centromere protein L	NA
100129387	GABPB1-AS1	GABPB1 antisense RNA 1	NA
54467	ANKIB1	ankyrin repeat and IBR domain containing 1	NA
7754	ZNF204P	zinc finger protein 204, pseudogene	NA
253832	ZDHHC20	zinc finger DHHC-type containing 20	NA
137872	ADHFE1	alcohol dehydrogenase, iron containing 1	NA
7691	ZNF132	zinc finger protein 132	NA
10231	RCAN2	regulator of calcineurin 2	NA
284323	ZNF780A	zinc finger protein 780A	NA
390980	ZNF805	zinc finger protein 805	NA
3097	HIVEP2	human immunodeficiency virus type I enhancer binding protein 2	NA
376940	ZC3H6	zinc finger CCCH-type containing 6	NA
90594	ZNF439	zinc finger protein 439	NA
143903	LAYN	layilin	NA
2218	FKTN	fukutin	NA
80823	BHLHB9	basic helix-loop-helix family member b9	NA
100462772	FTH1P22	ferritin heavy chain 1 pseudogene 22	NA
54901	CDKAL1	CDK5 regulatory subunit associated protein 1 like 1	NA
54832	VPS13C	vacuolar protein sorting 13 homolog C	NA
23705	CADM1	cell adhesion molecule 1	NA

100507062	PSMD6-AS2	PSMD6 antisense RNA 2	NA
148789	B3GALNT2	beta-1,3-N-acetylgalactosaminyltransferase 2	NA
153642	ARSK	arylsulfatase family member K	NA
64854	USP46	ubiquitin specific peptidase 46	NA
9651	PLCH2	phospholipase C eta 2	NA
64375	IKZF4	IKAROS family zinc finger 4	NA
147923	ZNF420	zinc finger protein 420	NA
2053	EPHX2	epoxide hydrolase 2	NA
5465	PPARA	peroxisome proliferator activated receptor alpha	NA
28683	TRAV8-3	T cell receptor alpha variable 8-3	NA
60312	AFAP1	actin filament associated protein 1	NA
29068	ZBTB44	zinc finger and BTB domain containing 44	NA
157567	ANKRD46	ankyrin repeat domain 46	NA
3631	INPP4A	inositol polyphosphate-4-phosphatase type I A	NA
5420	PODXL	podocalyxin like	NA
7294	TXK	TXK tyrosine kinase	NA
57476	GRAMD1B	GRAM domain containing 1B	NA
5578	PRKCA	protein kinase C alpha	NA
57639	CCDC146	coiled-coil domain containing 146	NA
146540	ZNF785	zinc finger protein 785	NA
162655	ZNF519	zinc finger protein 519	NA
84542	KIAA1841	KIAA1841	NA
25966	C2CD2	C2 calcium dependent domain containing 2	NA
79677	SMC6	structural maintenance of chromosomes 6	NA
4775	NFATC3	nuclear factor of activated T cells 3	NA
127933	UHMK1	U2AF homology motif kinase 1	NA
6558	SLC12A2	solute carrier family 12 member 2	NA
85444	LRRCC1	leucine rich repeat and coiled-coil centrosomal protein 1	NA
81571	MIR600HG	MIR600 host gene	NA
6563	SLC14A1	solute carrier family 14 member 1 (Kidd blood group)	NA
80350	LPAL2	lipoprotein(a) like 2, pseudogene	NA
4773	NFATC2	nuclear factor of activated T cells 2	NA
728340	GTF2H2C	GTF2H2 family member C	NA
79820	CATSPERB	cation channel sperm associated auxiliary subunit beta	NA
79848	CSPP1	centrosome and spindle pole associated protein 1	NA
101060200	ZNF891	zinc finger protein 891	NA
130026	ICA1L	islet cell autoantigen 1 like	NA
5800	PTPRO	protein tyrosine phosphatase, receptor type O	NA
285386	TPRG1	tumor protein p63 regulated 1	NA
283209	PGM2L1	phosphoglucomutase 2 like 1	NA
222223	KIAA1324L	KIAA1324 like	NA
5334	PLCL1	phospholipase C like 1 (inactive)	NA
79844	ZDHHC11	zinc finger DHHC-type containing 11	NA
114826	SMYD4	SET and MYND domain containing 4	NA
23015	GOLGA8A	golgin A8 family member A	NA

54841	BIVM	basic, immunoglobulin-like variable motif containing	NA
7273	TTN	titin	NA
28968	SLC6A16	solute carrier family 6 member 16	NA
7582	ZNF33B	zinc finger protein 33B	NA
8503	PIK3R3	phosphoinositide-3-kinase regulatory subunit 3	NA
29904	EEF2K	eukaryotic elongation factor 2 kinase	NA
55592	GOLGA2P5	GOLGA2 pseudogene 5	NA
8835	SOCS2	suppressor of cytokine signaling 2	NA
54549	SDK2	sidekick cell adhesion molecule 2	NA
9459	ARHGEF6	Rac/Cdc42 guanine nucleotide exchange factor 6	NA
79634	SCRN3	secernin 3	NA
339500	ZNF678	zinc finger protein 678	NA
22862	FNDC3A	fibronectin type III domain containing 3A	NA
7626	ZNF75D	zinc finger protein 75D	NA
63939	FAM217B	family with sequence similarity 217 member B	NA
10734	STAG3	stromal antigen 3	NA
23556	PIGN	phosphatidylinositol glycan anchor biosynthesis class N	NA
5205	ATP8B1	ATPase phospholipid transporting 8B1	NA
5422	POLA1	DNA polymerase alpha 1, catalytic subunit	NA
91526	ANKRD44	ankyrin repeat domain 44	NA
9497	SLC4A7	solute carrier family 4 member 7	NA
1352	COX10	cytochrome c oxidase assembly factor heme A farnesyltransferase COX10	NA
55840	EAF2	ELL associated factor 2	NA
131544	CRYBG3	crystallin beta-gamma domain containing 3	NA
1795	DOCK3	dedicator of cytokinesis 3	NA
5933	RBL1	RB transcriptional corepressor like 1	NA
55006	TRMT61B	tRNA methyltransferase 61B	NA
80321	CEP70	centrosomal protein 70	NA
387119	CEP85L	centrosomal protein 85 like	NA
56937	PMEPA1	prostate transmembrane protein, androgen induced 1	NA
25896	INTS7	integrator complex subunit 7	NA
64753	CCDC136	coiled-coil domain containing 136	NA
388558	ZNF808	zinc finger protein 808	NA
80205	CHD9	chromodomain helicase DNA binding protein 9	NA
6004	RGS16	regulator of G protein signaling 16	NA
51301	GCNT4	glucosaminyl (N-acetyl) transferase 4, core 2	NA
1528	CYB5A	cytochrome b5 type A	NA
54495	TMX3	thioredoxin related transmembrane protein 3	NA
440352	SNX29P2	sorting nexin 29 pseudogene 2	NA
162966	ZNF600	zinc finger protein 600	NA
4240	MFGE8	milk fat globule-EGF factor 8 protein	NA
79830	ZMYM1	zinc finger MYM-type containing 1	NA
63979	FIGNL1	fidgetin like 1	NA
79815	NIPAL2	NIPA like domain containing 2	NA

81856	ZNF611	zinc finger protein 611	NA
28558	TRBV29-1	T cell receptor beta variable 29-1	NA
115350	FCRL1	Fc receptor like 1	NA
9824	ARHGAP11A	Rho GTPase activating protein 11A	NA
254225	RNF169	ring finger protein 169	NA
165	AEBP1	AE binding protein 1	NA
55622	TTC27	tetratricopeptide repeat domain 27	NA
8924	HERC2	HECT and RLD domain containing E3 ubiquitin protein ligase 2	NA
969	CD69	CD69 molecule	NA
51361	HOOK1	hook microtubule tethering protein 1	NA
63901	FAM111A	family with sequence similarity 111 member A	NA
53405	CLIC5	chloride intracellular channel 5	NA
55164	SHQ1	SHQ1, H/ACA ribonucleoprotein assembly factor	NA
84947	SERAC1	serine active site containing 1	NA
57593	EBF4	EBF family member 4	NA
3908	LAMA2	laminin subunit alpha 2	NA
55722	CEP72	centrosomal protein 72	NA
54438	GFOD1	glucose-fructose oxidoreductase domain containing 1	NA
54497	HEATR5B	HEAT repeat containing 5B	NA
158219	TTC39B	tetratricopeptide repeat domain 39B	NA
353274	ZNF445	zinc finger protein 445	NA
7098	TLR3	toll like receptor 3	NA
120526	DNAJC24	DnaJ heat shock protein family (Hsp40) member C24	NA
6894	TARBP1	TAR (HIV-1) RNA binding protein 1	NA
401124	DTHD1	death domain containing 1	NA
79925	SPEF2	sperm flagellar 2	NA
10667	FARS2	phenylalanyl-tRNA synthetase 2, mitochondrial	NA
10198	MPHOSPH9	M-phase phosphoprotein 9	NA
6502	SKP2	S-phase kinase associated protein 2	NA
79809	TTC21B	tetratricopeptide repeat domain 21B	NA
2313	FLI1	Fli-1 proto-oncogene, ETS transcription factor	NA
29767	TMOD2	tropomodulin 2	NA
57697	FANCM	FA complementation group M	NA
221264	AK9	adenylate kinase 9	NA
64328	XPO4	exportin 4	NA
5189	PEX1	peroxisomal biogenesis factor 1	NA
84955	NUDCD1	NudC domain containing 1	NA
9949	AMMECR1	Alport syndrome, mental retardation, midface hypoplasia and elliptocytosis chromosomal region gene 1	NA
26586	CKAP2	cytoskeleton associated protein 2	NA
8681	JMJD7-PLA2G4B	JMJD7-PLA2G4B readthrough	NA
2323	FLT3LG	fms related tyrosine kinase 3 ligand	NA
29121	CLEC2D	C-type lectin domain family 2 member D	NA
53344	CHIC1	cysteine rich hydrophobic domain 1	NA

36	ACADSB	acyl-CoA dehydrogenase short/branched chain	NA
9465	AKAP7	A-kinase anchoring protein 7	NA
219854	TMEM218	transmembrane protein 218	NA
7220	TRPC1	transient receptor potential cation channel subfamily C member 1	NA
55703	POLR3B	RNA polymerase III subunit B	NA
148867	SLC30A7	solute carrier family 30 member 7	NA
1429	CRYZ	crystallin zeta	NA
9731	CEP104	centrosomal protein 104	NA
10283	CWC27	CWC27 spliceosome associated protein homolog	NA
149420	PDIK1L	PDLIM1 interacting kinase 1 like	NA
84315	MON1A	MON1 homolog A, secretory trafficking associated	NA
197131	UBR1	ubiquitin protein ligase E3 component n-recognin 1	NA
222484	LNX2	ligand of numb-protein X 2	NA
28978	TMEM14A	transmembrane protein 14A	NA
3708	ITPR1	inositol 1,4,5-trisphosphate receptor type 1	NA
9498	SLC4A8	solute carrier family 4 member 8	NA
10526	IPO8	importin 8	NA
23408	SIRT5	sirtuin 5	NA
219899	TBCEL	tubulin folding cofactor E like	NA
124411	ZNF720	zinc finger protein 720	NA
115752	DIS3L	DIS3 like exosome 3'-5' exoribonuclease	NA
89801	PPP1R3F	protein phosphatase 1 regulatory subunit 3F	NA
7267	TTC3	tetratricopeptide repeat domain 3	NA
286148	DPY19L4	dpy-19 like 4	NA
2820	GPD2	glycerol-3-phosphate dehydrogenase 2	NA
200316	APOBEC3F	apolipoprotein B mRNA editing enzyme catalytic subunit 3F	NA
79699	ZYG11B	zyg-11 family member B, cell cycle regulator	NA
60437	CDH26	cadherin 26	NA
27067	STAU2	staufen double-stranded RNA binding protein 2	NA
148206	ZNF714	zinc finger protein 714	NA
11093	ADAMTS13	ADAM metallopeptidase with thrombospondin type 1 motif 13	NA
5125	PCSK5	proprotein convertase subtilisin/kexin type 5	NA
10873	ME3	malic enzyme 3	NA
55728	N4BP2	NEDD4 binding protein 2	NA
64921	CASD1	CAS1 domain containing 1	NA
154796	AMOT	angiomotin	NA
140707	BRI3BP	BRI3 binding protein	NA
254048	UBN2	ubinuclein 2	NA
119	ADD2	adducin 2	NA
390595	UBAP1L	ubiquitin associated protein 1 like	NA
64318	NOC3L	NOC3 like DNA replication regulator	NA
1021	CDK6	cyclin dependent kinase 6	NA
28568	TRBV19	T cell receptor beta variable 19	NA
25780	RASGRP3	RAS guanyl releasing protein 3	NA

442578	STAG3L3	stromal antigen 3-like 3 (pseudogene)	NA
55070	DET1	DET1, COP1 ubiquitin ligase partner	NA
126068	ZNF441	zinc finger protein 441	NA
23317	DNAJC13	DnaJ heat shock protein family (Hsp40) member C13	NA
23137	SMC5	structural maintenance of chromosomes 5	NA
153364	MBLAC2	metallo-beta-lactamase domain containing 2	NA
92283	ZNF461	zinc finger protein 461	NA
3292	HSD17B1	hydroxysteroid 17-beta dehydrogenase 1	NA
549	AUH	AU RNA binding methylglutaconyl-CoA hydratase	NA
79956	ERMP1	endoplasmic reticulum metallopeptidase 1	NA
5664	PSEN2	presenilin 2	NA
27340	UTP20	UTP20, small subunit processome component	NA
9781	RNF144A	ring finger protein 144A	NA
23421	ITGB3BP	integrin subunit beta 3 binding protein	NA
10000	AKT3	AKT serine/threonine kinase 3	NA
9581	PREPL	prolyl endopeptidase like	NA
6683	SPAST	spastin	NA
317762	CCDC85C	coiled-coil domain containing 85C	NA
57134	MAN1C1	mannosidase alpha class 1C member 1	NA
6596	HLTF	helicase like transcription factor	NA
178	AGL	amylo-alpha-1, 6-glucosidase, 4-alpha-glucanotransferase	NA
51575	ESF1	ESF1 nucleolar pre-rRNA processing protein homolog	NA
3991	LIPE	lipase E, hormone sensitive type	NA
1266	CNN3	calponin 3	NA
55297	CCDC91	coiled-coil domain containing 91	NA
142940	TRUB1	TruB pseudouridine synthase family member 1	NA
80724	ACAD10	acyl-CoA dehydrogenase family member 10	NA
91433	RCCD1	RCC1 domain containing 1	NA
6984	TRGV10	T cell receptor gamma variable 10 (non-functional)	NA
79018	GID4	GID complex subunit 4 homolog	NA
26046	LTN1	listerin E3 ubiquitin protein ligase 1	NA
11101	ATE1	arginyltransferase 1	NA
57534	MIB1	mindbomb E3 ubiquitin protein ligase 1	NA
26235	FBXL4	F-box and leucine rich repeat protein 4	NA
10450	PPIE	peptidylprolyl isomerase E	NA
1565	CYP2D6	cytochrome P450 family 2 subfamily D member 6	NA
54989	ZNF770	zinc finger protein 770	NA
374879	ZNF699	zinc finger protein 699	NA
55188	RIC8B	RIC8 guanine nucleotide exchange factor B	NA
54760	PCSK4	proprotein convertase subtilisin/kexin type 4	NA
81610	FAM83D	family with sequence similarity 83 member D	NA
146330	FBXL16	F-box and leucine rich repeat protein 16	NA
100506084	ARL17B	ADP ribosylation factor like GTPase 17B	NA
84671	ZNF347	zinc finger protein 347	NA
23230	VPS13A	vacuolar protein sorting 13 homolog A	NA

23530	NNT	nicotinamide nucleotide transhydrogenase	NA
388228	SBK1	SH3 domain binding kinase 1	NA
10178	TENM1	teneurin transmembrane protein 1	NA
10396	ATP8A1	ATPase phospholipid transporting 8A1	NA
79600	TCTN1	tectonic family member 1	NA
57504	MTA3	metastasis associated 1 family member 3	NA
26057	ANKRD17	ankyrin repeat domain 17	NA
400673	VMAC	vimentin type intermediate filament associated coiled-coil protein	NA
84105	PCBD2	pterin-4 alpha-carbinolamine dehydratase 2	NA
6801	STRN	striatin	NA
7589	ZSCAN21	zinc finger and SCAN domain containing 21	NA
66035	SLC2A11	solute carrier family 2 member 11	NA
27241	BBS9	Bardet-Biedl syndrome 9	NA
23461	ABCA5	ATP binding cassette subfamily A member 5	NA
84312	BRMS1L	BRMS1 like transcriptional repressor	NA
113251	LARP4	La ribonucleoprotein domain family member 4	NA
57553	MICAL3	microtubule associated monooxygenase, calponin and LIM domain containing 3	NA
22982	DIP2C	disco interacting protein 2 homolog C	NA
54737	MPHOSPH8	M-phase phosphoprotein 8	NA
6641	SNTB1	syntrophin beta 1	NA
23051	ZHX3	zinc fingers and homeoboxes 3	NA
196074	METTL15	methyltransferase like 15	NA
84910	TMEM87B	transmembrane protein 87B	NA
284058	KANSL1	KAT8 regulatory NSL complex subunit 1	NA
286410	ATP11C	ATPase phospholipid transporting 11C	NA
7029	TFDP2	transcription factor Dp-2	NA
79731	NARS2	asparaginyl-tRNA synthetase 2, mitochondrial	NA
80313	LRRC27	leucine rich repeat containing 27	NA
729082	OIP5-AS1	OIP5 antisense RNA 1	NA
7587	ZNF37A	zinc finger protein 37A	NA
140775	SMCR8	Smith-Magenis syndrome chromosome region, candidate 8	NA
23368	PPP1R13B	protein phosphatase 1 regulatory subunit 13B	NA
84912	SLC35B4	solute carrier family 35 member B4	NA
3660	IRF2	interferon regulatory factor 2	NA
65981	CAPRIN2	caprin family member 2	NA
56987	BBX	BBX, HMG-box containing	NA
100507034	PITRM1-AS1	PITRM1 antisense RNA 1	NA
7625	ZNF74	zinc finger protein 74	NA
3141	HLCS	holocarboxylase synthetase	NA
51449	PCYOX1	prenylcysteine oxidase 1	NA
6671	SP4	Sp4 transcription factor	NA
26751	SH3YL1	SH3 and SYLF domain containing 1	NA
55342	STRBP	spermatid perinuclear RNA binding protein	NA
57646	USP28	ubiquitin specific peptidase 28	NA

285527	FRYL	FRY like transcription coactivator	NA
10026	PIGK	phosphatidylinositol glycan anchor biosynthesis class K	NA
59	ACTA2	actin, alpha 2, smooth muscle, aorta	NA
9840	TESPA1	thymocyte expressed, positive selection associated 1	NA
27087	B3GAT1	beta-1,3-glucuronyltransferase 1	NA
114134	SLC2A13	solute carrier family 2 member 13	NA
8396	PIP4K2B	phosphatidylinositol-5-phosphate 4-kinase type 2 beta	NA
79657	RPAP3	RNA polymerase II associated protein 3	NA
50650	ARHGEF3	Rho guanine nucleotide exchange factor 3	NA
64172	OSGEPL1	O-sialoglycoprotein endopeptidase like 1	NA
143888	KDELC2	KDEL motif containing 2	NA
10666	CD226	CD226 molecule	NA
23155	CLCC1	chloride channel CLIC like 1	NA
28673	TRAV12-2	T cell receptor alpha variable 12-2	NA
23515	MORC3	MORC family CW-type zinc finger 3	NA
8573	CASK	calcium/calmodulin dependent serine protein kinase	NA
3710	ITPR3	inositol 1,4,5-trisphosphate receptor type 3	NA
11123	RCAN3	RCAN family member 3	NA
157378	TMEM65	transmembrane protein 65	NA
56242	ZNF253	zinc finger protein 253	NA
22876	INPP5F	inositol polyphosphate-5-phosphatase F	NA
120	ADD3	adducin 3	NA
139285	AMER1	APC membrane recruitment protein 1	NA
6311	ATXN2	ataxin 2	NA
80746	TSEN2	tRNA splicing endonuclease subunit 2	NA
8759	ADAM1A	ADAM metallopeptidase domain 1A (pseudogene)	NA
3075	CFH	complement factor H	NA
8790	FPGT	fucose-1-phosphate guanylyltransferase	NA
153396	TMEM161B	transmembrane protein 161B	NA
54822	TRPM7	transient receptor potential cation channel subfamily M member 7	NA
148266	ZNF569	zinc finger protein 569	NA
6966	TRGC1	T cell receptor gamma constant 1	NA
85302	FBF1	Fas binding factor 1	NA
8994	LIMD1	LIM domains containing 1	NA
54482	TRMT13	tRNA methyltransferase 13 homolog	NA
55770	EXOC2	exocyst complex component 2	NA
3675	ITGA3	integrin subunit alpha 3	NA
9736	USP34	ubiquitin specific peptidase 34	NA
9923	ZBTB40	zinc finger and BTB domain containing 40	NA
347344	ZNF81	zinc finger protein 81	NA
54932	EXD3	exonuclease 3'-5' domain containing 3	NA
79573	TTC13	tetratricopeptide repeat domain 13	NA
29799	YPEL1	yippee like 1	NA
9321	TRIP11	thyroid hormone receptor interactor 11	NA

285343	TCAIM	T cell activation inhibitor, mitochondrial	NA
957	ENTPD5	ectonucleoside triphosphate diphosphohydrolase 5	NA
90317	ZNF616	zinc finger protein 616	NA
10750	GRAP	GRB2 related adaptor protein	NA
26009	ZZZ3	zinc finger ZZ-type containing 3	NA
159195	USP54	ubiquitin specific peptidase 54	NA
57547	ZNF624	zinc finger protein 624	NA
29899	GPSM2	G protein signaling modulator 2	NA
170958	ZNF525	zinc finger protein 525	NA
80067	DCAF17	DDB1 and CUL4 associated factor 17	NA
1024	CDK8	cyclin dependent kinase 8	NA
63926	ANKEF1	ankyrin repeat and EF-hand domain containing 1	NA
64770	CCDC14	coiled-coil domain containing 14	NA
4999	ORC2	origin recognition complex subunit 2	NA
7402	UTRN	utrophin	NA
84033	OBSCN	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	NA
8914	TIMELESS	timeless circadian regulator	NA
25894	PLEKHG4	pleckstrin homology and RhoGEF domain containing G4	NA
11260	XPOT	exportin for tRNA	NA
203427	SLC25A43	solute carrier family 25 member 43	NA
49855	SCAPER	S-phase cyclin A associated protein in the ER	NA
23731	TMEM245	transmembrane protein 245	NA
83860	TAF3	TATA-box binding protein associated factor 3	NA
7697	ZNF138	zinc finger protein 138	NA
2074	ERCC6	ERCC excision repair 6, chromatin remodeling factor	NA
79858	NEK11	NIMA related kinase 11	NA
56916	SMARCAD1	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1	NA
23262	PPIP5K2	diphosphoinositol pentakisphosphate kinase 2	NA
55066	PDPR	pyruvate dehydrogenase phosphatase regulatory subunit	NA
9258	MFHAS1	malignant fibrous histiocytoma amplified sequence 1	NA
65084	TMEM135	transmembrane protein 135	NA
57643	ZSWIM5	zinc finger SWIM-type containing 5	NA
388512	CLEC17A	C-type lectin domain containing 17A	NA
113263	GLCCI1	glucocorticoid induced 1	NA
114819	CROCCP3	ciliary rootlet coiled-coil, rootletin pseudogene 3	NA
79961	DENND2D	DENN domain containing 2D	NA
23008	KLHDC10	kelch domain containing 10	NA
2043	EPHA4	EPH receptor A4	NA
55863	TMEM126B	transmembrane protein 126B	NA
146857	SLFN13	schlafen family member 13	NA
7049	TGFBR3	transforming growth factor beta receptor 3	NA

23224	SYNE2	spectrin repeat containing nuclear envelope protein 2	NA
26275	HIBCH	3-hydroxyisobutyryl-CoA hydrolase	NA
128611	ZNF831	zinc finger protein 831	NA
10905	MAN1A2	mannosidase alpha class 1A member 2	NA
1739	DLG1	discs large MAGUK scaffold protein 1	NA
63035	BCORL1	BCL6 corepressor like 1	NA
118924	FRA10AC1	FRA10A associated CGG repeat 1	NA
10314	LANCL1	LanC like 1	NA
57492	ARID1B	AT-rich interaction domain 1B	NA
27148	STK36	serine/threonine kinase 36	NA
9183	ZW10	zw10 kinetochore protein	NA
27130	INVS	inversin	NA
5775	PTPN4	protein tyrosine phosphatase, non-receptor type 4	NA
583	BBS2	Bardet-Biedl syndrome 2	NA
115209	OMA1	OMA1 zinc metallopeptidase	NA
8631	SKAP1	src kinase associated phosphoprotein 1	NA
11064	CNTRL	centriolin	NA
84064	HDHD2	haloacid dehalogenase like hydrolase domain containing 2	NA
84327	ZBED3	zinc finger BED-type containing 3	NA
2909	ARHGAP35	Rho GTPase activating protein 35	NA
5019	OXCT1	3-oxoacid CoA-transferase 1	NA
57120	GOPC	golgi associated PDZ and coiled-coil motif containing	NA
57531	HACE1	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	NA
132241	RPL32P3	ribosomal protein L32 pseudogene 3	NA
158358	KIAA2026	KIAA2026	NA
58487	CREBZF	CREB/ATF bZIP transcription factor	NA
10111	RAD50	RAD50 double strand break repair protein	NA
9875	URB1	URB1 ribosome biogenesis homolog	NA
57448	BIRC6	baculoviral IAP repeat containing 6	NA
7776	ZNF236	zinc finger protein 236	NA
987	LRBA	LPS responsive beige-like anchor protein	NA
2967	GTF2H3	general transcription factor IIH subunit 3	NA
492311	IGIP	IgA inducing protein	NA
254251	LCORL	ligand dependent nuclear receptor corepressor like	NA
64699	TMPRSS3	transmembrane serine protease 3	NA
54677	CROT	carnitine O-octanoyltransferase	NA
84343	HPS3	HPS3, biogenesis of lysosomal organelles complex 2 subunit 1	NA
9706	ULK2	unc-51 like autophagy activating kinase 2	NA
1121	CHM	CHM, Rab escort protein 1	NA
56995	TULP4	tubby like protein 4	NA
11159	RABL2A	RAB, member of RAS oncogene family like 2A	NA
116842	LEAP2	liver enriched antimicrobial peptide 2	NA
727866	FAM156B	family with sequence similarity 156 member B	NA

375248	ANKRD36	ankyrin repeat domain 36	NA
9653	HS2ST1	heparan sulfate 2-O-sulfotransferase 1	NA
2176	FANCC	FA complementation group C	NA
64863	METTL4	methyltransferase like 4	NA
7728	ZNF175	zinc finger protein 175	NA
10390	CEPT1	choline/ethanolamine phosphotransferase 1	NA
27031	NPHP3	nephrocystin 3	NA
2592	GALT	galactose-1-phosphate uridylyltransferase	NA
11279	KLF8	Kruppel like factor 8	NA
9139	CBFA2T2	CBFA2/RUNX1 translocation partner 2	NA
115399	LRRC56	leucine rich repeat containing 56	NA
8515	ITGA10	integrin subunit alpha 10	NA
23431	AP4E1	adaptor related protein complex 4 subunit epsilon 1	NA
161145	TMEM229B	transmembrane protein 229B	NA
57465	TBC1D24	TBC1 domain family member 24	NA
54536	EXOC6	exocyst complex component 6	NA
5922	RASA2	RAS p21 protein activator 2	NA
5607	MAP2K5	mitogen-activated protein kinase kinase 5	NA
26278	SACS	sacsin molecular chaperone	NA
83544	DNAL1	dynein axonemal light chain 1	NA
5892	RAD51D	RAD51 paralog D	NA
10446	LRRN2	leucine rich repeat neuronal 2	NA
55213	RCBTB1	RCC1 and BTB domain containing protein 1	NA
56672	AKIP1	A-kinase interacting protein 1	NA
5584	PRKCI	protein kinase C iota	NA
57728	WDR19	WD repeat domain 19	NA
375743	PTAR1	protein prenyltransferase alpha subunit repeat containing 1	NA
23077	MYCBP2	MYC binding protein 2, E3 ubiquitin protein ligase	NA
6874	TAF4	TATA-box binding protein associated factor 4	NA
100101467	ZSCAN30	zinc finger and SCAN domain containing 30	NA
51409	HEMK1	HemK methyltransferase family member 1	NA
5980	REV3L	REV3 like, DNA directed polymerase zeta catalytic subunit	NA
197358	NLRC3	NLR family CARD domain containing 3	NA
53981	CPSF2	cleavage and polyadenylation specific factor 2	NA
163081	ZNF567	zinc finger protein 567	NA
64919	BCL11B	B cell CLL/lymphoma 11B	NA
55783	CMTR2	cap methyltransferase 2	NA
6904	TBCD	tubulin folding cofactor D	NA
10747	MASP2	mannan binding lectin serine peptidase 2	NA
100505648	RAD51-AS1	RAD51 antisense RNA 1	NA
29929	ALG6	ALG6, alpha-1,3-glucosyltransferase	NA
55100	WDR70	WD repeat domain 70	NA
113612	CYP2U1	cytochrome P450 family 2 subfamily U member 1	NA
94134	ARHGAP12	Rho GTPase activating protein 12	NA
5634	PRPS2	phosphoribosyl pyrophosphate synthetase 2	NA

8464	SUPT3H	SPT3 homolog, SAGA and STAGA complex component	NA
81790	RNF170	ring finger protein 170	NA
1154	CISH	cytokine inducible SH2 containing protein	NA
23499	MACF1	microtubule-actin crosslinking factor 1	NA
5073	PARN	poly(A)-specific ribonuclease	NA
55619	DOCK10	dedicator of cytokinesis 10	NA
2005	ELK4	ELK4, ETS transcription factor	NA
100418737	KRT8P33	keratin 8 pseudogene 33	NA
3931	LCAT	lecithin-cholesterol acyltransferase	NA
222236	NAPEPLD	N-acyl phosphatidylethanolamine phospholipase D	NA
151888	BTLA	B and T lymphocyte associated	NA
7514	XPO1	exportin 1	NA
157680	VPS13B	vacuolar protein sorting 13 homolog B	NA
3738	KCNA3	potassium voltage-gated channel subfamily A member 3	NA
1633	DCK	deoxycytidine kinase	NA
54799	MBTD1	mbt domain containing 1	NA
3749	KCNC4	potassium voltage-gated channel subfamily C member 4	NA
55825	PECR	peroxisomal trans-2-enoyl-CoA reductase	NA
100505894	TMEM161B-AS1	TMEM161B antisense RNA 1	NA
54477	PLEKHA5	pleckstrin homology domain containing A5	NA
50809	HP1BP3	heterochromatin protein 1 binding protein 3	NA
8863	PER3	period circadian regulator 3	NA
1519	CTSO	cathepsin O	NA
11127	KIF3A	kinesin family member 3A	NA
27332	ZNF638	zinc finger protein 638	NA
59338	PLEKHA1	pleckstrin homology domain containing A1	NA
6491	STIL	STIL, centriolar assembly protein	NA
54894	RNF43	ring finger protein 43	NA
10743	RAI1	retinoic acid induced 1	NA
9202	ZMYM4	zinc finger MYM-type containing 4	NA
9879	DDX46	DEAD-box helicase 46	NA
1182	CLCN3	chloride voltage-gated channel 3	NA
91754	NEK9	NIMA related kinase 9	NA
54014	BRWD1	bromodomain and WD repeat domain containing 1	NA
83988	NCALD	neurocalcin delta	NA
57212	TP73-AS1	TP73 antisense RNA 1	NA
11022	TDRKH	tudor and KH domain containing	NA
4297	KMT2A	lysine methyltransferase 2A	NA
91875	TTC5	tetratricopeptide repeat domain 5	NA
57583	TMEM181	transmembrane protein 181	NA
1676	DFFA	DNA fragmentation factor subunit alpha	NA
146691	TOM1L2	target of myb1 like 2 membrane trafficking protein	NA
23094	SIPA1L3	signal induced proliferation associated 1 like 3	NA
253959	RALGAPA1	Ral GTPase activating protein catalytic alpha subunit 1	NA

9063	PIAS2	protein inhibitor of activated STAT 2	NA
8028	MLLT10	MLLT10, histone lysine methyltransferase DOT1L cofactor	NA
80326	WNT10A	Wnt family member 10A	NA
10380	BPNT1	3'(2'), 5'-bisphosphate nucleotidase 1	NA
54664	TMEM106B	transmembrane protein 106B	NA
401303	ZNF815P	zinc finger protein 815, pseudogene	NA
10128	LRPPRC	leucine rich pentatricopeptide repeat containing	NA
7707	ZNF148	zinc finger protein 148	NA
100507495	SDCBP2-AS1	SDCBP2 antisense RNA 1	NA
55599	RNPC3	RNA binding region (RNP1, RRM) containing 3	NA
57189	KIAA1147	KIAA1147	NA
79072	FASTKD3	FAST kinase domains 3	NA
145389	SLC38A6	solute carrier family 38 member 6	NA
145957	NRG4	neuregulin 4	NA
11168	PSIP1	PC4 and SFRS1 interacting protein 1	NA
10464	PIBF1	progesterone immunomodulatory binding factor 1	NA
394	ARHGAP5	Rho GTPase activating protein 5	NA
10818	FRS2	fibroblast growth factor receptor substrate 2	NA
55023	PHIP	pleckstrin homology domain interacting protein	NA
84765	ZNF577	zinc finger protein 577	NA
10973	ASCC3	activating signal cointegrator 1 complex subunit 3	NA
57695	USP37	ubiquitin specific peptidase 37	NA
356	FASLG	Fas ligand	NA
6565	SLC15A2	solute carrier family 15 member 2	NA
51466	EVL	Enah/Vasp-like	NA
6480	ST6GAL1	ST6 beta-galactoside alpha-2,6-sialyltransferase 1	NA
23583	SMUG1	single-strand-selective monofunctional uracil-DNA glycosylase 1	NA
3619	INCENP	inner centromere protein	NA
56922	MCCC1	methylcrotonoyl-CoA carboxylase 1	NA
65985	AACS	acetoacetyl-CoA synthetase	NA
27347	STK39	serine/threonine kinase 39	NA
84532	ACSS1	acyl-CoA synthetase short chain family member 1	NA
4299	AFF1	AF4/FMR2 family member 1	NA
8526	DGKE	diacylglycerol kinase epsilon	NA
55005	RMND1	required for meiotic nuclear division 1 homolog	NA
55449	DHRS4-AS1	DHRS4 antisense RNA 1	NA
84527	ZNF559	zinc finger protein 559	NA
132001	TAMM41	TAM41 mitochondrial translocator assembly and maintenance homolog	NA
54517	PUS7	pseudouridine synthase 7	NA
283537	SLC46A3	solute carrier family 46 member 3	NA
64839	FBXL17	F-box and leucine rich repeat protein 17	NA
5832	ALDH18A1	aldehyde dehydrogenase 18 family member A1	NA
152926	PPM1K	protein phosphatase, Mg2+/Mn2+ dependent 1K	NA
56950	SMYD2	SET and MYND domain containing 2	NA

50804	MYEF2	myelin expression factor 2	NA
150383	CDPF1	cysteine rich DPF motif domain containing 1	NA
23002	DAAM1	dishevelled associated activator of morphogenesis 1	NA
51427	ZNF107	zinc finger protein 107	NA
57542	KLHL42	kelch like family member 42	NA
23169	SLC35D1	solute carrier family 35 member D1	NA
161823	ADAL	adenosine deaminase like	NA
55014	STX17	syntaxin 17	NA
3595	IL12RB2	interleukin 12 receptor subunit beta 2	NA
10785	WDR4	WD repeat domain 4	NA
373	TRIM23	tripartite motif containing 23	NA
7639	ZNF85	zinc finger protein 85	NA
51176	LEF1	lymphoid enhancer binding factor 1	NA
5469	MED1	mediator complex subunit 1	NA
10800	CYSLTR1	cysteinyl leukotriene receptor 1	NA
8801	SUCLG2	succinate-CoA ligase GDP-forming beta subunit	NA
2035	EPB41	erythrocyte membrane protein band 4.1	NA
23195	MDN1	midasin AAA ATPase 1	NA
5150	PDE7A	phosphodiesterase 7A	NA
26122	EPC2	enhancer of polycomb homolog 2	NA
23294	ANKS1A	ankyrin repeat and sterile alpha motif domain containing 1A	NA
11059	WWP1	WW domain containing E3 ubiquitin protein ligase 1	NA
27252	KLHL20	kelch like family member 20	NA
91869	RFT1	RFT1 homolog	NA
317	APAF1	apoptotic peptidase activating factor 1	NA
114803	MYSM1	Myb like, SWIRM and MPN domains 1	NA
197322	ACSF3	acyl-CoA synthetase family member 3	NA
3590	IL11RA	interleukin 11 receptor subunit alpha	NA
147694	ZNF548	zinc finger protein 548	NA
55109	AGGF1	angiogenic factor with G-patch and FHA domains 1	NA
1718	DHCR24	24-dehydrocholesterol reductase	NA
144132	DNHD1	dynein heavy chain domain 1	NA
9735	KNTC1	kinetochore associated 1	NA
11043	MID2	midline 2	NA
6936	GCFC2	GC-rich sequence DNA-binding factor 2	NA
84186	ZCCHC7	zinc finger CCHC-type containing 7	NA
10651	MTX2	metaxin 2	NA
64393	ZMAT3	zinc finger matrin-type 3	NA
133308	SLC9B2	solute carrier family 9 member B2	NA
79085	SLC25A23	solute carrier family 25 member 23	NA
400322	HERC2P2	hect domain and RLD 2 pseudogene 2	NA
1371	CPOX	coproporphyrinogen oxidase	NA
253769	WDR27	WD repeat domain 27	NA
54816	ZNF280D	zinc finger protein 280D	NA
8850	KAT2B	lysine acetyltransferase 2B	NA

8502	PKP4	plakophilin 4	NA
22835	ZFP30	ZFP30 zinc finger protein	NA
9710	KIAA0355	KIAA0355	NA
84614	ZBTB37	zinc finger and BTB domain containing 37	NA
9617	MTRF1	mitochondrial translation release factor 1	NA
54847	SIDT1	SID1 transmembrane family member 1	NA
159090	FAM122B	family with sequence similarity 122B	NA
5911	RAP2A	RAP2A, member of RAS oncogene family	NA
286101	ZNF252P	zinc finger protein 252, pseudogene	NA
8621	CDK13	cyclin dependent kinase 13	NA
23078	VWA8	von Willebrand factor A domain containing 8	NA
2186	BPTF	bromodomain PHD finger transcription factor	NA
29843	SENP1	SUMO specific peptidase 1	NA
55094	GPATCH1	G-patch domain containing 1	NA
4621	MYH3	myosin heavy chain 3	NA
7444	VRK2	VRK serine/threonine kinase 2	NA
8317	CDC7	cell division cycle 7	NA
9648	GCC2	GRIP and coiled-coil domain containing 2	NA
9946	CRYZL1	crystallin zeta like 1	NA
11158	RABL2B	RAB, member of RAS oncogene family like 2B	NA
55102	ATG2B	autophagy related 2B	NA
23518	R3HDM1	R3H domain containing 1	NA
1355	COX15	cytochrome c oxidase assembly homolog COX15	NA
554251	FBXO48	F-box protein 48	NA
64131	XYLT1	xylosyltransferase 1	NA
7072	TIA1	TIA1 cytotoxic granule associated RNA binding protein	NA
10741	RBBP9	RB binding protein 9, serine hydrolase	NA
200576	PIKFYVE	phosphoinositide kinase, FYVE-type zinc finger containing	NA
55364	IMPACT	impact RWD domain protein	NA
51196	PLCE1	phospholipase C epsilon 1	NA
23252	OTUD3	OTU deubiquitinase 3	NA
57464	STRIP2	striatin interacting protein 2	NA
80153	EDC3	enhancer of mRNA decapping 3	NA
140564	APOBEC3D	apolipoprotein B mRNA editing enzyme catalytic subunit 3D	NA
57479	PRR12	proline rich 12	NA
154075	SAMD3	sterile alpha motif domain containing 3	NA
4437	MSH3	mutS homolog 3	NA
100289414	PHBP9	prohibitin pseudogene 9	NA
253714	MMS22L	MMS22 like, DNA repair protein	NA
23310	NCAPD3	non-SMC condensin II complex subunit D3	NA
9779	TBC1D5	TBC1 domain family member 5	NA
3662	IRF4	interferon regulatory factor 4	NA
221294	NT5DC1	5'-nucleotidase domain containing 1	NA
5825	ABCD3	ATP binding cassette subfamily D member 3	NA

9662	CEP135	centrosomal protein 135	NA
55617	TASP1	taspase 1	NA
6645	SNTB2	syntrophin beta 2	NA
23112	TNRC6B	trinucleotide repeat containing 6B	NA
29102	DROSHA	drosha ribonuclease III	NA
284565	NBPF15	NBPF member 15	NA
57337	SENP7	SUMO specific peptidase 7	NA
23729	SHPK	sedoheptulokinase	NA
11112	HIBADH	3-hydroxyisobutyrate dehydrogenase	NA
390637	GDPGP1	GDP-D-glucose phosphorylase 1	NA
89845	ABCC10	ATP binding cassette subfamily C member 10	NA
23333	DPY19L1	dpy-19 like C-mannosyltransferase 1	NA
58473	PLEKHB1	pleckstrin homology domain containing B1	NA
100131067	CKMT2-AS1	CKMT2 antisense RNA 1	NA
100131755	ARMCX4	armadillo repeat containing X-linked 4	NA
29969	MDFIC	MyoD family inhibitor domain containing	NA
8888	MCM3AP	minichromosome maintenance complex component 3 associated protein	NA
23274	CLEC16A	C-type lectin domain containing 16A	NA
27245	AHDC1	AT-hook DNA binding motif containing 1	NA
3841	KPNA5	karyopherin subunit alpha 5	NA
81550	TDRD3	tudor domain containing 3	NA
23279	NUP160	nucleoporin 160	NA
84436	ZNF528	zinc finger protein 528	NA
134353	LSM11	LSM11, U7 small nuclear RNA associated	NA
9652	TTC37	tetratricopeptide repeat domain 37	NA
169200	TMEM64	transmembrane protein 64	NA
91775	NXPE3	neurexophilin and PC-esterase domain family member 3	NA
23648	SSBP3	single stranded DNA binding protein 3	NA
26505	CNNM3	cyclin and CBS domain divalent metal cation transport mediator 3	NA
9317	PTER	phosphotriesterase related	NA
7644	ZNF91	zinc finger protein 91	NA
11232	POLG2	DNA polymerase gamma 2, accessory subunit	NA
22847	ZNF507	zinc finger protein 507	NA
56648	EIF5A2	eukaryotic translation initiation factor 5A2	NA
100506001	NARF-IT1	NARF intronic transcript 1	NA
84301	DDI2	DNA damage inducible 1 homolog 2	NA
5324	PLAG1	PLAG1 zinc finger	NA
55556	ENOSF1	enolase superfamily member 1	NA
55578	SUPT20H	SPT20 homolog, SAGA complex component	NA
51319	RSRC1	arginine and serine rich coiled-coil 1	NA
55656	INTS8	integrator complex subunit 8	NA
5569	PKIA	cAMP-dependent protein kinase inhibitor alpha	NA
6872	TAF1	TATA-box binding protein associated factor 1	NA
165631	PARP15	poly(ADP-ribose) polymerase family member 15	NA

778	CACNA1F	calcium voltage-gated channel subunit alpha1 F	NA
10392	NOD1	nucleotide binding oligomerization domain containing 1	NA
10927	SPIN1	spindlin 1	NA
84260	TCHP	trichoplein keratin filament binding	NA
101362076	GVQW1	GVQW motif containing 1	NA
51227	PIGP	phosphatidylinositol glycan anchor biosynthesis class P	NA
4124	MAN2A1	mannosidase alpha class 2A member 1	NA
80208	SPG11	SPG11, spatacsin vesicle trafficking associated	NA
192669	AGO3	argonaute 3, RISC catalytic component	NA
55230	USP40	ubiquitin specific peptidase 40	NA
113	ADCY7	adenylate cyclase 7	NA
124540	MSI2	musashi RNA binding protein 2	NA
29083	GTPBP8	GTP binding protein 8 (putative)	NA
400986	ANKRD36C	ankyrin repeat domain 36C	NA
23522	KAT6B	lysine acetyltransferase 6B	NA
7182	NR2C2	nuclear receptor subfamily 2 group C member 2	NA
55193	PBRM1	polybromo 1	NA
10142	AKAP9	A-kinase anchoring protein 9	NA
11311	VPS45	vacuolar protein sorting 45 homolog	NA
283899	INO80E	INO80 complex subunit E	NA
27072	VPS41	VPS41, HOPS complex subunit	NA
7768	ZNF225	zinc finger protein 225	NA
6693	SPN	sialophorin	NA
84461	NEURL4	neuralized E3 ubiquitin protein ligase 4	NA
11073	TOPBP1	DNA topoisomerase II binding protein 1	NA
22900	CARD8	caspase recruitment domain family member 8	NA
9031	BAZ1B	bromodomain adjacent to zinc finger domain 1B	NA
60489	APOBEC3G	apolipoprotein B mRNA editing enzyme catalytic subunit 3G	NA
140710	SOGA1	suppressor of glucose, autophagy associated 1	NA
1111	CHEK1	checkpoint kinase 1	NA
889	KRIT1	KRIT1, ankyrin repeat containing	NA
112937	GLB1L3	galactosidase beta 1 like 3	NA
253143	PRR14L	proline rich 14 like	NA
84928	TMEM209	transmembrane protein 209	NA
1629	DBT	dihydrolipoamide branched chain transacylase E2	NA
27327	TNRC6A	trinucleotide repeat containing 6A	NA
23512	SUZ12	SUZ12, polycomb repressive complex 2 subunit	NA
84869	CBR4	carbonyl reductase 4	NA
23032	USP33	ubiquitin specific peptidase 33	NA
440823	MIAT	myocardial infarction associated transcript	NA
5195	PEX14	peroxisomal biogenesis factor 14	NA
9793	CKAP5	cytoskeleton associated protein 5	NA
144348	ZNF664	zinc finger protein 664	NA
80184	CEP290	centrosomal protein 290	NA
159091	FAM122C	family with sequence similarity 122C	NA

26034	IPCEF1	interaction protein for cytohesin exchange factors 1	NA
388591	RNF207	ring finger protein 207	NA
9663	LPIN2	lipin 2	NA
545	ATR	ATR serine/threonine kinase	NA
55729	ATF7IP	activating transcription factor 7 interacting protein	NA
92922	CCDC102A	coiled-coil domain containing 102A	NA
10051	SMC4	structural maintenance of chromosomes 4	NA
9678	PHF14	PHD finger protein 14	NA
5335	PLCG1	phospholipase C gamma 1	NA
64395	GMCL1	germ cell-less, spermatogenesis associated 1	NA
6941	TCF19	transcription factor 19	NA
56474	CTPS2	CTP synthase 2	NA
54965	PIGX	phosphatidylinositol glycan anchor biosynthesis class X	NA
2729	GCLC	glutamate-cysteine ligase catalytic subunit	NA
25	ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase	NA
55125	CEP192	centrosomal protein 192	NA
255394	TCP11L2	t-complex 11 like 2	NA
104472715	SNHG14	small nucleolar RNA host gene 14	NA
7188	TRAF5	TNF receptor associated factor 5	NA
8473	OGT	O-linked N-acetylglucosamine (GlcNAc) transferase	NA
9941	EXOG	exo/endonuclease G	NA
222194	RSBN1L	round spermatid basic protein 1 like	NA
55153	SDAD1	SDA1 domain containing 1	NA
9329	GTF3C4	general transcription factor IIIC subunit 4	NA
55663	ZNF446	zinc finger protein 446	NA
23386	NUDCD3	NudC domain containing 3	NA
26157	GIMAP2	GTPase, IMAP family member 2	NA
285331	CCDC66	coiled-coil domain containing 66	NA
79745	CLIP4	CAP-Gly domain containing linker protein family member 4	NA
1112	FOXN3	forkhead box N3	NA
324	APC	APC, WNT signaling pathway regulator	NA
7692	ZNF133	zinc finger protein 133	NA
79871	RPAP2	RNA polymerase II associated protein 2	NA
29909	GPR171	G protein-coupled receptor 171	NA
2058	EPRS	glutamyl-prolyl-tRNA synthetase	NA
79832	QSER1	glutamine and serine rich 1	NA
79612	NAA16	N(alpha)-acetyltransferase 16, NatA auxiliary subunit	NA
26260	FBXO25	F-box protein 25	NA
1786	DNMT1	DNA methyltransferase 1	NA
5613	PRKX	protein kinase X-linked	NA
55975	KLHL7	kelch like family member 7	NA
757	TMEM50B	transmembrane protein 50B	NA
55210	ATAD3A	ATPase family, AAA domain containing 3A	NA

84809	CROCCP2	ciliary rootlet coiled-coil, rootletin pseudogene 2	NA
171392	ZNF675	zinc finger protein 675	NA
146198	ZFP90	ZFP90 zinc finger protein	NA
9738	CCP110	centriolar coiled-coil protein 110	NA
8895	CPNE3	copine 3	NA
64432	MRPS25	mitochondrial ribosomal protein S25	NA
222255	ATXN7L1	ataxin 7 like 1	NA
23244	PDS5A	PDS5 cohesin associated factor A	NA
9743	ARHGAP32	Rho GTPase activating protein 32	NA
5557	PRIM1	DNA primase subunit 1	NA
388152	GOLGA2P7	GOLGA2 pseudogene 7	NA
89910	UBE3B	ubiquitin protein ligase E3B	NA
253461	ZBTB38	zinc finger and BTB domain containing 38	NA
83787	ARMC10	armadillo repeat containing 10	NA
23341	DNAJC16	DnaJ heat shock protein family (Hsp40) member C16	NA
10788	IQGAP2	IQ motif containing GTPase activating protein 2	NA
65125	WNK1	WNK lysine deficient protein kinase 1	NA
146057	TTBK2	tau tubulin kinase 2	NA
55268	ECHDC2	enoyl-CoA hydratase domain containing 2	NA
339318	ZNF181	zinc finger protein 181	NA
57678	GPAM	glycerol-3-phosphate acyltransferase, mitochondrial	NA
201633	TIGIT	T cell immunoreceptor with Ig and ITIM domains	NA
84795	PYROXD2	pyridine nucleotide-disulphide oxidoreductase domain 2	NA
59271	EVA1C	eva-1 homolog C	NA
84464	SLX4	SLX4 structure-specific endonuclease subunit	NA
4157	MC1R	melanocortin 1 receptor	NA
66008	TRAK2	trafficking kinesin protein 2	NA
23233	EXOC6B	exocyst complex component 6B	NA
56852	RAD18	RAD18, E3 ubiquitin protein ligase	NA
9698	PUM1	pumilio RNA binding family member 1	NA
2317	FLNB	filamin B	NA
84340	GFM2	G elongation factor mitochondrial 2	NA
129293	TRABD2A	TraB domain containing 2A	NA
22911	WDR47	WD repeat domain 47	NA
115201	ATG4A	autophagy related 4A cysteine peptidase	NA
51112	TRAPPC12	trafficking protein particle complex 12	NA
64682	ANAPC1	anaphase promoting complex subunit 1	NA
7153	TOP2A	DNA topoisomerase II alpha	NA
8831	SYNGAP1	synaptic Ras GTPase activating protein 1	NA
84456	L3MBTL3	L3MBTL3, histone methyl-lysine binding protein	NA
10499	NCOA2	nuclear receptor coactivator 2	NA
23248	RPRD2	regulation of nuclear pre-mRNA domain containing 2	NA
80817	CEP44	centrosomal protein 44	NA
23592	LEMD3	LEM domain containing 3	NA
51028	VPS36	vacuolar protein sorting 36 homolog	NA

670	BPHL	biphenyl hydrolase like	NA
9487	PIGL	phosphatidylinositol glycan anchor biosynthesis class L	NA
54441	STAG3L1	stromal antigen 3-like 1 (pseudogene)	NA
122060	SLAIN1	SLAIN motif family member 1	NA
64756	ATPAF1	ATP synthase mitochondrial F1 complex assembly factor 1	NA
11118	BTN3A2	butyrophilin subfamily 3 member A2	NA
54464	XRN1	5'-3' exoribonuclease 1	NA
3953	LEPR	leptin receptor	NA
55032	SLC35A5	solute carrier family 35 member A5	NA
8418	CMAHP	cytidine monophospho-N-acetylneuraminic acid hydroxylase, pseudogene	NA
2037	EPB41L2	erythrocyte membrane protein band 4.1 like 2	NA
100132707	PAXIP1-AS2	PAXIP1 antisense RNA 2	NA
145748	LYSMD4	LysM domain containing 4	NA
7181	NR2C1	nuclear receptor subfamily 2 group C member 1	NA
51095	TRNT1	tRNA nucleotidyl transferase 1	NA
7597	ZBTB25	zinc finger and BTB domain containing 25	NA
54780	NSMCE4A	NSE4 homolog A, SMC5-SMC6 complex component	NA
22841	RAB11FIP2	RAB11 family interacting protein 2	NA
57122	NUP107	nucleoporin 107	NA
3029	HAGH	hydroxyacylglutathione hydrolase	NA
89839	ARHGAP11B	Rho GTPase activating protein 11B	NA
151613	TTC14	tetratricopeptide repeat domain 14	NA
387357	THEMIS	thymocyte selection associated	NA
729234	FAHD2CP	fumarylacetoacetate hydrolase domain containing 2C, pseudogene	NA
54778	RNF111	ring finger protein 111	NA
51668	HSPB11	heat shock protein family B (small) member 11	NA
3572	IL6ST	interleukin 6 signal transducer	NA
54906	FAM208B	family with sequence similarity 208 member B	NA
9874	TLK1	tousled like kinase 1	NA
23305	ACSL6	acyl-CoA synthetase long chain family member 6	NA
5565	PRKAB2	protein kinase AMP-activated non-catalytic subunit beta 2	NA
81573	ANKRD13C	ankyrin repeat domain 13C	NA
442582	STAG3L2	stromal antigen 3-like 2 (pseudogene)	NA
79786	KLHL36	kelch like family member 36	NA
26191	PTPN22	protein tyrosine phosphatase, non-receptor type 22	NA
10194	TSHZ1	teashirt zinc finger homeobox 1	NA
132949	AASDH	aminoadipate-semialdehyde dehydrogenase	NA
23334	SZT2	SZT2, KICSTOR complex subunit	NA
388561	ZNF761	zinc finger protein 761	NA
79269	DCAF10	DDB1 and CUL4 associated factor 10	NA
5128	CDK17	cyclin dependent kinase 17	NA
153241	CEP120	centrosomal protein 120	NA

221035	REEP3	receptor accessory protein 3	NA
85007	PHYKPL	5-phosphohydroxy-L-lysine phospho-lyase	NA
162239	ZFP1	ZFP1 zinc finger protein	NA
57489	ODF2L	outer dense fiber of sperm tails 2 like	NA
26013	L3MBTL1	L3MBTL1, histone methyl-lysine binding protein	NA
10714	POLD3	DNA polymerase delta 3, accessory subunit	NA
11329	STK38	serine/threonine kinase 38	NA
10847	SRCAP	Snf2 related CREBBP activator protein	NA
9649	RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	NA
64766	S100PBP	S100P binding protein	NA
9354	UBE4A	ubiquitination factor E4A	NA
57194	ATP10A	ATPase phospholipid transporting 10A (putative)	NA
4548	MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	NA
5198	PFAS	phosphoribosylformylglycinamidine synthase	NA
9801	MRPL19	mitochondrial ribosomal protein L19	NA
84181	CHD6	chromodomain helicase DNA binding protein 6	NA
56204	FAM214A	family with sequence similarity 214 member A	NA
23081	KDM4C	lysine demethylase 4C	NA
54954	FAM120C	family with sequence similarity 120C	NA
149628	PYHIN1	pyrin and HIN domain family member 1	NA
57692	MAGEE1	MAGE family member E1	NA
11119	BTN3A1	butyrophilin subfamily 3 member A1	NA
155400	NSUN5P1	NOP2/Sun RNA methyltransferase family member 5 pseudogene 1	NA
11320	MGAT4A	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase A	NA
55276	PGM2	phosphoglucomutase 2	NA
7700	ZNF141	zinc finger protein 141	NA
101735302	STAG3L5P	stromal antigen 3-like 5 pseudogene	NA
8814	CDKL1	cyclin dependent kinase like 1	NA
64848	YTHDC2	YTH domain containing 2	NA
8725	URI1	URI1, prefoldin like chaperone	NA
106903081	DUSP8P5	dual specificity phosphatase 8 pseudogene 5	NA
6314	ATXN7	ataxin 7	NA
170960	ZNF721	zinc finger protein 721	NA
56957	OTUD7B	OTU deubiquitinase 7B	NA
81539	SLC38A1	solute carrier family 38 member 1	NA
63977	PRDM15	PR/SET domain 15	NA
55198	APPL2	adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 2	NA
57493	HEG1	heart development protein with EGF like domains 1	NA
56980	PRDM10	PR/SET domain 10	NA
26263	FBXO22	F-box protein 22	NA
55512	SMPD3	sphingomyelin phosphodiesterase 3	NA
91137	SLC25A46	solute carrier family 25 member 46	NA
27314	RAB30	RAB30, member RAS oncogene family	NA

23085	ERC1	ELKS/RAB6-interacting/CAST family member 1	NA
29063	ZCCHC4	zinc finger CCHC-type containing 4	NA
5994	RFXAP	regulatory factor X associated protein	NA
9847	C2CD5	C2 calcium dependent domain containing 5	NA
57446	NDRG3	NDRG family member 3	NA
23131	GPATCH8	G-patch domain containing 8	NA
2595	GANC	glucosidase alpha, neutral C	NA
9924	PAN2	poly(A) specific ribonuclease subunit PAN2	NA
57198	ATP8B2	ATPase phospholipid transporting 8B2	NA
162972	ZNF550	zinc finger protein 550	NA
56254	RNF20	ring finger protein 20	NA
23189	KANK1	KN motif and ankyrin repeat domains 1	NA
57097	PARP11	poly(ADP-ribose) polymerase family member 11	NA
26043	UBXN7	UBX domain protein 7	NA
375748	ERCC6L2	ERCC excision repair 6 like 2	NA
9931	HELZ	helicase with zinc finger	NA
8506	CNTNAP1	contactin associated protein 1	NA
57658	CALCOCO1	calcium binding and coiled-coil domain 1	NA
116064	LRRC58	leucine rich repeat containing 58	NA
284618	RUSC1-AS1	RUSC1 antisense RNA 1	NA
57091	CASS4	Cas scaffold protein family member 4	NA
6601	SMARCC2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin subfamily c member 2	NA
4287	ATXN3	ataxin 3	NA
51479	ANKFY1	ankyrin repeat and FYVE domain containing 1	NA
5310	PKD1	polycystin 1, transient receptor potential channel interacting	NA
254128	NIFK-AS1	NIFK antisense RNA 1	NA
23199	GSE1	Gse1 coiled-coil protein	NA
22920	KIFAP3	kinesin associated protein 3	NA
151313	FAHD2B	fumarylacetoacetate hydrolase domain containing 2B	NA
2738	GLI4	GLI family zinc finger 4	NA
57674	RNF213	ring finger protein 213	NA
23132	RAD54L2	RAD54 like 2	NA
59339	PLEKHA2	pleckstrin homology domain containing A2	NA
2214	FCGR3A	Fc fragment of IgG receptor IIIa	NA
79710	MORC4	MORC family CW-type zinc finger 4	NA
84245	MRI1	methylthioribose-1-phosphate isomerase 1	NA
55127	HEATR1	HEAT repeat containing 1	NA
55015	PRPF39	pre-mRNA processing factor 39	NA
25814	ATXN10	ataxin 10	NA
4522	MTHFD1	methylenetetrahydrofolate dehydrogenase, cyclohydrolase and formyltetrahydrofolate synthetase 1	NA
56339	METTL3	methyltransferase like 3	NA
86	ACTL6A	actin like 6A	NA
388403	YPEL2	yippee like 2	NA

4134	MAP4	microtubule associated protein 4	NA
7711	ZNF155	zinc finger protein 155	NA
54462	CCSER2	coiled-coil serine rich protein 2	NA
84319	CMSS1	cms1 ribosomal small subunit homolog	NA
51054	PLEKHA8P1	pleckstrin homology domain containing A8 pseudogene 1	NA
79022	TMEM106C	transmembrane protein 106C	NA
6595	SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	NA
91750	LIN52	lin-52 DREAM MuvB core complex component	NA
55425	GPALPP1	GPALPP motifs containing 1	NA
9585	KIF20B	kinesin family member 20B	NA
641	BLM	BLM RecQ like helicase	NA
389524	GTF2IRD2B	GTF2I repeat domain containing 2B	NA
3796	KIF2A	kinesin family member 2A	NA
9751	SNPH	syntaphilin	NA
79006	METRN	meteorin, glial cell differentiation regulator	NA
23373	CRTC1	CREB regulated transcription coactivator 1	NA
1456	CSNK1G3	casein kinase 1 gamma 3	NA
5576	PRKAR2A	protein kinase cAMP-dependent type II regulatory subunit alpha	NA
25851	TECPR1	tectonin beta-propeller repeat containing 1	NA
80063	ATF7IP2	activating transcription factor 7 interacting protein 2	NA
54956	PARP16	poly(ADP-ribose) polymerase family member 16	NA
3419	IDH3A	isocitrate dehydrogenase 3 (NAD(+)) alpha	NA
139818	DOCK11	dedicator of cytokinesis 11	NA
4214	MAP3K1	mitogen-activated protein kinase kinase kinase 1	NA
80143	SIKE1	suppressor of IKBKE 1	NA
3612	IMPA1	inositol monophosphatase 1	NA
399687	MYO18A	myosin XVIIIA	NA
399512	SLC25A35	solute carrier family 25 member 35	NA
8450	CUL4B	cullin 4B	NA
64062	RBM26	RNA binding motif protein 26	NA
9749	PHACTR2	phosphatase and actin regulator 2	NA
23607	CD2AP	CD2 associated protein	NA
7048	TGFBR2	transforming growth factor beta receptor 2	NA
23264	ZC3H7B	zinc finger CCCH-type containing 7B	NA
7024	TFCP2	transcription factor CP2	NA
9631	NUP155	nucleoporin 155	NA
51684	SUFU	SUFU negative regulator of hedgehog signaling	NA
10057	ABCC5	ATP binding cassette subfamily C member 5	NA
57798	GATAD1	GATA zinc finger domain containing 1	NA
128272	ARHGEF19	Rho guanine nucleotide exchange factor 19	NA
8295	TRRAP	transformation/transcription domain associated protein	NA
22869	ZNF510	zinc finger protein 510	NA
10225	CD96	CD96 molecule	NA

91612	CHURC1	churchill domain containing 1	NA
84146	ZNF644	zinc finger protein 644	NA
643836	ZFP62	ZFP62 zinc finger protein	NA
8111	GPR68	G protein-coupled receptor 68	NA
51362	CDC40	cell division cycle 40	NA
8508	NIPSNAP1	nipsnap homolog 1	NA
64793	CEP85	centrosomal protein 85	NA
64324	NSD1	nuclear receptor binding SET domain protein 1	NA
79982	DNAJB14	DnaJ heat shock protein family (Hsp40) member B14	NA
3709	ITPR2	inositol 1,4,5-trisphosphate receptor type 2	NA
51379	CRLF3	cytokine receptor like factor 3	NA
79866	BORA	bora, aurora kinase A activator	NA
4820	NKTR	natural killer cell triggering receptor	NA
51194	IPO11	importin 11	NA
687	KLF9	Kruppel like factor 9	NA
9692	KIAA0391	KIAA0391	NA
9759	HDAC4	histone deacetylase 4	NA
25992	SNED1	sushi, nidogen and EGF like domains 1	NA
9397	NMT2	N-myristoyltransferase 2	NA
55120	FANCL	FA complementation group L	NA
9690	UBE3C	ubiquitin protein ligase E3C	NA
83752	LONP2	lon peptidase 2, peroxisomal	NA
23683	PRKD3	protein kinase D3	NA
79828	METTL8	methyltransferase like 8	NA
55706	NDC1	NDC1 transmembrane nucleoporin	NA
9896	FIG4	FIG4 phosphoinositide 5-phosphatase	NA
7593	MZF1	myeloid zinc finger 1	NA
26985	AP3M1	adaptor related protein complex 3 subunit mu 1	NA
219749	ZNF25	zinc finger protein 25	NA
1353	COX11	cytochrome c oxidase copper chaperone COX11	NA
10238	DCAF7	DDB1 and CUL4 associated factor 7	NA
6095	RORA	RAR related orphan receptor A	NA
22883	CLSTN1	calsyntenin 1	NA
54487	DGCR8	DGCR8, microprocessor complex subunit	NA
55183	RIF1	replication timing regulatory factor 1	NA
7581	ZNF33A	zinc finger protein 33A	NA
84131	CEP78	centrosomal protein 78	NA
10564	ARFGEF2	ADP ribosylation factor guanine nucleotide exchange factor 2	NA
548644	POLR2J3	RNA polymerase II subunit J3	NA
10608	MXD4	MAX dimerization protein 4	NA
8050	PDHX	pyruvate dehydrogenase complex component X	NA
23443	SLC35A3	solute carrier family 35 member A3	NA
23228	PLCL2	phospholipase C like 2	NA
6938	TCF12	transcription factor 12	NA
55156	ARMC1	armadillo repeat containing 1	NA

57536	KIAA1328	KIAA1328	NA
64421	DCLRE1C	DNA cross-link repair 1C	NA
57404	CYP20A1	cytochrome P450 family 20 subfamily A member 1	NA
55133	SRBD1	S1 RNA binding domain 1	NA
55676	SLC30A6	solute carrier family 30 member 6	NA
133418	EMB	embigin	NA
6871	TADA2A	transcriptional adaptor 2A	NA
23170	TTLL12	tubulin tyrosine ligase like 12	NA
56052	ALG1	ALG1, chitobiosyldiphosphodolichol beta-mannosyltransferase	NA
8732	RNGTT	RNA guanylyltransferase and 5'-phosphatase	NA
728927	ZNF736	zinc finger protein 736	NA
51433	ANAPC5	anaphase promoting complex subunit 5	NA
8504	PEX3	peroxisomal biogenesis factor 3	NA
79175	ZNF343	zinc finger protein 343	NA
6560	SLC12A4	solute carrier family 12 member 4	NA
11128	POLR3A	RNA polymerase III subunit A	NA
284086	NEK8	NIMA related kinase 8	NA
283130	SLC25A45	solute carrier family 25 member 45	NA
80349	WDR61	WD repeat domain 61	NA
84939	MUM1	melanoma associated antigen (mutated) 1	NA
55246	CCDC25	coiled-coil domain containing 25	NA
57786	RBAK	RB associated KRAB zinc finger	NA
55274	PHF10	PHD finger protein 10	NA
7074	TIAM1	T cell lymphoma invasion and metastasis 1	NA
145508	CEP128	centrosomal protein 128	NA
83714	NRIP2	nuclear receptor interacting protein 2	NA
114836	SLAMF6	SLAM family member 6	NA
109731405	PDXDC2P	pyridoxal dependent decarboxylase domain containing 2, pseudogene	NA
80195	TMEM254	transmembrane protein 254	NA
7175	TPR	translocated promoter region, nuclear basket protein	NA
8289	ARID1A	AT-rich interaction domain 1A	NA
2052	EPHX1	epoxide hydrolase 1	NA
55253	TYW1	tRNA-yW synthesizing protein 1 homolog	NA
9684	LRRC14	leucine rich repeat containing 14	NA
23167	EFR3A	EFR3 homolog A	NA
79608	RIC3	RIC3 acetylcholine receptor chaperone	NA
23390	ZDHHC17	zinc finger DHHC-type containing 17	NA
84866	TMEM25	transmembrane protein 25	NA
55140	ELP3	elongator acetyltransferase complex subunit 3	NA
55670	PEX26	peroxisomal biogenesis factor 26	NA
255520	ELMOD2	ELMO domain containing 2	NA
23019	CNOT1	CCR4-NOT transcription complex subunit 1	NA
23259	DDHD2	DDHD domain containing 2	NA
129685	TAF8	TATA-box binding protein associated factor 8	NA

25839	COG4	component of oligomeric golgi complex 4	NA
26953	RANBP6	RAN binding protein 6	NA
9724	UTP14C	UTP14C, small subunit processome component	NA
54680	ZNHIT6	zinc finger HIT-type containing 6	NA
4008	LMO7	LIM domain 7	NA
9857	CEP350	centrosomal protein 350	NA
58491	ZNF71	zinc finger protein 71	NA
154807	VKORC1L1	vitamin K epoxide reductase complex subunit 1 like 1	NA
147991	DPY19L3	dpy-19 like C-mannosyltransferase 3	NA
23091	ZC3H13	zinc finger CCCH-type containing 13	NA
1911	PHC1	polyhomeotic homolog 1	NA
79184	BRCC3	BRCA1/BRCA2-containing complex subunit 3	NA
10928	RALBP1	ralA binding protein 1	NA
375775	PNPLA7	patatin like phospholipase domain containing 7	NA
84792	FAM220A	family with sequence similarity 220 member A	NA
51175	TUBE1	tubulin epsilon 1	NA
116984	ARAP2	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 2	NA
9095	TBX19	T-box 19	NA
51633	OTUD6B	OTU domain containing 6B	NA
129531	MITD1	microtubule interacting and trafficking domain containing 1	NA
7174	TPP2	tripeptidyl peptidase 2	NA
257160	RNF214	ring finger protein 214	NA
64779	MTHFSD	methenyltetrahydrofolate synthetase domain containing	NA
91949	COG7	component of oligomeric golgi complex 7	NA
23013	SPEN	spen family transcriptional repressor	NA
4957	ODF2	outer dense fiber of sperm tails 2	NA
115294	PCMTD1	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 1	NA
4338	MOCS2	molybdenum cofactor synthesis 2	NA
4507	MTAP	methylthioadenosine phosphorylase	NA
51018	RRP15	ribosomal RNA processing 15 homolog	NA
1107	CHD3	chromodomain helicase DNA binding protein 3	NA
8085	KMT2D	lysine methyltransferase 2D	NA
440145	MZT1	mitotic spindle organizing protein 1	NA
10513	APPBP2	amyloid beta precursor protein binding protein 2	NA
23049	SMG1	SMG1, nonsense mediated mRNA decay associated PI3K related kinase	NA
55870	ASH1L	ASH1 like histone lysine methyltransferase	NA
2475	MTOR	mechanistic target of rapamycin kinase	NA
8663	EIF3C	eukaryotic translation initiation factor 3 subunit C	NA
9203	ZMYM3	zinc finger MYM-type containing 3	NA
25981	DNAH1	dynein axonemal heavy chain 1	NA
117178	SSX2IP	SSX family member 2 interacting protein	NA
23035	PHLPP2	PH domain and leucine rich repeat protein phosphatase 2	NA

23038	WDTC1	WD and tetratricopeptide repeats 1	NA
25996	REXO2	RNA exonuclease 2	NA
100507321	ERVK13-1	endogenous retrovirus group K13 member 1	NA
54842	MFSD6	major facilitator superfamily domain containing 6	NA
8883	NAE1	NEDD8 activating enzyme E1 subunit 1	NA
79053	ALG8	ALG8, alpha-1,3-glucosyltransferase	NA
11201	POLI	DNA polymerase iota	NA
256471	MFSD8	major facilitator superfamily domain containing 8	NA
8243	SMC1A	structural maintenance of chromosomes 1A	NA
155435	RBM33	RNA binding motif protein 33	NA
63908	NAPB	NSF attachment protein beta	NA
4967	OGDH	oxoglutarate dehydrogenase	NA
114876	OSBPL1A	oxysterol binding protein like 1A	208158_s_at
51192	CKLF	chemokine like factor	219161_s_at
100529251	CKLF-CMTM1	CKLF-CMTM1 readthrough	219161_s_at
121512	FGD4	FYVE, RhoGEF and PH domain containing 4	227948_at
132864	CPEB2	cytoplasmic polyadenylation element binding protein 2	235479_at
84898	PLXDC2	plexin domain containing 2	236297_at
137886	UBXN2B	UBX domain protein 2B	238903_at
1241	LTB4R	leukotriene B4 receptor	236172_at
19	ABCA1	ATP binding cassette subfamily A member 1	203505_at
57132	CHMP1B	charged multivesicular body protein 1B	218177_at
79772	MCTP1	multiple C2 and transmembrane domain containing 1	220122_at
254013	ETFBKMT	electron transfer flavoprotein subunit beta lysine methyltransferase	226258 at
196394	AMN1	antagonist of mitotic exit network 1 homolog	226258_at
115548	FCHO2	FCH domain only 2	228220 at
58517	RBM25	RNA binding motif protein 25	236613_at
6173	RPL36A	ribosomal protein L36a	243201 at
2549	GAB1	GRB2 associated binding protein 1	214987_at
51366	UBR5	ubiquitin protein ligase E3 component n-recognin 5	208883 at
22848	AAK1	AP2 associated kinase 1	205434_s_at
57505	AARS2	alanyl-tRNA synthetase 2, mitochondrial	230024 at
23527	ACAP2	ArfGAP with coiled-coil, ankyrin repeat and PH domains 2	1552472 a at
51	ACOX1	acyl-CoA oxidase 1	209600_s_at
123876	ACSM2A	acyl-CoA synthetase medium chain family member 2A	244723 at
10097	ACTR2	actin related protein 2	200727_s_at
440888	ACTR3BP2	ACTR3B pseudogene 2	224425 x at
102	ADAM10	ADAM metallopeptidase domain 10	202604_x_at
79602	ADIPOR2	adiponectin receptor 2	201346_at
54812	AFTPH	aftiphilin	217939_s_at
56894	AGPAT3	1-acylglycerol-3-phosphate O-acyltransferase 3	223182 s at
202	CRYBG1	crystallin beta-gamma domain containing 1	212543_at
26993	AKAP8L	A-kinase anchoring protein 8 like	218064_s_at

224	ALDH3A2	aldehyde dehydrogenase 3 family member A2	202053_s_at
80216	ALPK1	alpha kinase 1	227438_at
262	AMD1	adenosylmethionine decarboxylase 1	201196_s_at
118429	ANTXR2	ANTXR cell adhesion molecule 2	225524_at
311	ANXA11	annexin A11	228727_at
10139	ARFRP1	ADP ribosylation factor related protein 1	215984_s_at
64333	ARHGAP9	Rho GTPase activating protein 9	224451_x_at
51222	ZNF219	zinc finger protein 219	227855_at
51329	ARL6IP4	ADP ribosylation factor like GTPase 6 interacting protein 4	218216_x_at
411	ARSB	arylsulfatase B	206129_s_at
51008	ASCC1	activating signal cointegrator 1 complex subunit 1	219336_s_at
22926	ATF6	activating transcription factor 6	203952_at
25923	ATL3	atlastin GTPase 3	223452_s_at
1822	ATN1	atrophin 1	40489_at
23200	ATP11B	ATPase phospholipid transporting 11B (putative)	1564063_a_at
27032	ATP2C1	ATPase secretory pathway Ca2+ transporting 1	209934_s_at
506	ATP5F1B	ATP synthase F1 subunit beta	201322_at
9551	ATP5MF	ATP synthase membrane subunit f	202961_s_at
101929716	ARMC2-AS1	ARMC2 antisense RNA 1	202961_s_at
51382	ATP6V1D	ATPase H+ transporting V1 subunit D	208898_at
6310	ATXN1	ataxin 1	242230_at
11273	ATXN2L	ataxin 2 like	207798_s_at
57099	AVEN	apoptosis and caspase activation inhibitor	219366_at
23080	AVL9	AVL9 cell migration associated	212475_at
22893	BAHD1	bromo adjacent homology domain containing 1	203051_at
149428	BNIPL	BCL2 interacting protein like	236534_at
54964	C1orf56	chromosome 1 open reading frame 56	236534_at
10902	BRD8	bromodomain containing 8	202227_s_at
9577	BABAM2	BRISC and BRCA1 A complex member 2	212645_x_at
11120	BTN2A1	butyrophilin subfamily 2 member A1	244063_at
414235	PRR26	proline rich 26	1557548_at
64115	VSIR	V-set immunoregulatory receptor	225372_at
746	TMEM258	transmembrane protein 258	218213_s_at
10944	C11orf58	chromosome 11 open reading frame 58	1555364_at
80017	DGLUCY	D-glutamate cyclase	218298_s_at
56851	EMC7	ER membrane protein complex subunit 7	217898_at
84326	METTL26	methyltransferase like 26	228114_x_at
9605	VPS9D1	VPS9 domain containing 1	205781_at
29035	C16orf72	chromosome 16 open reading frame 72	1568954_s_at
79086	SMIM7	small integral membrane protein 7	224717_s_at
199675	MCEMP1	mast cell expressed membrane protein 1	235568_at
91304	TMEM259	transmembrane protein 259	213985_s_at
55924	INKA2	inka box actin regulator 2	229608_at
101928718	LOC101928718	uncharacterized LOC101928718	229608_at
26097	CHTOP	chromatin target of PRMT1	202559_x_at
714	C1QC	complement C1q C chain	225353_s_at

57136	APMAP	adipocyte plasma membrane associated protein	206656_s_at
51507	RTF2	replication termination factor 2	217737_x_at
80215	RUNX1-IT1	RUNX1 intronic transcript 1	220918_at
90288	EFCAB12	EF-hand calcium binding domain 12	232639_at
375341	C3orf62	chromosome 3 open reading frame 62	241817_at
100422952	MIR4271	microRNA 4271	241817_at
84273	NOA1	nitric oxide associated 1	223157_at
401152	C4orf3	chromosome 4 open reading frame 3	224602_at
285636	C5orf51	chromosome 5 open reading frame 51	226159_at
115416	MALSU1	mitochondrial assembly of ribosomal large subunit 1	226385_s_at
78996	CYREN	cell cycle regulator of NHEJ	220949_s_at
541565	C8orf58	chromosome 8 open reading frame 58	227263_at
124583	CANT1	calcium activated nucleotidase 1	46323_at
726	CAPN5	calpain 5	226292_at
84433	CARD11	caspase recruitment domain family member 11	1562368_at
55259	CASC1	cancer susceptibility 1	220168_at
838	CASP5	caspase 5	207500_at
10367	MICU1	mitochondrial calcium uptake 1	216903_s_at
867	CBL	Cbl proto-oncogene	243475_at
133957	CCDC127	coiled-coil domain containing 127	226515_at
159686	CFAP58	cilia and flagella associated protein 58	1557544_at
149483	CCDC17	coiled-coil domain containing 17	236320_at
92558	BICDL1	BICD family like cargo adaptor 1	228320_x_at
440193	CCDC88C	coiled-coil domain containing 88C	215343_at
961	CD47	CD47 molecule	211075_s_at
966	CD59	CD59 molecule (CD59 blood group)	200983_x_at
974	CD79B	CD79b molecule	205297_s_at
56882	CDC42SE1	CDC42 small effector 1	222537_s_at
988	CDC5L	cell division cycle 5 like	209056_s_at
23097	CDK19	cyclin dependent kinase 19	212899_at
634	CEACAM1	carcinoembryonic antigen related cell adhesion molecule 1	209498_at
10659	CELF2	CUGBP Elav-like family member 2	1554569_a_at
22897	CEP164	centrosomal protein 164	204251_s_at
54480	CHPF2	chondroitin polymerizing factor 2	221799_at
51267	CLEC1A	C-type lectin domain family 1 member A	219761_at
1296	COL8A2	collagen type VIII alpha 2 chain	52651_at
1315	COPB1	coatomer protein complex subunit beta 1	201359_at
22820	COPG1	coatomer protein complex subunit gamma 1	217749_at
51117	COQ4	coenzyme Q4	218328_at
90639	COX19	cytochrome c oxidase assembly factor COX19	235533_at
10328	EMC8	ER membrane protein complex subunit 8	227363_s_at
54973	INTS11	integrator complex subunit 11	233563_s_at
102465435	MIR6727	microRNA 6727	233563_s_at
9586	CREB5	cAMP responsive element binding protein 5	205931_s_at
1488	CTBP2	C-terminal binding protein 2	210554_s_at

1486	CTBS	chitobiase	218924_s_at
1506	CTRL	chymotrypsin like	214377_s_at
55280	CWF19L1	CWF19 like cell cycle control factor 1	218787_x_at
143884	CWF19L2	CWF19 like cell cycle control factor 2	237040_at
1525	CXADR	CXADR Ig-like cell adhesion molecule	203917_at
124637	CYB5D1	cytochrome b5 domain containing 1	226833_at
9266	CYTH2	cytohesin 2	1555842_at
27071	DAPP1	dual adaptor of phosphotyrosine and 3-phosphoinositides 1	222858_s_at
11258	DCTN3	dynactin subunit 3	204246_s_at
1654	DDX3X	DEAD-box helicase 3 X-linked	212514_x_at
26082	LINC02249	long intergenic non-protein coding RNA 2249	1569476_at
8847	DLEU2	deleted in lymphocytic leukemia 2	1564443_at
406948	MIR15A	microRNA 15a	1563229_at
83475	DOHH	deoxyhypusine hydroxylase	208141_s_at
5977	DPF2	double PHD fingers 2	202116_at
10589	DRAP1	DR1 associated protein 1	203258_at
84062	DTNBP1	dystrobrevin binding protein 1	223445_at
196403	DTX3	deltex E3 ubiquitin ligase 3	49051_g_at
56940	DUSP22	dual specificity phosphatase 22	218845_at
55837	EAPP	E2F associated phosphoprotein	202623_at
54583	EGLN1	egl-9 family hypoxia inducible factor 1	221497_x_at
317649	EIF4E3	eukaryotic translation initiation factor 4E family member 3	225939_at
30817	ADGRE2	adhesion G protein-coupled receptor E2	207610_s_at
84658	ADGRE3	adhesion G protein-coupled receptor E3	210724_at
2072	ERCC4	ERCC excision repair 4, endonuclease catalytic subunit	210158_at
57222	ERGIC1	endoplasmic reticulum-golgi intermediate compartment 1	224576_at
10613	ERLIN1	ER lipid raft associated 1	202441_at
2081	ERN1	endoplasmic reticulum to nucleus signaling 1	235745_at
2130	EWSR1	EWS RNA binding protein 1	210011_s_at
51010	EXOSC3	exosome component 3	223490_s_at
474383	F8A2	coagulation factor VIII associated 2	203274_at
474384	F8A3	coagulation factor VIII associated 3	203274_at
11124	FAF1	Fas associated factor 1	218080_x_at
399665	FAM102A	family with sequence similarity 102 member A	212400_at
90268	OTULIN	OTU deubiquitinase with linear linkage specificity	229268_at
81926	ABHD17A	abhydrolase domain containing 17A	221267_s_at
10827	FAM114A2	family with sequence similarity 114 member A2	218588_s_at
158293	FAM120AOS	family with sequence similarity 120A opposite strand	225395_s_at
286499	FAM133A	family with sequence similarity 133 member A	239481_at
285512	FAM13A-AS1	FAM13A antisense RNA 1	1558711_at
80011	FAM192A	family with sequence similarity 192 member A	217896_s_at
253725	WASHC2C	WASH complex subunit 2C	212929_s_at
387680	WASHC2A	WASH complex subunit 2A	212929_s_at

286336	FAM78A	family with sequence similarity 78 member A	227002_at
81545	FBXO38	F-box protein 38	224369_s_at
55338	LINC01949	long intergenic non-protein coding RNA 1949	215187_at
80094	BIN3-IT1	BIN3 intronic transcript 1	207287_at
643977	FLJ32255	uncharacterized LOC643977	235292_at
400579	LINC02076	long intergenic non-protein coding RNA 2076	1557895_at
6426	SRSF1	serine and arginine rich splicing factor 1	229246_at
91010	FMNL3	formin like 3	232249_at
96459	FNIP1	folliculin interacting protein 1	228250_at
3607	FOXK2	forkhead box K2	203064_s_at
3344	FOXN2	forkhead box N2	206708_at
2589	GALNT1	polypeptide N-acetylgalactosaminyltransferase 1	1568618_a_at
54834	GDAP2	ganglioside induced differentiation associated protein 2	219473_at
2677	GGCX	gamma-glutamyl carboxylase	214005_at
79893	GGNBP2	gametogenetin binding protein 2	218079_s_at
2686	GGT7	gamma-glutamyltransferase 7	226471_at
170575	GIMAP1	GTPase, IMAP family member 1	1552315_at
9815	GIT2	GIT ArfGAP 2	204982_at
2713	GK3P	glycerol kinase 3 pseudogene	215966_x_at
56287	GKN1	gastrokine 1	220191_at
2717	GLA	galactosidase alpha	214430_at
29997	NOP53	NOP53 ribosome biogenesis factor	234339_s_at
692091	SNORD23	small nucleolar RNA, C/D box 23	234339_s_at
92292	GLYATL1	glycine-N-acyltransferase like 1	1562089_at
51292	GMPR2	guanosine monophosphate reductase 2	217990_at
2776	GNAQ	G protein subunit alpha q	224861_at
59345	GNB4	G protein subunit beta 4	223488_s_at
2794	GNL1	G protein nucleolar 1 (putative)	203307_at
10007	GNPDA1	glucosamine-6-phosphate deaminase 1	202382_s_at
2804	GOLGB1	golgin B1	201057_s_at
27198	HCAR1	hydroxycarboxylic acid receptor 1	224131_at
222487	ADGRG3	adhesion G protein-coupled receptor G3	220404_at
2976	GTF3C2	general transcription factor IIIC subunit 2	204366_s_at
3015	H2AFZ	H2A histone family member Z	200853_at
283254	HARBI1	harbinger transposase derived 1	241498_at
414777	HCG18	HLA complex group 18	232975_at
23593	HEBP2	heme binding protein 2	203430_at
57520	HECW2	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2	232080_at
3082	HGF	hepatocyte growth factor	210755_at
55662	HIF1AN	hypoxia inducible factor 1 subunit alpha inhibitor	218525_s_at
10114	HIPK3	homeodomain interacting protein kinase 3	207764_s_at
3010	HIST1H1T	histone cluster 1 H1 family member t	207982_at
8351	HIST1H3D	histone cluster 1 H3 family member d	214472_at
3013	HIST1H2AD	histone cluster 1 H2A family member d	214472_at
8350	HIST1H3A	histone cluster 1 H3 family member a	214472_at

8352	HIST1H3C	histone cluster 1 H3 family member c	214472_at
8353	HIST1H3E	histone cluster 1 H3 family member e	214472_at
8354	HIST1H3I	histone cluster 1 H3 family member i	214472_at
8355	HIST1H3G	histone cluster 1 H3 family member g	214472_at
8356	HIST1H3J	histone cluster 1 H3 family member j	214472_at
8357	HIST1H3H	histone cluster 1 H3 family member h	214472_at
8358	HIST1H3B	histone cluster 1 H3 family member b	214472_at
8968	HIST1H3F	histone cluster 1 H3 family member f	214472_at
8360	HIST1H4D	histone cluster 1 H4 family member d	208076_at
3184	HNRNPD	heterogeneous nuclear ribonucleoprotein D	227744_s_at
3191	HNRNPL	heterogeneous nuclear ribonucleoprotein L	35201_at
3192	HNRNPU	heterogeneous nuclear ribonucleoprotein U	216855_s_at
10855	HPSE	heparanase	219403_s_at
3274	HRH2	histamine receptor H2	220805_at
3284	HSD3B2	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2	206294_at
3297	HSF1	heat shock transcription factor 1	213756 s at
3304	HSPA1B	heat shock protein family A (Hsp70) member 1B	202581_at
3303	HSPA1A	heat shock protein family A (Hsp70) member 1A	202581 at
10553	HTATIP2	HIV-1 Tat interactive protein 2	209448_at
3454	IFNAR1	interferon alpha and beta receptor subunit 1	225669 at
3455	IFNAR2	interferon alpha and beta receptor subunit 2	204785_x_at
10320	IKZF1	IKAROS family zinc finger 1	205038 at
3611	ILK	integrin linked kinase	201234_at
125476	INO80C	INO80 complex subunit C	229582 at
65123	INTS3	integrator complex subunit 3	202809_s_at
55846	ITFG2	integrin alpha FG-GAP repeat containing 2	220590 at
84522	JAGN1	jagunal homolog 1	223104_at
3716	JAK1	Janus kinase 1	201648 at
10899	JTB	jumping translocation breakpoint	210434_x_at
84078	KBTBD7	kelch repeat and BTB domain containing 7	223412 at
56888	KCMF1	potassium channel modulatory factor 1	222471_s_at
90134	KCNH7	potassium voltage-gated channel subfamily H member 7	224099 at
3759	KCNJ2	potassium voltage-gated channel subfamily J member 2	206765_at
51780	KDM3B	lysine demethylase 3B	210878 s at
9870	AREL1	apoptosis resistant E3 ubiquitin protein ligase 1	202128_at
57654	UVSSA	UV stimulated scaffold protein A	233893 s at
90231	KIAA2013	KIAA2013	224708_at
57498	KIDINS220	kinase D interacting substrate 220	1557246 at
23095	KIF1B	kinesin family member 1B	225878_at
1316	KLF6	Kruppel like factor 6	224606 at
8609	KLF7	Kruppel like factor 7	1555420_a_at
23276	KLHL18	kelch like family member 18	212882 at
57563	KLHL8	kelch like family member 8	226874_at
3836	KPNA1	karyopherin subunit alpha 1	213741 s at
3838	KPNA2	karyopherin subunit alpha 2	201088_at

3840	KPNA4	karyopherin subunit alpha 4	213567_at
26524	LATS2	large tumor suppressor kinase 2	227013_at
353135	LCE1E	late cornified envelope 1E	1559224_at
84648	LCE3D	late cornified envelope 3D	224328_s_at
84458	LCOR	ligand dependent nuclear receptor corepressor	226520_at
54923	LIME1	Lck interacting transmembrane adaptor 1	219541_at
3985	LIMK2	LIM domain kinase 2	202193_at
100049716	LOC100049716	uncharacterized LOC100049716	236886_at
100127983	C8orf88	chromosome 8 open reading frame 88	1557961_s_at
54758	KLHDC4	kelch domain containing 4	234738_s_at
56853	CELF4	CUGBP Elav-like family member 4	243666_at
80218	NAA50	N(alpha)-acetyltransferase 50, NatE catalytic subunit	239138_at
11068	CYB561D2	cytochrome b561 family member D2	229636_at
8365	HIST1H4H	histone cluster 1 H4 family member h	232035_at
285692	LINC02112	long intergenic non-protein coding RNA 2112	1561650_s_at
344595	DUBR	DPPA2 upstream binding RNA	235606_at
400499	LOC400499	vitellogenin	244889_at
401324	LOC401324	uncharacterized LOC401324	1563586_at
440104	TMEM198B	transmembrane protein 198B (pseudogene)	227106_at
442075	EMC3-AS1	EMC3 antisense RNA 1	237005_at
645513	LOC645513	septin 7 pseudogene	1556257_at
731424	MIR3945HG	MIR3945 host gene	1559777_at
2846	LPAR4	lysophosphatidic acid receptor 4	206960_at
54947	LPCAT2	lysophosphatidylcholine acyltransferase 2	239598_s_at
9926	LPGAT1	lysophosphatidylglycerol acyltransferase 1	1555058_a_at
123355	LRRC28	leucine rich repeat containing 28	227423_at
64101	LRRC4	leucine rich repeat containing 4	223552_at
57819	LSM2	LSM2 homolog, U6 small nuclear RNA and mRNA degradation associated	209449_at
388695	LYSMD1	LysM domain containing 1	232283_at
10894	LYVE1	lymphatic vessel endothelial hyaluronan receptor 1	219059_s_at
10916	MAGED2	MAGE family member D2	208682_s_at
93487	MAPK1IP1L	mitogen-activated protein kinase 1 interacting protein 1 like	212644_s_at
4140	MARK3	microtubule affinity regulating kinase 3	202569_s_at
4141	MARS	methionyl-tRNA synthetase	201475_x_at
102465454	MIR6758	microRNA 6758	201475_x_at
8930	MBD4	methyl-CpG binding domain 4, DNA glycosylase	214047_s_at
4174	MCM5	minichromosome maintenance complex component 5	201755_at
55784	MCTP2	multiple C2 and transmembrane domain containing 2	1554833_at
9969	MED13	mediator complex subunit 13	201986_at
51003	MED31	mediator complex subunit 31	222867_s_at
59274	TLNRD1	talin rod domain containing 1	223264_at
23184	MESD	mesoderm development LRP chaperone	238679_at
4238	MFAP3	microfibril associated protein 3	1552312_a_at

56947	MFF	mitochondrial fission factor	223354_x_at
4242	MFNG	MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase	213783_at
4247	MGAT2	mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	211061_s_at
4258	MGST2	microsomal glutathione S-transferase 2	204168_at
9645	MICAL2	microtubule associated monooxygenase, calponin and LIM domain containing 2	236475_at
6945	MLX	MAX dimerization protein MLX	217910_x_at
27249	MMADHC	metabolism of cobalamin associated D	217883_at
22879	MON1B	MON1 homolog B, secretory trafficking associated	203644_s_at
10934	MORF4	mortality factor 4 (pseudogene)	221381_s_at
10933	MORF4L1	mortality factor 4 like 1	221381_s_at
158747	MOSPD2	motile sperm domain containing 2	64883_at
10205	MPZL2	myelin protein zero like 2	230518_at
3140	MR1	major histocompatibility complex, class I-related	207565_s_at
51073	MRPL4	mitochondrial ribosomal protein L4	223743_s_at
64975	MRPL41	mitochondrial ribosomal protein L41	227186_s_at
63931	MRPS14	mitochondrial ribosomal protein S14	203800_s_at
64960	MRPS15	mitochondrial ribosomal protein S15	223292_s_at
339287	MSL1	MSL complex subunit 1	224765_at
10943	MSL3	MSL complex subunit 3	207551_s_at
54516	MTRF1L	mitochondrial translational release factor 1 like	225206_s_at
4600	MX2	MX dynamin like GTPase 2	204994_at
29116	MYLIP	myosin regulatory light chain interacting protein	228098_s_at
10443	N4BP2L2	NEDD4 binding protein 2 like 2	214748_at
54187	NANS	N-acetylneuraminate synthase	218189_s_at
4683	NBN	nibrin	202907_s_at
4077	NBR1	NBR1 autophagy cargo receptor	1568857_a_at
644660	RAD21-AS1	RAD21 antisense RNA 1	1560207_at
100500914	MIR3610	microRNA 3610	1560207_at
84719	LINC00260	long intergenic non-protein coding RNA 260	210711_at
4704	NDUFA9	NADH:ubiquinone oxidoreductase subunit A9	208969_at
4726	NDUFS6	NADH:ubiquinone oxidoreductase subunit S6	203606_at
10725	NFAT5	nuclear factor of activated T cells 5	224984_at
4807	NHLH1	nescient helix-loop-helix 1	214628_at
51199	NIN	ninein	224304_x_at
81614	NIPA2	NIPA magnesium transporter 2	212129_at
4817	NIT1	nitrilase 1	202891_at
9520	NPEPPS	aminopeptidase puromycin sensitive	201455_s_at
2908	NR3C1	nuclear receptor subfamily 3 group C member 1	201865_x_at
4898	NRDC	nardilysin convertase	208709_s_at
4905	NSF	N-ethylmaleimide sensitive factor, vesicle fusing ATPase	202395_at
8439	NSMAF	neutral sphingomyelinase activation associated factor	232149_s_at
4926	NUMA1	nuclear mitotic apparatus protein 1	214250_at
8021	NUP214	nucleoporin 214	202155_s_at

55301	OLAH	oleoyl-ACP hydrolase	219975_x_at
5032	P2RY11	purinergic receptor P2Y11	214546_s_at
692312	PPAN-P2RY11	PPAN-P2RY11 readthrough	214546_s_at
5048	PAFAH1B1	platelet activating factor acetylhydrolase 1b regulatory subunit 1	200815_s_at
10914	PAPOLA	poly(A) polymerase alpha	212718_at
5069	PAPPA	pappalysin 1	224940_s_at
9061	PAPSS1	3'-phosphoadenosine 5'-phosphosulfate synthase 1	209043_at
64098	PARVG	parvin gamma	223562_at
51248	PDZD11	PDZ domain containing 11	223037_at
27043	PELP1	proline, glutamate and leucine rich protein 1	215354_s_at
283871	PGP	phosphoglycolate phosphatase	222622_at
51105	PHF20L1	PHD finger protein 20 like 1	222133_s_at
51317	PHF21A	PHD finger protein 21A	1554153_a_at
5256	PHKA2	phosphorylase kinase regulatory subunit alpha 2	209439_s_at
8554	PIAS1	protein inhibitor of activated STAT 1	217864_s_at
5279	PIGC	phosphatidylinositol glycan anchor biosynthesis class C	202846_s_at
55011	PIH1D1	PIH1 domain containing 1	217872_at
5289	PIK3C3	phosphatidylinositol 3-kinase catalytic subunit type 3	204297_at
5294	PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma	206369_s_at
54984	PINX1	PIN2 (TERF1) interacting telomerase inhibitor 1	223907_s_at
5337	PLD1	phospholipase D1	215724_at
122618	PLD4	phospholipase D family member 4	235802_at
11243	PMF1	polyamine modulated factor 1	202337_at
29774	POM121L9P	POM121 transmembrane nucleoporin like 9, pseudogene	222253_s_at
10585	POMT1	protein O-mannosyltransferase 1	218476_at
132160	PPM1M	protein phosphatase, Mg2+/Mn2+ dependent 1M	226074_at
5520	PPP2R2A	protein phosphatase 2 regulatory subunit Balpha	202313_at
5534	PPP3R1	protein phosphatase 3 regulatory subunit B, alpha	204506_at
5536	PPP5C	protein phosphatase 5 catalytic subunit	201979_s_at
5537	PPP6C	protein phosphatase 6 catalytic subunit	206174_s_at
9701	PPP6R2	protein phosphatase 6 regulatory subunit 2	202792_s_at
5547	PRCP	prolylcarboxypeptidase	201494_at
7799	PRDM2	PR/SET domain 2	205277_at
5573	PRKAR1A	protein kinase cAMP-dependent type I regulatory subunit alpha	200604_s_at
3276	PRMT1	protein arginine methyltransferase 1	206445_s_at
10419	PRMT5	protein arginine methyltransferase 5	1564521_x_at
114224	LOC114224	uncharacterized LOC114224	223797_at
26121	PRPF31	pre-mRNA processing factor 31	202407_s_at
79056	PRRG4	proline rich and Gla domain 4	238513_at
5693	PSMB5	proteasome subunit beta 5	208799_at
29893	PSMC3IP	PSMC3 interacting protein	205956_x_at
5707	PSMD1	proteasome 26S subunit, non-ATPase 1	201198_s_at
5725	PTBP1	polypyrimidine tract binding protein 1	212015_x_at

100616459	MIR4745	microRNA 4745	212015_x_at
401494	HACD4	3-hydroxyacyl-CoA dehydratase 4	244050_at
5788	PTPRC	protein tyrosine phosphatase receptor type C	1569830_at
29920	PYCR2	pyrroline-5-carboxylate reductase 2	231715_s_at
102466270	MIR6741	microRNA 6741	231715_s_at
10890	RAB10	RAB10, member RAS oncogene family	222981_s_at
8766	RAB11A	RAB11A, member RAS oncogene family	200863_s_at
84440	RAB11FIP4	RAB11 family interacting protein 4	225739_at
5906	RAP1A	RAP1A, member of RAS oncogene family	228548_at
22821	RASA3	RAS p21 protein activator 3	206221_at
158158	RASEF	RAS and EF-hand domain containing	1553186_x_at
5936	RBM4	RNA binding motif protein 4	200997_at
100526737	RBM14-RBM4	RBM14-RBM4 readthrough	200997_at
55696	RBM22	RNA binding motif protein 22	218134_s_at
27303	RBMS3	RNA binding motif single stranded interacting protein 3	238447_at
149041	RC3H1	ring finger and CCCH-type domains 1	228996_at
285613	RELL2	RELT like 2	1564031_a_at
11079	RER1	retention in endoplasmic reticulum sorting receptor 1	202296_s_at
57455	REXO1	RNA exonuclease 1 homolog	226144_at
100302210	MIR1909	microRNA 1909	226144_at
64326	COP1	COP1 E3 ubiquitin ligase	1552617_a_at
6002	RGS12	regulator of G protein signaling 12	1555022_at
11342	RNF13	ring finger protein 13	201779_s_at
81847	RNF146	ring finger protein 146	223886_s_at
10193	RNF41	ring finger protein 41	201961_s_at
9991	PTBP3	polypyrimidine tract binding protein 3	207223_s_at
80135	RPF1	ribosome production factor 1 homolog	234243_at
58490	RPRD1B	regulation of nuclear pre-mRNA domain containing 1B	225024_at
83861	RSPH3	radial spoke head 3	1568613_at
9667	SAFB2	scaffold attachment factor B2	32099_at
79685	SAP30L	SAP30 like	225509_at
51128	SAR1B	secretion associated Ras related GTPase 1B	223512_at
51150	SDF4	stromal cell derived factor 4	221972_s_at
6400	SEL1L	SEL1L adaptor subunit of ERAD E3 ubiquitin ligase	202063_s_at
10509	SEMA4B	semaphorin 4B	234725_s_at
22928	SEPHS2	selenophosphate synthetase 2	200961_at
6418	SET	SET nuclear proto-oncogene	200630_x_at
646817	SETSIP	SET like protein	200630_x_at
6421	SFPQ	splicing factor proline and glutamine rich	201585_s_at
6433	SFSWAP	splicing factor SWAP	202775_s_at
375035	SFT2D2	SFT2 domain containing 2	214838_at
8879	SGPL1	sphingosine-1-phosphate lyase 1	212321_at
9905	SGSM2	small G protein signaling modulator 2	212319_at
6457	SH3GL3	SH3 domain containing GRB2 like 3, endophilin A3	211565_at

128646	SIRPD	signal regulatory protein delta	232891_at
6507	SLC1A3	solute carrier family 1 member 3	202800_at
284439	SLC25A42	solute carrier family 25 member 42	226737_at
30061	SLC40A1	solute carrier family 40 member 1	233123_at
6546	SLC8A1	solute carrier family 8 member A1	241752_at
9748	SLK	STE20 like kinase	206874_s_at
4087	SMAD2	SMAD family member 2	235598_at
6597	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	214728_x_at
9342	SNAP29	synaptosome associated protein 29	222597_at
29916	SNX11	sorting nexin 11	53912_at
81609	SNX27	sorting nexin 27	221006_s_at
58533	SNX6	sorting nexin 6	217789_at
6646	SOAT1	sterol O-acyltransferase 1	244661_at
6651	SON	SON DNA binding protein	226465_s_at
6653	SORL1	sortilin related receptor 1	230707_at
6654	SOS1	SOS Ras/Rac guanine nucleotide exchange factor 1	230337_at
6667	SP1	Sp1 transcription factor	224760_at
6672	SP100	SP100 nuclear antigen	210985_s_at
3431	SP110	SP110 nuclear body protein	223980_s_at
6689	SPIB	Spi-B transcription factor	232739_at
9806	SPOCK2	SPARC (osteonectin), cwcv and kazal like domains proteoglycan 2	202523_s_at
8405	SPOP	speckle type BTB/POZ protein	238923_at
80725	SRCIN1	SRC kinase signaling inhibitor 1	232547_at
51593	SRRT	serrate, RNA effector molecule	201680_x_at
85464	SSH2	slingshot protein phosphatase 2	226080_at
29101	SSU72	SSU72 homolog, RNA polymerase II CTD phosphatase	223051_at
6487	ST3GAL3	ST3 beta-galactoside alpha-2,3-sialyltransferase 3	225905_s_at
256435	ST6GALNAC3	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	235334_at
10254	STAM2	signal transducing adaptor molecule 2	215044_s_at
10617	STAMBP	STAM binding protein	202811_at
6774	STAT3	signal transducer and activator of transcription 3	208992_s_at
6777	STAT5B	signal transducer and activator of transcription 5B	1555088_x_at
9262	STK17B	serine/threonine kinase 17b	205214_at
6789	STK4	serine/threonine kinase 4	205411_at
92335	STRADA	STE20 related adaptor alpha	52169_at
3703	STT3A	STT3 oligosaccharyltransferase complex catalytic subunit A	202223_at
8675	STX16	syntaxin 16	221499_s_at
6827	SUPT4H1	SPT4 homolog, DSIF elongation factor subunit	201483_s_at
6834	SURF1	SURF1 cytochrome c oxidase assembly factor	204295_at
64420	SUSD1	sushi domain containing 1	226264_at
51111	KMT5B	lysine methyltransferase 5B	222759_at
8867	SYNJ1	synaptojanin 1	207594_s_at

26115	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2	224952_at
6929	TCF3	transcription factor 3	209153_s_at
6949	TCOF1	treacle ribosome biogenesis factor 1	202384_s_at
9524	TECR	trans-2,3-enoyl-CoA reductase	208336_s_at
7039	TGFA	transforming growth factor alpha	205016_at
10618	TGOLN2	trans-golgi network protein 2	212043_at
8563	THOC5	THO complex 5	209418_s_at
1678	TIMM8A	translocase of inner mitochondrial membrane 8A	210800_at
114609	TIRAP	TIR domain containing adaptor protein	1552360_a_at
81793	TLR10	toll like receptor 10	223750_s_at
7100	TLR5	toll like receptor 5	210166_at
10548	TM9SF1	transmembrane 9 superfamily member 1	209149_s_at
7009	TMBIM6	transmembrane BAX inhibitor motif containing 6	200803_s_at
11322	TMC6	transmembrane channel like 6	204328_at
283578	TMED8	transmembrane p24 trafficking protein family member 8	225343_at
387521	TMEM189	transmembrane protein 189	223186_at
7335	UBE2V1	ubiquitin conjugating enzyme E2 V1	223186_at
387522	TMEM189-UBE2V1	TMEM189-UBE2V1 readthrough	223186_at
51259	TMEM216	transmembrane protein 216	223305_at
55161	TMEM33	transmembrane protein 33	235907_at
23585	TMEM50A	transmembrane protein 50A	222401_s_at
55217	TMLHE	trimethyllysine hydroxylase, epsilon	1554206_at
126282	TNFAIP8L1	TNF alpha induced protein 8 like 1	227420_at
7150	TOP1	DNA topoisomerase I	208901_s_at
26092	TOR1AIP1	torsin 1A interacting protein 1	212409_s_at
7169	TPM2	tropomyosin 2	204083_s_at
8459	TPST2	tyrosylprotein sulfotransferase 2	204079_at
7189	TRAF6	TNF receptor associated factor 6	205558_at
84231	TRAF7	TNF receptor associated factor 7	223031_s_at
5987	TRIM27	tripartite motif containing 27	212116_at
10475	TRIM38	tripartite motif containing 38	203610_s_at
140691	TRIM69	tripartite motif containing 69	1568592_at
51393	TRPV2	transient receptor potential cation channel subfamily V member 2	222855_s_at
57616	TSHZ3	teashirt zinc finger homeobox 3	223393_s_at
81619	TSPAN14	tetraspanin 14	221002_s_at
10100	TSPAN2	tetraspanin 2	214606_at
7265	TTC1	tetratricopeptide repeat domain 1	201434_at
283237	TTC9C	tetratricopeptide repeat domain 9C	226175_at
64852	TUT1	terminal uridylyl transferase 1, U6 snRNA-specific	218965_s_at
84959	UBASH3B	ubiquitin associated and SH3 domain containing B	228359_at
7322	UBE2D2	ubiquitin conjugating enzyme E2 D2	201345_s_at
100131816	UBE2DNL	ubiquitin conjugating enzyme E2 D N-terminal like (pseudogene)	1554763_at
7326	UBE2G1	ubiquitin conjugating enzyme E2 G1	209141_at

3093	UBE2K	ubiquitin conjugating enzyme E2 K	225179_at
23190	UBXN4	UBX domain protein 4	212007_at
56886	UGGT1	UDP-glucose glycoprotein glucosyltransferase 1	222569_at
7360	UGP2	UDP-glucose pyrophosphorylase 2	205480_s_at
23074	UHRF1BP1L	UHRF1 binding protein 1 like	213120_at
29761	USP25	ubiquitin specific peptidase 25	1563497_at
9960	USP3	ubiquitin specific peptidase 3	221654_s_at
57558	USP35	ubiquitin specific peptidase 35	229198_at
7874	USP7	ubiquitin specific peptidase 7	230967_s_at
7419	VDAC3	voltage dependent anion channel 3	208845_at
51652	CHMP3	charged multivesicular body protein 3	217837_s_at
100526767	RNF103-CHMP3	RNF103-CHMP3 readthrough	217837_s_at
55737	VPS35	VPS35 retromer complex component	217727_x_at
49856	WRAP73	WD repeat containing, antisense to TP73	236381_s_at
7516	XRCC2	X-ray repair cross complementing 2	207598_x_at
286451	YIPF6	Yip1 domain family member 6	212343_at
7528	YY1	YY1 transcription factor	224711_at
51776	MAP3K20	mitogen-activated protein kinase kinase kinase 20	1555259_at
79670	TUT7	terminal uridylyl transferase 7	236155_at
678	ZFP36L2	ZFP36 ring finger protein like 2	201367_s_at
7566	ZNF18	zinc finger protein 18	226787_at
7752	ZNF200	zinc finger protein 200	214706_at
7773	ZNF230	zinc finger protein 230	1570135_at
7572	ZNF24	zinc finger protein 24	1554045_at
23528	ZNF281	zinc finger protein 281	222619_at
126299	ZNF428	zinc finger protein 428	215429_s_at
57473	ZNF512B	zinc finger protein 512B	55872_at
8832	CD84	CD84 molecule	205988_at
55760	DHX32	DEAH-box helicase 32 (putative)	218198_at
84681	HINT2	histidine triad nucleotide binding protein 2	224415_s_at
93463	LOC93463	uncharacterized LOC93463	234861_at
81567	TXNDC5	thioredoxin domain containing 5	221253_s_at
100526836	BLOC1S5-TXNDC5	BLOC1S5-TXNDC5 readthrough (NMD candidate)	221253_s_at
23221	RHOBTB2	Rho related BTB domain containing 2	209441_at
6473	SHOX	short stature homeobox	207570_at
6847	SYCP1	synaptonemal complex protein 1	206740_x_at
8189	SYMPK	symplekin	202339_at
7027	TFDP1	transcription factor Dp-1	242939_at
23471	TRAM1	translocation associated membrane protein 1	201399_s_at
219699	UNC5B	unc-5 netrin receptor B	226899_at
22985	ACIN1	apoptotic chromatin condensation inducer 1	201715_s_at
51412	ACTL6B	actin like 6B	206014_at
11154	AP4S1	adaptor related protein complex 4 subunit sigma 1	210278_s_at
126792	B3GALT6	beta-1,3-galactosyltransferase 6	1553959_a_at
25874	MPC2	mitochondrial pyruvate carrier 2	202427_s_at
55009	C19orf24	chromosome 19 open reading frame 24	221587_s_at

57524	CASKIN1	CASK interacting protein 1	1552689_at
348254	CCDC144CP	coiled-coil domain containing 144C, pseudogene	1561271_at
1329	COX5B	cytochrome c oxidase subunit 5B	213736_at
51181	DCXR	dicarbonyl and L-xylulose reductase	217973_at
60528	ELAC2	elaC ribonuclease Z 2	201766_at
22936	ELL2	elongation factor for RNA polymerase II 2	226982_at
2067	ERCC1	ERCC excision repair 1, endonuclease non-catalytic subunit	203719_at
728215	FAM155A	family with sequence similarity 155 member A	214825_at
5348	FXYD1	FXYD domain containing ion transport regulator 1	205384_at
23151	GRAMD4	GRAM domain containing 4	212856_at
400581	GRAPL	GRB2 related adaptor protein like	1559688_at
2870	GRK6	G protein-coupled receptor kinase 6	210981_s_at
64423	INF2	inverted formin, FH2 and WH2 domain containing	224469_s_at
83737	ITCH	itchy E3 ubiquitin protein ligase	239101_at
10300	KATNB1	katanin regulatory subunit B1	203162_s_at
54442	KCTD5	potassium channel tetramerization domain containing 5	229837_s_at
51274	KLF3	Kruppel like factor 3	219657_s_at
414332	LCN10	lipocalin 10	238071_at
100129250	SMIM27	small integral membrane protein 27	221979_at
100130744	LOC100130744	uncharacterized LOC100130744	236228_at
374819	LRRC37A3	leucine rich repeat containing 37 member A3	244478_at
25804	LSM4	LSM4 homolog, U6 small nuclear RNA and mRNA degradation associated	202736_s_at
79903	NAA60	N(alpha)-acetyltransferase 60, NatF catalytic subunit	45526_g_at
57447	NDRG2	NDRG family member 2	214278_s_at
9148	NEURL1	neuralized E3 ubiquitin protein ligase 1	204889_s_at
8106	PABPN1	poly(A) binding protein nuclear 1	201545_s_at
100529063	BCL2L2-PABPN1	BCL2L2-PABPN1 readthrough	201545_s_at
10471	PFDN6	prefoldin subunit 6	222029_x_at
102465501	MIR6834	microRNA 6834	222029_x_at
11284	PNKP	polynucleotide kinase 3'-phosphatase	218961_s_at
283659	PRTG	protogenin	229073_at
57148	RALGAPB	Ral GTPase activating protein non-catalytic beta subunit	221738_at
5935	RBM3	RNA binding motif protein 3	222026_at
9986	RCE1	Ras converting CAAX endopeptidase 1	205332_at
253260	RICTOR	RPTOR independent companion of MTOR complex 2	228248_at
55197	RPRD1A	regulation of nuclear pre-mRNA domain containing 1A	228566_at
6869	TACR1	tachykinin receptor 1	210637_at
9277	WDR46	WD repeat domain 46	209196_at
23144	ZC3H3	zinc finger CCCH-type containing 3	213445_at
79038	ZFYVE21	zinc finger FYVE-type containing 21	219929_s_at
84622	ZNF594	zinc finger protein 594	235690_at
2182	ACSL4	acyl-CoA synthetase long chain family member 4	NA

23287	AGTPBP1	ATP/GTP binding protein 1	NA
64853	AIDA	axin interactor, dorsalization associated	NA
54882	ANKHD1	ankyrin repeat and KH domain containing 1	NA
100132420	ANKRD36P1	ankyrin repeat domain 36 pseudogene 1	NA
101927547	ANKRD44-IT1	ANKRD44 intronic transcript 1	NA
57584	ARHGAP21	Rho GTPase activating protein 21	NA
10092	ARPC5	actin related protein 2/3 complex subunit 5	NA
528	ATP6V1C1	ATPase H+ transporting V1 subunit C1	NA
10678	B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2	NA
8678	BECN1	beclin 1	NA
10904	BLCAP	BLCAP, apoptosis inducing factor	NA
2972	BRF1	BRF1, RNA polymerase III transcription initiation factor subunit	NA
9184	BUB3	BUB3, mitotic checkpoint protein	NA
84281	C2orf88	chromosome 2 open reading frame 88	NA
441108	C5orf56	chromosome 5 open reading frame 56	NA
10203	CALCRL	calcitonin receptor like receptor	NA
830	CAPZA2	capping actin protein of muscle Z-line subunit alpha 2	NA
100151661	CBX5P1	chromobox 5 pseudogene 1	NA
90693	CCDC126	coiled-coil domain containing 126	NA
952	CD38	CD38 molecule	NA
916	CD3E	CD3e molecule	NA
1039	CDR2	cerebellar degeneration related protein 2	NA
24146	CLDN15	claudin 15	NA
4849	CNOT3	CCR4-NOT transcription complex subunit 3	NA
729530	CNOT6LP1	CCR4-NOT transcription complex subunit 6-like pseudogene 1	NA
10087	COL4A3BP	collagen type IV alpha 3 binding protein	NA
80315	CPEB4	cytoplasmic polyadenylation element binding protein 4	NA
7812	CSDE1	cold shock domain containing E1	NA
80344	DCAF11	DDB1 and CUL4 associated factor 11	NA
1642	DDB1	damage specific DNA binding protein 1	NA
54514	DDX4	DEAD-box helicase 4	NA
8662	EIF3B	eukaryotic translation initiation factor 3 subunit B	NA
2131	EXT1	exostosin glycosyltransferase 1	NA
57700	FAM160B1	family with sequence similarity 160 member B1	NA
100128011	FAR1P1	fatty acyl-CoA reductase 1 pseudogene 1	NA
114907	FBXO32	F-box protein 32	NA
2319	FLOT2	flotillin 2	NA
23360	FNBP4	formin binding protein 4	NA
8939	FUBP3	far upstream element binding protein 3	NA
23193	GANAB	glucosidase II alpha subunit	NA
115361	GBP4	guanylate binding protein 4	NA
388646	GBP7	guanylate binding protein 7	NA
2764	GMFB	glia maturation factor beta	NA

100499251	GMFBP1	glia maturation factor beta pseudogene 1	NA
440270	GOLGA8B	golgin A8 family member B	NA
9402	GRAP2	GRB2 related adaptor protein 2	NA
645506	GYG1P3	glycogenin 1 pseudogene 3	NA
100873938	GYG2-AS1	GYG2 antisense RNA 1	NA
2994	GYPB	glycophorin B (MNS blood group)	NA
441484	HAUS1P2	HAUS augmin like complex subunit 1 pseudogene 2	NA
3069	HDLBP	high density lipoprotein binding protein	NA
100421389	HNRNPA1P9	heterogeneous nuclear ribonucleoprotein A1 pseudogene 9	NA
10075	HUWE1	HECT, UBA and WWE domain containing 1, E3 ubiquitin protein ligase	NA
3376	IARS	isoleucyl-tRNA synthetase	NA
3422	IDI1	isopentenyl-diphosphate delta isomerase 1	NA
3566	IL4R	interleukin 4 receptor	NA
3575	IL7R	interleukin 7 receptor	NA
3609	ILF3	interleukin enhancer binding factor 3	NA
3695	ITGB7	integrin subunit beta 7	NA
3702	ITK	IL2 inducible T cell kinase	NA
133746	JMY	junction mediating and regulatory protein, p53 cofactor	NA
84162	KIAA1109	KIAA1109	NA
11275	KLHL2	kelch like family member 2	NA
388818	KRTAP26-1	keratin associated protein 26-1	NA
83746	L3MBTL2	L3MBTL2, polycomb repressive complex 1 subunit	NA
23367	LARP1	La ribonucleoprotein domain family member 1	NA
3945	LDHB	lactate dehydrogenase B	NA
101929559	LINC00211	long intergenic non-protein coding RNA 211	NA
441123	LINC00266-3	long intergenic non-protein coding RNA 266-3	NA
100507612	LINC00402	long intergenic non-protein coding RNA 402	NA
100506190	LINC00963	long intergenic non-protein coding RNA 963	NA
3998	LMAN1	lectin, mannose binding 1	NA
8425	LTBP4	latent transforming growth factor beta binding protein 4	NA
4154	MBNL1	muscleblind like splicing regulator 1	NA
100616493	MIR4441	microRNA 4441	NA
4300	MLLT3	MLLT3, super elongation complex subunit	NA
9611	NCOR1	nuclear receptor corepressor 1	NA
283820	NOMO2	NODAL modulator 2	NA
408050	NOMO3	NODAL modulator 3	NA
30833	NT5C	5', 3'-nucleotidase, cytosolic	NA
51667	NUB1	negative regulator of ubiquitin like proteins 1	NA
114882	OSBPL8	oxysterol binding protein like 8	NA
5087	PBX1	PBX homeobox 1	NA
84333	PCGF5	polycomb group ring finger 5	NA
5164	PDK2	pyruvate dehydrogenase kinase 2	NA
5824	PEX19	peroxisomal biogenesis factor 19	NA

65979	PHACTR4	phosphatase and actin regulator 4	NA
5509	PPP1R3D	protein phosphatase 1 regulatory subunit 3D	NA
5527	PPP2R5C	protein phosphatase 2 regulatory subunit B'gamma	NA
5583	PRKCH	protein kinase C eta	NA
439949	PRKCQ-AS1	PRKCQ antisense RNA 1	NA
54899	PXK	PX domain containing serine/threonine kinase like	NA
9230	RAB11B	RAB11B, member RAS oncogene family	NA
5984	RFC4	replication factor C subunit 4	NA
55177	RMDN3	regulator of microtubule dynamics 3	NA
106480046	RNU6-1092P	RNA, U6 small nuclear 1092, pseudogene	NA
106481249	RNU6-226P	RNA, U6 small nuclear 226, pseudogene	NA
6102	RP2	RP2, ARL3 GTPase activating protein	NA
100271133	RPS4XP10	ribosomal protein S4X pseudogene 10	NA
6240	RRM1	ribonucleotide reductase catalytic subunit M1	NA
677770	SCARNA22	small Cajal body-specific RNA 22	NA
152579	SCFD2	sec1 family domain containing 2	NA
51460	SFMBT1	Scm like with four mbt domains 1	NA
94097	SFXN5	sideroflexin 5	NA
57568	SIPA1L2	signal induced proliferation associated 1 like 2	NA
6504	SLAMF1	signaling lymphocytic activation molecule family member 1	NA
56996	SLC12A9	solute carrier family 12 member 9	NA
117247	SLC16A10	solute carrier family 16 member 10	NA
29957	SLC25A24	solute carrier family 25 member 24	NA
84255	SLC37A3	solute carrier family 37 member 3	NA
54407	SLC38A2	solute carrier family 38 member 2	NA
121456	SLC9A7P1	solute carrier family 9 member 7 pseudogene 1	NA
162394	SLFN5	schlafen family member 5	NA
338427	SNORD108	small nucleolar RNA, C/D box 108	NA
6655	SOS2	SOS Ras/Rho guanine nucleotide exchange factor 2	NA
11262	SP140	SP140 nuclear body protein	NA
93349	SP140L	SP140 nuclear body protein like	NA
23514	SPIDR	scaffold protein involved in DNA repair	NA
10929	SRSF8	serine and arginine rich splicing factor 8	NA
6772	STAT1	signal transducer and activator of transcription 1	NA
10388	SYCP2	synaptonemal complex protein 2	NA
11138	TBC1D8	TBC1 domain family member 8	NA
25976	TIPARP	TCDD inducible poly(ADP-ribose) polymerase	NA
153339	TMEM167A	transmembrane protein 167A	NA
23043	TNIK	TRAF2 and NCK interacting kinase	NA
9537	TP53I11	tumor protein p53 inducible protein 11	NA
28745	TRAJ10	T cell receptor alpha joining 10	NA
28754	TRAJ1	T cell receptor alpha joining 1 (non-functional)	NA
28744	TRAJ11	T cell receptor alpha joining 11	NA
28742	TRAJ13	T cell receptor alpha joining 13	NA
28741	TRAJ14	T cell receptor alpha joining 14	NA
28738	TRAJ17	T cell receptor alpha joining 17	NA

28737	TRAJ18	T cell receptor alpha joining 18	NA
28736	TRAJ19	T cell receptor alpha joining 19 (non-functional)	NA
28753	TRAJ2	T cell receptor alpha joining 2 (non-functional)	NA
28733	TRAJ22	T cell receptor alpha joining 22	NA
28731	TRAJ24	T cell receptor alpha joining 24	NA
28728	TRAJ27	T cell receptor alpha joining 27	NA
28726	TRAJ29	T cell receptor alpha joining 29	NA
28725	TRAJ30	T cell receptor alpha joining 30	NA
28752	TRAJ3	T cell receptor alpha joining 3	NA
28723	TRAJ32	T cell receptor alpha joining 32	NA
28721	TRAJ34	T cell receptor alpha joining 34	NA
28719	TRAJ36	T cell receptor alpha joining 36	NA
28718	TRAJ37	T cell receptor alpha joining 37	NA
28717	TRAJ38	T cell receptor alpha joining 38	NA
28713	TRAJ42	T cell receptor alpha joining 42	NA
28711	TRAJ44	T cell receptor alpha joining 44	NA
28710	TRAJ45	T cell receptor alpha joining 45	NA
28707	TRAJ48	T cell receptor alpha joining 48	NA
28703	TRAJ52	T cell receptor alpha joining 52	NA
28698	TRAJ57	T cell receptor alpha joining 57	NA
28749	TRAJ6	T cell receptor alpha joining 6	NA
28748	TRAJ7	T cell receptor alpha joining 7	NA
28746	TRAJ9	T cell receptor alpha joining 9	NA
28681	TRAV8-5	T cell receptor alpha variable 8-5 (pseudogene)	NA
6958	TRBV29OR9-2	T cell receptor beta variable 29/OR9-2 (non-functional)	NA
7706	TRIM25	tripartite motif containing 25	NA
286144	TRIQK	triple QxxK/R motif containing	NA
7327	UBE2G2	ubiquitin conjugating enzyme E2 G2	NA
7357	UGCG	UDP-glucose ceramide glucosyltransferase	NA
55245	UQCC1	ubiquinol-cytochrome c reductase complex assembly factor 1	NA
8237	USP11	ubiquitin specific peptidase 11	NA
23358	USP24	ubiquitin specific peptidase 24	NA
84749	USP30	ubiquitin specific peptidase 30	NA
80146	UXS1	UDP-glucuronate decarboxylase 1	NA
730239	ZDHHC20P4	zinc finger DHHC-type containing 20 pseudogene 4	NA
10308	ZNF267	zinc finger protein 267	NA
11194	ABCB8	ATP binding cassette subfamily B member 8	NA
11332	ACOT7	acyl-CoA thioesterase 7	NA
9255	AIMP1	aminoacyl tRNA synthetase complex interacting multifunctional protein 1	NA
208	AKT2	AKT serine/threonine kinase 2	NA
212	ALAS2	5'-aminolevulinate synthase 2	NA
23600	AMACR	alpha-methylacyl-CoA racemase	NA
161	AP2A2	adaptor related protein complex 2 subunit alpha 2	NA
1173	AP2M1	adaptor related protein complex 2 subunit mu 1	NA

23780	APOL2	apolipoprotein L2	NA
84334	APOPT1	apoptogenic 1, mitochondrial	NA
55082	ARGLU1	arginine and glutamate rich 1	NA
9181	ARHGEF2	Rho/Rac guanine nucleotide exchange factor 2	NA
79065	ATG9A	autophagy related 9A	NA
91647	ATPAF2	ATP synthase mitochondrial F1 complex assembly factor 2	NA
282569	BACE2-IT1	BACE2 intronic transcript 1	NA
80341	BPIFB2	BPI fold containing family B member 2	NA
25855	BRMS1	BRMS1, transcriptional repressor and anoikis regulator	NA
56673	C11orf16	chromosome 11 open reading frame 16	NA
79703	C11orf80	chromosome 11 open reading frame 80	NA
100131378	C11orf91	chromosome 11 open reading frame 91	NA
728568	C12orf73	chromosome 12 open reading frame 73	NA
759	CA1	carbonic anhydrase 1	NA
23261	CAMTA1	calmodulin binding transcription activator 1	NA
23066	CAND2	cullin associated and neddylation dissociated 2 (putative)	NA
387707	CC2D2B	coiled-coil and C2 domain containing 2B	NA
920	CD4	CD4 molecule	NA
942	CD86	CD86 molecule	NA
996	CDC27	cell division cycle 27	NA
5129	CDK18	cyclin dependent kinase 18	NA
1120	CHKB	choline kinase beta	NA
1129	CHRM2	cholinergic receptor muscarinic 2	NA
10256	CNKSR1	connector enhancer of kinase suppressor of Ras 1	NA
25904	CNOT10	CCR4-NOT transcription complex subunit 10	NA
1307	COL16A1	collagen type XVI alpha 1 chain	NA
131034	CPNE4	copine 4	NA
55861	DBNDD2	dysbindin domain containing 2	NA
100874229	DCUN1D2-AS	DCUN1D2 antisense RNA	NA
64798	DEPTOR	DEP domain containing MTOR interacting protein	NA
9162	DGKI	diacylglycerol kinase iota	NA
81704	DOCK8	dedicator of cytokinesis 8	NA
54344	DPM3	dolichyl-phosphate mannosyltransferase subunit 3	NA
25778	DSTYK	dual serine/threonine and tyrosine protein kinase	NA
6990	DYNLT3	dynein light chain Tctex-type 3	NA
5610	EIF2AK2	eukaryotic translation initiation factor 2 alpha kinase 2	NA
8666	EIF3G	eukaryotic translation initiation factor 3 subunit G	NA
51614	ERGIC3	ERGIC and golgi 3	NA
112479	ERI2	ERI1 exoribonuclease family member 2	NA
105379270	FAM197Y6	family with sequence similarity 197 Y-linked member 6	NA
9855	FARP2	FERM, ARH/RhoGEF and pleckstrin domain protein 2	NA
100128733	FCF1P3	FCF1 pseudogene 3	NA
2257	FGF12	fibroblast growth factor 12	NA

2297	FOXD1	forkhead box D1	NA
84253	GARNL3	GTPase activating Rap/RanGAP domain like 3	NA
23163	GGA3	golgi associated, gamma adaptin ear containing, ARF binding protein 3	NA
26058	GIGYF2	GRB10 interacting GYF protein 2	NA
2778	GNAS	GNAS complex locus	NA
2959	GTF2B	general transcription factor IIB	NA
54676	GTPBP2	GTP binding protein 2	NA
3026	HABP2	hyaluronan binding protein 2	NA
55027	HEATR3	HEAT repeat containing 3	NA
8342	HIST1H2BM	histone cluster 1 H2B family member m	NA
55355	HJURP	Holliday junction recognition protein	NA
3098	HK1	hexokinase 1	NA
3189	HNRNPH3	heterogeneous nuclear ribonucleoprotein H3	NA
3270	HRC	histidine rich calcium binding protein	NA
51763	INPP5K	inositol polyphosphate-5-phosphatase K	NA
386617	KCTD8	potassium channel tetramerization domain containing 8	NA
23366	KIAA0895	KIAA0895	NA
57691	KIAA1586	KIAA1586	NA
3897	L1CAM	L1 cell adhesion molecule	NA
100128782	LINC00476	long intergenic non-protein coding RNA 476	NA
104472718	LINC00508	long intergenic non-protein coding RNA 508	NA
4026	LPP	LIM domain containing preferred translocation partner in lipoma	NA
284348	LYPD5	LY6/PLAUR domain containing 5	NA
5605	MAP2K2	mitogen-activated protein kinase kinase 2	NA
55777	MBD5	methyl-CpG binding domain protein 5	NA
79892	MCMBP	minichromosome maintenance complex binding protein	NA
442907	MIR339	microRNA 339	NA
693122	MIR421	microRNA 421	NA
100616109	MIR4464	microRNA 4464	NA
574481	MIR521-2	microRNA 521-2	NA
100500919	MIR548Y	microRNA 548y	NA
724028	MIR658	microRNA 658	NA
100313778	MIR759	microRNA 759	NA
65258	MPPE1	metallophosphoesterase 1	NA
57129	MRPL47	mitochondrial ribosomal protein L47	NA
128308	MRPL55	mitochondrial ribosomal protein L55	NA
4638	MYLK	myosin light chain kinase	NA
4739	NEDD9	neural precursor cell expressed, developmentally down-regulated 9	NA
4815	NINJ2	ninjurin 2	NA
4900	NRGN	neurogranin	NA
55593	OTUD5	OTU deubiquitinase 5	NA
26986	PABPC1	poly(A) binding protein cytoplasmic 1	NA

25891	PAMR1	peptidase domain containing associated with muscle regeneration 1	NA
56965	PARP6	poly(ADP-ribose) polymerase family member 6	NA
192111	PGAM5	PGAM family member 5, mitochondrial serine/threonine protein phosphatase	NA
27315	PGAP2	post-GPI attachment to proteins 2	NA
51131	PHF11	PHD finger protein 11	NA
728233	PI4KAP1	phosphatidylinositol 4-kinase alpha pseudogene 1	NA
10908	PNPLA6	patatin like phospholipase domain containing 6	NA
23126	POGZ	pogo transposable element derived with ZNF domain	NA
5430	POLR2A	RNA polymerase II subunit A	NA
5432	POLR2C	RNA polymerase II subunit C	NA
23759	PPIL2	peptidylprolyl isomerase like 2	NA
5539	PPY	pancreatic polypeptide	NA
9055	PRC1	protein regulator of cytokinesis 1	NA
10549	PRDX4	peroxiredoxin 4	NA
5678	PSG9	pregnancy specific beta-1-glycoprotein 9	NA
5706	PSMC6	proteasome 26S subunit, ATPase 6	NA
9491	PSMF1	proteasome inhibitor subunit 1	NA
391356	PTRHD1	peptidyl-tRNA hydrolase domain containing 1	NA
11137	PWP1	PWP1 homolog, endonuclein	NA
26064	RAI14	retinoic acid induced 14	NA
5930	RBBP6	RB binding protein 6, ubiquitin ligase	NA
100874290	RERG-IT1	RERG intronic transcript 1	NA
80010	RMI1	RecQ mediated genome instability 1	NA
100873423	RNA5SP157	RNA, 5S ribosomal pseudogene 157	NA
100873363	RNA5SP31	RNA, 5S ribosomal pseudogene 31	NA
7732	RNF112	ring finger protein 112	NA
9604	RNF14	ring finger protein 14	NA
106481522	RNU6-1087P	RNA, U6 small nuclear 1087, pseudogene	NA
106481940	RNU6-1171P	RNA, U6 small nuclear 1171, pseudogene	NA
106480138	RNU6-1311P	RNA, U6 small nuclear 1311, pseudogene	NA
106480568	RNU6-308P	RNA, U6 small nuclear 308, pseudogene	NA
106480570	RNU6-320P	RNA, U6 small nuclear 320, pseudogene	NA
106480572	RNU6-344P	RNA, U6 small nuclear 344, pseudogene	NA
106481892	RNU6-399P	RNA, U6 small nuclear 399, pseudogene	NA
106479818	RNU6-567P	RNA, U6 small nuclear 567, pseudogene	NA
100873774	RNU6-71P	RNA, U6 small nuclear 71, pseudogene	NA
106481421	RNU6-763P	RNA, U6 small nuclear 763, pseudogene	NA
100873779	RNU6-79P	RNA, U6 small nuclear 79, pseudogene	NA
106479923	RNU6-807P	RNA, U6 small nuclear 807, pseudogene	NA
101447996	RNVU1-8	RNA, variant U1 small nuclear 8	NA
6124	RPL4	ribosomal protein L4	NA
100129306	RPS20P32	ribosomal protein S20 pseudogene 32	NA
6251	RSU1	Ras suppressor protein 1	NA
51493	RTCB	RNA 2',3'-cyclic phosphate and 5'-OH ligase	NA

22908	SACM1L	SAC1 like phosphatidylinositide phosphatase	NA
23034	SAMD4A	sterile alpha motif domain containing 4A	NA
256380	SCML4	Scm polycomb group protein like 4	NA
92745	SLC38A5	solute carrier family 38 member 5	NA
6540	SLC6A13	solute carrier family 6 member 13	NA
767610	SNORD114-29	small nucleolar RNA, C/D box 114-29	NA
26785	SNORD63	small nucleolar RNA, C/D box 63	NA
27131	SNX5	sorting nexin 5	NA
6733	SRPK2	SRSF protein kinase 2	NA
727837	SSX2B	SSX family member 2B	NA
9705	ST18	ST18, C2H2C-type zinc finger	NA
8869	ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	NA
6764	ST5	suppression of tumorigenicity 5	NA
128989	TANGO2	transport and golgi organization 2 homolog	NA
23216	TBC1D1	TBC1 domain family member 1	NA
122046	TEX26	testis expressed 26	NA
100507064	TEX26-AS1	TEX26 antisense RNA 1	NA
7041	TGFB1I1	transforming growth factor beta 1 induced transcript 1	NA
100499177	THAP9-AS1	THAP9 antisense RNA 1	NA
7089	TLE2	transducin like enhancer of split 2	NA
137695	TMEM68	transmembrane protein 68	NA
7145	TNS1	tensin 1	NA
114034	TOE1	target of EGR1, exonuclease	NA
100874024	TRPC7-AS1	TRPC7 antisense RNA 1	NA
54902	TTC19	tetratricopeptide repeat domain 19	NA
10382	TUBB4A	tubulin beta 4A class IVa	NA
619519	TWF1P1	twinfilin 1 pseudogene 1	NA
25828	TXN2	thioredoxin 2	NA
55236	UBA6	ubiquitin like modifier activating enzyme 6	NA
55833	UBAP2	ubiquitin associated protein 2	NA
9898	UBAP2L	ubiquitin associated protein 2 like	NA
9101	USP8	ubiquitin specific peptidase 8	NA
7414	VCL	vinculin	NA
200403	VWA3B	von Willebrand factor A domain containing 3B	NA
55717	WDR11	WD repeat domain 11	NA
26100	WIPI2	WD repeat domain, phosphoinositide interacting 2	NA
7473	WNT3	Wnt family member 3	NA
55135	WRAP53	WD repeat containing antisense to TP53	NA
11060	WWP2	WW domain containing E3 ubiquitin protein ligase 2	NA
83719	YPEL3	yippee like 3	NA
53349	ZFYVE1	zinc finger FYVE-type containing 1	NA
25850	ZNF345	zinc finger protein 345	NA
346157	ZNF391	zinc finger protein 391	NA
90589	ZNF625	zinc finger protein 625	NA
401509	ZNF658B	zinc finger protein 658B (pseudogene)	NA

646864	ZNF723	zinc finger protein 723	NA
664701	ZNF826P	zinc finger protein 826, pseudogene	NA
83857	TMTC1	transmembrane and tetratricopeptide repeat containing 1	226931_at
10335	MRVI1	murine retrovirus integration site 1 homolog	230214_at
1824	DSC2	desmocollin 2	226817_at
79937	CNTNAP3	contactin associated protein like 3	244065_at
728577	CNTNAP3B	contactin associated protein like 3B	244065_at
3772	KCNJ15	potassium voltage-gated channel subfamily J member 15	230585_at
8825	LIN7A	lin-7 homolog A, crumbs cell polarity complex component	227929_at
399972	GSEC	G-quadruplex forming sequence containing lncRNA	227925_at
552891	DNAJC25-GNG10	DNAJC25-GNG10 readthrough	201921_at
2790	GNG10	G protein subunit gamma 10	201921_at
60675	PROK2	prokineticin 2	232629_at
643827	CNTNAP3P2	CNTNAP3 pseudogene 2	223796_at
2123	EVI2A	ecotropic viral integration site 2A	204774_at
79768	KATNBL1	katanin regulatory subunit B1 like 1	222745_s_at
399844	LINC01002	long intergenic non-protein coding RNA 1002	225899_x_at
402483	LINC01000	long intergenic non-protein coding RNA 1000	225899_x_at
729737	LOC729737	uncharacterized LOC729737	225899_x_at
100132062	LOC100132062	uncharacterized LOC100132062	225899_x_at
100133161	LINC01001	long intergenic non-protein coding RNA 1001	225899_x_at
100133331	LOC100133331	uncharacterized LOC100133331	225899_x_at
101929819	LOC101929819	uncharacterized LOC101929819	225899_x_at
4697	NDUFA4	NDUFA4 mitochondrial complex associated	217773_s_at
57612	ERV3-2	endogenous retrovirus group 3 member 2	222139_at
5743	PTGS2	prostaglandin-endoperoxide synthase 2	1554997_a_at
10096	ACTR3	actin related protein 3	228603_at
79660	PPP1R3B	protein phosphatase 1 regulatory subunit 3B	222662_at
723790	HIST2H2AA4	histone cluster 2 H2A family member a4	214290_s_at
9765	ZFYVE16	zinc finger FYVE-type containing 16	203651_at
101928429	LOC101928429	uncharacterized LOC101928429	242714_at
166929	SGMS2	sphingomyelin synthase 2	227038_at
11216	AKAP10	A-kinase anchoring protein 10	213396_s_at
8530	CST7	cystatin F	210140_at
64332	NFKBIZ	NFKB inhibitor zeta	223218_s_at
8339	HIST1H2BG	histone cluster 1 H2B family member g	214455_at
8343	HIST1H2BF	histone cluster 1 H2B family member f	214455_at
8344	HIST1H2BE	histone cluster 1 H2B family member e	214455_at
8346	HIST1H2BI	histone cluster 1 H2B family member i	214455_at
84188	FAR1	fatty acyl-CoA reductase 1	224866_at
55284	UBE2W	ubiquitin conjugating enzyme E2 W	222657_s_at
9778	KIAA0232	KIAA0232	232366_at
10161	LPAR6	lysophosphatidic acid receptor 6	218589_at
3250	HPR	haptoglobin-related protein	208470_s_at

3552	IL1A	interleukin 1 alpha	208200_at
660	BMX	BMX non-receptor tyrosine kinase	206464_at
140849	LINC00266-1	long intergenic non-protein coding RNA 266-1	232953_at
100129518	SOD2-OT1	SOD2 overlapping transcript 1	215078_at
2850	GPR27	G protein-coupled receptor 27	227769_at
55016	Mar-01	membrane associated ring-CH-type finger 1	235385_at
51290	ERGIC2	ERGIC and golgi 2	218135_at
134957	STXBP5	syntaxin binding protein 5	243904_at
102465472	MIR6787	microRNA 6787	202856_s_at
5926	ARID4A	AT-rich interaction domain 4A	230141_at
4664	NAB1	NGFI-A binding protein 1	209272_at
7096	TLR1	toll like receptor 1	210176_at
22856	CHSY1	chondroitin sulfate synthase 1	203044_at
196264	MPZL3	myelin protein zero like 3	227747_at
79971	WLS	Wnt ligand secretion mediator	228949_at
2529	FUT7	fucosyltransferase 7	210506_at
116496	NIBAN1	niban apoptosis regulator 1	217966_s_at
50484	RRM2B	ribonucleotide reductase regulatory TP53 inducible subunit M2B	223342_at
4090	SMAD5	SMAD family member 5	225219_at
9821	RB1CC1	RB1 inducible coiled-coil 1	202033_s_at
79730	NSUN7	NOP2/Sun RNA methyltransferase family member 7	238983_at
83853	ROPN1L	rhophilin associated tail protein 1 like	223609_at
84803	GPAT3	glycerol-3-phosphate acyltransferase 3	224480_s_at
100507217	LINC01578	long intergenic non-protein coding RNA 1578	244443_at
9236	CCPG1	cell cycle progression 1	222156_x_at
100533483	DNAAF4-CCPG1	DNAAF4-CCPG1 readthrough (NMD candidate)	222156_x_at
100506328	LINC01127	long intergenic non-protein coding RNA 1127	1560679_at
55589	BMP2K	BMP2 inducible kinase	226853_at
8794	TNFRSF10C	TNF receptor superfamily member 10c	206222_at
9867	PJA2	praja ring finger ubiquitin ligase 2	201133_s_at
401233	HTATSF1P2	HIV-1 Tat specific factor 1 pseudogene 2	1558882_at
203274	LINC00537	long intergenic non-protein coding RNA 537	232034_at
152100	CMC1	C-X9-C motif containing 1	228283_at
719	C3AR1	complement C3a receptor 1	209906_at
10797	MTHFD2	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	201761_at
121457	IKBIP	IKBKB interacting protein	227295_at
100288432	IL10RB-DT	IL10RB divergent transcript	230631_s_at
54790	TET2	tet methylcytosine dioxygenase 2	227624_at
84984	CEP19	centrosomal protein 19	1553158_at
4707	NDUFB1	NADH:ubiquinone oxidoreductase subunit B1	206790_s_at
9334	B4GALT5	beta-1,4-galactosyltransferase 5	221485_at
100505702	LINC01094	long intergenic non-protein coding RNA 1094	229635_at
284759	SIRPB2	signal regulatory protein beta 2	1559034_at
7110	TMF1	TATA element modulatory factor 1	242243_at

7091	TLE4	TLE family member 4, transcriptional corepressor	216997_x_at
731275	LINC01347	long intergenic non-protein coding RNA 1347	234664_at
388572	LOC388572	uncharacterized LOC388572	234664_at
26502	NARF	nuclear prelamin A recognition factor	219862_s_at
51279	C1RL	complement C1r subcomponent like	218983_at
3553	IL1B	interleukin 1 beta	39402_at
9693	RAPGEF2	Rap guanine nucleotide exchange factor 2	215992_s_at
5591	PRKDC	protein kinase, DNA-activated, catalytic subunit	208694_at
5586	PKN2	protein kinase N2	212629_s_at
118987	PDZD8	PDZ domain containing 8	213549_at
114294	LACTB	lactamase beta	1552485_at
120425	JAML	junction adhesion molecule like	228094_at
120224	TMEM45B	transmembrane protein 45B	230323_s_at
80232	WDR26	WD repeat domain 26	224897_at
84263	HSDL2	hydroxysteroid dehydrogenase like 2	209512_at
100616393	MIR4657	microRNA 4657	226762_at
23325	WASHC4	WASH complex subunit 4	212794_s_at
80035	ANP32A-IT1	ANP32A intronic transcript 1	220710_at
10746	MAP3K2	mitogen-activated protein kinase kinase kinase 2	226979_at
647070	LOC647070	uncharacterized LOC647070	215468_at
57688	ZSWIM6	zinc finger SWIM-type containing 6	226208_at
27439	TMEM121B	transmembrane protein 121B	224393_s_at
2591	GALNT3	polypeptide N-acetylgalactosaminyltransferase 3	203397_s_at
100287569	LINC00173	long intergenic non-protein coding RNA 173	237591_at
8301	PICALM	phosphatidylinositol binding clathrin assembly protein	212511_at
1452	CSNK1A1	casein kinase 1 alpha 1	1556007_s_at
254896	LOC254896	uncharacterized LOC254896	210483_at
6814	STXBP3	syntaxin binding protein 3	1560486_at
51155	JPT1	Jupiter microtubule associated homolog 1	217755_at
2015	ADGRE1	adhesion G protein-coupled receptor E1	207111_at
2157	F8	coagulation factor VIII	205756_s_at
837	CASP4	caspase 4	213596_at
5494	PPM1A	protein phosphatase, Mg2+/Mn2+ dependent 1A	210407_at
10144	FAM13A	family with sequence similarity 13 member A	202973_x_at
157697	ERICH1	glutamate rich 1	1558371_a_at
378938	MALAT1	metastasis associated lung adenocarcinoma transcript 1	224559_at
338758	ATP2B1-AS1	ATP2B1 antisense RNA 1	238893_at
10673	TNFSF13B	TNF superfamily member 13b	223502_s_at
58475	MS4A7	membrane spanning 4-domains A7	223344_s_at
23023	TMCC1	transmembrane and coiled-coil domain family 1	237943_at
56929	FEM1C	fem-1 homolog C	213341_at
115825	WDFY2	WD repeat and FYVE domain containing 2	243514_at
541471	MIR4435-2HG	MIR4435-2 host gene	225799_at
112597	CYTOR	cytoskeleton regulator RNA	225799_at
5166	PDK4	pyruvate dehydrogenase kinase 4	1562321_at

404636	FAM45A	family with sequence similarity 45 member A	222955_s_at
55855	FAM45BP	family with sequence similarity 45, member A pseudogene	222955_s_at
152007	GLIPR2	GLI pathogenesis related 2	225604_s_at
441951	ZFAS1	ZNFX1 antisense RNA 1	229899_s_at
140809	SRXN1	sulfiredoxin 1	225252_at
51809	GALNT7	polypeptide N-acetylgalactosaminyltransferase 7	222587_s_at
1347	COX7A2	cytochrome c oxidase subunit 7A2	201597_at
9556	ATP5MPL	ATP synthase membrane subunit 6.8PL	202279_at
6850	SYK	spleen associated tyrosine kinase	244023_at
100505746	ITGB2-AS1	ITGB2 antisense RNA 1	229041_s_at
88455	ANKRD13A	ankyrin repeat domain 13A	238851_at
221037	JMJD1C	jumonji domain containing 1C	228793_at
79101	TAF1D	TATA-box binding protein associated factor, RNA polymerase I subunit D	218750_at
26268	FBXO9	F-box protein 9	1559096_x_at
56655	POLE4	DNA polymerase epsilon 4, accessory subunit	1553587_a_at
59348	ZNF350	zinc finger protein 350	233169_at
254170	FBXO33	F-box protein 33	226970_at
100129550	LINC02035	long intergenic non-protein coding RNA 2035	229699_at
403341	ZBTB34	zinc finger and BTB domain containing 34	227111_at
440836	ODF3B	outer dense fiber of sperm tails 3B	243934_at
137835	TMEM71	transmembrane protein 71	238429_at
8036	SHOC2	SHOC2 leucine rich repeat scaffold protein	202777_at
81846	SBF2	SET binding factor 2	226169_at
5500	PPP1CB	protein phosphatase 1 catalytic subunit beta	201407_s_at
4659	PPP1R12A	protein phosphatase 1 regulatory subunit 12A	201602_s_at
158401	SHOC1	shortage in chiasmata 1	1553920_at
5966	REL	REL proto-oncogene, NF-kB subunit	239486_at
79415	CYBC1	cytochrome b-245 chaperone 1	218130_at
10640	EXOC5	exocyst complex component 5	225084_at
55331	ACER3	alkaline ceramidase 3	222687_s_at
1050	CEBPA	CCAAT enhancer binding protein alpha	204039_at
10724	OGA	O-GlcNAcase	235868_at
9658	ZNF516	zinc finger protein 516	203604_at
60559	SPCS3	signal peptidase complex subunit 3	222753_s_at
6670	SP3	Sp3 transcription factor	232529_at
138151	NACC2	NACC family member 2	212993_at
23218	NBEAL2	neurobeachin like 2	212443_at
1379	CR1L	complement C3b/C4b receptor 1 like	239205_s_at
51714	SELENOT	selenoprotein T	217811_at
57162	PELI1	pellino E3 ubiquitin protein ligase 1	232304_at
84628	NTNG2	netrin G2	233072_at
441124	GTF2IP20	general transcription factor IIi pseudogene 20	227922_x_at
284751	LINC01270	long intergenic non-protein coding RNA 1270	1556896_at
102724364	SEC22B4P	SEC22 homolog B4, pseudogene	209206_at
163486	DENND1B	DENN domain containing 1B	238787_at

30	ACAA1	acetyl-CoA acyltransferase 1	214274_s_at
933	CD22	CD22 molecule	217422_s_at
100506779	TSPOAP1-AS1	TSPOAP1, SUPT4H1 and RNF43 antisense RNA 1	228826_at
220001	VWCE	von Willebrand factor C and EGF domains	242957_at
643837	LINC01128	long intergenic non-protein coding RNA 1128	1557055_s_at
5046	PCSK6	proprotein convertase subtilisin/kexin type 6	207414_s_at
100507472	LOC100507472	uncharacterized LOC100507472	207414_s_at
388011	LINC01550	long intergenic non-protein coding RNA 1550	1559097_at
100129528	MUC8	mucin 8	216671_x_at
8848	TSC22D1	TSC22 domain family member 1	215111_s_at
78992	YIPF2	Yip1 domain family member 2	212512_s_at
10498	CARM1	coactivator associated arginine methyltransferase 1	212512_s_at
3081	HGD	homogentisate 1,2-dioxygenase	205221_at
8131	NPRL3	NPR3 like, GATOR1 complex subunit	210672_s_at
3212	HOXB2	homeobox B2	205453_at
81027	TUBB1	tubulin beta 1 class VI	230690_at
4352	MPL	MPL proto-oncogene, thrombopoietin receptor	207550_at
57699	CPNE5	copine 5	227189_at
221895	JAZF1	JAZF zinc finger 1	225800_at
973	CD79A	CD79a molecule	1555779_a_at
5026	P2RX5	purinergic receptor P2X 5	210448_s_at
157285	PRAG1	PEAK1 related, kinase-activating pseudokinase 1	235085_at
29802	VPREB3	V-set pre-B cell surrogate light chain 3	220068_at
114884	OSBPL10	oxysterol binding protein like 10	219073_s_at
5450	POU2AF1	POU class 2 homeobox associating factor 1	205267_at
9241	NOG	noggin	231798_at
3507	IGHM	immunoglobulin heavy constant mu	212827_at
6799	SULT1A2	sulfotransferase family 1A member 2	211385_x_at
118932	ANKRD22	ankyrin repeat domain 22	238439_at
55432	YOD1	YOD1 deubiquitinase	227309_at
8851	CDK5R1	cyclin dependent kinase 5 regulatory subunit 1	204995_at
57458	TMCC3	transmembrane and coiled-coil domain family 3	235146_at
56255	TMX4	thioredoxin related transmembrane protein 4	201581_at
23161	SNX13	sorting nexin 13	213292_s_at
29994	BAZ2B	bromodomain adjacent to zinc finger domain 2B	203080_s_at
120892	LRRK2	leucine rich repeat kinase 2	229584_at
84986	ARHGAP19	Rho GTPase activating protein 19	37577_at
497661	C18orf32	chromosome 18 open reading frame 32	224957_at
100526842	RPL17-C18orf32	RPL17-C18orf32 readthrough	224957_at
6940	ZNF354A	zinc finger protein 354A	205427_at
355	FAS	Fas cell surface death receptor	216252_x_at
768211	RELL1	RELT like 1	226430_at
3930	LBR	lamin B receptor	201795_at
9100	USP10	ubiquitin specific peptidase 10	209137_s_at
101927221	UBR5-AS1	UBR5 antisense RNA 1	1555888_at
9958	USP15	ubiquitin specific peptidase 15	210681_s_at

102465133	MIR6125	microRNA 6125	210681_s_at
23766	GABARAPL3	GABA type A receptor associated protein like 3 pseudogene	211458_s_at
23710	GABARAPL1	GABA type A receptor associated protein like 1	211458_s_at
94241	TP53INP1	tumor protein p53 inducible nuclear protein 1	225912_at
55196	RESF1	retroelement silencing factor 1	227152_at
23347	SMCHD1	structural maintenance of chromosomes flexible hinge domain containing 1	1558747_at
22990	PCNX1	pecanex 1	238792_at
51314	NME8	NME/NM23 family member 8	220384_at
80254	CEP63	centrosomal protein 63	219242_at
9208	LRRFIP1	LRR binding FLII interacting protein 1	223492_s_at
89849	ATG16L2	autophagy related 16 like 2	225883_at
9107	MTMR6	myotubularin related protein 6	214429_at
6696	SPP1	secreted phosphoprotein 1	209875_s_at
284757	MIR646HG	MIR646 host gene	236846_at
10512	SEMA3C	semaphorin 3C	203789_s_at
192670	AGO4	argonaute RISC component 4	219190_s_at
134510	UBLCP1	ubiquitin like domain containing CTD phosphatase 1	243916_x_at
7750	ZMYM2	zinc finger MYM-type containing 2	210281_s_at
9818	NUP58	nucleoporin 58	204435_at
1130	LYST	lysosomal trafficking regulator	203518_at
51635	DHRS7	dehydrogenase/reductase 7	210788_s_at
4215	MAP3K3	mitogen-activated protein kinase kinase kinase 3	243505_at
23345	SYNE1	spectrin repeat containing nuclear envelope protein 1	232027_at
4193	MDM2	MDM2 proto-oncogene	225160_x_at
10762	NUP50	nucleoporin 50	218294_s_at
283131	NEAT1	nuclear paraspeckle assembly transcript 1	214657_s_at
693197	MIR612	microRNA 612	214657_s_at
79663	HSPBAP1	HSPB1 associated protein 1	219284_at
100129960	SDCBPP2	syndecan binding protein pseudogene 2	1569952_x_at
9488	PIGB	phosphatidylinositol glycan anchor biosynthesis class B	205452_at
130502	TTC32	tetratricopeptide repeat domain 32	226838_at
54807	ZNF586	zinc finger protein 586	219711_at
10472	ZBTB18	zinc finger and BTB domain containing 18	207164_s_at
57708	MIER1	MIER1 transcriptional regulator	1555105_a_at
58488	PCTP	phosphatidylcholine transfer protein	218676_s_at
55062	WIPI1	WD repeat domain, phosphoinositide interacting 1	213836_s_at
4660	PPP1R12B	protein phosphatase 1 regulatory subunit 12B	1557553_at
22921	MSRB2	methionine sulfoxide reductase B2	219451_at
5525	PPP2R5A	protein phosphatase 2 regulatory subunit B'alpha	202187_s_at
84280	BTBD10	BTB domain containing 10	223174_at
92482	BBIP1	BBSome interacting protein 1	213220_at
143384	CACUL1	CDK2 associated cullin domain 1	227257_s_at
9967	THRAP3	thyroid hormone receptor associated protein 3	234942_s_at

407004	MIR22	microRNA 22	214696_at
752	FMNL1	formin like 1	1569257_at
3423	IDS	iduronate 2-sulfatase	210666_at
57580	PREX1	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 1	224925_at
388677	NOTCH2NLA	notch 2 N-terminal like A	227067_x_at
100996717	NOTCH2NLC	notch 2 N-terminal like C	227067_x_at
100996763	NOTCH2NLB	notch 2 N-terminal like B	227067_x_at
101929796	NOTCH2NLR	notch 2 N-terminal like R	227067_x_at
121274	ZNF641	zinc finger protein 641	229897_at
200315	APOBEC3A	apolipoprotein B mRNA editing enzyme catalytic subunit 3A	210873_x_at
100913187	APOBEC3A_B	APOBEC3A and APOBEC3B deletion hybrid	210873_x_at
10521	DDX17	DEAD-box helicase 17	213998_s_at
283876	LINC00921	long intergenic non-protein coding RNA 921	243547_at
10127	ZNF263	zinc finger protein 263	243547_at
54518	APBB1IP	amyloid beta precursor protein binding family B member 1 interacting protein	1554571_at
79587	CARS2	cysteinyl-tRNA synthetase 2, mitochondrial	218153_at
573	BAG1	BCL2 associated athanogene 1	211475_s_at
1879	EBF1	EBF transcription factor 1	232204_at
22872	SEC31A	SEC31 homolog A, COPII coat complex component	200945_s_at
64087	MCCC2	methylcrotonoyl-CoA carboxylase 2	209623_at
924	CD7	CD7 molecule	214049_x_at
10133	OPTN	optineurin	202074_s_at
10730	YME1L1	YME1 like 1 ATPase	211902_x_at
10838	ZNF275	zinc finger protein 275	225382_at
55278	QRSL1	glutaminyl-tRNA amidotransferase subunit QRSL1	218948_at
7157	TP53	tumor protein p53	201746_at
3537	IGLC1	immunoglobulin lambda constant 1	217148_x_at
3546	IGLV@	immunoglobulin lambda variable cluster	217148_x_at
28831	IGLJ3	immunoglobulin lambda joining 3	217148_x_at
864	RUNX3	RUNX family transcription factor 3	204197_s_at
4302	MLLT6	MLLT6, PHD finger containing	224784_at
9588	PRDX6	peroxiredoxin 6	200844_s_at
1235	CCR6	C-C motif chemokine receptor 6	206983_at
149345	SHISA4	shisa family member 4	226674_at
3111	HLA-DOA	major histocompatibility complex, class II, DO alpha	226878_at
27042	UTP25	UTP25 small subunit processor component	204700_x_at
23476	BRD4	bromodomain containing 4	202102_s_at
23450	SF3B3	splicing factor 3b subunit 3	200687_s_at
6925	TCF4	transcription factor 4	203753_at
28663	TRAV20	T cell receptor alpha variable 20	210972_x_at
3514	IGKC	immunoglobulin kappa constant	214768_x_at
79993	ELOVL7	ELOVL fatty acid elongase 7	227180_at
4541	ND6	NADH dehydrogenase, subunit 6 (complex I)	1553575_at

199786	NIBAN3	niban apoptosis regulator 3	230983_at
10158	PDZK1IP1	PDZK1 interacting protein 1	219630_at
3326	HSP90AB1	heat shock protein 90 alpha family class B member 1	214359_s_at
91523	PCED1B	PC-esterase domain containing 1B	228298_at
84824	FCRLA	Fc receptor like A	235400_at
653082	ZDHHC11B	zinc finger DHHC-type containing 11B	1552283_s_at
200407	CREG2	cellular repressor of E1A stimulated genes 2	1552714_at
169693	TMEM252	transmembrane protein 252	1553809_a_at
63947	DMRTC1	DMRT like family C1	1553998_at
728656	DMRTC1B	DMRT like family C1B	1553998_at
388387	LINC00671	long intergenic non-protein coding RNA 671	1556737_at
100861555	LINC00565	long intergenic non-protein coding RNA 565	1558847_at
112724	RDH13	retinol dehydrogenase 13	1559190_s_at
100506229	LINC01093	long intergenic non-protein coding RNA 1093	1559573_at
399716	LINC02656	long intergenic non-protein coding RNA 2656	1559756_at
100507404	TMLHE-AS1	TMLHE antisense RNA 1	1560797_s_at
101927830	LOC101927830	uncharacterized LOC101927830	1560797_s_at
161497	STRC	stereocilin	1561306_s_at
101928361	LOC101928361	uncharacterized LOC101928361	1561489_at
55843	ARHGAP15	Rho GTPase activating protein 15	1561489_at
285429	DCAF4L1	DDB1 and CUL4 associated factor 4 like 1	1562209_at
404201	WDFY3-AS2	WDFY3 antisense RNA 2	1562953_s_at
731779	LINC01300	long intergenic non-protein coding RNA 1300	1569297_at
3728	JUP	junction plakoglobin	201015_s_at
6035	RNASE1	ribonuclease A family member 1, pancreatic	201785_at
10123	ARL4C	ADP ribosylation factor like GTPase 4C	202206_at
5617	PRL	prolactin	205445_at
11341	SCRG1	stimulator of chondrogenesis 1	205475_at
10677	AVIL	advillin	205539_at
9256	TSPOAP1	TSPO associated protein 1	205839_s_at
8292	COLQ	collagen like tail subunit of asymmetric acetylcholinesterase	206073_at
11039	SMA4	glucuronidase beta pseudogene	206565_x_at
101927057	LOC101927057	uncharacterized LOC101927057	207093_s_at
3604	TNFRSF9	TNF receptor superfamily member 9	207536_s_at
27284	SULT1B1	sulfotransferase family 1B member 1	207601_at
114088	TRIM9	tripartite motif containing 9	209859_at
5243	ABCB1	ATP binding cassette subfamily B member 1	209993_at
5321	PLA2G4A	phospholipase A2 group IVA	210145_at
6788	STK3	serine/threonine kinase 3	211078_s_at
113146	AHNAK2	AHNAK nucleoprotein 2	212992_at
1028	CDKN1C	cyclin dependent kinase inhibitor 1C	213183_s_at
1314	COPA	coatomer protein complex subunit alpha	214337_at
11042	SMA5	glucuronidase beta pseudogene	215043_s_at
653188	GUSBP3	GUSB pseudogene 3	215043_s_at
643332	ECRP	ribonuclease A family member 2 pseudogene	216667_at

103752589	TMEM92-AS1	TMEM92 antisense RNA 1	216789_at
2064	ERBB2	erb-b2 receptor tyrosine kinase 2	216836_s_at
10417	SPON2	spondin 2	218638_s_at
100130872	LOC100130872	uncharacterized LOC100130872	218638_s_at
51762	RAB8B	RAB8B, member RAS oncogene family	219210_s_at
79623	GALNT14	polypeptide N-acetylgalactosaminyltransferase 14	219271_at
30009	TBX21	T-box 21	220684_at
101928290	PLBD1-AS1	PLBD1 antisense RNA 1	222639_s_at
101928317	LOC101928317	uncharacterized LOC101928317	222639_s_at
84419	C15orf48	chromosome 15 open reading frame 48	223484_at
84290	CAPNS2	calpain small subunit 2	223832_s_at
9509	ADAMTS2	ADAM metallopeptidase with thrombospondin type 1 motif 2	226311 at
135114	HINT3	histidine triad nucleotide binding protein 3	226537_at
5802	PTPRS	protein tyrosine phosphatase receptor type S	226571 s at
8997	KALRN	kalirin RhoGEF kinase	227750_at
130355	C2orf76	chromosome 2 open reading frame 76	227840 at
5212	VIT	vitrin	227899_at
51023	MRPS18C	mitochondrial ribosomal protein S18C	228019 s at
57571	CARNS1	carnosine synthase 1	228984_at
921	CD5	CD5 molecule	230489 at
135927	LLCFC1	LLLL and CFNLAS motif containing 1	231435_at
200010	SLC5A9	solute carrier family 5 member 9	232378 at
100379345	MIR181A2HG	MIR181A2 host gene	232478_at
284367	SIGLEC17P	sialic acid binding Ig like lectin 17, pseudogene	232686 at
57232	ZNF630	zinc finger protein 630	233082_at
11145	PLAAT3	phospholipase A and acyltransferase 3	235110 at
79026	AHNAK	AHNAK nucleoprotein	235281_x_at
4648	MYO7B	myosin VIIB	235383 at
777	CACNA1E	calcium voltage-gated channel subunit alpha1 E	236013_at
90249	UNC5A	unc-5 netrin receptor A	236448 at
387755	INSC	INSC spindle orientation adaptor protein	237056_at
145447	ABHD12B	abhydrolase domain containing 12B	237974 at
100616234	MIR4454	microRNA 4454	237974_at
100130733	LRRC70	leucine rich repeat containing 70	238488 at
140947	DCANP1	dendritic cell associated nuclear protein	239529_at
497189	TIFAB	TIFA inhibitor	239529 at
100289274	DNAJC3-DT	DNAJC3 divergent transcript	240574_at
161582	DNAAF4	dynein axonemal assembly factor 4	241713 s at
103352539	LINC01410	long intergenic non-protein coding RNA 1410	243824_at
102723526	LINC02218	long intergenic non-protein coding RNA 2218	244269 at
407024	MIR29B1	microRNA 29b-1	244322_at
64788	LMF1	lipase maturation factor 1	46142 at
253827	MSRB3	methionine sulfoxide reductase B3	1554127_s_at
105372881	LOC105372881	uncharacterized LOC105372881	1556107 at
283521	TMEM272	transmembrane protein 272	1557465_at
285847	LOC285847	uncharacterized LOC285847	1557724 a at

101928988	LOC101928988	uncharacterized LOC101928988	1560257_at
105376037	LOC105376037	uncharacterized LOC105376037	1560424_at
401612	SLC25A53	solute carrier family 25 member 53	1561660_at
11328	FKBP9	FKBP prolyl isomerase 9	212169_at
51207	DUSP13	dual specificity phosphatase 13	219963_at
11107	PRDM5	PR/SET domain 5	220792_at
140	ADORA3	adenosine A3 receptor	223660_at
57121	LPAR5	lysophosphatidic acid receptor 5	230252_at
59269	HIVEP3	human immunodeficiency virus type I enhancer binding protein 3	235122_at
129563	DIS3L2	DIS3 like 3'-5' exoribonuclease 2	244304 at
157983	DOCK8-AS1	DOCK8 antisense RNA 1	1552757_s_at
57609	DIP2B	disco interacting protein 2 homolog B	1553271 at
285848	PNPLA1	patatin like phospholipase domain containing 1	1553364_at
285154	CYP1B1-AS1	CYP1B1 antisense RNA 1	1553829 at
85003	BFSP2-AS1	BFSP2 antisense RNA 1	1553936_a_at
8672	EIF4G3	eukaryotic translation initiation factor 4 gamma 3	1554309 at
57501	KIAA1257	KIAA1257	1554852_a_at
54084	TSPEAR	thrombospondin type laminin G domain and EAR repeats	1555049 at
4337	MOCS1	molybdenum cofactor synthesis 1	1555127_at
100500907	MIR3150B	microRNA 3150b	1556755 s at
100130357	LOC100130357	uncharacterized LOC100130357	1559507_at
51130	ASB3	ankyrin repeat and SOCS box containing 3	1560219 at
101928386	LOC101928386	uncharacterized LOC101928386	1560538_at
101928513	CATIP-AS1	CATIP antisense RNA 1	1560625 s at
554279	LINC00862	long intergenic non-protein coding RNA 862	1562909_at
101929007	LOC101929007	WAS/WASL-interacting protein family member 1	1568894 at
692247	LOC692247	uncharacterized LOC692247	1569453_a_at
340527	NHSL2	NHS like 2	1569932 at
101927851	LOC101927851	uncharacterized LOC101927851	1570285_at
4072	EPCAM	epithelial cell adhesion molecule	201839 s at
4609	MYC	MYC proto-oncogene, bHLH transcription factor	202431_s_at
9780	PIEZO1	piezo type mechanosensitive ion channel component 1	202771 at
9619	ABCG1	ATP binding cassette subfamily G member 1	204567_s_at
4261	CIITA	class II major histocompatibility complex transactivator	205101 at
1602	DACH1	dachshund family transcription factor 1	205471_s_at
9447	AIM2	absent in melanoma 2	206513 at
8529	CYP4F2	cytochrome P450 family 4 subfamily F member 2	206514_s_at
57096	RPGRIP1	RPGR interacting protein 1	206608 s at
6335	SCN9A	sodium voltage-gated channel alpha subunit 9	206950_at
314	AOC2	amine oxidase copper containing 2	207064 s at
7101	NR2E1	nuclear receptor subfamily 2 group E member 1	207443_at
1663	DDX11	DEAD/H-box helicase 11	208149 x at
27181	SIGLEC8	sialic acid binding Ig like lectin 8	208253_at
80303	EFHD1	EF-hand domain family member D1	209343 at

9901	SRGAP3	SLIT-ROBO Rho GTPase activating protein 3	209794_at
5218	CDK14	cyclin dependent kinase 14	211502_s_at
31	ACACA	acetyl-CoA carboxylase alpha	212186_at
4628	MYH10	myosin heavy chain 10	212372_at
158471	PRUNE2	prune homolog 2 with BCH domain	212806_at
161291	TMEM30B	transmembrane protein 30B	213285_at
1521	CTSW	cathepsin W	214450_at
6498	SKIL	SKI like proto-oncogene	217591_at
25984	KRT23	keratin 23	218963_s_at
57823	SLAMF7	SLAM family member 7	219159_s_at
10170	DHRS9	dehydrogenase/reductase 9	219799_s_at
27156	RSPH14	radial spoke head 14 homolog	220105_at
25790	CFAP45	cilia and flagella associated protein 45	220308_at
9501	RPH3AL	rabphilin 3A like (without C2 domains)	221614_s_at
25825	BACE2	beta-secretase 2	222446_s_at
84662	GLIS2	GLIS family zinc finger 2	223378_at
114880	OSBPL6	oxysterol binding protein like 6	223805_at
54602	NDFIP2	Nedd4 family interacting protein 2	224801_at
55529	PIP4P2	phosphatidylinositol-4,5-bisphosphate 4-phosphatase 2	226338_at
3672	ITGA1	integrin subunit alpha 1	226731_at
57633	LRRN1	leucine rich repeat neuronal 1	226884_at
101929693	RARA-AS1	RARA antisense RNA 1	228037_at
343990	KIAA1211L	KIAA1211 like	228067_at
25885	POLR1A	RNA polymerase I subunit A	228515_at
4094	MAF	MAF bZIP transcription factor	229327_s_at
79717	PPCS	phosphopantothenoylcysteine synthetase	230897_at
728621	CCDC30	coiled-coil domain containing 30	230897_at
92749	DRC1	dynein regulatory complex subunit 1	231133_at
400680	LINC00664	long intergenic non-protein coding RNA 664	231470_at
27235	COQ2	coenzyme Q2, polyprenyltransferase	232126_at
10251	SPRY3	sprouty RTK signaling antagonist 3	232157_at
55582	KIF27	kinesin family member 27	232514_at
388799	FAM209B	family with sequence similarity 209 member B	233086_at
101930164	LINC02555	long intergenic non-protein coding RNA 2555	233591_at
79683	ZDHHC14	zinc finger DHHC-type containing 14	235133_at
7094	TLN1	talin 1	236132_at
23639	LRRC6	leucine rich repeat containing 6	236587_at
207063	DHRSX	dehydrogenase/reductase X-linked	238053_at
8330	HIST1H2AK	histone cluster 1 H2A family member k	239041_at
147947	ZNF542P	zinc finger protein 542, pseudogene	239250_at
283152	CCDC153	coiled-coil domain containing 153	240293_at
4033	LRMP	lymphoid restricted membrane protein	240718_at
338811	TAFA2	TAFA chemokine like family member 2	241399_at
200772	LOC200772	uncharacterized LOC200772	241525_at
100507557	LOC100507557	uncharacterized LOC100507557	241745_at
168451	THAP5	THAP domain containing 5	244190_at

22859	ADGRL1	adhesion G protein-coupled receptor L1	47560_at

Appendix 4 – Chapter 3, Supplementary Data 3
List of proteins from untargeted stroke proteomics studies

ENTREZID	UNIPROT	SYMBOL	GENENAME
1192	O00299	CLIC1	chloride intracellular channel 1
10059	O00429	DNM1L	dynamin 1 like
10109	O15144	ARPC2	actin related protein 2/3 complex subunit 2
6001	O43665	RGS10	regulator of G protein signaling 10
6451	O75368	SH3BGRL	SH3 domain binding glutamate rich protein like
3939	P00338	LDHA	lactate dehydrogenase A
6647	P00441	SOD1	superoxide dismutase 1
3045	P02042	HBD	hemoglobin subunit delta
2243	P02671	FGA	fibrinogen alpha chain
4057	P02788	LTF	lactotransferrin
847	P04040	CAT	catalase
226	P04075	ALDOA	aldolase, fructose-bisphosphate A
301	P04083	ANXA1	annexin A1
6279	P05109	S100A8	S100 calcium binding protein A8
4353	P05164	MPO	myeloperoxidase
6280	P06702	S100A9	S100 calcium binding protein A9
2821	P06744	GPI	glucose-6-phosphate isomerase
5660	P07602	PSAP	prosaposin
353	P07741	APRT	adenine phosphoribosyltransferase
3948	P07864	LDHC	lactate dehydrogenase C
1991	P08246	ELANE	elastase, neutrophil expressed
1511	P08311	CTSG	cathepsin G
7431	P08670	VIM	vimentin
2950	P09211	GSTP1	glutathione S-transferase pi 1
2539	P11413	G6PD	glucose-6-phosphate dehydrogenase
128	P11766	ADH5	alcohol dehydrogenase 5 (class III), chi polypeptide
306	P12429	ANXA3	annexin A3
3936	P13796	LCP1	lymphocyte cytosolic protein 1
140465	P14649	MYL6B	myosin light chain 6B
4830	P15531	NME1	NME/NM23 nucleoside diphosphate kinase 1
5175	P16284	PECAM1	platelet and endothelial cell adhesion molecule 1
3009	P16401	HIST1H1B	histone cluster 1 H1 family member b
3006	P16403	HIST1H1C	histone cluster 1 H1 family member c
948	P16671	CD36	CD36 molecule
5223	P18669	PGAM1	phosphoglycerate mutase 1
566	P20160	AZU1	azurocidin 1
5657	P24158	PRTN3	proteinase 3
6286	P25815	S100P	S100 calcium binding protein P
4478	P26038	MSN	moesin
2040	P27105	STOM	stomatin
7086	P29401	TKT	transketolase
25824	P30044	PRDX5	peroxiredoxin 5
1992	P30740	SERPINB1	serpin family B member 1

7529	P31946	YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein beta
7001	P32119	PRDX2	peroxiredoxin 2
978	P32320	CDA	cytidine deaminase
4046	P33241	LSP1	lymphocyte specific protein 1
4628	P35580	MYH10	myosin heavy chain 10
88	P35609	ACTN2	actinin alpha 2
6888	P37837	TALDO1	transaldolase 1
4332	P41218	MNDA	myeloid cell nuclear differentiation antigen
5226	P52209	PGD	phosphogluconate dehydrogenase
397	P52566	ARHGDIB	Rho GDP dissociation inhibitor beta
3306	P54652	HSPA2	heat shock protein family A (Hsp70) member 2
8775	P54920	NAPA	NSF attachment protein alpha
3017	P58876	HIST1H2BD	histone cluster 1 H2B family member d
1667	P59665	DEFA1	defensin alpha 1
60	P60709	ACTB	actin beta
998	P60953	CDC42	cell division cycle 42
11034	P60981	DSTN	destrin, actin depolymerizing factor
4218	P61006	RAB8A	RAB8A, member RAS oncogene family
7334	P61088	UBE2N	ubiquitin conjugating enzyme E2 N
387	P61586	RHOA	ras homolog family member A
4069	P61626	LYZ	lysozyme
7531	P62258	YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon
8294	P62805	HIST1H4I	histone cluster 1 H4 family member i
5478	P62937	PPIA	peptidylprolyl isomerase A
7311	P62987	UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1
2885	P62993	GRB2	growth factor receptor bound protein 2
5879	P63000	RAC1	Rac family small GTPase 1
7171	P67936	TPM4	tropomyosin 4
10376	P68363	TUBA1B	tubulin alpha 1b
3043	P68871	HBB	hemoglobin subunit beta
3039	P69905	HBA1	hemoglobin subunit alpha 1
9446	P78417	GSTO1	glutathione S-transferase omega 1
3934	P80188	LCN2	lipocalin 2
10409	P80723	BASP1	brain abundant membrane attached signal protein 1
381	P84085	ARF5	ADP ribosylation factor 5
1213	Q00610	CLTC	clathrin heavy chain
10487	Q01518	CAP1	cyclase associated actin cytoskeleton regulatory protein 1
89	Q08043	ACTN3	actinin alpha 3 (gene/pseudogene)
23406	Q14019	COTL1	coactosin like F-actin binding protein 1
23193	Q14697	GANAB	glucosidase II alpha subunit
10130	Q15084	PDIA6	protein disulfide isomerase family A member 6
11140	Q16543	CDC37	cell division cycle 37
8290	Q16695	HIST3H3	histone cluster 3 H3

259215	Q5SQ64	LY6G6F	lymphocyte antigen 6 family member G6F
445582	Q6S8J3	POTEE	POTE ankyrin domain family member E
94239	Q71UI9	H2AFV	H2A histone family member V
339122	Q86YS6	RAB43	RAB43, member RAS oncogene family
246181	Q8NHP1	AKR7L	aldo-keto reductase family 7 like (gene/pseudogene)
1265	Q99439	CNN2	calponin 2
11343	Q99685	MGLL	monoglyceride lipase
8331	Q99878	HIST1H2AJ	histone cluster 1 H2A family member j
10938	Q9H4M9	EHD1	EH domain containing 1
29766	Q9NYL9	TMOD3	tropomodulin 3
58527	Q9P1F3	ABRACL	ABRA C-terminal like
50848	Q9Y624	F11R	F11 receptor
25932	Q9Y696	CLIC4	chloride intracellular channel 4
8773	O00161	SNAP23	synaptosome associated protein 23
487	O14983	ATP2A1	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 1
10092	O15511	ARPC5	actin related protein 2/3 complex subunit 5
9555	O75367	H2AFY	H2A histone family member Y
1431	O75390	CS	citrate synthase
8676	O75558	STX11	syntaxin 11
10476	O75947	ATP5PD	ATP synthase peripheral stalk subunit d
23052	O94919	ENDOD1	endonuclease domain containing 1
2771	P04899	GNAI2	G protein subunit alpha i2
3690	P05106	ITGB3	integrin subunit beta 3
3689	P05107	ITGB2	integrin subunit beta 2
292	P05141	SLC25A5	solute carrier family 25 member 5
5757	P06454	PTMA	prothymosin alpha
5232	P07205	PGK2	phosphoglycerate kinase 2
7169	P07951	TPM2	tropomyosin 2
7057	P07996	THBS1	thrombospondin 1
389	P08134	RHOC	ras homolog family member C
3674	P08514	ITGA2B	integrin subunit alpha 2b
3956	P09382	LGALS1	galectin 1
3178	P09651	HNRNPA1	heterogeneous nuclear ribonucleoprotein A1
3015	P0C0S5	H2AFZ	H2A histone family member Z
7295	P10599	TXN	thioredoxin
8288	P11678	EPX	eosinophil peroxidase
6037	P12724	RNASE3	ribonuclease A family member 3
87	P12814	ACTN1	actinin alpha 1
488	P16615	ATP2A2	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 2
3310	P17066	HSPA6	heat shock protein family A (Hsp70) member 6
239	P18054	ALOX12	arachidonate 12-lipoxygenase, 12S type
7414	P18206	VCL	vinculin
8348	P23527	HIST1H2BO	histone cluster 1 H2B family member o
5777	P29350	PTPN6	protein tyrosine phosphatase, non-receptor type 6
6136	P30050	RPL12	ribosomal protein L12

11151	P31146	CORO1A	coronin 1A
4627	P35579	MYH9	myosin heavy chain 9
5630	P41219	PRPH	peripherin
820	P49913	CAMP	cathelicidin antimicrobial peptide
2776	P50148	GNAQ	G protein subunit alpha q
6156	P62888	RPL30	ribosomal protein L30
72	P63267	ACTG2	actin, gamma 2, smooth muscle, enteric
5720	Q06323	PSME1	proteasome activator subunit 1
961	Q08722	CD47	CD47 molecule
23157	Q14141	Sep-06	septin 6
3927	Q14847	LASP1	LIM and SH3 protein 1
5094	Q15366	PCBP2	poly(rC) binding protein 2
345651	Q562R1	ACTBL2	actin, beta like 2
160287	Q6ZMR3	LDHAL6A	lactate dehydrogenase A like 6A
7846	Q71U36	TUBA1A	tubulin alpha 1a
81567	Q8NBS9	TXNDC5	thioredoxin domain containing 5
51109	Q8TC12	RDH11	retinol dehydrogenase 11
51762	Q92930	RAB8B	RAB8B, member RAS oncogene family
50	Q99798	ACO2	aconitase 2
51154	Q9UKD2	MRTO4	MRT4 homolog, ribosome maturation factor
373156	Q9Y2Q3	GSTK1	glutathione S-transferase kappa 1
58472	Q9Y6N5	SQOR	sulfide quinone oxidoreductase
728378	A5A3E0	POTEF	POTE ankyrin domain family member F
5874	O00194	RAB27B	RAB27B, member RAS oncogene family
9352	O43396	TXNL1	thioredoxin like 1
8436	O95810	CAVIN2	caveolae associated protein 2
80739	O95866	MPIG6B	megakaryocyte and platelet inhibitory receptor G6b
3848	P04264	KRT1	keratin 1
3688	P05556	ITGB1	integrin subunit beta 1
2934	P06396	GSN	gelsolin
302	P07355	ANXA2	annexin A2
823	P07384	CAPN1	calpain 1
3320	P07900	HSP90AA1	heat shock protein 90 alpha family class A member 1
5341	P08567	PLEK	pleckstrin
2773	P08754	GNAI3	G protein subunit alpha i3
2026	P09104	ENO2	enolase 2
3146	P09429	HMGB1	high mobility group box 1
653781	P0CG39	POTEJ	POTE ankyrin domain family member J
1191	P10909	CLU	clusterin
6515	P11169	SLC2A3	solute carrier family 2 member 3
6714	P12931	SRC	SRC proto-oncogene, non-receptor tyrosine kinase
5315	P14618	PKM	pyruvate kinase M1/2
7430	P15311	EZR	ezrin
1890	P19971	TYMP	thymidine phosphorylase
1465	P21291	CSRP1	cysteine and glycine rich protein 1
7416	P21796	VDAC1	voltage dependent anion channel 1

5742	P23219	PTGS1	prostaglandin-endoperoxide synthase 1
7453	P23381	WARS	tryptophanyl-tRNA synthetase
1937	P26641	EEF1G	eukaryotic translation elongation factor 1 gamma
10961	P30040	ERP29	endoplasmic reticulum protein 29
10935	P30048	PRDX3	peroxiredoxin 3
1892	P30084	ECHS1	enoyl-CoA hydratase, short chain 1
2923	P30101	PDIA3	protein disulfide isomerase family A member 3
2664	P31150	GDI1	GDP dissociation inhibitor 1
471	P31939	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
3187	P31943	HNRNPH1	heterogeneous nuclear ribonucleoprotein H1
3305	P34931	HSPA1L	heat shock protein family A (Hsp70) member 1 like
5245	P35232	PHB	prohibitin
3857	P35527	KRT9	keratin 9
2879	P36969	GPX4	glutathione peroxidase 4
3313	P38646	HSPA9	heat shock protein family A (Hsp70) member 9
908	P40227	CCT6A	chaperonin containing TCP1 subunit 6A
4190	P40925	MDH1	malate dehydrogenase 1
1399	P46109	CRKL	CRK like proto-oncogene, adaptor protein
22948	P48643	CCT5	chaperonin containing TCP1 subunit 5
5878	P51148	RAB5C	RAB5C, member RAS oncogene family
7879	P51149	RAB7A	RAB7A, member RAS oncogene family
4673	P55209	NAP1L1	nucleosome assembly protein 1 like 1
136319	P58546	MTPN	myotrophin
5901	P62826	RAN	RAN, member RAS oncogene family
2280	P62942	FKBP1A	FK506 binding protein 1A
10399	P63244	RACK1	receptor for activated C kinase 1
71	P63261	ACTG1	actin gamma 1
7332	P68036	UBE2L3	ubiquitin conjugating enzyme E2 L3
7277	P68366	TUBA4A	tubulin alpha 4a
5049	P68402	PAFAH1B2	platelet activating factor acetylhydrolase 1b catalytic subunit 2
391	P84095	RHOG	ras homolog family member G
3020	P84243	H3F3A	H3 histone family member 3A
5250	Q00325	SLC25A3	solute carrier family 25 member 3
22915	Q13201	MMRN1	multimerin 1
10954	Q14554	PDIA5	protein disulfide isomerase family A member 5
4735	Q15019	Sep-02	septin 2
22919	Q15691	MAPRE1	microtubule associated protein RP/EB family member 1
6813	Q15833	STXBP2	syntaxin binding protein 2
29094	Q3ZCW2	LGALSL	galectin like
11344	Q6IBS0	TWF2	twinfilin actin binding protein 2
55679	Q7Z4I7	LIMS2	LIM zinc finger domain containing 2
340205	Q86YW5	TREML1	triggering receptor expressed on myeloid cells like 1
441531	Q8N0Y7	PGAM4	phosphoglycerate mutase family member 4
26578	Q92882	OSTF1	osteoclast stimulating factor 1
56681	Q9NR31	SAR1A	secretion associated Ras related GTPase 1A

22908	Q9NTJ5	SACM1L	SAC1 like phosphatidylinositide phosphatase
51079	Q9P0J0	NDUFA13	NADH:ubiquinone oxidoreductase subunit A13
9218	Q9P0L0	VAPA	VAMP associated protein A
9124	O00151	PDLIM1	PDZ and LIM domain 1
10095	O15143	ARPC1B	actin related protein 2/3 complex subunit 1B
10094	O15145	ARPC3	actin related protein 2/3 complex subunit 3
81	O43707	ACTN4	actinin alpha 4
51596	O60888	CUTA	cutA divalent cation tolerance homolog
9948	O75083	WDR1	WD repeat domain 1
3417	O75874	IDH1	isocitrate dehydrogenase (NADP(+)) 1, cytosolic
2936	P00390	GSR	glutathione-disulfide reductase
2162	P00488	F13A1	coagulation factor XIII A chain
4860	P00491	PNP	purine nucleoside phosphorylase
5230	P00558	PGK1	phosphoglycerate kinase 1
760	P00918	CA2	carbonic anhydrase 2
2244	P02675	FGB	fibrinogen beta chain
2266	P02679	FGG	fibrinogen gamma chain
213	P02768	ALB	albumin
1476	P04080	CSTB	cystatin B
826	P04632	CAPNS1	calpain small subunit 1
3315	P04792	HSPB1	heat shock protein family B (small) member 1
6185	P04844	RPN2	ribophorin II
6176	P05386	RPLP1	ribosomal protein lateral stalk subunit P1
6181	P05387	RPLP2	ribosomal protein lateral stalk subunit P2
506	P06576	ATP5F1B	ATP synthase F1 subunit beta
2876	P07203	GPX1	glutathione peroxidase 1
2811	P07359	GP1BA	glycoprotein Ib platelet subunit alpha
203068	P07437	TUBB	tubulin beta class I
6678	P09486	SPARC	secreted protein acidic and cysteine rich
7168	P09493	TPM1	tropomyosin 1
230	P09972	ALDOC	aldolase, fructose-bisphosphate C
3107	P10321	HLA-C	major histocompatibility complex, class I, C
2098	P10768	ESD	esterase D
2812	P13224	GP1BB	glycoprotein Ib platelet subunit beta
6352	P13501	CCL5	C-C motif chemokine ligand 5
5224	P15259	PGAM2	phosphoglycerate mutase 2
10627	P19105	MYL12A	myosin light chain 12A
4001	P20700	LMNB1	lamin B1
4831	P22392	NME2	NME/NM23 nucleoside diphosphate kinase 2
1072	P23528	CFL1	cofilin 1
515	P24539	ATP5PB	ATP synthase peripheral stalk-membrane subunit b
10398	P24844	MYL9	myosin light chain 9
821	P27824	CANX	calnexin
5686	P28066	PSMA5	proteasome subunit alpha 5
9588	P30041	PRDX6	peroxiredoxin 6
5313	P30613	PKLR	pyruvate kinase L/R

10963	P31948	STIP1	stress induced phosphoprotein 1
4629	P35749	MYH11	myosin heavy chain 11
8407	P37802	TAGLN2	transgelin 2
4191	P40926	MDH2	malate dehydrogenase 2
830	P47755	CAPZA2	capping actin protein of muscle Z-line subunit alpha 2
539	P48047	ATP5PO	ATP synthase peripheral stalk subunit OSCP
3987	P48059	LIMS1	LIM zinc finger domain containing 1
3418	P48735	IDH2	isocitrate dehydrogenase (NADP(+)) 2, mitochondrial
7284	P49411	TUFM	Tu translation elongation factor, mitochondrial
2665	P50395	GDI2	GDP dissociation inhibitor 2
7408	P50552	VASP	vasodilator stimulated phosphoprotein
3295	P51659	HSD17B4	hydroxysteroid 17-beta dehydrogenase 4
3185	P52597	HNRNPF	heterogeneous nuclear ribonucleoprotein F
829	P52907	CAPZA1	capping actin protein of muscle Z-line subunit alpha 1
7873	P55145	MANF	mesencephalic astrocyte derived neurotrophic factor
51552	P61106	RAB14	RAB14, member RAS oncogene family
10096	P61158	ACTR3	ARP3 actin related protein 3 homolog
567	P61769	B2M	beta-2-microglobulin
7534	P63104	YWHAZ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta
375	P84077	ARF1	ADP ribosylation factor 1
5214	Q01813	PFKP	phosphofructokinase, platelet
800	Q05682	CALD1	caldesmon 1
5052	Q06830	PRDX1	peroxiredoxin 1
3611	Q13418	ILK	integrin linked kinase
10981	Q13637	RAB32	RAB32, member RAS oncogene family
2273	Q13642	FHL1	four and a half LIM domains 1
3837	Q14974	KPNB1	karyopherin subunit beta 1
6251	Q15404	RSU1	Ras suppressor protein 1
10982	Q15555	MAPRE2	microtubule associated protein RP/EB family member 2
7791	Q15942	ZYX	zyxin
10235	Q7LDG7	RASGRP2	RAS guanyl releasing protein 2
11315	Q99497	PARK7	Parkinsonism associated deglycase
3028	Q99714	HSD17B10	hydroxysteroid 17-beta dehydrogenase 10
81876	Q9H0U4	RAB1B	RAB1B, member RAS oncogene family
83442	Q9H299	SH3BGRL3	SH3 domain binding glutamate rich protein like 3
81027	Q9H4B7	TUBB1	tubulin beta 1 class VI
30845	Q9NZN3	EHD3	EH domain containing 3
51411	Q9UBW5	BIN2	bridging integrator 2
83699	Q9UJC5	SH3BGRL2	SH3 domain binding glutamate rich protein like 2
55573	Q9UKY7	CDV3	CDV3 homolog
5036	Q9UQ80	PA2G4	proliferation-associated 2G4
7419	Q9Y277	VDAC3	voltage dependent anion channel 3
335	P02647	APOA1	apolipoprotein A1
3273	P04196	HRG	histidine rich glycoprotein
7448	P04004	VTN	vitronectin

718	P01024	C3	complement C3
350	P02749	APOH	apolipoprotein H
259	P02760	AMBP	alpha-1-microglobulin/bikunin precursor
5196	P02776	PF4	platelet factor 4
1356	P00450	CP	ceruloplasmin
733	P07360	C8G	complement C8 gamma chain
5473	P02775	PPBP	pro-platelet basic protein
5265	P01009	SERPINA1	serpin family A member 1
336	P02652	APOA2	apolipoprotein A2
5004	P02763	ORM1	orosomucoid 1
1	P04217	A1BG	alpha-1-B glycoprotein
348	P02649	APOE	apolipoprotein E
462	P01008	SERPINC1	serpin family C member 1
8542	O14791	APOL1	apolipoprotein L1
3697	P19827	ITIH1	inter-alpha-trypsin inhibitor heavy chain 1
3827	P01042	KNG1	kininogen 1
344	P02655	APOC2	apolipoprotein C2
5950	P02753	RBP4	retinol binding protein 4
7276	P02766	TTR	transthyretin
710	P05155	SERPING1	serpin family G member 1
563	P25311	AZGP1	alpha-2-glycoprotein 1, zinc-binding
5345	P08697	SERPINF2	serpin family F member 2
5005	P19652	ORM2	orosomucoid 2
183	P01019	AGT	angiotensinogen
735	P02748	C9	complement C9
114770	Q96PD5	PGLYRP2	peptidoglycan recognition protein 2
3512	P01591	JCHAIN	joining chain of multimeric IgA and IgM
325	P02743	APCS	amyloid P component, serum
116844	P02750	LRG1	leucine rich alpha-2-glycoprotein 1
6291	P35542	SAA4	serum amyloid A4, constitutive
715	P00736	C1R	complement C1r
922	O43866	CD5L	CD5 molecule like
727	P01031	C5	complement C5
4060	P51884	LUM	lumican
3053	P05546	SERPIND1	serpin family D member 1
5627	P07225	PROS1	protein S
6401	P16581	SELE	selectin E
341	P02654	APOC1	apolipoprotein C1
1264	P51911	CNN1	calponin 1
2161	P00748	F12	coagulation factor XII
1401	P02741	CRP	C-reactive protein
3479	P05019	IGF1	insulin like growth factor 1
721	P0C0L5	C4B	complement C4B (Chido blood group)
720	P0C0L4	C4A	complement C4A (Rodgers blood group)
5444	P27169	PON1	paraoxonase 1
2147	P00734	F2	coagulation factor II, thrombin

5340	P00747	PLG	plasminogen
5446	Q15166	PON3	paraoxonase 3
5624	P04070	PROC	protein C, inactivator of coagulation factors Va and VIIIa
8858	P22891	PROZ	protein Z, vitamin K dependent plasma glycoprotein
2153	P12259	F5	coagulation factor V
347	P05090	APOD	apolipoprotein D
2160	P03951	F11	coagulation factor XI
3486	P17936	IGFBP3	insulin like growth factor binding protein 3
5104	P05154	SERPINA5	serpin family A member 5
6283	P80511	S100A12	S100 calcium binding protein A12
1956	P00533	EGFR	epidermal growth factor receptor
5617	P01236	PRL	prolactin
6288	P0DJI8	SAA1	serum amyloid A1
1742	P78352	DLG4	discs large MAGUK scaffold protein 4
3929	P18428	LBP	lipopolysaccharide binding protein
9370	A8K660	ADIPOQ	adiponectin, C1Q and collagen domain containing
7450	P04275	VWF	von Willebrand factor
4059	A0A087WXM8	BCAM	basal cell adhesion molecule (Lutheran blood group)
12	A0A024R6P0	SERPINA3	serpin family A member 3
3484	P08833	IGFBP1	insulin like growth factor binding protein 1
3383	P05362	ICAM1	intercellular adhesion molecule 1
3833	A0A024RCS7	KIFC1	kinesin family member C1
6462	B0FWH2	SHBG	sex hormone binding globulin
4151	A0A1K0FU49	MB	myoglobin
3105	P30443	HLA-A	major histocompatibility complex, class I, A
8857	Q9Y6R7	FCGBP	Fc fragment of IgG binding protein
1621	P09172	DBH	dopamine beta-hydroxylase
3959	A0A0S2Z3Y1	LGALS3BP	galectin 3 binding protein
3699	Q06033	ITIH3	inter-alpha-trypsin inhibitor heavy chain 3
27445	Q9Y6V0	PCLO	piccolo presynaptic cytomatrix protein
7060	E7ES19	THBS4	thrombospondin 4
165055	Q96M89	CCDC138	coiled-coil domain containing 138
124565	Q9HBR0	SLC38A10	solute carrier family 38 member 10
5297	B4DYG5	PI4KA	phosphatidylinositol 4-kinase alpha
23216	Q86TI0	TBC1D1	TBC1 domain family member 1
117159	P81605	DCD	dermcidin
10463	A0A0S2Z514	SLC30A9	solute carrier family 30 member 9
51188	A0A024R2Q8	SS18L2	SS18 like 2
10678	Q9NY97	B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2
84818	Q8NAC3	IL17RC	interleukin 17 receptor C
337	NA	APOA4	apolipoprotein A4
375616	Q6ZWJ8	KCP	kielin cysteine rich BMP regulator
221711	B4DFB8	SYCP2L	synaptonemal complex protein 2 like
23581	B2CIS9	CASP14	caspase 14
57578	Q9P2D8	UNC79	unc-79 homolog, NALCN channel complex subunit

83394	A1A5C9	PITPNM3	PITPNM family member 3
3782	A0A087WYJ0	KCNN3	potassium calcium-activated channel subfamily N member 3
6278	P31151	S100A7	S100 calcium binding protein A7

Appendix 5 – Ethics Addition: La Trobe Ethics

LA TROBE UNIVERSITY

Research and Graduate Studies Committee
University Human Ethics Committee
College Human Ethics Sub-Committees
www.latrobe.edu.au/researchers/ethics/human-ethics

Research Office

MODIFICATION FORM – HUMAN RESEARCH ETHICS

1. Approval Number	Ethics Application HEC10-071
2. Project Title	Prediction and Prevention to Achieve Optimal Recovery Endpoints after stroke (PrePARE)
3. Chief Investigator / Supervisor: (academic staff members only)	Name: Leeanne Carey Email address: [redacted]
4. Student (if applicable)	Name: Vinh Nguyen Email address: [redacted]
5. Project Duration: (subject to annual review)	Project commenced: 19/09/2010 Project concludes: 31/12/2015

PLEASE NOTE THAT THE MODIFICATIONS PROPOSED IN THIS FORM MUST NOT COMMENCE WITHOUT PRIOR WRITTEN APPROVAL FROM THE UHEC OR RELEVANT CHESC

6. MODIFICATIONS PROPOSED: modifications may include minor changes to the study, such as the aims, direction, procedures, personnel, duration, recruitment methods or numbers of participants, in addition to alterations of support documents. The UHEC or appropriate CHESC will review the proposed modifications and reserve the right to determine if a new application is required. Please itemise the changes you are requesting.

Addition of research personnel; Vinh Nguyen

7. REASONS FOR THE MODIFICATIONS: please summarise your reasons for requesting the above changes and indicate whether to date, any ethically significant incidents have arisen or any complaints have been received in connection with this project.

Vinh Nguyen is enrolled in Masters of Science (by research) and will be undertaking his research within the Neurorehabilitation and Recovery Team at the Florey Institute where the PrePARE study is being conducted. Vinh will be conducting an investigation into the trajectory of protemomics of stroke using plasma and tandem mass spectrometry collected through the START-PrePARE study. Prof Leeannce Carey is one of Vinh's supervisors.

For new personnel please complete an Investigator Template for each new investigator.

NEW INVESTIGATOR For database purposes please ensure that all details are up to date and correct.			
Name	Vinh Nguyen	Phone	[redacted]
		Email	[redacted]
School/Institute Position	Department of Psychology and Counseling, La Trobe University and The Florey Institute of Neuroscience and Mental Health. Post-graduate student	Staff/ Student No.	
Academic Title / Qualification	Bachelor of Psychological Science	Signature	Click here to enter text.
Position / Other affiliations. If Student provide details on Level and Course of Study	Post graduate student undertaking a Masters of Science (By Research)		

The report must be submitted electronically by the Chief investigator from the La Trobe University staff email account.