Triton X-114 Phase Separation in the Isolation and Purification of Mouse Liver Microsome Membrane Proteins

Rommel A. Mathias[1#], Yuan-Shou Chen[1,2#], Eugene A. Kapp[1], David W. Greening[1], Suresh Mathivanan[1], and Richard J. Simpson[1*]

[1] Ludwig Institute for Cancer Research, Parkville, Victoria, Australia.

[2] Department of Surgery (RMH), The University of Melbourne, Parkville, Victoria, Australia.

[#] These authors contributed equally to this work

*** To whom correspondence should be addressed:**

Professor Richard J. Simpson

Ludwig Institute for Cancer Research

PO Box 2008, Royal Melbourne Hospital

Royal Parade, Parkville, Victoria 3050, Australia

Tel: +61 03 9341 3155

Fax: +61 03 9341 3192

Email: **Richard.Simpson@ludwig.edu.au**

Running title: Triton X-114 membrane protein enrichment

Keywords: microsomes, liver, proteomics, mass spectrometry, Triton X-114, cytochrome P450, detergent

Abbreviations:

TX-114	Triton X-114
IMP	integral membrane protein
IPI	International Protein Index
MF	microsome fraction
DP	detergent phase
AP	aqueous phase
PF	pellet fraction
TMHMM	transmembrane hidden markov model
GRAVY	grand average of hydropathy
TMD	transmembrane domain
LC-MS/MS	liquid chromatography tandem mass spectrometry
CYP	cytochrome P450
2-D	two-dimensional
3-D	three-dimensional
ER	endoplasmic reticulum
CMC	critical micelle concentration
MGF	MASCOT generic file

ABSTRACT

Integral membrane proteins (IMPs) mediate several cellular functions including cell adhesion, ion and nutrient transport, and cell signalling. IMPs are typically hard to isolate and purify due to their hydrophobic nature and low cellular abundance, however, microsomes are small lipid vesicles rich in IMPs, which form spontaneously when cells are mechanically disrupted. To examine membrane protein expression in mouse liver, samples were homogenised, and microsomes collected by differential centrifugation, washed with sodium carbonate, and subjected to GeLC-MS/MS analysis. A total of 1124 proteins were identified in the microsome fraction, with 47% (524/1124) predicted to contain at least one transmembrane domain (TMD) according to TMHMM. The ability of the detergent Triton X-114 (TX-114) to further enrich for membrane proteins was also evaluated. Microsomes were subjected to successive rounds of solubility-based phase separation, with proteins partitioning into the aqueous phase, detergent phase, or pellet fraction. GeLC-MS/MS analysis of the three TX-114 fractions identified 1212 proteins, of which 146 were not detected in the un-fractionated microsome sample. Conspicuously, IMPs partitioned to the detergent phase, with 56% (435/770) of proteins identified containing one or more TMDs. Thus TX-114 phase separation further enriched IMPs by 9%. GO Slim characterisation of microsome proteins revealed cytoplasm, endoplasmic reticulum, mitochondria, Golgi apparatus, and endosome origin. Further, microsomes were enriched with enzymes including oxidoreductases, with monooxygenases well represented through the identification of 35 cytochrome P450s. Evaluation of biological processes showed enrichment of proteins involved in fatty acid biosynthesis and elongation, and steroid synthesis. In addition, transport proteins were enriched in the microsome dataset including 24 members of the Rab family of GTPases. Comparison of this dataset with the current microsome proteome adds an additional 648 protein identifications, of which 50% (326/648) contain at least one TMD. Application of sodium carbonate to remove peripherally-associated proteins and intracellular contaminants, followed by TX-114 phase partitioning, is an effective strategy to isolate and purify microsome IMPs.

1. INTRODUCTION

Integral membrane proteins (IMPs) are found in all cellular membranes and perform many of the functions essential for life including cell-cell communication, adhesion, and cell signalling {Wu, 2003 #267; Chou, 2005 #207}. IMPs are a valuable class of proteins since they are the target of 70-80% of all drugs and comprise three major protein categories: – 1) transmembrane proteins, the most common type of IMPs, which span, either single pass or multipass, the lipid bilayer in which it is embedded, - 2) cytosolic membrane-associated proteins that non-covalently attach to the internal membrane surface via an amphipathic α helix, and – 3) lipid-linked proteins that localise outside the lipid bilayer but which connect to the bilayer via one or more of their covalently C-terminally attached lipid moieties {Levental, 2010 #881}. Based on the amino acid sequences from several genomes, IMPs are predicted to represent 20-30% of eukaryotic proteomes {Wallin, 1998 #249; Arai, 2003 #183; Krogh, 2001 #210; Eisenhaber, 2004 #253; Punta, 2007 #221}. However, their identification at the protein level remains under-represented {Tan, 2008 #262} due, primarily, to their low cellular abundance and hydrophobicity {Simpson, 2000 #372; Wu, 2003 #97}. The Human Proteome Project will demand more facile and reliable methods for membrane isolation and characterisation {Nilsson, 2010 #882; Rabilloud, 2010 #883}.

Purification and identification of IMPs is generally tedious requiring three major steps; -1) extraction from their native membrane environment into a detergent buffer that only partially mimics the physical and chemical properties of a lipid membrane, -2) protein separation, and -3) protein identification, typically by LC-MS/MS, and characterisation {Michelsen, 2009 #608}. Needless to say, the first two steps are the crucial towards achieving a comprehensive and successful IMP characterisation. Given that most membrane proteins have relatively low expression levels compared to cytosolic proteins (e.g., actin, vimentin), their detection in highly complex unfractionated mixtures is difficult due to the limitation of protein abundance dynamic range issues for most protein separation methods and MS instrumentation {Macher, 2007 #113; Helbig, 2010 #114}. In addition, the hydrophobicity of membrane proteins often cause them to aggregate and precipitate upon detergent removal, leading to purification complications {Tan, 2008 #594; Helbig, 2010 #114}. To overcome these limitations, several isolation and

purification strategies have been developed to enhance membrane protein solubilisation and enrichment {Gilmore, 2010 #317}.

Techniques enriching membrane proteins from whole cell lysates include centrifugation-based, affinity capture, and hydrophilic protein depletion methods {Gilmore, 2010 #317}. Centrifugation-based methods, including density gradient and differential centrifugation, enrich membranes on the basis of their size, shape and density {Moebius, 2005 #303; Simpson, 2000 #271}. Density gradient centrifugation typically utilises sucrose or iodoxinol (OptiPrep^{TM}), gradient mediums to separate membrane-bound cellular organelles {Castle, 2001 #728; Simpson, 2009 #292}. Affinity capture of membrane proteins is commonly performed using antibodies directed towards a cell surface protein {Mathivanan, 2010 #320}, biotinylation labelling of cell surface proteins {Scheurer, 2005 #618; Yu, 2006 #339; Elia, 2008 #620; Wollscheid, 2009 #294}, or lectin chemistry to isolate glycosylated cell surface proteins {Wei, 2009 #721; Ghosh, 2004 #623}. Depletion strategies which target hydrophilic proteins (for example, strong ionic solutions such as sodium or potassium chloride) can also be used to enrich hydrophobic membrane proteins, while high pH solutions (e.g., sodium carbonate) can be used to disrupt membrane vesicles, resulting in release of contaminating luminal proteins and peripherally-associated proteins {Fujiki, 1982 #110; Zhao, 2004 #714}.

Detergents such as SDS, Triton, and CHAPS {Prive, 2007 #686; Arnold, 2001 #685; Gilmore, 2010 #868; Arnold, 2007 #884} and to a lesser extent, organic solvent mixtures such as aqueous methanol (60%), methanol/chloroform, and methanol/trifluoroethanol, {Blonder, 2002 #632; Ruth, 2006 #633; Zuobi-Hasona, 2005 #634; Zhang, 2007 #635} have been used extensively in membrane proteome profiling studies. Another detergent-based method, detergent phase separation, provides a powerful alternative membrane purification method, due to its simplicity, cheapness and efficiency {Arnold, 2007 #884}. However, this method has not been widely used in the proteomics community for membrane proteome (also referred to as surfaceome) studies. Phase separation exploits the amphiphilic properties of membrane proteins. Detergents are amphipathic molecules containing both polar (charged head groups) and non-polar regions (extended hydrophobic hydrocarbon chains). At low concentrations, detergents are soluble in water as monomers and with increasing concentrations, above the so-called critical micelle

concentration (CMC), they form aggregates with defined size distribution called micelles (for a review on detergent biophysical properties, see {Arnold, 2007 #884}). Factors that influence the size of the micelle and CMC include the ionic strength and temperature of the detergent solution and, to a lesser extent, pressure, pH and impurities. Hence, by increasing detergent concentration or by changing temperature or salt concentration, the micellar solution becomes turbid (cloud point), and the micelles become immiscible with water and form large aggregates that separate from the aqueous phase; these aggregates form a separate detergent phase. Proteins in solution will partition into the aqueous or detergent phase, based upon their relative hydrophobicity (with membrane proteins typically partitioning to the detergent phase). Of the detergents used for phase separation in the enrichment of membrane proteins, Triton X-114 (TX-114) has been the most widely exploited {Bordier, 1981 #111; Brusca, 1994 #107}.

Given that microsomes are a rich source of IMPs {Wong, 2009 #875}, and have been extensively studied at the protein level {Ghosh, 2008 #475; Ghosh, 2010 #476; Peng, 2008 #478; Stevens, 2008 #477; Zgoda, 2009 #474; Zgoda, 2006 #479}, mouse liver microsomes were selected as a model to evaluate the efficacy of TX-114 to enrich for membrane proteins. Microsomes, small lipid bilayer vesicles that form spontaneously when cells are mechanically disrupted {Dallner, 1974 #468}, can be classified as rough or smooth depending on the recruitment of ribosomes {Dallner, 1974 #468}. They contain proteins of endoplasm reticulum (ER), plasma membrane, mitochondria, Golgi apparatus origin, as well as other cytosolic organelles. However, their organelle composition varies according to tissue source and isolation procedure {Dallner, 1974 #468; Dreger, 2003 #469}. Microsomes are a rich source of the family of cytochrome P450 (CYP) {Seliskar, 2007 #472}, a family of membrane associated enzymes that govern the metabolism of xenobiotics and fat-soluble vitamins, oxidation of unsaturated fatty acids, and hormone synthesis {Hasler, 1999 #473}. For this reason, liver microsomes are extensively used pharmacokinetic and drug metabolic studies {Hasler, 1999 #473}.

In the current manuscript, we investigated the use of TX-114 phase separation in the isolation and purification of mouse liver microsomal IMPs. GeLC-MS/MS analysis of microsomes purified with sodium carbonate identified 1124 proteins. Using a combination of TX-114 phase separation followed by GeLC-MS/MS revealed an additional 146 proteins. Overall, 648 protein

identifications have not been seen in previously published mouse microsome proteome studies. These findings suggest that TX-114 phase separation will play an important role in unravelling the membrane component of the Human Proteome.

2. MATERIAL AND METHODS

2.1 Preparation of mouse liver microsomes

Mouse liver microsome preparations {Peng, 2010 #848} were kindly provided by Prof. Bill Jordan (Victoria University, Wellington, New Zealand). Briefly, livers were excised from C57BL/6J male mice (10-11 weeks old), and immediately placed in ice-cold homogenization medium (5 mM Tris-HCl, pH 7.4 containing 0.25 M sucrose, 1 mM tetrasodium EGTA, 1 mM sodium orthovanadate, 2 mM sodium fluoride, and 1% (v/v) protease inhibitor cocktail (Sigma-Aldrich)). Livers were diced into 5-10 mm pieces, washed twice with homogenisation medium, and homogenised using a Polytron PT10/35 homogeniser. The homogenate was then centrifuged at 12,000 × *g* for 15 min to sediment large organelles. The supernatant was collected and centrifuged at 100,000 × *g* for 60 min. The microsomal pellet was resuspended on ice with 0.1 M Na_2CO_3, pH 11.5 containing 1% protease inhibitors for 60 min. The suspension was centrifuged at 100,000 × *g* for 60 min and resuspended in ice-cold water. All procedures were performed at 4°C.

2.2 Pre-condensation of Triton X-114

Before use, Triton X-114 (TX-114) (Fluka) was pre-condensed as described {Bordier, 1981 #111}, to remove the more hydrophilic components from the commercial reagent. 20 µg of TX-114 was dissolved in 980 mL of 10 mM Tris HCl buffer, pH 7.5 containing 150 mM NaCl at 4°C. The mixture was then heated in a 30°C water bath and incubated at the same temperature overnight for phase separation. The upper aqueous phase was discarded and replaced by the same volume of Tris-HCL buffer. The lower detergent phase was collected and stored at 4°C. All procedures were repeated twice. The concentration of TX-114 used was 11.4% (w/v) {Bordier, 1981 #111}.

2.3 Triton X-114 phase partitioning of mouse liver microsomes

TX-114 phase partitioning was performed using the multiple washing method reported by Brusca *et al.* {Brusca, 1994 #107}. Briefly, the carbonate-washed microsome fraction (MF), was resuspended in aqueous 2% (w/v) TX-114, containing 10 mM Tris-HCl, pH 7.5, and 150 mM NaCl. The suspension was incubated on ice for 30 min with frequent vortexing, and then

centrifuged at 10,000 × *g* at 4°C for 20 min to sediment the pellet fraction (PF). The supernatant was collected, and incubated for 15 min at 37°C to achieve phase partitioning. The mixture was centrifuged at 5,000 × *g* at 25°C for 30 min, and the upper aqueous phase (AP) and lower detergent phase (DP) carefully collected. The AP was further purified by adjusting the TX-114 concentration to 2% (w/v) and repeating the phase partitioning, as above. Similarly, the DP was further purified by adding an equal volume of aqueous 0.06% (w/v) TX-114 and repeating the phase partitioning process, as above. Proteins in each of the AP, DP, and PF fractions were harvested by acetone precipitation, performed twice (three volumes ice-cold acetone: one volume sample {Jiang, 2004 #92}), before being air dried and stored at -80°C.

2.4 Bio-Rad protein concentration assay

The colormetric-based Bio-Rad protein assay was used, according to manufacturer's instructions (http://www3.biorad.com/LifeScience/pdf/Bulletin_9004.pdf), to estimate protein concentration of microsome samples, Briefly, varying concentrations of BSA (0.2, 0.4, 0.6, 0.8 and 1.0 mg/mL) were used to generate a standard curve. Equal volumes (3 μL) of BSA standard and unknown samples were carefully mixed with 100 μL Bio-Rad dye reagent (five-time-diluted) in a 96-well plate. The absorbance (595 nm) of each sample was measured to determine protein concentration. Samples were analysed in triplicate.

2.4 SDS-PAGE

Aliquots (30 μg) of each of the four membrane fractions (MF, DP, AP, PF) were resuspended in 1× NuPAGE LDS sample buffer (Invitrogen) containing 50 mM DTT, boiled at 95°C for 10 min, and loaded onto a 4-12% NuPAGE Novex Bis-Tris gel (Invitrogen). Electrophoresis was performed in MES running buffer (Invitrogen) at 150 V (constant voltage) until the tracking dye reached the bottom of the gel. Proteins were visualised using Imperial Protein Stain (PIERCE).

2.5 In-gel protein digestion

Individual gel lanes were excised into 24 2-3 mm gel slices, and gel slices were reduced, alkylated, and subjected to in-gel tryptic digestion, as previously described {Moritz, 1996 #122}. Briefly, gel slices were reduced with 10 mM DTT for 30 min, alkylated with 25 mM iodoacetamide for 20 min, and trypsinised (150 ng) (Worthington Biochemical Corporation) for

5 h at 37ºC. Peptide digests were extracted and dried by centrifugal lyophilisation (SpeedVac, Savant) to a volume of ~10 μL. Digests were then subjected to MS/MS analysis.

2.6 Nano LC-MS/MS analysis

A 96-well plate containing peptide digests was loaded into the autosampler for injection and fractionation by nano flow reverse-phase-HPLC (Model 1200, Agilent). Fractionation was performed using a nano-Acquity (C18) 150 mm x 0.15 mm I.D. RP-UPLC column (Waters) developed with a linear 60-min gradient from 0-100% Buffer B (0.1% (v/v) aqueous formic acid / 60% (v/v) acetonitrile) with a flow rate of 0.8 μL/min at 45°C, where Buffer A was 0.1% (v/v) aqueous formic acid. The capillary HPLC was coupled on-line to the LTQ-Orbitrap mass spectrometer equipped with a nano-electrospray ion source (Thermo Fisher Scientific). Positive ion mode was used for data-dependent acquisition. Survey MS scans were acquired with the resolution set to 30,000. Each scan was recalibrated in real time by co-injecting an internal standard from ambient air into the C-trap {Olsen, 2005 #7}. Up to the five most intense ions per cycle were fragmented and analysed in the linear trap. Target ions already selected for MS/MS were dynamically excluded for 180 s.

2.7 Protein identification and bioinformatic analysis

Peak lists were generated using *extract-msn*, with the following parameters: minimum mass 700; maximum mass 5,000; grouping tolerance 0.01 Da; intermediate scans 200; minimum group count 1; 10 peaks minimum and TIC of 100. Peak lists for each LC-MS/MS run were merged into a single MASCOT Generic File (MGF) for searches. Automatic charge-state recognition was used because of the high resolution survey scan (30,000). Database search parameters were as follows: S-carboxymethylation of cysteine was set as a fixed modification (+58 Da), as well as variable modifications consisting of NH_2-terminal acetylation (+42 Da) and oxidation of methionine (+16 Da), and the allowance for up to three missed tryptic cleavages. Peptide mass tolerance was ±20 ppm and #13C defined as 1. MGFs were searched against the International Protein Index (IPI) mouse protein sequence database (version 3.36 containing 51,326 entries), using the MASCOT search algorithm (v2.2.04, Matrix Science, U.K.) {Perkins, 1999 #8}.

An in-house software program (MSPro) was used for parsing and summarizing the output files from MASCOT searches, as previously described {Greening, 2008 #12}. Peptide identifications were deemed significant if the Ion score (IS) was ≥ the Homology score (or Identity score if there was no Homology score). False positive protein identifications were estimated by searching MS/MS spectra against the corresponding reversed-sequence (decoy) database {Kapp, 2005 #9}. Proteins with a Protein Score above 47 (false discovery rate of 1%) and contained at least 2 peptides identifications were deemed significant.

Transmembrane-spanning alpha helices were predicted using the web-based prediction program transmembrane hidden markov model (TMHMM) v2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0) {Sonnhammer, 1998 #245; Krogh, 2001 #210}. Grand average of hydropathy (GRAVY) scores were calculated according to the Kyte and Doolittle hydropathy scoring system {Kyte, 1982 #109}, using in-house software (FASTA2GRAVY). Protein Information Resource (http://pir.georgetown.edu) was used to obtain gene ontology annotations. GO Slim gene ontology annotation were retrieved for each of the 51,326 entries in the mouse IPI database (version 3.36), termed the mouse proteome. Category enrichment analysis was performed by comparing the total microsome dataset against the mouse proteome, and Chi square testing used to reveal classes that were statistically significant.

3. RESULTS AND DISCUSSION

3.1 GeLC-MS/MS analysis of TX-114 phase-partitioned mouse liver microsomal proteins

The purified carbonate-washed mouse liver microsome fraction (MF) was subjected to 2% TX-114 extraction to yield an aqueous phase (AP), detergent phase (DP) and pellet fraction (PF) (Figure 1). Protein aliquots (30 μg) from each fraction, were separated by 1D-SDS-PAGE (Figure 2A). To determine the protein composition of each fraction, gel bands from each of the four lanes were excised, trypsinised, and subjected to LC-Orbitrap MS/MS (Figure 2A). Acquired MS/MS spectra were searched against the mouse International Protein Index (IPI) protein database (v3.36) using the MASCOT search algorithm. Protein datasets for each fraction were processed individually (Supplementary Tables S2-S5), and merged as a master protein list for further analysis (Supplementary Table S1). As indicated in the two-way Venn diagram (Figure 2B), a total of 1270 unique protein identifications (termed the microsome proteome) were obtained from the four datasets (Supplementary Table S1), with 1124 proteins from the MF, and 1212 from the combined TX-114 fractions (AP, DP and PF); 1066 protein identifications were common to both analyses, and 58 and 146 proteins unique to the MF and combined TX-114 fractions, respectively.

It has been previously reported that washing of microsomes with sodium carbonate enhanced the detection of proteins containing TMDs, primarily through the depletion of soluble proteins from the membrane-enclosed compartments of the ER and other organelles {Peng, 2010 #848}. Thus, TMHMM {Krogh, 2001 #101} was used to predict the proportion of proteins in the MF that contained a TMD, and evaluate whether TX-114 was able to further enrich for this class of proteins. Of the 1124 proteins identified in the MF, an impressive 47% (524/1124) contained at least one TMD. Surprisingly, 32 TMD proteins were lost (out of 58) and 21 gained (out of 146) as a consequence of TX-114 phase separation of the MF. Rather, TX-114 extraction enhanced the detection of proteins in the AP.

< INSERT FIGURES 1 AND 2 >

3.2 Efficacy of TX-114 phase partitioning

Analysis of proteins identified in the three TX-114 fractions revealed 1034 in the AP and 292 unique, 770 in the DP and 58 unique, and 720 in the PF with 57 unique identifications; 507 proteins were common to the three TX-114 fractions (Figure 2C). To assess the ability of TX-114 to partition hydrophobic IMPs into the DP, TMHMM {Krogh, 2001 #101} and GRAVY {Kyte, 1982 #109} analysis was performed for all 1270 proteins in the microsome proteome (Figure 3).

TMHMM predicted that 56% (435/770) of proteins in the DP, 46% (328/720) in the PF, and 14% (145/1034) in the AP contained at least one TMD (Table 1). The percentage of proteins containing a TMD (1-12 spanning domains) was generally the highest in the DP compared to all other fractions (Figure 3A). These data indicate that the DP following TX-114 extraction was further enriched for IMPs compared to the MF which had 47% (524/1124) of proteins containing a TMD. Although the DP had a higher percentage, the overall number of TMD proteins identified was reduced. This is presumably due to incomplete separation in which IMPs partition to either the AP or PF. Nonetheless, examination of the 58 proteins unique to the DP (Figure 2C) further highlights the preferential partitioning of IMPs to this phase, as 84% (49/58) were predicted to contain at least one TMD (Table 1). For example, alkylglycerol monooxygenase, and integrin alpha-1 are proteins that contain five and two TMDs respectively, which were only identified in the DP.

GRAVY scores can be used as a relative index to calculate the hydrophobicity of a protein {Kyte, 1982 #109}. Developed by Kyte and Doolittle, essentially each amino acid is given a hydropathy score, and the total protein score is the average of the summed amino acids over the protein sequence. For example, the more positive a GRAVY score, the more hydrophobic the protein. Based on this definition, membrane proteins usually contain GRAVY scores greater than -0.4, which is the average score for soluble proteins {Kyte, 1982 #109}, and proteins with a GRAVY score greater than 0.3 are considered to be highly hydrophobic {Wilkins, 1998 #642}. Analysis of the GRAVY score distribution for all of the proteins identified in each of the fractions revealed that all fractions had a hydrophobic average (i.e., >-0.4) (Figure 3B). However, the DP had the highest average of -0.15, followed by the PF with -0.21 and the AP

with -0.25. The original starting MF had an average GRAVY score of -0.21. GRAVY analysis also confirmed the partitioning of hydrophobic proteins to the DP following TX-114 phase separation, as it contained the highest percentage of hydropobic proteins with GRAVY scores above -0.4 (77%), as well as highly hydrophobic proteins with scores greater than 0.3 (10%) compared to the PF, AP and MF (Figure 3B). 75 proteins in the DP had GRAVY scores greater than 0.3, with 9 of those unique to that fraction including vesicle transport proteins GOT1B and SFT2B, ninjurin-1, myeloid-associated differentiation marker, transmembrane protein 56, PRA1 family protein 3, immediate early response 3-interacting protein 1, claudin-1, and uncharacterized protein C9orf7 homolog.

Collectively, these data indicate three major points: -1) the MF is rich in IMPs (47% TMD) due to the carbonate wash step that removes peripherally-associated and intracellular proteins. While performing additional TX-114 phase separation increases the percentage of TMD and hydrophobic proteins that partition to the DP (56% TMD), some are unavoidably lost during the purification process, resulting in a lower number overall, -2) IMPs containing a TMD appear to partition to the DP, while the AP appears to contain the highest number of hydrophilic proteins, and the PF contains a mixture of hydrophobic and hydrophilic constituents, -3) the phase partitioning process is incomplete as there are hydrophobic proteins, highly hydrophobic proteins, and proteins containing a TMD that were identified in the AP and PFs. This may be due to the extraction procedure only being performed twice to obtain the final fractions. Thus, additional rounds of phase partitioning (up to 5) would yield fractions with higher purity, which could then be combined prior to LC-MS/MS identification, and -4) TMD prediction by TMHMM and GRAVY hydrophobicity prediction do not correlate well. For example, 69% of proteins in the AP were deemed hydrophobic, while only 14% in this phase were predicted to contain at least one TMD. This may result from the fact that TMHMM predicts short amino acid sequences that may span the membrane, whereas the GRAVY score is averaged over the entire length of the protein which can sometimes be misleading.

< INSERT FIGURE 3 >

3.3 Characterisation of the mouse liver microsome proteome

To gain insights into the proteome profile of our mouse liver microsomal preparation we merged the MF, DP, AP, and PF datasets into a single microsome dataset (1270 proteins total) for analysis (Supplementary Table S1). GO Slim annotations for '*molecular function*' and '*biological process*' were obtained for the microsome proteome from the Protein Information Resource database and compared with the corresponding information for the entire mouse proteome (51,326 entries – version 3.36 – http://www.ebi.ac.uk/ipi/ipimouse.html) (Figure 4).

It can be seen in Figure 4A that the cellular components that were enriched in the microsome proteome include membrane proteins, as well as proteins from the cytoplasm, ER, mitochondria, Golgi apparatus, and endosome. Notably, 50% of all proteins identified in our proteomic study had membrane annotation. Nuclear proteins were the only class that were under-represented in the microsome proteome compared to the entire mouse proteome.

Examination of the GO Slim molecular function annotations revealed several classes of enzymes to be statistically enriched in our microsome proteome dataset (Figure 4B). Enzymes can be classified into six major groups based on chemical reactions, which include oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases (http://www.chem.qmul.ac.uk/iubmb/enzyme/). Most classes of enzymes were significantly enriched in the microsome dataset when compared to total mouse entries in the IPI database. A salient feature of this analysis was the statistically significant enrichment of the monooxygenases cytochrome P450s (CYPs), which also belong to the oxidoreductase family of enzymes. Eukaryotic CYPs are localised to the ER and mitochondrial membrane, and contain TMDs or hydrophobic sequences that anchor them in the membrane {Omura, 2006 #869}. CYPs are important modulators of drug metabolism {Nelson, 2004 #870}, and in mouse, 102 full-length *CYP* genes 88 pseudogenes have been identified {Nelson, 2004 #870}. In our microsome study, 35 CYP proteins were identified (Table 2), including CYP51, an evolutionary conserved integral enzyme involved in cholesterol biosynthesis in various tissues/organs {Seliskar, 2007 #472; Lepesheva, 2007 #647}. CYPs - 2c29, 2c44, 2c50, 2c54, 2c70, 2e1, 2j5, 2u1, 4a10, 4a12a, 4a12b, and 4f14, which are implicated in arachidonic acid and linoleic acid metabolism {Capdevila, 2000 #707; Kroetz, 2002 #709; Moran, 2000 #710; Zeldin, 2001 #708}, were also

identified, along with CYPs - 3a11, 3a13, and 3a25 that are involved in steroid hormone biosynthesis {Ghayee, 2007 #712}. Several other monooxygenases including flavin-containing monooxygenases (Fmo1 and Fmo5), 14-3-3 protein q, and phenylalanine-4-hydroxylase were also identified.

Conspicuously, proteins containing GO Slim biological annotation 'lipid metabolic process' were enriched in the microsome proteome (Figure 4C, Table 2). For example, several proteins involved in steroid and steroid hormone biosynthesis were identified including 3 b-hydroxysteroid dehydrogenase type 5, 3-b-hydroxysteroid-Delta(8)-Delta(7) isomerase, 3-keto-steroid reductase, 7-dehydrocholesterol reductase, corticosteroid 11-b-dehydrogenase isozyme 1, CYP51, estradiol 17-b-dehydrogenase 2 and 12, hydroxysteroid 17-b dehydrogenase 6 precursor, lanosterol synthase, steroid 5 a-reductase 1, Sterol-4-a- carboxylate 3-dehydrogenase, and steryl-sulfatase. In addition, proteins modulating unsaturated fatty acid biosythesis (acyl-CoA desaturase 1, elongation of very long chain fatty acids protein 5, estradiol 17-b-dehydrogenase 12, fatty acid desaturase 1 and 2, peroxisomal trans-2-enoyl-CoA reductase, peroxisomal 3-ketoacyl-CoA thiolase A, and synaptic glycoprotein SC2) and fatty acid elongation (3-ketoacyl-CoA thiolase, enoyl-CoA hydratase, and hydroxyacyl-coenzyme A dehydrogenase) were strongly represented in the proteomic datasets. These findings are of particular significance given that several attempts have been made to purify enzymes involved in the fatty acid elongation process, the majority of which have been unsuccessful due to the hydrophobic properties of the proteins {Jakobsson, 2006 #872}. Thus, the microsome preparation and TX-114 purification strategy employed in this study may be a method used to enrich these molecules of interest.

Proteins involved in various aspects of transport such as vesicle-mediated, ion, protein, lipid, and hydrogen transport are significantly enriched in microsomes (Figure 4C). These include clathrin, adapter proteins, ADP-ribosylation factors, vesicle transport proteins, vesicle-associated membrane proteins, syntaxins, and Rab GTPases (Table 2). The Rab family of proteins has least 60 members {Stenmark, 2009 #650}, 24 of which were identified in the total microsome dataset (Table 2). Rabs localise to distinct intracellular membranes, and mediate several processes including inter-organelle trafficking, exocytosis, endocytosis, and phagocytosis {Schwartz, 2007

#873; Zerial, 2001 #874}. Given that Rab pathways and intracellular vesicle trafficking can be hijacked by pathogens and genetic diseases {Stenmark, 2009 #650}, understanding Rab regulation, function, and localisation may provide new insights and mechanisms for defence. The TX-114 strategy, described in this study, provides a facile method for isolating the Rab family of proteins.

< INSERT FIGURE 4 >

3.4 Extending the proteome of mouse liver microsomes

Although mouse liver microsomes have been employed in thousands of metabolic, pharmacological, and biochemical studies (over 5,000 articles in PubMed), up until recently only a minority of these studies incorporated proteomic technologies {Breuza, 2004 #876; Kanaeva, 2005 #667; Knoblach, 2003 #877; Taylor, 2000 #878; Wu, 2004 #879}. Zgoda *et al.* studied the proteomic differences between control and phenobarbital-treated mouse liver microsomes {Zgoda, 2006 #479}, and the most comprehensive microsome dataset to date was a study also by Zgoda *et al*. where they compared three separation workflows to explore the mouse liver microsome proteome {Zgoda, 2009 #474}, namely -1) SDS-PAGE followed by LC-MS/MS (519 protein identifications), -2) 2-D LC-MS (1410 protein identifications), and -3) 3-D LC-MS (3703 protein identifications). Collectively, these three separation approaches revealed 4142 proteins identifications. As expected, the highest number of protein identifications came from the 3-D-LC separation strategy, which significantly reduced sample complexity prior to MS analysis. Owing to the significant differences in the samples analysed (Phenobarbital-treated versus untreated livers), detergents used for protein extraction, and protein separation methods, making meaningful comparisons is difficult. Of the 4241 proteins reported in the Zgoda dataset only 622 were in common with the 1270 proteins reported in our microsome proteome. Strikingly, there were 648 proteins uniquely identified in our microsome dataset, and of these, 50% (326/648) contained a TMD.

Additionally, comparison of mouse liver microsome studies with respect to the number of CYPs yields our identification of 35 the most extensive. A study by Lane and colleagues focused on the identification of CYPs in immuno-deficient mice using isotope labelling quantitative

proteomics and reported 17 CYPs {Lane, 2007 #666}. Sutton *et al.* reported 26 CYPs analysing mouse liver microsomes by GeLC-MS/MS {Sutton, 2010 #711}, and Peng and co-workers reported 25 CYPs in a study recently published in Proteomics {Peng, 2010 #848}. Zgoda *et al.* identified 29 CYPs in their combined three-part analysis.

4 CONCLUSION

FIGURE LEGENDS

Figure 1. Triton X-114 solubility-based phase partitioning of mouse liver microsomes

A) Preparation of mouse liver microsomes, as described by Jordan and colleagues {Peng, 2010 #848}. Mouse liver was dissected, diced, and homogenised using a Polytron PT10/35 homogeniser. The homogenate was then subjected to centrifugation to sediment large organelles and cell debris, and the supernatant further centrifuged at higher speed to collect the crude microsome sample. The crude sample was then washed with sodium carbonate to strip peripherally-associated proteins, and the purified microsome fraction (MF) collected by centrifugation. B) The MF was then subjected to phase partitioning using 2% Triton X-114 and the detergent phase (DP), aqueous phase (AP), and pellet fraction (PF) collected for proteomic analysis.

Figure 2. Proteomic analysis of Trion X-114 fractionated mouse liver microsomes

A) 30 µg of each protein sample was separated by SDS-PAGE and stained with Coomassie blue. Each gel lane was excised into 24 gel slices (2-3 mm), and gel bands reduced, alkylated, and trypsinised. Extracted tryptic peptides were subjected to LC-MS/MS, and acquired spectra searched against a protein database to reveal protein identifications. B) Venn diagram showing the number of proteins identified in the un-fractionated MF, compared to the combined TX-114 extracted fractions. C) Comparison of proteins identified in each of the three TX-114 fractions.

Figure 3. Transmembrane domain analysis and hydrophobicity distribution of proteins identified in mouse liver microsomes

A) Proteins containing one or more transmembrane domains (TMDs) were predicted using the TMHMM algorithm for all four datasets (MF, AP, DP, and PF). The DP contained the highest number of proteins with TMDs. B) GRAVY scores of proteins identified in each fraction was calculated based on Kyte and Doolittle algorithms, and are represented by a dot. The two dotted lines indicate the thresholds required for hydrophobic (-0.4) and highly hydrophobic (0.3) determination, and the thick black line the average GRAVY score for that dataset. The DP contains the highest number of hydrophobic proteins, indicated by the highest GRAVY score and number of proteins with GRAVY score greater than -0.4.

Figure 4. GO Slim characterisation of mouse liver microsome proteins

A) Cellular component, B) Molecular function, and C) Biological process annotation was obtained from the Protein Information Resource database for proteins identified in the total microsome dataset of 1270 proteins, and compared with the annotation of the entire mouse proteome. Chi-square statistical testing was used to reveal the subcellular localisation of proteins that were enriched in the total microsome dataset. Significance (*).

REFERENCES

ACKNOWLEDGEMENTS

This work was supported, in part, by the National Health & Medical Research Council of Australia (program grant #487922 (R.J.S), the Australian government Endeavour International Postgraduate Research Scholarship and the University of Melbourne International Research Scholarship (Y-S.C). Analysis of proteomic data described in this work was supported using the *Australian Proteomics Computational Facility* funded by the National Health & Medical Research Council of Australia grant #381413. This work was supported by funds from the Operational Infrastructure Support Program provided by the Victorian Government Australia. We acknowledge the Australian Cancer Research Foundation for providing funds to purchase the Orbitrap™ mass spectrometer. The authors thank Dr. Lifeng Peng and Prof. Bill Jordan for the microsome samples, and Robert Goode for access to the GRAVY software.

The authors declare no conflict of interest

<u>Table 1.</u> Membrane protein characteristics of mouse liver microsomes following Triton X-114 phase separation

	MF (1124)	unique (58)	AP (1034)	unique (292)	DP (770)	unique (58)	PF (720)	unique (57)
Transmembrane domain (TMHMM ≥1)	47% (524)	55% (32)	14% (145)	14% (42)	56% (435)	84% (49)	46% (328)	35% (20)
Hydrophobic (GRAVY >-0.4)	71% (793)	67% (39)	69% (710)	61% (178)	77% (592)	83% (48)	70% (507)	61% (35)

Table 2. Mouse liver microsome proteins involved in transport and enzyme activity

Accession No	Gene name	Protein Description
Monooxygenase (GO:0004497)		
IPI00128287.1	*CYP1A2*	Cytochrome P450 1A2
IPI00652237.1	*CYP2A5*	Cytochrome P450, family 2, subfamily a, polypeptide 5
IPI00122634.1	*CYP2A12*	Cytochrome P450, family 2, subfamily a, polypeptide 12
IPI00264925.3	*CYP2B10*	57 kDa protein
IPI00134503.1	*CYP2C29*	Cytochrome P450 2C29
IPI00114778.1	*CYP2C37*	Cytochrome P450 2C37
IPI00114780.4	*CYP2C39*	Cytochrome P450 2C39
IPI00229835.4	*CYP2C44*	Cytochrome P450 CYP2C44
IPI00128489.5	*CYP2C50*	Isoform 1 of Cytochrome P450 2C50
IPI00409800.1	*CYP2C54*	Cytochrome P450 2C54
IPI00409412.3	*CYP2C67*	Cytochrome P450, family 2, subfamily c, polypeptide 67
IPI00405136.5	*CYP2C68*	Cytochrome P450, family 2, subfamily c, polypeptide 68
IPI00269265.6	*CYP2C70*	Cytochrome P450 2C70
IPI00323908.1	*CYP2D10*	Cytochrome P450 2D10
IPI00225828.5	*CYP2D11*	Similar to P45016a-ms2
IPI00123619.6	*CYP2D22*	Cytochrome P450, family 2, subfamily d, polypeptide 22
IPI00321644.3	*CYP2D26*	Cytochrome P450 2D26
IPI00116572.1	*CYP2D9*	Cytochrome P450 2D9
IPI00110556.1	*CYP2E1*	Cytochrome P450 2E1
IPI00308328.3	*CYP2F2*	Cytochrome P450 2F2
IPI00117741.1	*CYP2J5*	Cytochrome P450 2J5
IPI00850984.1	*CYP2U1*	Isoform 1 of Cytochrome P450 2U1
IPI00134504.1	*CYP3A11*	Cytochrome P450 3A11
IPI00134519.1	*CYP3A13*	Cytochrome P450 3A13
IPI00754568.1	*CYP3A25*	Cytochrome P450 3A25
IPI00322218.4	*CYP4A10*	Cytochrome P450 4A10
IPI00828590.1	*CYP4A12A*	Cytochrome P450, family 4, subfamily a, polypeptide 12
IPI00350781.4	*CYP4A12B*	Cytochrome P450 4A12A
IPI00117996.2	*CYP4F13*	Cytochrome P450, family 4, subfamily f, polypeptide 13
IPI00109257.1	*CYP4F14*	Cytochrome P450 4F14
IPI00117994.1	*CYP4F15*	Cytochrome P450 CYP4F15
IPI00120197.1	*CYP4V3*	Cytochrome P450 4V3
IPI00458711.2	*CYP51*	Cytochrome P450, family 51
IPI00122818.1	*CYP7B1*	Cytochrome P450 7B1
IPI00316822.1	*CYP8B1*	Cytochrome P450 8B1
IPI00116432.1	*FMO1*	Dimethylaniline monooxygenase [N-oxide-forming] 1
IPI00352124.2	*FMO5*	Flavin containing monooxygenase 5
IPI00133549.3	*PAH*	Phenylalanine-4-hydroxylase
IPI00408378.4	*YWHAQ*	Isoform 1 of 14-3-3 protein theta

Vesicle-mediated transport (GO:0016192)		
IPI00112614.2	*ABCA1*	ATP-binding cassette sub-family A member 1
IPI00118899.1	*ACTN4*	Alpha-actinin-4
IPI00329843.4	*ANKFY1*	Isoform 1 of Ankyrin repeat and FYVE domain-containing protein 1
IPI00346965.5	*AP1B1*	Adaptor protein complex AP-1, beta 1 subunit
IPI00621460.2	*AP1G1*	AP-1 complex subunit gamma-1
IPI00119680.3	*AP1M1*	AP-1 complex subunit mu-1
IPI00118026.1	*AP1S1*	AP-1 complex subunit sigma-1A
IPI00753468.1	*AP2A2*	Adaptor protein complex AP-2, alpha 2 subunit, full insert sequence
IPI00221613.5	*ARF1*	ADP-ribosylation factor 1
IPI00221615.5	*ARF5*	ADP-ribosylation factor 5
IPI00230422.6	*BCAP31*	B-cell receptor-associated protein 31
IPI00132685.1	*BET1*	BET1 homolog
IPI00278462.4	*BNIP1*	Vesicle transport protein SEC20
IPI00323624.3	*C3*	Isoform Long of Complement C3 precursor (Fragment)
IPI00124830.3	*CD47*	Isoform 2 of Leukocyte surface antigen CD47 precursor
IPI00229680.5	*CHP*	Calcium-binding protein p22
IPI00330594.2	*CLEC4F*	C-type lectin domain family 4 member F
IPI00228978.2	*CLTB*	Isoform 1 of Clathrin light chain B
IPI00169916.11	*CLTC*	Clathrin heavy chain
IPI00129304.7	*COLEC12*	Collectin sub-family member 12, full insert sequence
IPI00453776.1	*EEA1*	Early endosome antigen 1
IPI00321744.3	*ERGIC1*	Endoplasmic reticulum-Golgi intermediate compartment protein 1
IPI00129485.3	*FCGR2B*	Fc receptor, IgG, low affinity IIb isoform 1
IPI00338854.3	*GNAI3*	Guanine nucleotide-binding protein G
IPI00316682.2	*GOLGA5*	Golgin subfamily A member 5
IPI00132923.1	*GOSR1*	Golgi SNAP receptor complex member 1
IPI00311726.3	*HIP1R*	Huntingtin-interacting protein 1-related protein
IPI00225550.4	*KALRN*	Similar to kalirin, RhoGEF kinase
IPI00132475.2	*LMAN1*	Lectin, mannose-binding, 1, full insert sequence
IPI00119063.2	*LRP*	Prolow-density lipoprotein receptor-related protein 1 precursor
IPI00281011.7	*MARCKSL1*	MARCKS-related protein
IPI00113457.1	*MBL2*	Mannose-binding protein C precursor
IPI00845556.1	*MIA3*	Isoform 2 of Melanoma inhibitory activity protein 3 precursor
IPI00126186.1	*MRC1*	Macrophage mannose receptor 1 precursor
IPI00620550.1	*MSR1*	Macrophage scavenger receptor 1
IPI00656325.2	*NSF*	Vesicle-fusing ATPase
IPI00121277.1	*PI4K2A*	Phosphatidylinositol 4-kinase type 2-alpha
IPI00264501.8	*PICALM*	Isoform 1 of Phosphatidylinositol-binding clathrin assembly protein
IPI00133119.1	*RER1*	Protein RER1
IPI00112948.2	*RTN3*	Isoform 3 of Reticulon-3
IPI00115644.2	*SAR1A*	SAR1a gene homolog 1 (S. cerevisiae), full insert sequence
IPI00278804.1	*SCAMP1*	Secretory carrier-associated membrane protein 1

IPI00114368.3	*SEC22B*	Vesicle-trafficking protein SEC22b
IPI00123349.2	*SEC23A*	Protein transport protein Sec23A
IPI00317604.1	*SEC23B*	Protein transport protein Sec23B
IPI00222225.2	*SEC24A*	Isoform 1 of Protein transport protein Sec24A
IPI00229483.3	*SEC24C*	SEC24 related gene family, member C
IPI00420955.5	*SORT1*	Isoform 1 of Sortilin precursor
IPI00111416.1	*STX12*	Syntaxin-12
IPI00621076.2	*STX16*	Syntaxin-16
IPI00857890.1	*STX16*	Syntaxin 16 isoform c
IPI00112000.2	*STX18*	Isoform 1 of Syntaxin-18
IPI00109335.1	*STX4A*	Syntaxin-4
IPI00785417.1	*STX5*	Isoform 1 of Syntaxin-5
IPI00109506.1	*STX6*	Isoform 2 of Syntaxin-6
IPI00329953.1	*STX7*	Syntaxin-7
IPI00466570.4	*TMED10*	Transmembrane emp24 domain-containing protein 10 precursor
IPI00229703.6	*VAMP2*	Vesicle-associated membrane protein 2
IPI00132276.1	*VAMP3*	Vesicle-associated membrane protein 3
IPI00118372.1	*VAMP4*	Vesicle-associated membrane protein 4
IPI00137647.1	*VAMP7*	Synaptobrevin-like protein
IPI00453589.1	*VAMP8*	Vesicle-associated membrane protein 8
IPI00329942.4	*VPS26A*	Isoform 2 of Vacuolar protein sorting-associated protein 26A
IPI00124291.1	*VPS45*	Vacuolar protein sorting-associated protein 45
IPI00128941.1	*YIF1A*	Protein YIF1A
IPI00111420.1	*YIPF5*	Protein YIPF5
IPI00114560.5	*RAB1*	Ras-related protein Rab-1A
IPI00130118.1	*RAB10*	Ras-related protein Rab-10
IPI00323897.3	*RAB11A*	Ras-related protein Rab-11A
IPI00169699.4	*RAB12*	RAB12, member RAS oncogene family
IPI00126042.3	*RAB14*	Ras-related protein Rab-14
IPI00116768.1	*RAB17*	Ras-related protein Rab-17
IPI00116770.1	*RAB18*	Ras-related protein Rab-18
IPI00133706.1	*RAB1B*	Ras-related protein Rab-1B
IPI00337980.5	*RAB21*	Ras-related protein Rab-21
IPI00116729.2	*RAB22A*	Ras-related protein Rab-22A
IPI00137227.1	*RAB2A*	Ras-related protein Rab-2A
IPI00222632.3	*RAB31*	Similar to RAB31, member RAS oncogene family
IPI00112587.3	*RAB32*	Ras-related protein Rab-32
IPI00130489.1	*RAB35*	Ras-related protein Rab-35
IPI00116688.1	*RAB3D*	Ras-related protein Rab-3D
IPI00113127.7	*RAB4A*	RAB4A, member RAS oncogene family
IPI00271059.2	*RAB4B*	Ras-related protein Rab-4B
IPI00132410.1	*RAB5A*	Ras-related protein Rab-5A
IPI00116563.2	*RAB5B*	Rab5B
IPI00224518.2	*RAB5C*	Ras-related protein Rab-5C

IPI00116697.1	*RAB6A*	Isoform 1 of Ras-related protein Rab-6A
IPI00408892.2	*RAB7A*	Ras-related protein Rab-7a
IPI00331128.3	*RAB8A*	Ras-related protein Rab-8A
IPI00127820.1	*RAB9A*	Ras-related protein Rab-9
Lipid metabolic process (GO:0006629)		
IPI00121833.3	*ACAA1A*	3-ketoacyl-CoA thiolase A, peroxisomal precursor
IPI00226430.2	*ACAA2*	3-ketoacyl-CoA thiolase, mitochondrial
IPI00112549.1	*ACSL1*	Long-chain-fatty-acid--CoA ligase 1
IPI00169772.2	*ACSL3*	Long-chain-fatty-acid--CoA ligase 3
IPI00111950.1	*AKR1C6*	Estradiol 17 beta-dehydrogenase 5
IPI00377351.2	*APOA4*	Apolipoprotein A-IV precursor
IPI00666034.2	*APOB*	Apolipoprotein B homolog
IPI00312058.5	*CAT*	Catalase
IPI00132076.1	*COMT1*	Isoform Membrane-bound of Catechol O-methyltransferase
IPI00121079.3	*CYB5R3*	Isoform 1 of NADH-cytochrome b5 reductase 3
IPI00458711.2	*CYP51*	Cytochrome P450, family 51
IPI00122818.1	*CYP7B1*	Cytochrome P450 7B1
IPI00130988.1	*DHCR7*	7-dehydrocholesterol reductase
IPI00137471.3	*EBP*	3-β-hydroxysteroid-Delta (8), Delta(7)-isomerase
IPI00315582.3	*EBPL*	Emopamil-binding protein-like
IPI00130804.1	*ECH1*	Mitochondrial Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase
IPI00454049.4	*ECHS1*	Enoyl-CoA hydratase, mitochondrial precursor
IPI00308278.1	*ELOVL5*	Elongation of very long chain fatty acids protein 5
IPI00468859.3	*FADS1*	Fatty acid desaturase 1
IPI00129362.1	*FADS2*	Fatty acid desaturase 2
IPI00134836.1	*FDFT1*	Squalene synthetase
IPI00119645.1	*G6PC*	Glucose-6-phosphatase
IPI00126184.7	*GC*	Vitamin D-binding protein precursor
IPI00262743.6	*GPSN2*	Synaptic glycoprotein SC2
IPI00121105.2	*HADH*	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial precursor
IPI00115599.6	*HSD11B1*	Corticosteroid 11-beta-dehydrogenase isozyme 1
IPI00110224.1	*HSD17B11*	Dehydrogenase/reductase SDR family member 8 precursor
IPI00119219.2	*HSD17B12*	Isoform 1 of Estradiol 17-beta-dehydrogenase 12
IPI00125076.4	*HSD17B2*	Estradiol 17-beta-dehydrogenase 2
IPI00331628.5	*HSD17B4*	Peroxisomal multifunctional enzyme type 2
IPI00127016.1	*HSD17B6*	Hydroxysteroid 17-beta dehydrogenase 6 precursor
IPI00316067.5	*HSD17B7*	3-keto-steroid reductase
IPI00314189.8	*HSD3B5*	3 beta-hydroxysteroid dehydrogenase type 5
IPI00111009.1	*HSD3B7*	3 beta-hydroxysteroid dehydrogenase type 7
IPI00126253.1	*LASS2*	LAG1 longevity assurance homolog 2
IPI00785217.1	*LDLR*	Low density lipoprotein receptor
IPI00117834.3	*LSR*	Isoform 1 of Lipolysis-stimulated lipoprotein receptor precursor
IPI00169958.3	*LSS*	Lanosterol synthase

IPI00309073.1	*MTTP*	Microsomal triglyceride transfer protein large subunit precursor
IPI00132600.1	*NPC1*	Niemann-Pick C1 protein precursor
IPI00128692.1	*NSDHL*	Sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating
IPI00116750.2	*OPRS1*	Isoform 1 of Sigma 1-type opioid receptor
IPI00461212.1	*OSBPL8*	Oxysterol-binding protein-like protein 8 isoform b
IPI00114710.2	*PCX*	Pyruvate carboxylase, full insert sequence
IPI00331596.6	*PECR*	Peroxisomal trans-2-enoyl-CoA reductase
IPI00317356.10	*PON1*	Serum paraoxonase/arylesterase 1
IPI00129433.4	*RDH1*	Retinol dehydrogenase type 1
IPI00116921.1	*SCARB1*	Scavenger receptor class B member 1
IPI00322530.1	*SCD1*	Acyl-CoA desaturase 1
IPI00114818.1	*SEC14L2*	SEC14-like protein 2
IPI00313236.3	*SLC27A5*	Bile acyl-CoA synthetase
IPI00130671.5	*SLC37A4*	Glucose 6-phosphate transporter
IPI00470988.3	*SRD5A1*	Steroid 5 alpha-reductase 1
IPI00118038.1	*STS*	Steryl-sulfatase precursor
IPI00751362.2	*TM7SF2*	Transmembrane 7 superfamily member 2
IPI00467833.5	*TPI1*	Triosephosphate isomerase