# The role of ZYG11 family members in development and disease

Charlene Mae Magtoto

A thesis submitted in the complete fulfilment of the requirements
of the degree of Doctor of Philosophy

June 2020

La Trobe Institute for Molecular Science
Department of Biochemistry and Genetics

La Trobe University

Victoria, Australia

# Table of Contents

# List of tables and figures

# Abstract

At the basis of organised epithelial tissues is the ability of cells to maintain normal cell shape and precisely orientate themselves at both a single-cell level and in relation to neighbouring cells. These properties are largely intertwined with a mechanism known as epithelial cell polarity, whereby disruptions in this process hallmark the initiation of epithelial cancers. High-content screening identified ZYG11A as a novel regulator of cell shape – a protein belonging to an E3 ubiquitin ligase family known to recognise N-terminal glycine degrons that, along with its conserved counterparts ZYG11B and ZER1, has been linked to invertebrate cell polarity yet has not been studied in this context in mammalian systems.

Through the use of *in vitro* mammalian cell lines, I have found that the individual loss of these genes is morphologically and functionally divergent. I show that ZYG11A is most important for cell migration and maintaining normal cell morphology, and these alterations are associated with the upregulation of the known polarity protein, Scribble. In addition, I have used two in vivo models – *D. melanogaster* and *D. rerio* – to provide the first evidence that the fundamental requirement for the ZYG11 family in development is context dependent and may be limited to invertebrate systems.

I have used siRNA rescue screening and proximity biotinylation to explore the known role of ZYG11A in protein degradation and identified novel ZYG11A genetic interactions, as well as key degradation, adhesion, and importantly, polarity proteins as putative targets of ZYG11A ubiquitination. These findings highlight a novel ZYG11A interactome, providing new mechanisms for how the ZYG11 family regulates cell polarity and morphology, and how this may ultimately impact the initiation and progression of epithelial cancers.

# Statement of authorship

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 or in the following preface. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

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

Charlene Mae Magtoto

28th June 2020

# Preface

I acknowledge the assistance of the following people who contributed to the work described in this thesis. Dr. Lee Miles (Monash University) for the generation and initial sequencing of all *zyg11* and *zer1* zebrafish CRISPR-Cas9 knockout lines used in Chapter 3. Daniel Thomas and Jennii Luu at the Victorian Centre for Functional Genomics (VCCC) for their assistance with cherry picking and robotic transfection of the siRNA library for screening conducted in Chapter 4. Dr. Pierre Faou from the Comprehensive Proteomics Platform (CPP) at La Trobe Institute for Molecular Science (LIMS) for processing and running the mass spectrometry samples in Chapter 4, and Dr. Harinda Rajapaksha (CPP at LIMS), who ran the statistical analysis of the resulting datasets.

# Abbreviations

| | |
|---|---|
| 26S | 26-subunit |
| 7-AAD | 7-aminoactinomycin D |
| AB | Apical basal |
| ABC | Ammonium Bicarbonate |
| ACD | Asymmetric cell division |
| AMP | Adenosine monophosphate |
| AP | Anterior posterior |
| APC/C | Anaphase promoting complex/Cyclosome |
| aPKC | Atypical protein kinase C |
| ARM | Armadillo |
| ATP | Adenosine triphosphate |
| BrdU | Bromodeoxyuridine |
| BSA | Bovine serum albumin |
| CCNE | Cyclin E |
| CD | Conductivity |
| CDC | Cell division cycle |
| Co-IP | Co-immunoprecipitation |
| Crb | Crumbs |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CRL | Cullin-ring ligase |
| CT | C-terminal |
| CUL | Cullin |
| CYB | Cyclin B |
| DAPI | 4′,6-diamidino-2-phenylindole |
| DLG | Discs large |
| DMEM | Dulbecco's Modified Eagle Medium |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| ECL | Enhanced chemiluminescence |
| EDTA | Ethylenediaminetetraacetic acid |
| EGF | Epidermal growth factor |
| ELM | Eukaryotic linear motif |
| Elo-B/-C | Elongin B/C |
| EMI | Early mitotic inhibitor |
| EMT | Epithelial to mesenchymal transition |
| FACS | Fluorescent activated cell sorting |
| FEM-1 | Feminisation of XX and XO animals' protein 1 |
| FITC | Fluorescein isothiocyanate |
| FL | Full length |
| Gal4 | Galactosidase transcription factor 4 |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| GC | Genetic compensation |
| GFP | Green fluorescent protein |
| GH | General hardness |
| GO | Gene ontology |
| GSK3β | Glycogen Synthase Kinase 3 Beta |
| HECT | Homologous to E6AP C-terminus |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HPV | Human papilloma virus |

| | |
|---|---|
| ICD | Intracellular domain |
| INX | In-cross |
| KD | Knockdown |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| KH | Carbonate Hardness |
| KO | Knockout |
| LGL | Lethal giant larvae |
| LMB | Leptomycin B |
| LRR | Leucine-rich repeats |
| MAPK | Mitogen-activated proteins kinase |
| MEMI | Meiosis to mitosis transition associated |
| mg | Milligram |
| mL | Millilitre |
| mM | Millimolar |
| MS | Mass spectrometry |
| MSCV | Murine stem cell virus |
| NES | Nuclear export signal |
| ng | Nanogram |
| NMD | Nonsense-mediated decay |
| NSCLC | Non-small cell lung cancer |
| NT | N-terminal |
| NUB1 | NEDD8 ultimate buster 1 |
| OTP-NT | On-target plus-non-targeting |
| Pals | Polarity protein associated with lin seven 1 |
| PAR | Partitioning defective |
| PARC | Cullin-9 |
| Patj | Protein associated to tight junctions |
| PBS | Phosphate buffered saline |
| PCA | Principal component analysis |
| PCP | Planar cell polarity |
| PCR | Polymerase chain reaction |
| PFA | Paraformaldehyde |
| PI | Propidium iodide |
| PI3K | Phosphoinositide 3-kinases |
| PPI | Protein-protein interaction |
| PRAME | Preferentially expressed antigen in melanoma |
| PTM | Post-translational modification |
| PVDF | Polyvinylidene fluoride or polyvinylidene difluoride |
| qRT-PCR | Quantitative real-time PCR |
| RB1 | Retinoblastoma 1 |
| RBR | Ring-between-ring |
| RBX1 | Ring-box 1 |
| RING | Really interesting new gene |
| RIPA | Radioimmunoprecipitation assay |
| RNP | Ribonucleoprotein |
| SDS-PAGE | Sodium dodecyl sulphate-polyacrylamide gel electrophoresis |
| sgRNA | Single guide RNA |
| Shh | Sonic hedgehog |
| siRNA | Small interfering RNA |
| SRS | Substrate recognition subunit |
| TBS | Tris buffered saline |
| TRiC | T-complex protein-1 ring complex |

| | |
|---|---|
| UAE | Ubiquitin activating enzyme |
| UAS | Upstream activation sequence |
| Ub | Ubiquitin |
| UBD | Ubiquitin binding domain |
| UPS | Ubiquitin-proteasome system |
| USP | Ubiquitin specific peptidase |
| VHL | von Hippel-Lindau |
| WCL | Whole cell lysate |
| ZER1 | Zyg-11 related-1 |
| ZO-1 | Zonula occludens-1 |
| ZYG11 | Zygotic defective-11 |
| ZYG11BL | ZYG11B-like |
| β-ME | β-mercaptoethanol |
| µg | Microgram |
| µl | Microlitre |
| µM | Micromolar |

# Acknowledgements

Oh wow, I can't believe I'm at this point! Where do I begin? I think really have to start by thanking my supervisor and mentor, Patrick Humbert, who taught me that there are countless ways to approach a question and that a great scientist should never have tunnel vision. I can honestly say that I could not have achieved this without your guidance. Thank you for taking a chance on an undergrad with a bad GPA who had never used a pipette before.

To the Humbert lab who have been my science family for the last 5 years – Tanja, Sam, Rebecca, Krystle, Lucas, Andrew, Michal and Bree (and honorary member Natasha). I could not have gotten through the highs and lows of this PhD without the hour-long lunches and trips to the Agora for a coffee. Cheers for all the karaoke, beers, and conference boogies we got to share. I am really going to miss tearing up the dancefloor with all of you. Special shout out to Sam and Tanja who have really been there since the Peter Mac days – well, looks like this is it from the old crew!

There are so many people who have helped me through this PhD, so I'll try and keep this short. Helena and Marta, everything I know about flies is thanks to you. You are both constant reminders of what strong women in science look like. Kaylene and the team at the VCFG – I really could not have gotten all of those screens done without your help. Everyone down at the CCP, especially Pierre and Harinda who honestly carried me through all of the Proteomics work. Thanks to Seb, Lee and Nishanthi who were all so patient with me when I was still learning how to be a fish person. I also have to thank Krystle who taught me how to drive the FACS machines… how do you set up the stream again?

I have to mention all the people who have given me something other than science to talk about during my PhD. Pete and Amir, who are always up for a chat about good food or 'that great restaurant I went to for dinner'. To Rebecca who was always keen to talk trash about tv and look for the doughnut man on Wednesdays. All the way on the other side of Australia, I have to thank Jeffreena for all the long chats and VCCC lunches.

A HUGE thank you to my family who have really supported me through the last 4 years. For always understanding when I say, 'I have to work this weekend' or 'don't ask me about my project/thesis'. I would have gone mad without the trips to the movies and big family lunches/dinners.

Finally, to my better half, Efthimios. You have been my constant through this rollercoaster of a PhD. Thank you for always believing in me, even when I didn't believe in myself. Looks like this is it!!!

# Chapter 1

Introduction

Cancer is an extensively widespread disease and a leading cause of death worldwide. On the global spectrum, 1 in 6 deaths is the result of cancer-related diseases, killing an estimated 8.8 million people per year (World Health Organisation, 2017). By the age of 85, half of the Australian population is diagnosed with cancer with more than 48,000 of these cases resulting in patient mortality on a yearly basis. Cancers initiated in epithelial tissues remain to be one of the most commonly diagnosed cancer types, where breast, prostate, lung, and skin cancer represent a large proportion of cancer-related deaths. One hallmark that epithelial tumours share is the loss of normal tissue architecture. Epithelial cell polarity plays an important role in the maintenance of tissue organisation and its disruption has been implicated in not only the initiation of primary cancerous lesions but also the progression of cancer into a more aggressive and metastatic state to form secondary lesions. Many polarity proteins from the Scribble, Pars and Crumbs complexes are known to act as neoplastic tumour suppressors with the ability to prevent tissue overgrowth and metastasis in epithelia (reviewed by Humbert *et al.*, 2008). Although these genes have been well characterised individually, there is still much to learn about how these genes are directly regulated and how novel regulators of cell polarity and cell morphology can interplay with the existing polarity networks. In this thesis, I will explore the role of a poorly characterised family of ubiquitin ligases, the ZYG11 family, and investigate how this family of proteins may regulate cell polarity and cell shape.

## 1.1 Cell polarity and the maintenance of cell shape

### 1.1.1 Types of cell polarity and how cell polarity is established

All tissues within the human body are the result of a precisely organised collection of cells that are capable of correctly orientating themselves in various planes individually, as well as in relation to neighbouring cells. How these cells maintain tissue homeostasis and respond to external and developmental clues is largely intertwined with the internal composition and asymmetric distribution of cellular determinants. Known as cell polarity, this fundamental characteristic whereby cells are able to actively create gradients of intracellular components (including proteins, lipids and organelles) is essential for the morphological and functional characteristics of most cells. Alterations in the spatial configuration of cellular polarity determinants is a defining factor between the many

flavours of polarity. *Asymmetric cell division (ACD)* is a type of cell-autonomous polarity typically observed in the context of stem cells and terminal differentiation. Here, the cellular determinants required to lose or maintain stemness are asymmetrically segregated during cell division in order to produce two different cells. *Front and rear polarity*, which is a type of anterior-posterior polarity, occurs during directed cell migration to distinguish the leading edge and the trailing edge of cells. Epithelial tissues are comprised of highly polarised networks of cells that exhibit two main forms of cell polarity. *Planar cell polarity (PCP)* coordinates the alignment of epithelial cells across a plane, for example, in the scales of a fish. *Apicobasal polarity*, defines top to bottom asymmetry within a cell and is important for the correct orientation and organisation of epithelial sheets (Figure 1.1). Other more specialised types of polarity also exist in other cells like neurons (for axon guidance) and T-lymphocytes (for immune synapse formation).

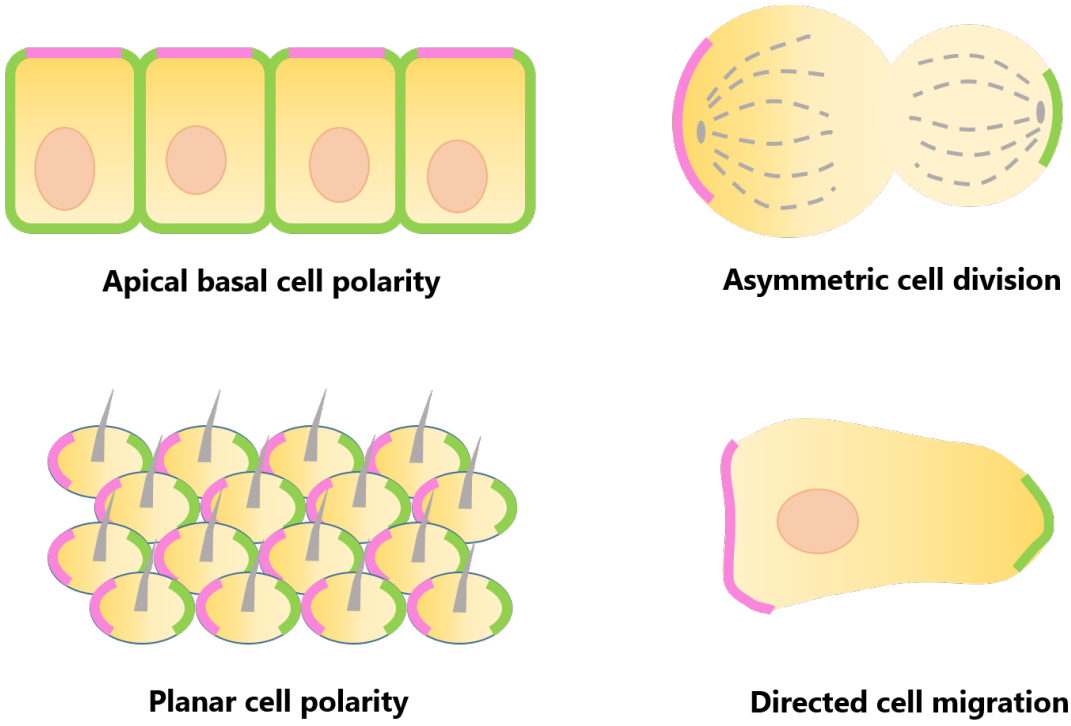



**Figure 1.1 The types of cell polarity**

The defining characteristic of a polarised cell is the result of an asymmetric distribution of cellular components (depicted in pink and green). Cell polarity is required for an array of cellular processes including cell orientation (apicobasal), asymmetric cell division, and directed cell migration (anterior-posterior). The effects of cell polarity are not confined to a single cell and can often facilitate the propagation of cell polarity signals and cell communication between a plane of cells through cell-cell connections. Planar cell polarity occurs across a plane of cells – for example, the scales of a fish or *Drosophila* bristles.

#### 1.1.1.1 Apical-basal cell polarity

In order to form organised and complex tissue structures, epithelial cells rely on the demarcation of the apical and basolateral membranes, which is achieved through the spatial restriction of polarity complexes within the cell. These principles were first identified in *D. melanogaster* and have since been mirrored in numerous organisms both *in vitro* and *in vivo*, and in invertebrate and vertebrate species (reviewed extensively by Rodriguez-Boulan and Macara, 2014; Humbert, Russell and Richardson, 2015; Campanale, Sun and Montell, 2017). One important trigger for polarity establishment in unpolarised cells comes from outside-in signalling from the extracellular matrix (ECM) to the actin and tubulin networks via transmembrane integrins (Howlett *et al.*, 1995; O'Brien *et al.*, 2001; Myllymäki, Teräväinen and Manninen, 2011). The removal of apical determinants from the maturing cell-ECM interface occurs through endocytosis and is accompanied by the secretion of laminins and other adhesive proteins to form the basement membrane (Wang, Ojakian and Nelson, 1990; O'Brien *et al.*, 2001; Li *et al.*, 2003; Yu *et al.*, 2005; Akhtar and Streuli, 2013). Concurrent to the formation of the basement membrane is the formation of adherens junctions, which is initiated by the recruitment of E-cadherin, α-catenin and β-catenin to facilitate cell-cell adhesion (Yonemura *et al.*, 1995). The maturation of adherens junctions prompts the synchronised formation of tight junctions that integrate occludins, claudins and ZO-1 (among many other proteins) at the apical interface of cells. Tight junctions are located above the adherens junctions and thus act as the primary line of barrier protection between the inner lumen of epithelia and the extracellular environment. The formation of these junctions both require the polarisation of proteins along the apicobasal axis of the cell. At the apex of cell polarity control are three highly conserved polarity complexes: 1) The Scribble complex, consisting of Scribble (Scrib), Discs large (Dlg) and Lethal giant larvae (Lgl), 2) The Par complex, which includes Partitioning defection protein 3 (Par3), Partitioning defective protein 6 (Par6) and Atypical protein kinase C (aPKC), and 3) The Crumbs complex, involving Crumbs (Crb1), Polarity protein associated with lin seven 1 (Pals1) and Protein associated to tight junctions (Patj). At the apical domain, the kinase activity of aPKC regulates the spatial restriction of the Par and Crumbs complexes to specific regions at the apical membrane and apical-lateral cortex near tight junctions. This works in a mutually restrictive manner with the kinase Par1, and the basolaterally located Scribble complex near the adherens junctions to establish and maintain apicobasal polarity. Many of these proteins act as scaffolds that recruit additional junctional components to coordinate Rho GTPase signalling, and the actin and tubulin networks. Together, these features define three important epithelial

membrane domains – the basal, apical, and lateral membranes – and form the basis of cell shape, and cell adhesion and barrier formation in epithelial tissues.

#### 1.1.1.2 Anterior-posterior (AP) polarity and directed migration

Anterior-posterior polarity has been extensively studied in the developing *C. elegans* embryo following the discovery of PAR proteins in genetic lethality screens (Kemphues *et al.*, 1988). In response to egg fertilisation, cytoskeletal remodelling and cortical flow breaks embryonic symmetry to redistribute PAR proteins to anterior (PAR-3 and -6) and posterior (PAR-1 and -2) regions. These findings have since been translated into numerous aspects of polarity in vertebrate and invertebrate systems (reviewed by Nance and Zallen, 2011; Motegi and Seydoux, 2013; Roth and Lynch, 2013). Other than apicobasal polarity (see above), the interplay between Par proteins and the Scribble polarity complex has also been described in a form of anterior-posterior polarity – front and rear polarity. During directed migration in mammalian cells, Scribble localises to the leading edge where it recruits βPix, Cdc42 and Rac1 to promote actin polymerisation and cellular protrusions in response to extracellular cues via integrins (Dow *et al.*, 2007). These proteins subsequently localise the Par complex to the leading edge via Cdc42. Here, the E3 ligase SMURF1 is recruited by the Par complex to degrade RhoA, restricting actomyosin contraction to the trailing edge of the cell (Wang, 2003; Zhang, Wang and Wrana, 2004; Wang *et al.*, 2006). This ultimately reinforces the formation of lamellipodia and focal adhesions at the front of the cell, and the concurrent contraction and detachment at the rear in order to ensure coordinated forward movement.

#### 1.1.1.3 Asymmetric cell division

Asymmetric cell division is a type of polarity that is largely observed during the differentiation of stem cell populations, where the asymmetric distribution of different cell fate determinants is localised to opposing poles (Figure 1.1, pink and green). In development, this has been well characterised in the developing *Drosophila* brain where the progenitor cells (neuroblasts) divide by establishing the spindle orientation along the apicobasal axis. This gives rise to two different cells - one self-renewing neuroblast often containing Par polarity proteins, and a smaller ganglion mother cell (daughter cell), which contains proteins like Miranda, Prospero and Numb that promote terminal differentiation

into various neuronal lineages (Harding and White, 2018). Indeed, the regulation of this process has been described in many other stem cell populations that harbour the same principles and has been extensively studied and reviewed (Santoro *et al.*, 2016; Venkei and Yamashita, 2018; Manzano-López and Monje-Casas, 2020).

#### 1.1.1.4 Planar cell polarity

Planar cell polarity (PCP) in epithelial tissues is often thought of in terms of two axes, the apicobasal axis and the proximal-distal axis. The regulators of these two axes cooperate in order to coordinate processes like tissue morphogenesis. This has been best described in *Drosophila* systems, for example in fly hair orientation, however many of the fundamental principles have been conserved in mammalian systems. At the core of planar cell polarity signalling are a number of transmembrane proteins including Frizzled at the proximal cortex (which responds to Wnt signalling ligands), and Vangl and Celsr1 (at the distal cortex). In order to facilitate communication between adjacent cells, these membrane proteins interact with other PCP proteins in the cytoplasm. At the distal cortex for example, Scribble and Dlg interact with Vangl2, whereas Lgl and Dishevelled interact with Frizzled (reviewed by Milgrom-Hoffman and Humbert, 2018). The polarised interactions at these opposing cortexes creates a gradient along the proximal-distal axis and permits the propagation of a polarity signal across a plane of cells.

### 1.1.2 Cell polarity and cancer

The deregulation of normal cell and tissue homeostasis is at the apex of cancer establishment and progression, with the key hallmarks of cancer having been defined by Hanahan and Weinberg in 2000 (and later revised in Hanahan and Weinberg, 2011). Cell polarity disruptions can be associated with a number of key steps in cancer development including aberrant tissue overgrowth, tissue disorganization, and epithelial to mesenchymal transitions (EMT), which are crucial for cancer invasion and metastasis (Thiery *et al.*, 2009; Godde *et al.*, 2010; Shapiro *et al.*, 2011; Savagner, 2015). We now know that the deregulation of polarity determinants, both the loss and overexpression are diagnostic criteria in cancer and that the signalling pathways affected have the capacity to initiate cancers and promote aggressive tumours (Pearson *et al.*, 2011; Royer and Lu, 2011;

Martin-Belmonte and Perez-Moreno, 2012; Macara and McCaffrey, 2013; Saito, Desai and Muthuswamy, 2018; Stephens *et al.*, 2018).

Perhaps one of the most obvious consequences of polarity deregulation in terms of cancer development and progression is linked to the loss of apicobasal polarity and tissue disorganisation. Polarity proteins are inherently required for the formation and maintenance of cell-cell and cell-matrix adhesion (see Section 1.1.1.1). However, defects in these core polarity complexes can weaken cell-cell junctions, and connections to the basement membrane, often leading to a switch from apicobasal polarity to front-rear polarity. This not only affects overall tissue organisation, but also creates an opportunity for defective cells to spread to distant sites where they can integrate into other tissues by mesenchymal to epithelial transition (MET), and form secondary lesions, i.e. metastasise. This plasticity between EMT and MET has been observed in number of solid tumours with polarity protein alterations, e.g. PATJ, PAR3 and tight junction protein Claudin-7 in breast cancer (Ozdamar, 2005; Godde *et al.*, 2010; McCaffrey *et al.*, 2012; Xue *et al.*, 2013) and aPKC and Lgl2 in lung adenocarcinoma (Schimanski *et al.*, 2005; Imamura *et al.*, 2013).

Proper asymmetric cell divisions are essential for stem cell maintenance. Therefore, the disruption of cell asymmetry in this context can result in either long-lived populations of defective stem cells (cancer stem cell pools) that have the capacity to self-renew, or aberrant cell differentiation. These cancer stem cells have been implicated in tumour over-proliferation (Sugiarto *et al.*, 2011; Gómez-López, Lerner and Petritsch, 2014) and indeed, this unregulated ability to self-renew has also been defined as one of the major hallmarks of cancer (Hanahan and Weinberg, 2011). Finally, defects in Wnt/planar cell polarity (PCP) signalling are widespread across a number of cancer types. Non-canonical Wnt signalling defects have been identified in many epithelial cancers including but not limited to breast, ovarian and prostate cancer. Because PCP signalling incorporates polarity cues from other polarity types, defects in these pathways can also lead to alterations in tissue organisation, EMT, migration, and stemness (Daulat and Borg, 2017).

Cell polarity genes have impacts on a plethora of signalling pathways important for cancer progression. Due to the interconnection between polarity proteins, alterations in these proteins affect processes including cell organisation, cell migration, and cell division and proliferation. However, many questions still remain. For example, it is not well understood how cell polarity control and cell cycle/proliferation control overlap. Whether the loss of cell proliferation is a secondary effect due to loss of cell polarity, or whether cell polarity

regulators play direct roles in cell cycle progression is still unclear. Defining the boundaries of where each process is a primary or secondary consequence remain to be important distinctions to make.

### 1.2 Protein ubiquitination and degradation

Protein synthesis is a constantly occurring and tightly regulated process within all living cells. In order to maintain a homeostatic environment, protein degradation functions in parallel to remove proteins that have accumulated or are no longer required by the cell through a set of dynamic regulatory mechanisms that incorporate two major pathways. One of these involves the lysosome, which is an organelle enriched with proteases that is the site of degradation for an array of proteins. The formation of small intracellular vesicles, or endosomes that enclose proteins destined for lysosomal processing is required prior to the fusion of these vesicles and their contents to the lysosome. The break-down of these proteins functions as a method of protein recycling, ultimately redistributing these peptides back into the cell for re-use.

In most cases however, the typical mode of protein degradation is via the ubiquitin-proteasome system (UPS), whereby unwanted proteins are covalently tagged with ubiquitin to trigger the degradation of these proteins by the 26S proteasome. This form of protein degradation is reliant on the function of three major components – the proteasome, a ubiquitin ligase complex, and ubiquitin itself. Occurring in a stepwise cascade, the ubiquitin-proteasome system initiates a series of enzymatic reactions in the ubiquitin ligase complex, leading to the attachment of ubiquitin to the target protein. This attachment of ubiquitin (known as ubiquitination) is a form of post-translational modification that essentially flags the substrate for recognition by the proteasome where degradation of the substrate ultimately occurs. Ubiquitin machinery is constantly active, facilitating the turnover of thousands of proteins across all cellular processes including cell cycle, DNA repair and epithelial to mesenchymal transition (EMT). Unsurprisingly, the deregulation of proteostasis has been implicated in a myriad of diseases including the development of many cancers that is typically owing to alterations in the function of E3 ubiquitin ligases (Hu *et al.*, 2011; Wang *et al.*, 2014).

### 1.2.1 The ubiquitin-proteasome system

Various methods of ubiquitin-substrate attachment have been described with each providing a distinct signal. Depending on the number of ubiquitin molecules to be attached, the subsequent cellular functions following ubiquitination can often be determined. Downstream signalling, DNA repair and endocytosis are among the cellular processes linked to monoubiquitylation (single ubiquitin attachment), whereas the degradation of lysosomes and endosomal sorting are linked to multiple monoubiquitylation sites on a protein. These linkages can be built upon by forming polyubiquitin chains to add more complexity to the ubiquitin signalling cascade. For example, some of these processes – like DNA repair and endocytosis – can also be associated with polyubiquitination (Hoeller *et al.*, 2007). Not only is the number of attached ubiquitin important, but so is the specific lysine residue on the ubiquitin protein through which these polyubiquitin chains form. In ubiquitin, there are 7 lysine residues: K63, K48, K11, K6, K27, K29 and K33, and this, together with the varying ubiquitin chain lengths, dictates the future function of the protein. Many of these attachments have been associated with specific downstream processes with the K63 and K48 linked chains forming a large proportion of characterised linkages (reviewed extensively in Komander and Rape, 2012). While Lys-63 ubiquitin chains are often associated with lysosomal degradation and other signalling pathways like kinase activity (Yang, Zhang and Lin, 2010; Grumati and Dikic, 2018), Lys-48 and Lys-11 linkages largely translate into the recognition and degradation of the targeted protein by the 26S proteasome, which will be focus of this study.

Complex ubiquitin attachments occur in a well-orchestrated fashion and central to the ubiquitin-proteasome system are ubiquitin ligase complexes where the process of ubiquitination occurs. It is between the E1, E2 and E3 enzymatic components of this complex that the catalysis of key enzymatic reactions facilitates the conjugation of ubiquitin to a target substrate (reviewed by Hershko and Ciechanover, 1998). The ubiquitination process begins with the activation of ubiquitin via the E1 complex. Also known as the Ubiquitin Activating Enzyme (UAE), this complex hydrolyses adenosine triphosphate (ATP) to a high energy adenosine monophosphate (AMP) in order to promote the formation of a thioester between the active cystine of the E1 at the C-terminal glycine of ubiquitin (Ciechanover *et al.*, 1981). Activated ubiquitin is transferred along the cascade to the E2, otherwise known as the UBC or Ubiquitin Conjugating Enzyme, by forming a thioester bond in a similar manner with the ubiquitin protein. Indeed, these ubiquitin loaded E2s play an integral role in the ubiquitin-proteasome system and in some cases,

ubiquitination can occur directly between the E2 ubiquitin binding domain (UBD) and the substrate – albeit, this process has been largely associated with low specificity monoubiquitylation (Hoeller *et al.*, 2007). Precise ubiquitination activity often requires a third party known as the E3 ubiquitin ligase that coordinates the recognition of a specific substrate prior to ubiquitin transfer from the E2.

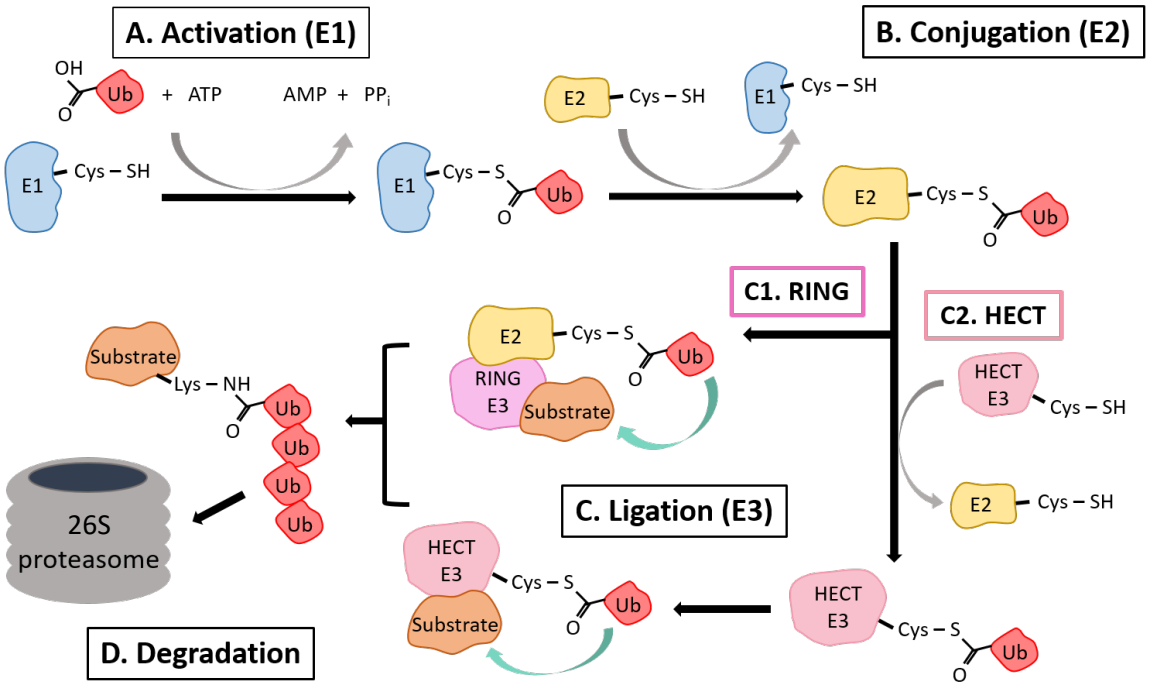



**Figure 1.2 The ubiquitin transfer cascade via RING or HECT E3 ubiquitin ligases complexes**

**A.** At the E1, the first enzymatic step is initiated where ubiquitin is activated. ATP (adenosine triphosphate) hydrolysed AMP (adenosine monophosphate) and the C-terminal of ubiquitin is adenylated. This allows the formation of a thioester between the active cysteine on the E1 and ubiquitin. **B.** A similar reaction to A, another thioester is formed as the ubiquitin is transferred from the E1 to the E2 in a step known as ubiquitin conjugation. **C.** Following conjugation, ubiquitin is ligated to the substrate by one of two methods depending on which E3 is used. **C1.** A RING E3 ligase facilitates direct ubiquitin transfer to the substrate by attaching to both the E2 and the substrate to bring them into close proximity so transfer can occur. **C2.** HECT E3s transfer ligate ubiquitin in a two-step transfer. Like in A and B, ubiquitin is attached to the E3 by forming an intermediate thioester. The E3-ubiquitin conjugate then transfers ubiquitin to the substrate. **D.** Once the substrate is ubiquitinated, the tag is recognised by the 26S proteasome where it is degraded, and ubiquitin is recycled.

### 1.2.2 E3 Ubiquitin ligases

Physical interaction between the ubiquitin charged E2 and the E3 ligase complex is often required for the transfer of ubiquitin to selected substrates. Unsurprisingly, a very diverse range of E3 ubiquitin ligases are required for specificity of substrate ubiquitination. In fact, genome wide analysis of the human genome has identified over 600 genes encoding putative E3 ubiquitin ligases (Li *et al.*, 2008). Unlike both the E1 and E2 protein components that consist of only a single protein, the majority of E3 ubiquitin ligases require multiple subunits to function. The modular composition of the E3 is responsible for a high degree of interchange between each component and for this reason, E3 ligases heavily dictate the ubiquitination activity within the cell.

#### 1.2.2.1 Mechanisms of ubiquitin transfer by different E3 ubiquitin ligases

Classified by a central E3 ligase, there are three major E3 families that are found in eukaryotes. They are distinguished by their respective conserved motifs that are the HECT (Homologous to the E6AP Carboxyl Terminus), the RING (Really Interesting New Gene), the RBR (RING-Between-RING), and the U-box domains (a modified version of the RING motif). Due to the difference in domains, the second key difference between these families resides in the process through which ubiquitin is transferred from an E2 to the substrate. When ubiquitination occurs via a HECT or RBR domain containing ligase, the process is similar to what occurs in the E1 and E2 ubiquitin transfer steps. In a two-step process, the ubiquitin is first transferred to the E3 by forming an intermediate thioester between the E2 conjugated ubiquitin and an active cystine on the E3 before being passed on to the substrate (Huibregtse et al., 1995; Scheffner, Nuber and Huibregtse, 1995) (Figure 1.2, C2). The RING family on the other hand, along with RING-related E3s, are the predominant ubiquitin ligase family and account for the vast majority of E3 ligases in the human genome (Metzger, Hristova and Weissman, 2012). In contrast to the HECT E3s that play an active enzymatic role in ubiquitination, the RING E3 ligases are often made up of a multi-subunit complex, functioning largely as a scaffold. Indeed, although some RING E3 ligases such as c-Cbl have both a catalytic domain and a substrate-recruiting module as a single polypeptide, many others such as BRCA1/BARD1, Cullin–RING ligases (CRLs), and the anaphase-promoting complex (APC), keep these as separate subunits of a multicomponent E3 complex (Zheng and Shabek, 2017). Both RING and U-box domain containing E3s do not require thioester formation – rather, they can associate with the E2 and the substrate

simultaneously to facilitate a direct transfer to the substrate (Budhidarmo, Nakatani and Day, 2012; Berndsen and Wolberger, 2014) (Figure 1.2, C1). From here, these tagged proteins can then be recognised and degraded by the 26S proteasome (Figure 1.2, C and D).

### 1.2.3 The Cullin-RING family of E3 ubiquitin ligases

Being the largest group of E3 ligases has afforded RINGs a considerable amount of variability and therefore, have been further characterised into individual RING subtypes. Among these is the Cullin-RING ligase family that characteristically assembles with other proteins to carry out ubiquitination. One of these is a RING finger protein that links to the C-terminal of the Cullin-RING ligase to act as a docking station for the E2. Adaptor proteins at the opposite N-terminal connect the Cullin-RING ligase to the substrate recognition subunit (SRS), which determines substrate specificity. Collectively, these proteins act in unison to draw the substrate and the E2 into a close enough proximity so that the ubiquitin transfer can occur (Figure 1.3, A). Nine members of the Cullin-RING ligase family are present in the human genome including CUL1, CUL2, CUL3, CUL4A/4B, CUL5, CUL7, PARC and APC2, with each Cullin-RING ligase distinguishable by the adaptor proteins that can bind to its N-terminal (Sarikas, Hartmann and Pan, 2011).

### 1.2.4 Cullin-2 and its various substrate recognition subunits

One of the most well studied Cullin-RING ligases is Cullin-2 (Cul2). This ligase is characterised by the presence a binding site for a substrate recognition subunit, as well as a binding site for Elongin C (Elo-C) – an adaptor protein that subsequently binds to Elongin B (Elo-B). Together, these two proteins link Cul2 to the substrate recognition subunit through two highly conserved specific binding motifs or 'box' domains present on the substrate recognition subunit. The first motif through which a direct attachment between Cul2 and the substrate recognition subunit forms is via the Cullin box, identified by the sequence ϕPXXϕXXXϕ, where ϕ represents a hydrophobic residue. The second point of attachment occurs indirectly via the adaptor proteins Elongin C and Elongin B, which are both associated with to the N-terminal of Cul2. Here, the adaptors bind to the VHL box motif ((S,T,P)LXXX(C,S,A)XXXϕ) (Kamura *et al.*, 2004). This sequence, originally identified in the von Hippel Lindau (VHL) protein, is derived from a stretch of 35 amino

acid residues that have become synonymous with Cul2 substrate recognition partners (Kamura *et al.*, 1998, 2004). Currently, there are nine mammalian substrate recognition subunits that are known to complex with Cul2/Elongin B/Elongin C through the VHL box (or a close variant) (Cai and Yang, 2016a; S. Wang *et al.*, 2016). In addition to VHL, there is also Leucine Rich Repeat protein 1 (LRR-1), FEM-1, PRAME, BAF250, RACK1, and the focus of this thesis, the ZYG11 family of Cullin-2 substrate recognition subunits that include ZYG11A, ZYG11B and ZER1 (Figure 1.3).

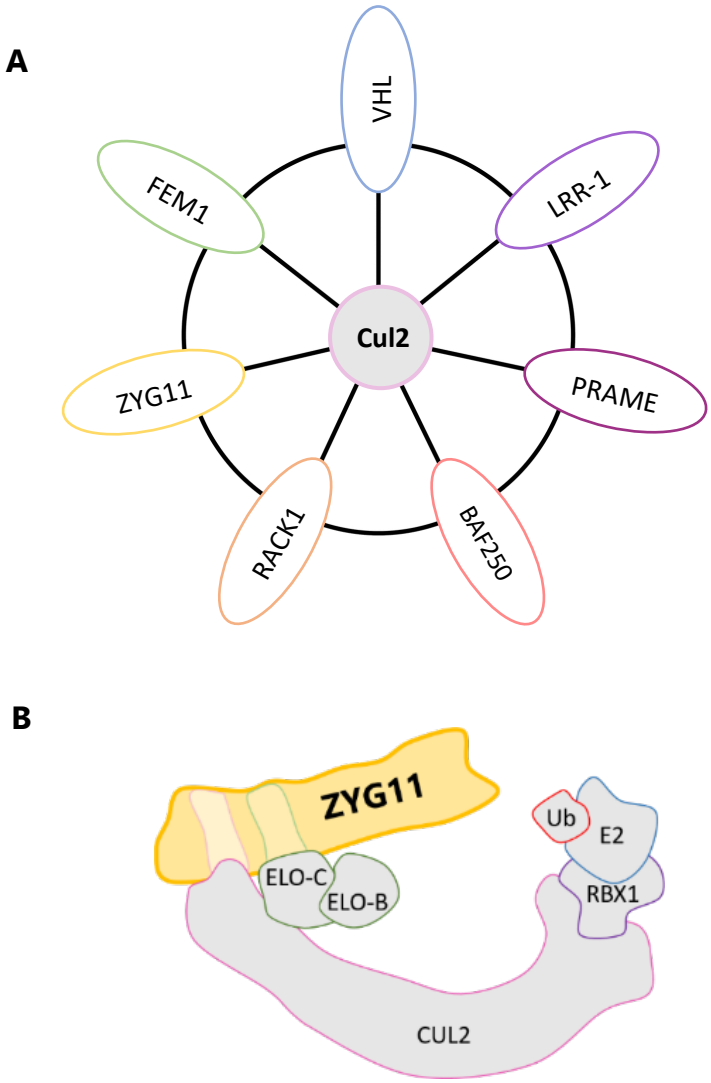



**Figure 1.3 The Cullin-2 complex and the various related substrate recognition subunits**

**A.** Currently identified Cullin-2 substrate recognition subunits (SRS) include von Hippel Lindau (VHL), Leucine Rich Repeat protein 1 (LRR-1), FEM-1, PRAME, BAF250, RACK1 and ZYG11 family members, ZYG11A, ZYG11B and ZER1. **B.** The Cullin-2 complex is composed of various subunits. At the centre is the scaffold CUL2 that is associated with a RING-box protein, most commonly RBX1, at the C-terminal. This serves

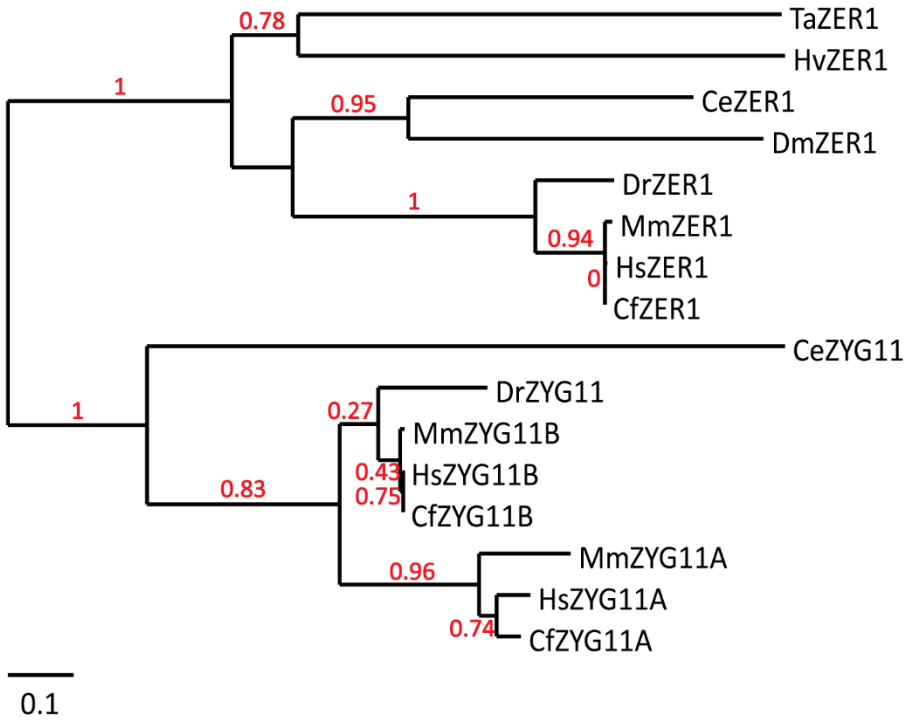
a docking station for a ubiquitin charged E2. The N-terminal is linked to the adaptor proteins Elongin B/C via a VHL box motif whilst a CUL2 box on an SRS provides a direct point of contact with CUL2. The SRS is a variable component that acts as the interface between the substrate and the complex, allowing it to be brought into close proximity of the E2 for ubiquitin transfer.

## 1.3 The ZYG11 family

### 1.3.1 Evolutionary conservation of the ZYG11 family

The evolutionary conservation of genes across a number of species is often an indicator that a gene fulfils a specific and fundamental role, especially in cases where certain motifs are present in all orthologues. The ZYG11 family is a set of E3 ubiquitin ligases that are widespread across metazoans, present from humans to the fly, to even one of the simplest and most ancient multicellular organisms like the placozoan, *Trichoplax adhaerens* (Figure 1.4). The number of ZYG11 family members can vary between organisms. In lower organisms for example, the nematode *C. elegans* (roundworm) and *D. rerio* (zebrafish) contain two family members, Zyg11 and Zer1. A couple of exceptions to this are *Trichoplax adhaerens* and *D. melanogaster* (vinegar fly), which encode only a single orthologue, *zer1*. Whether this is indicative of a ZER1 based ancestral line, or simply the loss of the ZYG11 lineage in insects and placozoans is still unclear. Complex mammalian genomes harbour three ZYG11 family members, two of which are the result of a recent gene duplication in the ZYG11 branch. ZER1 is also present in mammalian systems, although it is sometimes referred to as ZYG11B-like (ZYG11BL). The designation of ZER1 as ZYG11B-like rather than ZYG11A-like was likely to be based on sequence similarity as until recently, there were no functional similarities between the two genes. In fact, what is known about ZYG11 family function is based largely on cell cycle and embryonic polarity defects identified in the *C. elegans* model where ZYG-11 is known to regulate these processes via its E3 ubiquitin ligase activity.

A


0.78
TaZER1
HvZER1
1
0.95
CeZER1
DmZER1
1
DrZER1
0.94
MmZER1
0
HsZER1
CfZER1
CeZYG11
1
0.27
DrZYG11
MmZYG11B
0.43
HsZYG11B
0.83
0.75
CfZYG11B
0.96
MmZYG11A
HsZYG11A
0.74
CfZYG11A
0.1

B

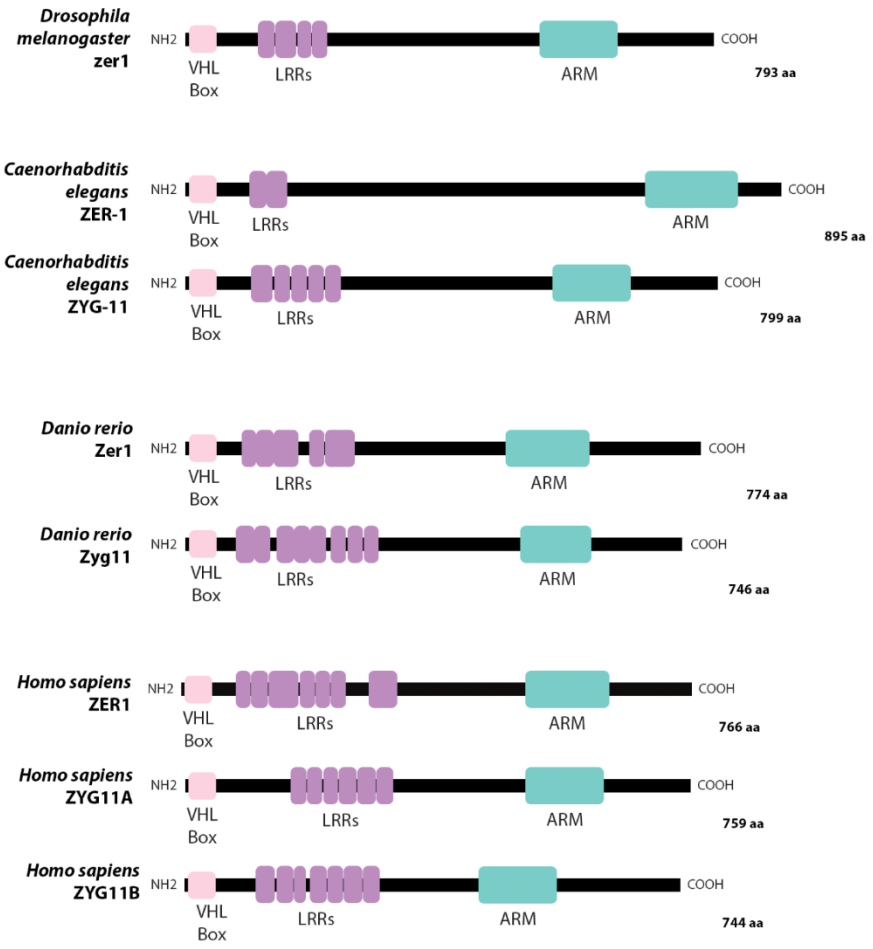

Drosophila melanogaster zer1
NH2
VHL Box
LRRs
ARM
COOH
793 aa
Caenorhabditis elegans ZER-1
895 aa
Caenorhabditis elegans ZYG-11
799 aa
Danio rerio Zer1
774 aa
Danio rerio Zyg11
746 aa
Homo sapiens ZER1
766 aa
Homo sapiens ZYG11A
759 aa
Homo sapiens ZYG11B
744 aa

**Figure 1.4 The ZYG11 family is highly conserved in multicellular organisms**

**A.** Phylogenetic tree representing the evolution of ZYG11 family members across a number of metazoans: *Trichoplax adhaerens* (Ta), *Hydra vulgaris* (Hv), *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Danio rerio* (Dr), *Mus musculus* (Mm), *Canis familiaris* (Cf), and *Homo sapiens* (Hs). Bootstrap values are displayed in red. The scale bar represents the mean number of amino acid substitutions. There are two main clades – ZER1 and ZYG11 where ZYG11 is further subdivided into the ZYG11A and the ZYG11B branches in mammalian systems. **B.** Human ZYG11 family members all possess a VHL box, an LRR domain and an ARM domain.

### 1.3.2 The ZYG11 family in invertebrate organisms

#### 1.3.2.1 The ZYG11 family in *C. elegans* embryonic cell cycle

The presence of the ZYG11 family is spread across a number of multicellular organisms but has been described primarily in *C. elegans* where the focus has been on ZYG-11 rather than ZER-1. First described in 1976 by David Hirsh and Rebecca Vanderslice, the *zyg-11* gene was identified as a temperature sensitive mutant that when lost, resulted in fertilised worm eggs that failed to yield viable progeny and showed cleavage defects at the single cell stage. Phenotypically, this embryonic lethal outcome is referred to as 'zygote defective', hence 'ZYG'-11 (as it was the 11$^{th}$ mutant that displayed this outcome) (Hirsh and Vanderslice, 1976). The Hirsh group further characterised ZYG-11, reporting that embryonic *zyg-11* mutants exhibit a symmetric first cleavage as opposed to the typical asymmetric division (Wood *et al.*, 1980). Mutant *zyg-11* embryos have also been shown to exhibit a slew of additional phenotypes including delayed metaphase during meiosis II, a lack of polar body II extrusion, plasma membrane blebbing, DNA fragmentation, variable nuclei numbers, irregular cell division and P granule mislocalisation – all of which can be rescued by *zyg-11* cDNA. However, despite the strong maternal effects of *zyg-11* loss, the transcription of this gene is not restricted to the female germ-line (Kemphues *et al.*, 1986; Carter, Roos and Kemphues, 1990).

The mechanisms through which ZYG-11 regulates cell cycle, cell differentiation and embryonic cell polarity have been linked to its role in protein ubiquitination, with early indications of ZYG-11 functioning as a substrate recognition subunit discovered again in *C. elegans*. Here, the *zyg-11* mutants shared phenotypes with *cul-2* that were reminiscent of those previously reported, including meiosis II delay and arrest, multinuclei, and cytoplasmic extensions during mitosis (Kemphues *et al.*, 1986; Feng *et al.*, 1999). Most importantly however, was the requirement of both CUL-2 and ZYG-11 for the degradation

of the *C. elegans* cyclin B protein (CYB-1) prior to metaphase II in order for the cell cycle to progress normally (Liu, Vasudevan and Kipreos, 2004). This process was likely to occur via the CUL-2$^{ZYG-11}$ E3 ubiquitin ligase complex, which was further supported by the established physical interaction between the ZYG-11 VHL box and CUL-2/Elongin C, as well as a range of similar phenotypes indicating that these proteins might function in the same pathway (Vasudevan, Starostina and Kipreos, 2007). Moreover, many of these cell cycle defects caused by ZYG-11 loss can be rescued by CYB-1 downregulation (Balachandran *et al.*, 2016). Degradation of CYB-1 is required for the transition from metaphase to anaphase and is the result of ubiquitination through the Anaphase Promoting Complex/Cyclosome (APC/C) in mitotic cells. However, in meiosis, which occurs in two stages during the development of the *C. elegans* embryo, APC/C has only been shown to be important during this transition in meiosis I, with CUL-2$^{ZYG-11}$ filling this role during meiosis II. Indeed, APC/C mediated degradation of CYB-1 remains to be the canonical pathway in mitotic cells, however recent studies have found that ZYG-11 is also able to degrade Cyclin B during mammalian mitosis as a backup mechanism, albeit redundantly in normal situations (Balachandran *et al.*, 2016).

In a mechanism unrelated to CYB-1 degradation, it has also been proposed that ZYG-11 is required degrade a group of novel proteins involved in meiosis II progression known as Meiosis-to-Mitosis transition defect (MEMI) proteins, MEMI-1, -2, and -3. A lack of polar body II extrusion, defects in meiosis II chromatids and irregular mitotic cell divisions in the early embryonic cells are among the phenotypes present in both *memi* gain-of-function embryos and *zyg-11* mutant embryos (Ataeian *et al.*, 2016). The similarities in phenotypes support the idea that ZYG-11 functions as a substrate recognition subunit, considering that upon the loss of ZYG-11 its target substrates (in this case, MEMI proteins) are likely to accumulate. However, the MEMI proteins are yet to be validated as bona fide targets of ZYG-11 mediated degradation.

#### 1.3.2.2 The ZYG11 family in *C. elegans* embryonic polarity establishment

Another role of ZYG-11 is in the establishment of embryonic anterior-posterior polarity. At meiosis II, PAR proteins and P granules are usually evenly distributed throughout the cytoplasm and the cell cortex of the zygote, only acquiring asymmetry at the transition between meiosis and the first mitotic cell cycle (Sonneville and Gonczy, 2004). Accordingly, polarised PAR proteins and mislocalised P granules during meiosis II, which

are both seen in *zyg-11* mutants, indicate defective cell polarity regulation (Kemphues *et al.*, 1986; Sonneville and Gonczy, 2004). Important in the establishment of anterior-posterior embryonic polarity, P granules, PAR-1 and PAR-2 normally localise to the posterior cortex of the embryo whereas PAR-3 and PAR-6 localise to the anterior cortex. When there is a loss of *zyg-11* or *cul-2*, the onset of polarity is not only premature, but it is also inverted. Beginning at meiosis II and sustained to meiosis exit, this inversion of anterior-posterior polarity occurs independently to the previously described cell cycle defects, despite the coinciding timeframes (Sonneville and Gonczy, 2004). Relatedly, this inverted polarity is similarly seen at metaphase I when APC/C is lost, suggesting that there may a degree of overlap between the substrates of the two proteins beyond CYB-1 (Wallenfang and Seydoux, 2000). Other than P granule mislocalisation, severe cortical membrane ruffling is another polarity and cell shape related phenotype that is separate from CYB-1 degradation and that no further studies have characterised (Vasudevan, Starostina and Kipreos, 2007). The literature indicates that ZYG-11 and CUL-2 are able to dictate polarity cues as a function separate from cell cycle control, at least in the developing *C. elegans* embryo, yet how exactly this is achieved remains unclear. The function of the ZYG11 family member, ZER-1 in *C. elegans* is currently unknown.

#### 1.3.2.3 The ZYG11 family in *Drosophila*

Unlike in the *C. elegans* model where ZYG-11 has been the primary focus, the *Drosophila* ZYG11 family only contains a single orthologue that is more closely related to ZER1. Therefore, I will use *zer1* when referring to the *Drosophila* gene orthologue. To date there have been no studies directly addressing the function of *zer1*, although there are a handful of poorly characterised genetic interactions that have been reported in the context of larger genetic screens. For example, the overexpression of *zer1* has been shown to enhance thoracic development defects caused by alterations in *Drosophila* dpp/BMP signalling (transcription factor *pannier*), and external sensory organ development defects caused by alterations in Notch signalling (*notch* and its antagonist *hairless*) (Abdelilah-Seyfried *et al.*, 2000; Peña-Rangel, Rodriguez and Riesgo-Escovar, 2002). In the developing eye, the overexpression of *zer1* slightly reduces eye size, which is rescued by the overexpression of *cyclin E* (Tseng and Hariharan, 2002). For all of these interactions the precise mechanisms have not been further characterised, and this is also the case for its known protein-protein interactions. Indeed, these implicate *zer1* in a handful of developmental processes, albeit only in the context of overexpression. Currently, the effects of *zer1* loss

in the fly have not been studied, further highlighting a lack in basic research on the relevance of this gene family in *in vivo* development.

### 1.3.3 The ZYG11 family in mammalian systems

The mammalian function of the ZYG11 family remains poorly understood in comparison with its *C. elegans* counterpart. All of the *C. elegans* studies that have been translated into mammalian systems have been conducted in human cell lines, heavily focusing on cell cycle. Despite the similarity between ZYG11 proteins, most current studies have concentrated on only one or two family members, particularly dividing the family along the ZYG11 and ZER1 branches rather than studying the family as a whole. Understanding whether each family member acts separately or whether they show some degree of similarity remains to be a largely unanswered question.

#### 1.3.3.1 The function of the ZYG11 family in cell cycle regulation

Building on studies where *C. elegans* mutant *zyg-11* phenotypes were found to be rescued by CYB-1 expression (described in 1.3.2.1), it was shown that mammalian ZYG11A/B can regulate G2 to Mitosis (G2-M) phase transition in the osteosarcoma line, U2OS (Balachandran *et al.*, 2016). The progression from G2 to M phase is initiated by the activation of Cyclin B1, which remains active throughout the start of M phase, known as prophase. During this transition, Cyclin B1-Cdk1 phosphorylates and inactivates Separase, which is also inhibited by Securin activity. Together, this prevents Separase from cleaving Cohesin, which is required to prevent sister chromatid separation during prophase. During this phase, the E3 ligase APC/$C^{Cdc20}$ complex is inhibited by the mitotic checkpoint complex (MCC) (Figure 1.5, SAC active). At the onset of metaphase and once the spindle assembly checkpoints (SAC) have been achieved, Cyclin B1 and Securin are ubiquitinated and degraded by APC/$C^{Cdc20}$ (Figure 1.5, SAC inactive). This activates Separase, which can then cleave Cohesin and allow the separation of sister chromatids (Takizawa and Morgan, 2000). Functioning in a parallel and independent pathway, ZYG11A and ZYG11B are also capable of targeting and degrading Cyclin B1 in a ubiquitin-proteasome dependent manner, facilitating mitotic slippage. Mitotic slippage refers to the progression of metaphase and the rest of the cell cycle despite the proper spindle assembly checkpoints (Figure 1.5, SAC continuously active), resulting in the improper segregation of sister

chromatids. This occurs when APC/C is constitutively inhibited. Here, ZYG11A/B have been found to degrade Cyclin B1 to permit cell cycle progression and the cell cycle exit of tetraploid cells, i.e. mitotic slippage (Balachandran *et al.*, 2016; Brandeis, 2016).

In humans, ZYG11A and ZYG11B are generally investigated as a pair, however both have demonstrated the capacity to function separately. *In vitro* knockdown of *ZYG11A* using the non-small cell cancer (NSCLC) cell lines, H1299 and SPC-A1, has displayed various phenotypes such as a decrease in cell proliferation, migration, invasion and colony forming ability. An increase in G1 cells and an increase in Cyclin E1 has also been reported. Moreover, many of these phenotypes were mirrored in mice injected with the same cell lines harbouring *ZYG11A* knockdown. These mice developed smaller tumours which were accompanied by a reduction in tumour proliferation and reduced expression of Cyclin E1 (X. Wang *et al.*, 2016). Cyclin E1 is an essential cell cycle regulator that is activated at the G1-S phase transition (Sherr and Roberts, 1995). However, the experiments reported by Wang et al (2016) did not clarify whether there is a direct relationship between ZYG11A and Cyclin E protein turnover.

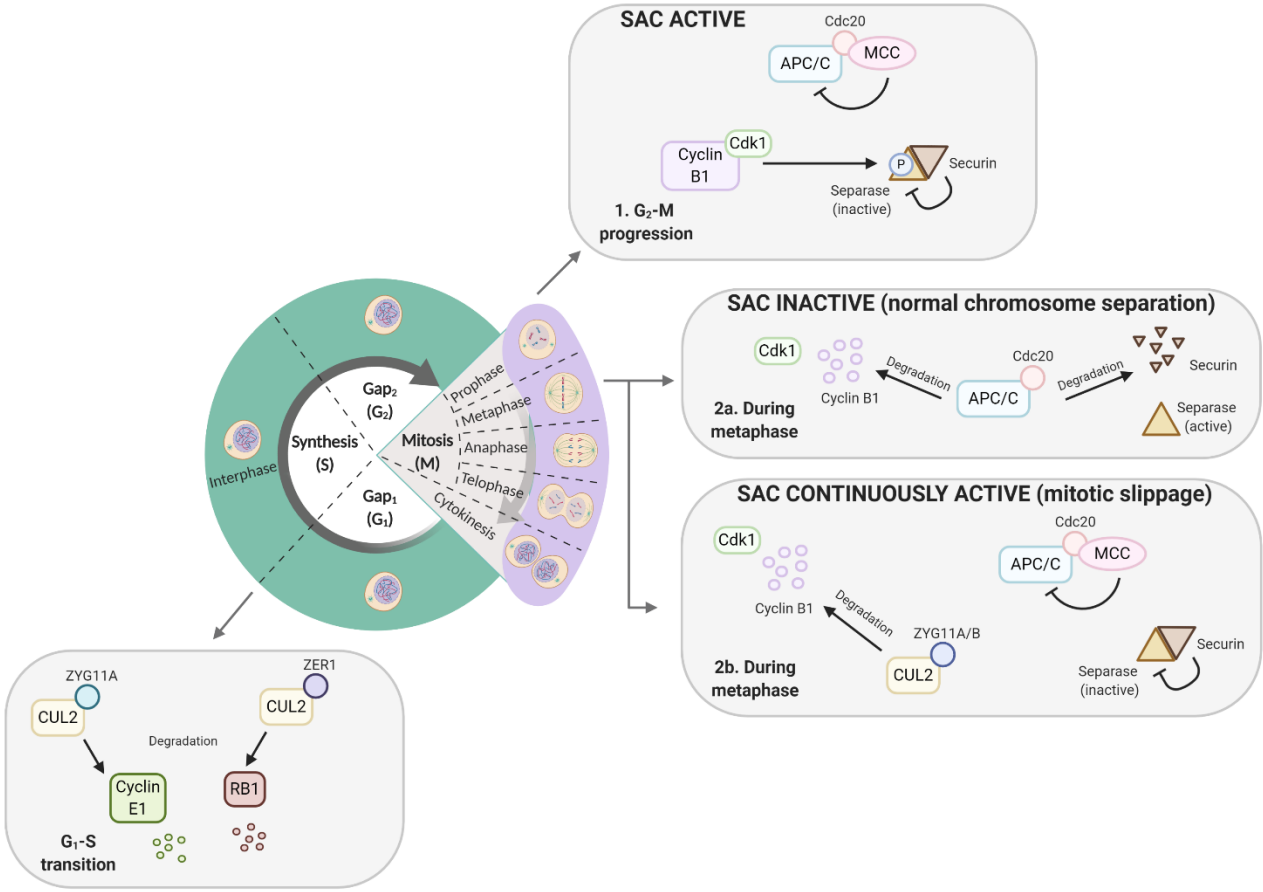



**Figure 1.5 Known functions of the ZYG11 family in mammalian cell cycle regulation**

At the G2 to M phase transition, ZYG11A/B degrade Cyclin B1 in an APC/C independent manner. Under APC/C inhibition, the normally redundant role of ZYG11A/B allows for mitotic cell cycle progression regardless of proper chromosomal segregation, resulting tetraploid cells (mitotic slippage). At the G1 to S phase transition, ZYG11A has been linked to Cyclin E1 degradation and ZER1 has been found to be required for the destabilisation of RB1 in HPV infected cells.

The spotlight has typically been placed on the ZYG11 branch of the family, both in nematode and mammalian studies, despite the similarities between the ZER1 and ZYG11A/B protein domains. In Human Papilloma Viruses (HPV), the HPV16 E7 subtype is known to associate with the Cullin-2 E3 ubiquitin ligase complex through a specific interaction that is facilitated by ZER1. HPV16 E7 expressing cells show the destabilisation of Retinoblastoma tumour suppressor-1 (RB1), a cell cycle regulator of G1 to S phase transition. This protein is phosphorylated and deactivated during G1 to allow S phase progression (Dyson *et al.*, 1989; Munger *et al.*, 1989). Here, ZER1 and Cullin-2 are required for the degradation of RB1, also implicating ZER1 in cell cycle regulation, albeit only in a viral context (White *et al.*, 2012) (Figure 1.5).

#### 1.3.3.2 The regulation of cell polarity by E3 ubiquitin ligases

The signalling pathways in epithelial cell polarity are tightly regulated in order to maintain normal tissue architecture. While each polarity complex fulfils a defined role in terms of their localisation and subsequent scaffolding properties, many downstream interactions are constantly being identified. Moreover, how this system is regulated both up- and downstream of these central polarity complexes in terms of ubiquitin ligase activity and protein turnover is still unclear. Apicobasal polarity regulates cell-cell adhesion and cell-ECM adhesion through coordinated movements of intracellular polarity proteins (as described above). Of the ubiquitin ligases that have been implicated in polarity regulation, many of those found to be relevant in mammalian systems have been linked to the activity of Human Papillomavirus (HPV) viral oncoproteins, specifically HPV E6 or E7 proteins. These include the targeted degradation of tumour suppressors like Scribble, Dlg, MAGI-1,-2 and -3, and Par3, as well as other important polarity and signalling proteins including Patj, PSD95, TIP-2 and MUPP1 (reviewed by Thomas *et al.*, 2008).

Outside of human viral infection models, one of the more well characterized E3 ubiquitin ligases in cell polarity regulation is Smurf1, which targets multiple polarity regulators for proteasomal degradation. Talin is an Actin and β-integrin binding protein whose phosphorylation by Cdk5 is required for cell migration, focal adhesion disassembly and lamellipodia stabilisation. However, without this phosphorylation, Talin is turned over by Smurf1 leading to alterations in focal adhesion formation and cell migration (Huang *et al.*, 2009). Similarly, the β-integrin coactivator Kindlin-2 is also ubiquitinated and degraded by Smurf1, further illustrating how this E3 ligase is able to control focal adhesion

formation (Wei *et al.*, 2017). Another important target for Smurf1 ubiquitination and degradation is the GTPase RhoA (see section 1.1.1.2). The effects of RhoA degradation are widely distributed within the cell and can have consequences on both apicobasal polarity and front-rear polarity regulation. This is because RhoA affects cytoskeletal rearrangements which in turn regulate tight junction integrity, EMT, and cell migration (Wang et al., 2003, 2006; Ozdamar et al., 2005; Tian et al., 2011). Smurf2 also affects apicobasal polarity through the targeting of EphrinB to inactivate the Par complex (Hwang *et al.*, 2013). The diversity of Smurfs in cell polarity are also reflected in the developing mouse cochlea during PCP development. Here, Smurf1/2 are recruited by Par6 and Dishevelled to degrade the planar cell polarity protein, Prickle1, and the loss of this process causes defects in neural tube closure (Narimatsu *et al.*, 2009). The E3 ligase for another integrin, β1 has also been identified as Cbl (Kaabeche *et al.*, 2005), further demonstrating how E3 ligases affect cell-ECM interactions. Cell-cell interactions can also be regulated by E3 ligases. One example is β-catenin, an adherens junction protein, which during canonical Wnt signalling is degraded by the APC E3 ligase complex (Stamos and Weis, 2013). Depending on the context and cell type, it is tightly regulated by different ligases including Skp1-Cul1-FWD1 in transformed kidney and colon cells, and $Cul5^{Ozz}$ in differentiating mouse muscles (Kitagawa *et al.*, 1999; Winston *et al.*, 1999). Tight junction protein, E-cadherin has also been linked to two E3 ligases, Hakai and MDM2 (Yang *et al.*, 2006; Nejsum and Nelson, 2007). In invertebrates, the *Drosophila* Crumbs complex protein Stardust (Pals1 in mammals) is degraded by the E3 ligase, Neuralised. Through this association, another apical polarity protein, Crumbs is directed towards the endosome for degradation (Perez-Mockus *et al.*, 2017). The Par6 complex is also targeted for degradation by Skp1-Cul1-Slmb in fly oocyte follicle cells (Eurico-de-Sá *et al.*, 2014). This same complex also acts to restrict aPKC to the apex of fly epithelia, acting in parallel with Scribble mediated antagonism of Par/aPKC proteins (Skwarek *et al.*, 2014).

Indeed, ubiquitin ligases in cell polarity are not strictly limited to the degradation pathway (as just described). In the developing mouse cochlea, the Nedd4 and Nedd4-2 E3 ligases are required for the monoubiquitination of Dlg3, where the attachment of ubiquitin is required to recruit Dlg3 to tight junctions (Van Campenhout *et al.*, 2011). PCP proteins also rely on non-degrative forms of ubiquitination. E3s like Nrdp1 and Huwe1 ubiquitinate Dishevelled to either recruit it to plasma membrane or to regulate Wnt signalling by oligomerisation (De Groot *et al.*, 2014; Wald *et al.*, 2017).

While protein ubiquitination is clearly important across several epithelial cell polarity processes, this picture is far from complete. Many ubiquitin targets in cell polarity have only been validated in invertebrate systems and there are still a number of polarity proteins whose protein turnovers have not yet been described. Therefore, how the proteome is manipulated to regulate these polarity proteins remains to be an important question in mammalian systems.

#### 1.3.3.3 The ZYG11 family in mammalian cell shape and polarity

Despite compelling evidence in *C. elegans* that ZYG-11 is important for embryonic polarity establishment, no studies have directly looked at this role in mammalian systems. Observations in our lab have emphasised a strong link between Scribble and β-catenin localisation with a study by Smith et al. revealing only a small subset of gene knockdowns that are able to mislocalise these two proteins away from the epithelial membrane (Smith *et al.*, 2016). Alongside MYH9 and IGF1R, ZYG11A was one of these proteins. Here, the loss of ZYG11A induced severe morphological alterations and the loss of contact inhibition (Lorey Smith, our unpublished findings). In addition, ZYG11A was also identified in a $Ras^{V12}$ bypass screen. In normal mammalian epithelial cells, the activation of oncogenic Ras ($Ras^{V12}$) causes cell anchorage-independent growth that can be suppressed by the overexpression of the polarity protein, Scribble. This suppression by Scribble can be prevented by the loss of ZYG11A, suggesting that it is important for Scribble tumour suppression activity (Lorey Smith, our unpublished findings). Although preliminary, these findings implicate ZYG11A in mammalian cell shape regulation for the first time. Determining the relationship between the entire ZYG11 family and mammalian cell morphology will be essential in defining the complex regulatory mechanisms on the polarity protein Scribble, as well as the broader polarity regulatory network.

## 1.4 Summary and aims of the thesis

In metazoans, the ZYG11 family is a highly conserved E3 ubiquitin ligase family with varying numbers of homologues between species. There is a fundamental requirement for ZYG-11 in *C. elegans* development due in part to its ability to target CYB-1 for degradation during meiotic cell cycle progression. In mammalian cell lines, ZYG11A/B are similarly able to target Cyclin B1, and each family member has also been linked to cell cycle progression. In the nematode, ZYG-11 is also essential for the establishment of embryonic polarity, although the mechanisms for this remain poorly defined. Given ZYG11A's genetic interaction with the polarity protein, Scribble, and its requirement for the maintenance of normal cell shape, it is likely that its role in polarity regulation is conserved in mammalian systems. Taken together I hypothesise that ZYG11A and potentially ZYG11B and ZER1, are novel regulators of cell shape and polarity in mammalian epithelia, and by virtue may influence the initiation and progression of cancer. At the beginning of this study, the three mammalian ZYG11 family members had not been investigated in parallel, nor had the relevance of this E3 ubiquitin ligase family in other *in vivo* models been tested.

In this thesis, the overarching questions that I have investigated are: 1. Are the individual human ZYG11 family members phenotypically and functionally alike, specifically in terms of cell shape and polarity regulation? 2. Do other *in vivo* models display the same requirement for the ZYG11 family during development? 3. What are the mechanisms behind the cell shape alterations following ZYG11A loss?

Chapter 3 provides a comprehensive and comparative study of the ZYG11 family across species and focuses on how the loss of individual ZYG11 family members affect normal cell function in human epithelial cells. It also uses two *in vivo* models to assess the developmental requirement for the ZYG11 family in vertebrate (*D. rerio*) and invertebrate (*D. melanogaster*) systems. Chapter 4 probes the ZYG11A genetic and protein interaction network to reveal novel ubiquitination targets and pathways that ZYG11A may be involved in regulating.

This work has highlighted divergent roles for each ZYG11 family member, not only in between family members, but also between species. It has also linked ZYG11A to the regulation of a number of important signalling pathways in cell shape and polarity as well as early cancer development.

# Chapter 2

## Materials and Methods

## 2.1 Drosophila maintenance and genetic manipulation

### 2.1.1 Fly maintenance and genetic manipulation

*Drosophila* stocks were routinely maintained at either ambient room temperature or 18˚C on standard *Drosophila* media (see Table 2.1). Genetic crosses were propagated at either 25˚C or 29˚C depending on the stock used and level of gene expression required.

### 2.1.2 Drosophila strains

The *Drosophila* lines used for all genetic crosses as well as the generation of new strains are listed in Table 2.2.

### 2.1.3 Imaging of *Drosophila* eye and wing *zer1* knockdown phenotypes

The crossed *Drosophila* adults were collected between 8-10 days after crossing and scored for any wing or eye defects. Examples of each wing phenotype was captured by first detaching the wing from the torso and mounting onto a glass slide with methyl salicylate and Canada Balsam (1:1 ratio) (Sigma Aldrich, St. Louis, MO, USA). The wings were scored based on the presence or absence of an ectopic vein or signs necrotic wing darkening. The number of flies scored and analysed are stated in the corresponding figure legends. Adult eyes were similarly collected and imaged. Scoring was based on the presence or absence of an enhanced rough eye phenotype.

All images were captured using an Olympus SZX7 microscope (2.5X magnification) on the INFINITY camera and software system. The eye size was calculated using the pixel area by Adobe Photoshop.

## 2.2 Mammalian cell culture

### 2.2.1 Cell maintenance

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) media with 10% foetal bovine serum and 15 mM HEPES. MCF10A cells were maintained in DMEM/F-12 supplemented with 5% donor horse serum (Life Technologies), 20 ng/ml

epidermal growth factor (Peprotech, Rehovot, Israel), 100 ng/ml Cholera toxin (List Biological Laboratories, Inc., CA, USA), 0.5 µg/ml hydrocortisone, 10 µg/ml insulin and 15 nM HEPES. For proteasome inhibition, the proteasome and calpain inhibitor MG132 (Selleck Chemicals, Houston, TX, USA) was supplemented into the media at the indicated concentrations and time points. For leptomycin B (LMB) (Cayman Chemical Company, MI, USA) treatment, cells were treated with 25 nM for 6 hours prior to imaging. All cells were maintained at 37˚C with 5% $CO_2$.

### 2.2.2 Genetic silencing by reverse siRNA transfection

Knockdown of ZYG11 family members was achieved using commercially available small interfering RNA (siRNA) multiplexes from the Dharmacon siGENOME range (GE Healthcare Dharmacon, Inc., CO, USA). Following the optimisation of transfection efficiency, MCF10A and HEK293T cells were seeded at 1.1 x $10^4$ cells per $cm^2$ in a 12-well plate. Per well, 1.5 µl of Dharmafect 3 (for MCF10As) or Dharmafect 1 (for HEK293Ts) (GE Healthcare Dharmacon, Inc. CO, USA) and 193.5 µl of Opti-MEM I (Life Technologies, CA, USA) were vortexed and allowed to complex for 5 minutes at room temperature. siRNA was added, such that when the cells were added the final concentration would be 25 nM. The mixture was vortexed briefly, incubated for 20 minutes at room temperature, then added to the seeded cells. Ensuring that the cells were dispersed evenly, they were then incubated for 24 hours before the media was refreshed and allowed to grow. The endpoint for most morphological and protein analysis was at 72 hours post-transfection, whereas functional assays were conducted at 48 hours post-transfection. The siRNA sequences used for 12-well transfections are listed in Table 2.5.

### 2.2.3 Immunofluorescent staining and imaging

Cells were fixed for 10 minutes in 4% PFA then permeabilised for 10 minutes in 0.3% Triton-X 100 in PBS, followed by blocking in 3% BSA in PBS for 1 hour. Primary antibodies against Scribble and β-catenin were then diluted in 3% BSA in PBS (see concentrations in Table 2.3). Primary antibodies were incubated with the cells overnight at 4˚C. Plates were then washed 2 x in 0.01% Tween20 in PBS (PBST) and 1 x in PBS prior to the manual addition of secondary antibodies. Following a 1-hour incubation, cells were washed as above and stained with the nuclear dye DAPI for 10 minutes. Cells were

imaged on a Zeiss LSM 780 laser scanning confocal. Nuclear to cytoplasmic analysis in localisation studies were performed using the ImageJ Image Analysis Software (Intensity_Ratio_Nuclei_Cytoplasm.ijm) (Rueden *et al.*, 2017).

### 2.2.4 Generation of stable cell lines via transfection and viral transduction

Transfection of 293T cells was performed using the calcium phosphate precipitation method. 24 hours prior to transfection, 293T cells were plated at a density of 1.7 x $10^4$ cells per $cm^2$. To generate stable cell lines, either a retroviral or lentiviral vector was used. For retrovirus, transfer DNA (10 μg) and the amphotropic packaging vector RD114 was combined with $CaCl_2$ (0.25 M final) and added dropwise to an equal volume of Hepes buffered saline (0.3 M NaCl, 0.06 M Hepes, 0.45 mM $Na_2HPO_4$, pH 7.10), with gentle agitation. Lentiviral transfection was achieved in a similar manner by replacing the RD114 vector for the packaging vector ΔCMV (10 μg) and a human lentiviral envelope vector (10 μg). The mixture was incubated for 15 minutes at room temperature, then dropped evenly over each 293T plate. At 16 hours post-transfection, the media was replaced with fresh MCF10A media. Concurrent with the 293T media change, MCF10A cells were seeded at 7.5 x $10^4$ cells/well in a 6-well plate and allowed to settle over 24 hours. For transduction, the virus-containing supernatant was 0.45 μm filter sterilised and supplemented with polybrene (8 μg/ml final) and EGF (20 μg/ml final). Media on the previously plated MCF10A cells was replaced with the viral supernatant before centrifugation at 1000 rcf, 37˚C for 45 minutes. Fresh viral supernatant was applied at the 8-hours and 16-hours post-transfection timepoint with spin incubation for retroviral cultures. A single viral spin was applied for lentivirus samples. Media was refreshed to normal media 8 hours post final viral spin and incubated for at least 24 hours to allow the cells to recover. Selection of transduced cells by fluorescence-activated cell sorting (FACS) was conducted on an Aria III for all MSCV (GFP), FgH1UTG (GFP) and Cas9 (mCherry) cells.

## 2.3 Nucleic acid

### 2.3.1 Generation of fusion protein plasmids and site mutants

All vectors with the pcDNA3.1 backbone that contained either ZYG11A-GFP, GFP-ZYG11A, GFP-ZYG11B and GFP-ZER1 fusion proteins were commercially synthesised by GenScript (Piscataway, NJ, USA). The pcDNA3.1 myc-BioID-MCS vector

(Kim *et al.*, 2016) was purchased from Addgene (MA, USA). Fusion proteins for ZYG11A, ZYG11B and ZER1 including N- and C-terminal truncations were subcloned into MSCV-IRES-GFP and MSCV-myc-BioID-IRES-GFP vectors by Gibson Assembly® (New England Biolabs, Ipswich, MA, USA). All GFP-ZYG11A full length site mutants were generated using the NEB Site Directed Mutagenesis kit® (New England Biolabs, Ipswich, MA, USA) as per the manufacturer's instructions.

### 2.3.2 Bacterial transformation

Chemically competent XL1 Blue bacterial cells were transformed using KCM heat shock on ice for 15 minutes, then at room temperature for a further 15 minutes. Transformed bacteria were selected overnight on Luria Broth agar plates containing Ampicillin (100 μg/ml) and colonies were screened for positive plasmids using small-scale plasmid preparations and restriction digests. Positive clones were further validated by service Sanger Sequencing by the AGRF (Melbourne, Australia).

### 2.3.3 DNA plasmid purification minipreps

Plasmid DNA was purified from bacterial colonies using the Wizard® Plus SV Minipreps DNA Purification System as per the manufacturer's instructions (Promega, WI, USA). To identify positive clones each plasmid preparation was restriction digested with one or a combination of HpaI, EcoRI or NotI (New England Biolabs, Ipswich, MA, USA or Promega, WI, USA). Digested plasmid DNA was visualised by gel electrophoresis.

### 2.3.4 DNA plasmid purification midi- and maxipreps

Large scale plasmid DNA purification from 250 mL overnight bacterial cultures were conducted according to the manufacturer's instructions from the PureLink™ HiPure Plasmid Filter Midiprep or Maxiprep Kits (Invitrogen, CA, USA). Plasmid DNA quality and concentration was determined by a Nanodrop UV-Vis spectrophotometer.

### 2.3.5 RNA isolation and quantitative real-time PCR

Adherent cells were scraped into TRIzol reagent and RNA was isolated according to the manufacturer's instructions (Invitrogen, CA, USA). RNA pellets were resuspended in 20 µl of diethyl pyrocarbonate (Sigma Aldrich, St. Louis, MO, USA) treated water. Concentration and quality of RNA was determined by Nanodrop. Removal of genomic DNA and the synthesis of cDNA was achieved using the QuantiTect Reverse Transcription Kit following the manufacturer's instructions (Qiagen, Hilden, Germany). For quantitative analysis of mRNA expression, cDNA was normalised to 10 ng/µl and amplified in triplicate wells using the Fast SYBR® Green Master Mix (Applied Biosystems, CA, USA). All primers used for RT-PCR are listed in Table 2.4. Each experiment was performed using samples from duplicate siRNA wells and RT-PCR was conducted in triplicate wells for each sample. The mRNA expression levels were calculated and normalised relative to *GAPDH* using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

### 2.3.6 CRISPR-Cas9 sgRNA generation and knockout validation

*ZYG11A* knockout lines were generated using a two-vector system with one vector containing the Cas9 (with an mCherry reporter) and another vector containing the sgRNA (with a GFP reporter) (Aubrey *et al.*, 2015). Guide arms targeted towards the start of *ZYG11A* with low off targets were selected to generate sgRNA and were determined by the CRISPR MIT site (http://crispr.mit.edu/). The synthesised guide oligonucleotides (Bioneer Pacific, VIC, AUS) were then annealed together by placing the complementary guide reaction into 95°C water and allowing to cool overnight. The recipient FgH1UTG vector was fused to the annealed guides via a BsmBI cut site. These vectors were introduced into MCF10A cells via lentiviral transduction as described in **2.2.3** and CRISPR-Cas9 editing was induced by doxycycline treatment at 1 µg/ml for 3-5 days. Stably transfected MCF10A populations were sorted based on GFP and mCherry fluorescence into a single heterogeneous population. Since ZYG11A protein levels were undetectable by western blot due to low expression and unsuitable antibodies, the CRISPR induced cuts were validated using the Alt-R Genome Editing Detection Kit (IDT Technologies, IA, USA). Briefly, genomic DNA from the sgRNA targeted regions was first amplified by high fidelity PCR. The DNA mismatches produced by CRISPR-Cas9 nicking were then recognised by the T7 endonuclease I, resulting in DNA cleavage and

distinct bands that were then analysed by gel electrophoresis. All guide arm and primer sequences used are listed in Tables 2.6 and 2.7, respectively.

## 2.4 Protein analysis

### 2.4.1 Protein isolation

Cells were washed twice with PBS and harvested by scraping directly into ice-cold PBS supplemented with 1 μM MG132 and pelleted. Cell lysis was achieved using RIPA lysis buffer (150 mM NaCl, 5 mM EDTA pH 8.0, 50 mM Tris-HCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with cOmplete mini protease inhibitors (Roche, Basel, Switzerland) and PhosStop (Roche, Basel, Switzerland) just prior to use. Lysates were incubated for 30 minutes on ice and isolated based on the protein fractions required. For RIPA soluble fractions, the insoluble fractions were removed by centrifugation and the supernatant of soluble proteins was used. For the remaining RIPA insoluble protein pellets, they were solubilised by adding an appropriate volume of 8M urea lysis buffer (8M urea, 0.5% sodium deoxycholate, 300 mM NaCl), also supplemented with the protease and phosphatase inhibitors listed above. Viscosity caused by DNA release was removed by passing each sample through a 27G syringe 5-10 times. Where required, this viscosity was also minimised by the addition of benzonase (125 U/mL) in the lysis buffers. For whole cell lysates, following RIPA lysis, an equivalent volume of 8M urea lysis buffer was added to further solubilise the protein lysates. Protein concentrations were determined by the Lowry assay method using the DC Protein Assay Kit (Bio-Rad Laboratories, CA, USA).

### 2.4.2 SDS-PAGE and western blotting

Proteins concentrations were normalised, and equal amounts of protein were mixed with β-ME loading buffer (1:2 ratio of 2% bromophenol blue and β-mercaptoethanol, 15% final concentration). For SDS-PAGE, proteins were resolved on 4-12% Bolt Bis-Tris Plus gels (Bio-Rad Laboratories, CA, USA) and transferred to a PVDF membrane (Merck Millipore, MA, USA). To reduce background, membranes were blocked for 1 hour at room temperature in a BSA blocking buffer (3% BSA dissolved in TBST, which is 0.2% Tween 20 in PBS) on a shaker. Primary antibodies were diluted in blocking buffer and immunoblotting was performed overnight at 4˚C. The following day membranes were

washed 3 times for 10-minute intervals with TBST before being incubated with the appropriate secondary antibody diluted in blocking buffer for 1 hour at room temperature. The membranes were washed at least 3 times over 30 minutes before they were visualised using either the Odyssey CLx (LI-COR Biosciences, NE, USA) or ECL (GE Healthcare). All antibodies and the dilutions used are listed in Table 2.3

## 2.5 Apoptosis, cell cycle and proliferation

### 2.5.1 Apoptosis analysis

Cell death, specifically apoptosis, was measured by Annexin V binding and PI staining. siRNA transfected cells were analysed at 48 hours post-transfection, and at the same time post-seed for stable cell lines. Cells were harvested by trypsinisation and resuspended in Annexin V binding buffer that was diluted as per the manufacturer's instructions (BD Biosciences, CA, USA). 5 µl of Annexin V conjugated with FITC (A13199, Invitrogen, CA, USA) or APC (550475, BD Biosciences, CS, USA) was added to the cells that were then incubated for 20 minutes at room temperature, protected from light. PI or DAPI diluted in Annexin V binding buffer was then added. For stable cell lines expressing GFP and/or mCherry, the combination of Annexin V and DAPI was used. Fluorescence was measured by flow cytometry on the Cytoflex S (Beckman Coulter, CA, USA) and analysis was done using FlowLogic™ (Inivai Technologies, VIC, AUS).

### 2.5.2 Cell cycle and proliferation analysis

Cell cycle phasing and proliferation in the cells were analysed using the BD APC-BrdU Flow Kit as per the manufacturer's instructions (BD Biosciences, CA, USA). For siRNA transfected cells, cell cycle analysis was performed at 48 hours post-transfection. Stable cell lines were similarly seeded and analysed at 50-60% confluence. Briefly, cells were pulse labelled by adding BrdU to the cells (final concentration 10 µM) and allowed to incubate under normal culture conditions for 1 hour. Following a PBS rinse, the cells were then harvested by trypsinisation, resuspended in the Cytofix/Cytoperm buffer and allowed to fix at room temperature for 30 minutes, then washed with Perm/Wash buffer. The cells were permeabilised in Permeabilisation Buffer Plus for 10 minutes at room temperature, washed with Perm/Wash Buffer, and re-fixed with Cytofix/Cytoperm buffer for 5 minutes

on ice. To expose the incorporated BrdU, the DNA was denatured by incubating the cells in DNase I (30 µg of DNase/$10^6$ cells) for 1 hour at 37˚C. The cells were washed with Perm/Wash buffer before the addition of anti-BrdU-APC and a 20-minute incubation at room temperature, protected from light. Following one final wash in Perm/Wash buffer, the cells were resuspended in the provided 7-AAD solution to stain total DNA, then diluted in PBS. For stable cell lines containing GFP and/or mCherry fluorescent markers, 7-AAD was replaced by DAPI. In this situation the cells were resuspended in DAPI diluted in PBS with a final concentration of 1 mg/ml. Fluorescence was measured by flow cytometry on the Cytoflex S (Beckman Coulter, CA, USA) and all cell cycle and proliferation analysis was conducted using FlowLogic™ (Inivai Technologies, VIC, AUS).

### 2.6 Transwell migration and chemoinvasion assays

Transwell migration was performed on both siRNA transfected lines and stable cell lines. At 32 hours post-transfection, or post seed for stable cell lines, cells were EGF starved overnight in assay media (DMEM/F-12 supplemented with 1% donor horse serum, 100 ng/ml Cholera toxin, 0.5 µg/ml hydrocortisone, 10 µg/ml insulin and 15 nM HEPES). In the upper chamber of a 24-well plate of transwells (8 µm pore size), the starved cells were seeded at 5 x $10^4$ cells per well in 200 µl of assay media and allowed to settle for 10 minutes. Complete media containing EGF was added to the bottom of the chamber and cells were allowed to migrate for 24 hours. For invasion assays, 30 µl of human Matrigel (BD Biosciences, CA, USA) was used to coat the bottom of the transwells, which was then allowed to set for 30 minutes prior to seeding the cells. To visualise migrated cells the transwells were fixed in 10% buffered formalin for 20 minutes at room temperature then washed with PBS. Cells were then permeabilised with 0.3% TritonX-100 for 20 minutes, washed with PBS, then incubated with PBS diluted DAPI protected from light (2 µg/ml final concentration). The transwells were then washed with PBS and prepared for mounting by removing any cells remaining on the upper side of the membrane with a cotton tip. The membranes were mounted and imaged using a Zeiss LSM 780 (Carl Zeiss, AG, Germany). With the 10x objective, 4 non-overlapping fields of view were captured per membrane and analysed using the ImageJ program.

## 2.7 Zebrafish maintenance and genetic manipulation

### 2.7.1 Zebrafish maintenance conditions and strain sources

Fish were maintained at the La Trobe University animal house by LARTF staff under the conditions listed in Table 2.8. Embryos were maintained in petri dishes in a 28˚C incubator until 5 days post-fertilisation. For timepoints after 5 days, fry and adult fish were raised in standard tanks with the size depending on the number and age of fish.

### 2.7.2 Generation of CRISPR zebrafish lines

CRISPR lines used in this study were generated and sequenced by Dr. Lee Miles with the single guide (sg) RNAs listed in Table 2.9. Briefly, a crRNA (final concentration 360 ng/µl) and tracrRNA (670 ng/µl) preparation was made into a gRNA solution (final 30 µM) as per the Alt-R™ CRISPR-Cas9 system. RNP complexes were prepared as per the manufacturers suggestions with the following final concentrations: Cas9 protein (644 µg/µl), previously prepared gRNA (206 µg/µl), KCl (300 mM), Hbalt guide (50 ng/µl) and Hbalt targeting plasmid (40 ng/µl), and phenol red (0.005%). Embryos at the 1-cell stage were injected with 2 nl of the RNP complex and raised to adulthood before genotyping and sequencing (see 2.7.3) to identify any genomic changes. All Sanger Sequencing was performed by the Australian Genome Research Facility (AGRF).

### 2.7.3 Genotyping of embryos and adult fish

Genomic DNA from embryos and adult fish fin clips was extracted using 50 mM NaOH at 95˚C for 10 minutes followed by neutralisation using 1M Tris-HCl (pH 8.0). CRISPR-Cas9 edited regions were amplified using the GoTaq Green PCR master mix (Promega). For *zyg11* +/- 1bp changes, the amplified DNA was digested with EcoRV for 2 hours prior to gel electrophoresis analysis of DNA fragments on a 3% agarose gel. The sample was prepared for sequencing using the EXO-SAP method. Primers sets for genotyping and sequencing are listed in Table 2.10.

## 2.8 High-content immunofluorescence functional (HCIF) screening

### 2.8.1 siRNA boutique library candidates

The boutique library of siRNA candidates (listed in Table 2.11) was cherry picked from the Dharmacon siGENOME SmartPool range of reagents (GE Healthcare Dharmacon, Inc. CO, USA). SmartPool siRNAs consisted of four individual siRNA oligonucleotides designed to broadly target specific regions in a single gene. The library was organised in a 384-well plate format as per Figure 2.1. The candidate list consisted of the known ZYG11A ubiquitination target substrates curated from the literature, as well as genes known to form protein-protein interactions with ZYG11A, including those known to interact with ZYG11A orthologues.

### 2.8.2 Transfection of siRNA for HCIF

Functional genomic screening was based on protocols previously published by our lab (Smith *et al.*, 2016). Prior to screening, passage matched MCF10A cells were cryopreserved at passage 10 for all screening purposes. Briefly, 22 µl of a pre-prepared Dharmafect 3/OptiMEM master mix was dispensed into a black walled 384-well plate (Plate A) using a BioTek 406 liquid handling robot, such that per well there was 10.94 µl of OptiMEM and 0.06 µl of Dharmafect 3. This master mix was complexed for 5 minutes prior to the addition of 3 µl of 1 µM siRNA from the arrayed boutique screening library by the Caliper Sciclone ALH3000 robot, which then mixed the siRNA/lipid/OptiMEM before transferring 12.5 µl to Plate B. Following a 20-minute complex at room temperature, MCF10A cells were dispensed into each well using the BioTek 406 at a density of 800 cells per well in 25 µl of complete MCF10A media. The final concentration of siRNA per well was 40 nM. All BioTek 406 liquid handing steps were conducted using the 5 µl cassette with a high flow rate. Each plate was then centrifuged briefly for 30 seconds at 500 x g then left on an even surface for 10 minutes to allow the cells to settle. All transfection plates were then incubated overnight at 37˚C with 5% $CO_2$ in the Liconic STX200 automated humidified incubator (Applied Biosystems, Foster City, USA). At 24 hours post-transfection, the transfection media was replaced using the BioTek 406 (50 µl per well) and returned to the Liconic STX200 incubator until 72 hours post-transfection.

For rescue screening where two siRNAs were used per well, similar robotic handling methods were implemented, with minor adjustments to screening volumes. 1.88 µl of the base siRNA (OTP-NT or siZYG11A) at 1 µM was manually pipetted into Plate A prior to

screening. In terms of lipid/OptiMEM, 21 µl of a pre-prepared mixture was dispensed using the BioTek (final 0.06 µl Dharmafect 3 + 10.56 µl OptiMEM per well). The boutique library was added at 3.76 µl of 1 µM siRNA per well. To Plate B, 12.4 µl was transferred prior to the addition of MCF10A cells at the concentrations mentioned above. The final siRNA concentration was increased 50 nM (25 nM of each individual siRNA) to maintain the robustness of the base siRNA phenotype.

### 2.8.3 High-content immunofluorescent imaging

At 72 hours post-transfection, cells were prepared for high-content immunofluorescent imaging using the BioTek 406 for all liquid handling steps unless otherwise stated. For dispensing, the 5 µl cassette and medium flow rate was used. For washing steps, a travel rate of 1CW with a 200 msec delay and a flow rate of 3 was used. All incubation steps were performed with gentle shaking and at room temperature unless otherwise stated. Cells were fixed for 10 minutes in 4% PFA then permeabilised for 10 minutes in 0.3% Triton-X 100 in PBS, followed by blocking in 3% BSA in PBS for 1 hour. Primary antibodies against Scribble and β-catenin were then diluted in 3% BSA in PBS (see concentrations in Table 2.3) and manually added to each plate using an automated pipette capable of dispensing 20 µl volumes into a 384-well plate format. Primary antibodies were incubated with the cells overnight at 4˚C. Plates were then washed 2 x in 0.01% Tween20 in PBS (PBST) and 1 x in PBS prior to the manual addition of secondary antibodies. Following a 1-hour incubation, plates were washed as above before staining with the nuclear dye DAPI (10 µg/ml) for 10 minutes (added manually). Plates were then washed with PBS briefly, filled with PBS (80 µl per well) and thermosealed with foil seals for imaging.

Cells were imaged using the Cellomics ArrayScan automated microscope using the 20x air objective. The DAPI channel was used as the reference channel for the autofocus and exposure for each well. 25 fields were imaged per well, taken from the centre of the well, outwards. Individual images were montaged post-capture to observe the morphology of cells within the entire well. Nuclei counts were obtained based on DAPI counts acquired by the Cellomics Software, as were the cell cycle profiles.

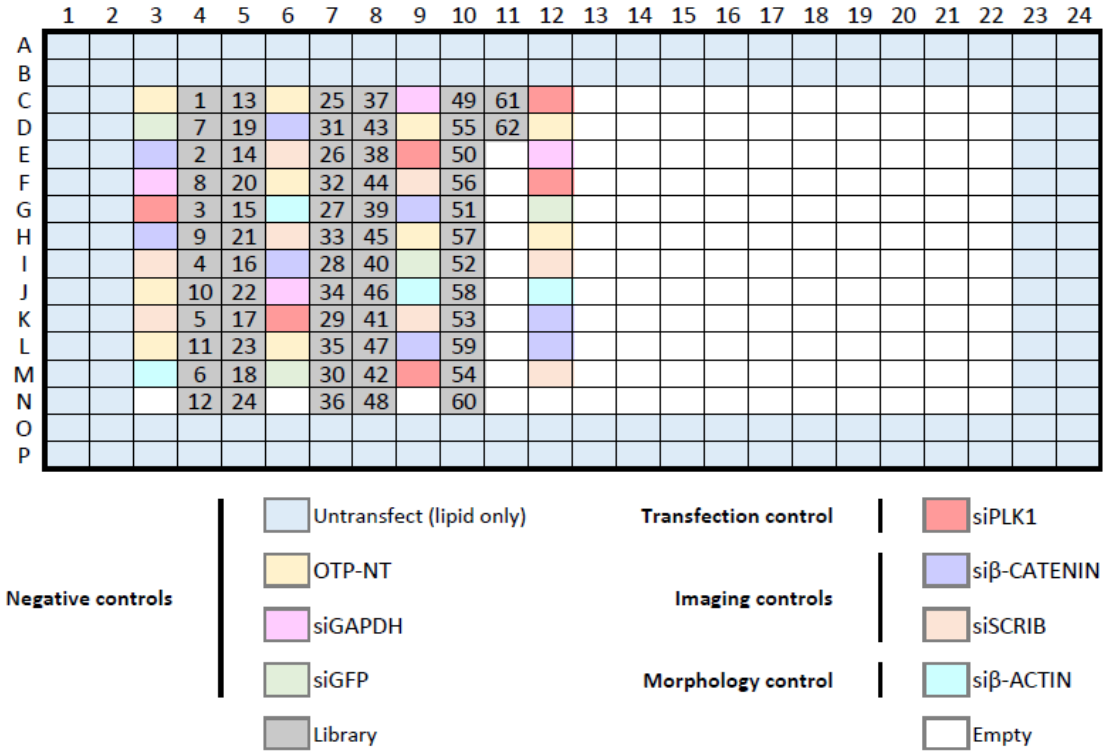



**Figure 2.1 High-content screening 384-well plate screening layout (numbers correspond to siRNA library listed in Table 2.11)**

## 2.9 Proximity biotinylation (BioID) and mass spectrometry

### 2.9.1 BioID

Each stable MCF10A cell line expressing BioID fusion proteins (including controls) were seeded into 6 x 10 cm plates at 5 x $10^5$ cells in 10 ml of complete MCF10A media and allowed to reach ~70-80% confluence. The media was then replaced with complete media supplemented with 50 µM of biotin and either a DMSO control or 10 µM of MG132 (3 plates/stable line) and incubated for 16 hours. Cells were then washed three times with 5 ml PBS and harvested (on ice) by scraping into ice cold PBS. Cells were then pelleted and snap frozen on dry ice and stored before protein isolation and normalisation steps (see 2.4.1, whole cell lysate preparation). Pierce streptavidin coated beads were equilibrated in 4M urea lysis buffer (4M urea, 150 mM NaCl) and added to the protein lysate (20 µl equilibrated beads/500 µg protein. Proteins pull downs were then incubated overnight at 4˚C with rotation. The next day, beads were washed twice with TAP lysis buffer (50 mM

HEPES-KOH, 2 mM EDTA, 100 mM KCl, 0.1% NP-40, 10% glycerol and 4 M urea), three times with ABC buffer (50 mM ammonium bicarbonate), and resuspended in ABC buffer before being sent to the La Trobe Comprehensive Proteomics Platform for mass spectrometry analysis.

### 2.9.2 Mass spectrometry

After the washing steps, beads were resuspended in 8M Urea, 50 mM Tris pH=8.3 followed by reduction and alkylation. Solubilised proteins were submitted to trypsin digestion overnight and the resulting peptides purified using the C18 stage tips procedure. For the analysis: peptides were reconstituted in 0.1% formic acid and 2% acetonitrile, and 500 µg peptides loaded onto C18 PepMap 100 µm ID × 2 cm trapping column (Thermo-Fisher Scientific) at 5 µl/min for 6 min, and washed for 6 minutes before switching the pre-column in line with the analytical column (Acquity BHE C18, 1.7 µm, 130 Å and 75 µm ID × 25 cm, Waters). The separation of peptides was performed at 250 nl/min using a 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 35 % over 90 minutes, then 50 % B in 15 min followed by 95 % B in 5 min, The column is then cleaned for 5 min at 95 % B following a 5 min equilibrated step (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 60 000 resolution, HCD MS/MS spectra were collected for the 10 most intense ions per MS scan at 15 000 resolution with a normalised collision energy of 28% and an isolation window of 1.4 m/z. 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 $3 \times 10^6$, MSMS for a maximum injection time of 110 ms with AGT target of $1.1 \times 10^4$.

### 2.9.3 Label free quantitation

MS data were analysed using MaxQuant (version 1.6.14) (Cox and Mann, 2008) with label-free quantitation (LFQ) turned on. Uniprot Homo sapiens database last modified on November 5, 2019 was used for the peptide and protein identification with the proteomics search engine Andromeda (Cox and Mann, 2008). Search engine settings used were parent

mass tolerance: 10 ppm; fragment mass tolerance: 20 ppm; allowed missed cleavages: 2; enzyme: trypsin; fixed modifications: carbamidomethyl C; variable modifications: oxidation of methionine and acetylation at the protein N-terminus. Minimum peptide length: 7 residues and the maximum peptide mass: 4600 Da. The window for the match between runs was 0.7 min. Protein groups were retained at 1% false discovery rate (FDR). Large LFQ ratios were stabilised to reduce the sensitivity for outliers. The cRAP database (Mellacheruvu *et al.*, 2013) was used for the identification of common laboratory contaminants.

### 2.9.4 Data analysis and data visualisation

Data analysis and visualisation were performed using the statistical programming language R (R Core team, 2015). Zero intensity values from the MaxQuant output were substituted with NA and were normalised using the 'Cyclicloess' normalisation method (M. Ritchie *et al.*, 2015). The principal comportment analysis (PCA) was performed using the base package "princomp". Heatmap was generated using the package pheatmap. Differential abundance analysis was performed using the empirical Bayes method available in the package limma (M. E. Ritchie *et al.*, 2015). The resulting probabilities were adjusted with the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) and those genes with an adjusted probability > 0.05 were considered as significant.

# Chapter 2

Tables

**Table 2.1 *Drosophila* food media recipe**

| Reagent | Amount | Final conc. |
|---|---|---|
| Molasses | 1400 g | 9.3 w/v% |
| Fresh baker's yeast | 900 g | 6 w/v% |
| Agar | 75 g | 0.5 w/v% |
| Glucose | 160 g | 1.07 w/v% |
| Semolina | 1320 g | 8.8 w/v% |
| $H_2O$ | 15 L | - |
| Reagents to add once cooled | | |
| **Tegosept solution** (10 w/v% methyl 4-hydroxybenzoate, 0.5 v/v% $H_2O$, 95 v/v% 100% EtOH) | 262 ml | |
| **Propionic acid mix** (41.2 v/v% 99% propionic acid, 4.2 v/v% 85% phosphoric acid) | 138 ml | |
| $H_2O$ | 546 ml | |

**Table 2.2 *Drosophila* strains used**

| Strain | Source | Strain # |
|---|---|---|
| *zer1*$^{RNAi}$ (KK line) (P{KK100351}VIE-260B) | VDRC | v108923 |
| *zer1*$^{RNAi}$ (GD line) ($w^{1118}$; P{GD7474}v31646) | VDRC | v31646 |
| *zer1 allele* (*y*$^{1}$ $w^{67c23}$; P{SUPor-P}CG12084$^{KG02879}$ $ry^{506}$/TM3, $Sb^{1}$) | Bloomington | 12996 |
| *GMR-GAL4* | HR stocks | - |
| *UAS-dicer; nub-GAL4* | HR stocks | - |

**Table 2.3 Antibody sources and concentrations used (primary, secondary and stains)**

| Antigen | Clone | Source | Supplier | Western blot | IF |
|---|---|---|---|---|---|
| GAPDH | | rabbit polyclonal IgG | Cell Signalling Technology (#2118) | 1:1000 | - |
| GFP | | rabbit polyclonal IgG, serum | Invitrogen (A-6455) | 1:1000 | 1:1000 |
| GM130 | 35 | mouse monoclonal $IgG_1$, κ | BD Biosciences (#610822) | - | 1:500 |
| Lamin A/C | 4C11 | mouse monoclonal $IgG_{2a}$ | Cell Signalling Technology (#4777) | 1:1000 | - |
| LAMP-2 | H4B4 | mouse monoclonal $IgG_1$, κ | Santa Cruz Biotechnology (sc-18822) | - | 1:500 |
| Myc | 9B11 | mouse monoclonal $IgG_{2a}$ | Cell Signalling Technology (#2276) | 1:1000 | 1:4000 |
| Scribble | | mouse monoclonal | Humbert lab | 1:100 | - |
| Scribble (C-20) | | goat polyclonal | Santa Cruz Biotechnology (sc-11049) | - | 1:200 |
| Ubiquitin | P4D1 | mouse monoclonal $IgG_1$ | Santa Cruz Biotechnology (sc-8017) | - | 1:500 |
| α-tubulin | B512 | mouse monoclonal $IgG_1$ | Sigma-Aldrich (T5168) | 1:20,000 | 1:5000 |
| β-catenin | 14 | mouse monoclonal $IgG_1$ | BD Biosciences (610153) | 1:2000 | 1:1000 |
| Integrin α6 | | rat monoclonal | Chemicon International (#MAB1378) | - | 1:500 |
| Rb1 | 4H1 | mouse monoclonal $IgG_{2a}$ | Cell Signalling Technology (#9309) | 1:1000 | - |
| p21 | 12D1 | rabbit monoclonal IgG | Cell Signalling Technology (#2947) | 1:500 | - |
| p53 | | rabbit polyclonal IgG | Cell Signalling Technology (#9282) | 1:500 | - |
| DAPI | | | Invitrogen (D1306) | - | 0.5 μg.ml |
| Phalloidin 568 | | | Molecular Probes (A12380) | - | 1:1000 |
| Phalloidin-iFluor 647 | | | Abcam (ab176759) | - | 1:1000 |
| AlexaFluor 488, 568, 633, 647 | | | Abcam | - | 1:1000 |

**Table 2.4 Real-time PCR primers**

| Target | Forward | Reverse |
|---|---|---|
| *CCNE1* | CGGCTCGCTCCAGGAA | TCATCTGGATCCTGCAAAAAAA |
| *GAPDH* | AAGGTGAAGGTCGGAGTCAAC | GAGTTAAAAGCAGCCCTGGTG |
| *P21* | CATGTGGACCTGTCACTGTCTTGTA | GAAGATCAGCCGGCGTTTG |
| *P53* | CAGCACATGACGGAGGTTGT | TCATCCAAATACTCCACACGC |
| *RB1* | TTGGATCACAGCGATACAAACTT | AGCGCACGCCAATAAAGACAT |
| *SCRIB* | AGCTGCCCAAGCCTTTTTTC | AACCGCTGGATCTCGTTGTC |
| *ZER1* | CGGAGAAACATCAATTACAGGTCA | TCTTTGATCAGCAGAGGGCA |
| *ZYG11A* | GTTGTCAGAGGTCACCTGTCT | GCAGCATCAAACCTGTCAAATG |
| *ZYG11B* | TGGCTGCCAAGCTTTCTACA | AGGTTCCAAAGTGCACTCAAAG |

**Table 2.5 siRNA sequences for 12-well transfections**

| Target | Product # | SMARTpool oligo sequences (Dharmacon) |
|---|---|---|
| ON-TARGETplus Non-targeting Pool (OTP-NT) | D-001810-10 | UGGUUUACAUGUCGACUAA<br>UGGUUUACAUGUUGUGUGA<br>UGGUUUACAUGUUUUCUGA<br>UGGUUUACAUGUUUUCCUA |
| *ZYG11A* (440590) | M-032396-01 | AAUCAAGGAUUGCAAAUCU<br>CAGCAUUGGUGACCUAUAG<br>UAUCAUAGCCCACCUGACA<br>GAAUGAGGAAUCACCCAUU |
| *ZYG11B* (79699) | M-021798-01 | GAAACCACCCUAUGAAUUU<br>GCUUGGACAUUGAGUCGUA<br>GAUUAGAGAGCUUGGAUAU<br>CAUGGUCUAUUGAAUGAUG |
| *ZER1* (10444) | M-019424-00 | GCAGUAACUUUAAAGAGGA<br>GACAGGAACAUUCAAGUGA<br>GAACAUGGACACGUCUAGA<br>CAUCAUACCUUUCCGGGCU |

**Table 2.6 *ZYG11A* CRISPR guide target sequences**

| Target exon | Target sequence (5' → 3') |
|---|---|
| Exon 1 | CGTCCCTCCTGACGCTCAGA |
| Exon 2 | AAGTAGCCGAGCGATTTCTC |
| Exon 3 | GGACTATTGGCCACGGATGC |
| Exon 4 | CAGAAGCACTGAGCCGATAC |
| Exon 5 | CGCTCTCAACCTAACACGCC |

**Table 2.7 *ZYG11A* Alt-R Genomic Editing Detection PCR primers**

| Target exon | Forward | Reverse |
|---|---|---|
| Exon 1 | CTCGCCGGGCATTTCAA | GCACCTACGTCTAAGCTCAC |
| Exon 2 | ACTTGAGCATCTGTGGATACTG | ATAGCAGGGCTAGCCTAGAA |
| Exon 3 | CAAGATTACTGTTCTTTAGTCAGCTC | TTATCTGTCTCTGTAACCCTAATACC |
| Exon 4 | TCTTGAATCCTGTATATCTCTGC | AGCACTTGACTCACTATAGAAT |
| Exon 5 | AAACATGCACAAGACAGCATAG | AACACAAATTGTCTGCCGTATG |

**Table 2.8 Zebrafish maintenance parameters**

| Parameter | Range |
|---|---|
| Temperature | 28°C (+/- 2°C) |
| pH | Stable between 7.0 – 8.0 |
| Dissolved Oxygen | 7.8mg/L at 28°C |
| Ammonia $NH_3$, $NH_4^+$ | <0.25 ppm |
| Nitrite $NO_3^-$ | 0 ppm |
| Nitrate $NO_2^-$ | < 50 ppm |
| Conductivity (CD) | 1000µS (+/- 200 µS) |
| General Hardness (GH) | ~100ppm |
| Carbonate Hardness (KH) | ~50 PPM |
| Copper (Cu) | 0ppm |
| Iron (Fe) | 0ppm |
| Phosphates (P) | <20ppm |

**Table 2.9 *zyg11* and *zer1* zebrafish CRISPR-Cas9 guide target sequences**

| Target Gene | Target sequence |
|---|---|
| *zyg11* | GGAGAGTCCACGGAGGATAT |
| *zer1* | GGGATCTTCTTCCGCAAGAG |

**Table 2.10 Zebrafish genotyping PCR and sequencing primers**

| Target gene | Forward | Reverse |
|---|---|---|
| *zyg11* | CAGTTCCGCTTGCGTCATG | TCAGGCCGTTCAGAACCAG |
| *zer1* | TCGTCTCGAGCTCTGCGT | GCCCACATTCAGCACACG |

**Table 2.11 siRNA screening library**

| Pick # | Entrez ID | Entrez Name |
|---|---|---|
| 1 | 8065 | CUL5 |
| 2 | 85417 | CCNB3 |
| 3 | 6950 | TCP1 |
| 4 | 10576 | CCT2 |
| 5 | 898 | CCNE1 |
| 6 | 7203 | CCT3 |
| 7 | 22948 | CCT5 |
| 8 | 908 | CCT6A |
| 9 | 10575 | CCT4 |
| 10 | 10694 | CCT8 |
| 11 | 332 | BIRC5 |
| 12 | 891 | CCNB1 |
| 13 | 10574 | CCT7 |
| 14 | 3920 | LAMP2 |
| 15 | 10693 | CCT6B |
| 16 | 84132 | USP42 |
| 17 | 6923 | TCEB2 |
| 18 | 29882 | ANAPC2 |
| 19 | 6921 | TCEB1 |
| 20 | 84858 | ZNF503 |
| 21 | 9423 | NTN1 |
| 22 | 55625 | ZDHHC7 |
| 23 | 4897 | NRCAM |
| 24 | 8878 | SQSTM1 |
| 25 | 169841 | ZNF169 |
| 26 | 7428 | VHL |
| 27 | 4917 | NTN2L |
| 28 | 4685 | NCAM2 |
| 29 | 10752 | CHL1 |
| 30 | 3916 | LAMP1 |
| 31 | 9422 | ZNF264 |

| Pick # | Entrez ID | Entrez Name |
|---|---|---|
| 32 | 4684 | NCAM1 |
| 33 | 23513 | SCRIB |
| 34 | 3897 | L1CAM |
| 35 | 9638 | FEZ1 |
| 36 | 23114 | NFASC |
| 37 | 161 | AP2A2 |
| 38 | 158399 | ZNF483 |
| 39 | 80139 | ZNF703 |
| 40 | 79699 | ZYG11B |
| 41 | 255762 | PDZD9 |
| 42 | 11051 | CPSF5 |
| 43 | 10444 | ZER1 (C9ORF60) |
| 44 | 57116 | SBZF3 |
| 45 | 51667 | NYREN18 (NUB1) |
| 46 | 8453 | CUL2 |
| 47 | 9820 | CUL7 |
| 48 | 8452 | CUL3 |
| 49 | 8454 | CUL1 |
| 50 | 8451 | CUL4A |
| 51 | 8450 | CUL4B |
| 52 | 26271 | FBXO5 |
| 53 | 286151 | FBXO43 |
| 54 | 9978 | RBX1 |
| 55 | 9616 | RNF7 |
| 56 | 23113 | PARC |
| 57 | 387755 | INSCB |
| 58 | 1499 | CTNNB |
| 59 | 440590 | ZYG11A |
| 60 | 60 | ACTB |
| 61 | 9637 | FEZ2 |

# Chapter 3

# Comparative phenotypic and functional characterisation of human ZYG11 family members

## 3.1 Introduction

Evolutionarily conserved families of genes are often a reflection of the requirement for those genetic families between the species. Phenotypically and functionally, particularly in cases where there are high levels of sequence similarity, conserved family members can exhibit parallels that ultimately lend themselves to increases in genetic compensation and potential redundancy within gene families. In contrast, small genetic variations can provide enough change to a protein's function or expression so that each family member has a distinguishing and more diverse subset of functions, leading to a more diverse set of proteins. The ZYG11 family of genes is one such family that possesses high conservation both across species and at the level of each family member. After first being described in the roundworm, *Caenorhabditis Elegans* (*C. elegans*), ZYG11 has been shown to be important in the establishment of embryonic cell polarity, and to regulate early mitotic and meiotic cell divisions (Kemphues *et al.*, 1986; Carter, Roos and Kemphues, 1990; Liu, Vasudevan and Kipreos, 2004; Vasudevan, Starostina and Kipreos, 2007). Rescue experiments have provided evidence that the cell cycle defects seen in ZYG-11 deficient worms can be alleviated without any effect on the cell polarity defects (Liu, Vasudevan and Kipreos, 2004; Vasudevan, Starostina and Kipreos, 2007) suggesting that these two functions are regulated by separate molecular pathways. Cell cycle has been at the centre of the majority of studies on the ZYG11 family and it has now been established that two of the mammalian counterparts of the ZYG11 family, ZYG11A and ZYG11B, play a role in mammalian cell cycle regulation at the transition of G2 to M phase (Balachandran *et al.*, 2016). However, how these genes regulate mammalian cell polarity, and how they interact and function in relation to each other remains unstudied.

As mentioned, the basis of ZYG11 function *in vivo* has been largely defined by cell cycle and polarity defects in *zyg-11* mutant *C. elegans* embryos (Carter, Roos and Kemphues, 1990; Liu, Vasudevan and Kipreos, 2004; Sonneville and Gonczy, 2004). Conversely, nothing is known about the role of the second roundworm ZYG11 counterpart, *zer-1*, and despite their close relationship no studies have considered *zer-1* activity during *zyg-11* loss. Outside of the *C. elegans* field none of these findings have been translated into other *in vivo* genetic models, and consequently little is known about how this family of genes function in other systems. The zebrafish provides an invaluable system where any defects caused by ZYG11 family loss can be easily observed – specifically those that occur during early vertebrate development. To date the ZYG11 gene family has not yet been studied as a whole *in vivo*, providing a number of open questions regarding the individual role of

these genes in embryonic polarity, cell cycle and overall development. Likewise, this is also an interesting question to address in other genetic model organisms like *D. melanogaster*, which unlike *C. elegans* or *D. rerio*, only contain a single ZYG11 orthologue (*zer1*). There are no studies specifically focusing on the function of *zer1*, however potential protein and genetic interactions with *zer1* have been captured through other genetic screening studies. These two models are both well characterised in terms of genetic amenability and provide a number of advantages to studying these genes *in vivo* including quick and trackable *ex vivo* development (*D. rerio*), and rapid development and overall life cycles (*D. melanogaster*).

In this Chapter, I explore the phenotypic and functional characteristics of the human ZYG11 family in mammalian cell culture systems using combinations of gain and loss of function approaches. I will also assess the phenotypic and functional consequences of ZYG11 family loss in two widely used genetic model organisms, *D. melanogaster* and *D. rerio*. Through the knockdown and knockout methods of RNAi and CRISPR-Cas9 technologies, I will investigate whether the loss of this gene family is detrimental in other *in vitro* and *in vivo* models, as it is in *C. elegans*.

## 3.2 Results

### 3.2.1 Expression and subcellular localisation of the ZYG11 family

#### 3.2.1.1 The ZYG11 family localises to the cytoplasm and is excluded from the nucleus

To begin to address the conservation between these human ZYG11 family members, I expressed and assessed the localisation of ZYG11A and other ZYG11 family members in the polarised human mammary epithelial cell line, MCF10A. I generated MCF10A cell lines stably expressing full length ZYG11A, ZYG11B and ZER1, N-terminally fused to a GFP fusion protein, as well as their N or C terminal domains (Figures 3.1 and 3.2). Cell lines were sorted by FACS analysis to ensure they all expressed comparable levels of GFP-fusion proteins. Using GFP as a marker, the GFP-ZYG11A fusion was localised throughout the cytoplasm and was largely excluded from the nucleus compared to the control GFP only signal (Figure 3.1B, two left columns). I attempted to confirm that the endogenous protein showed similar localisation but unfortunately, despite trying a number of commercially available antibodies and optimisation methods, I was unable to detect the endogenous ZYG11A by immunofluorescence microscopy.

ZYG11 family proteins harbour distinct protein domains – most notably a VHL box, which permits association with the Cullin-2 E3 ubiquitin ligase, leucine-rich repeats (LRRs) at the N-terminal, and an armadillo (ARM) domain at the C-terminal. To identify whether either of these domains were important for the localisation of ZYG11A, I subcloned two truncated forms of ZYG11A into GFP fusion proteins to separate the N- and C-terminal domains. The ZYG11A N-terminal protein had a localisation identical to the full-length protein, whereas the C-terminal protein was able to enter the nucleus and was localised throughout the cell (Figure 3.1A-B). Interestingly, similar patterns were observed when identical constructs for both ZYG11B and ZER1 were expressed (Figure 3.2A-B), and this is consistent with previous studies that have shown a similar localisation for ZYG11B overexpression constructs (Balachandran *et al.*, 2016). Taken together, it is apparent that in an overexpression setting, this N-terminal dependent exclusion from the nucleus is a property shared amongst all three ZYG11 family proteins.

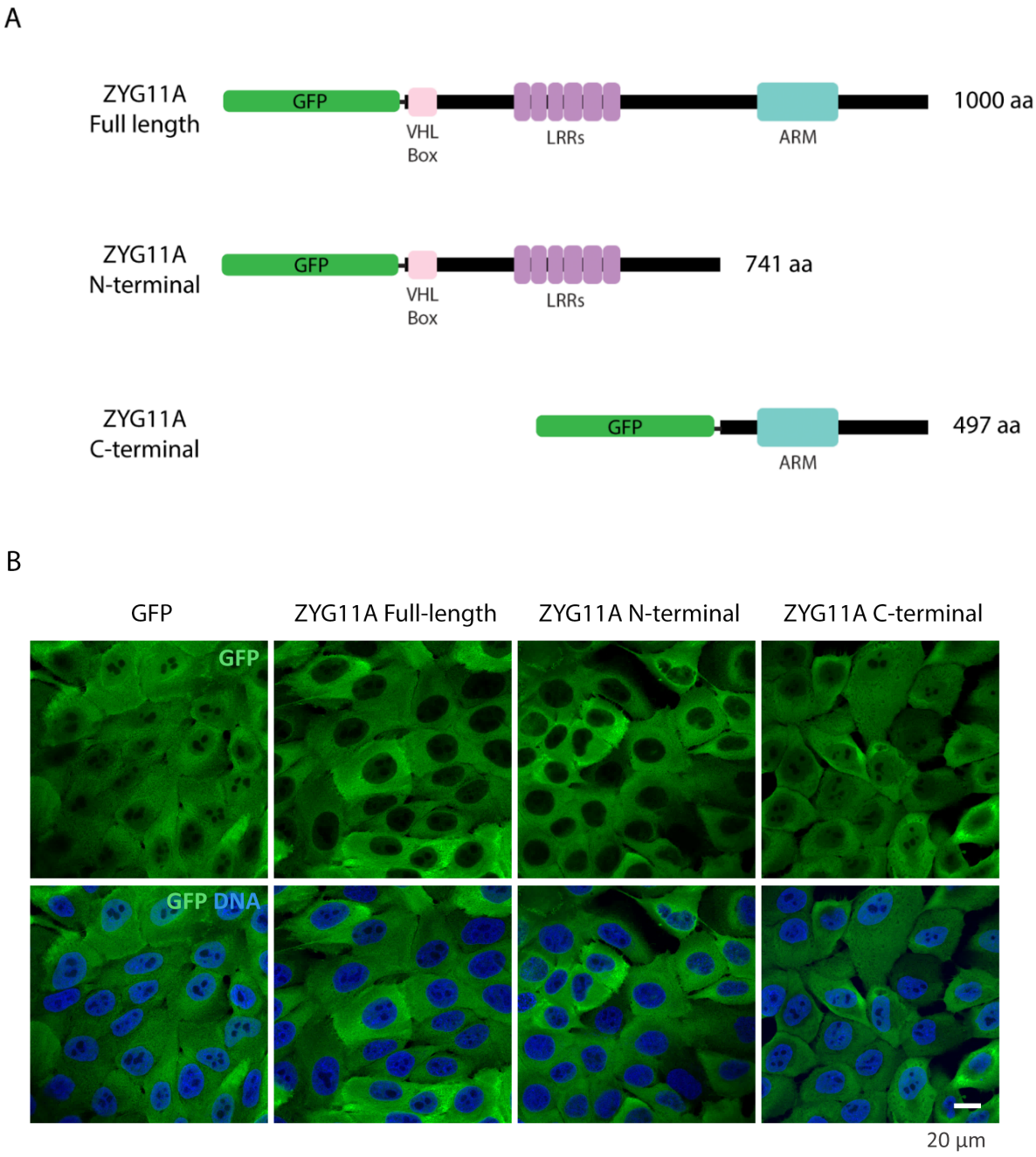



**Figure 3.1 Generation of ZYG11A full length and truncated fusion proteins**

**A.** Diagramatic depiction of the GFP-ZYG11A fusion proteins that were stably expressed in MCF10A cells. The full length contains all three known ZYG11 protein motifs. The N-terminal truncation harbours the von Hippel-Lindau (VHL, pink) and leucine rich repeats (LRR, magenta), whereas the C-terminal truncation contains the armadillo domains (ARM, teal). Each fusion protein is N-terminally tagged with a GFP fusion (green). Predicted molecular weights are shown at the end of each protein. **B.** Representative confocal images of MCF10A cells stably expressing either a GFP only, or the full length, N-terminal or C-terminal ZYG11A GFP fusion proteins. Fusion proteins are marked by GFP and the DNA is visualised using DAPI. In control cells, GFP is present throughout the cell whereas the full length and N-terminal fusions localise to the cytoplasm. The C-terminal localisation is similar to the GFP only control, present throughout the cell.

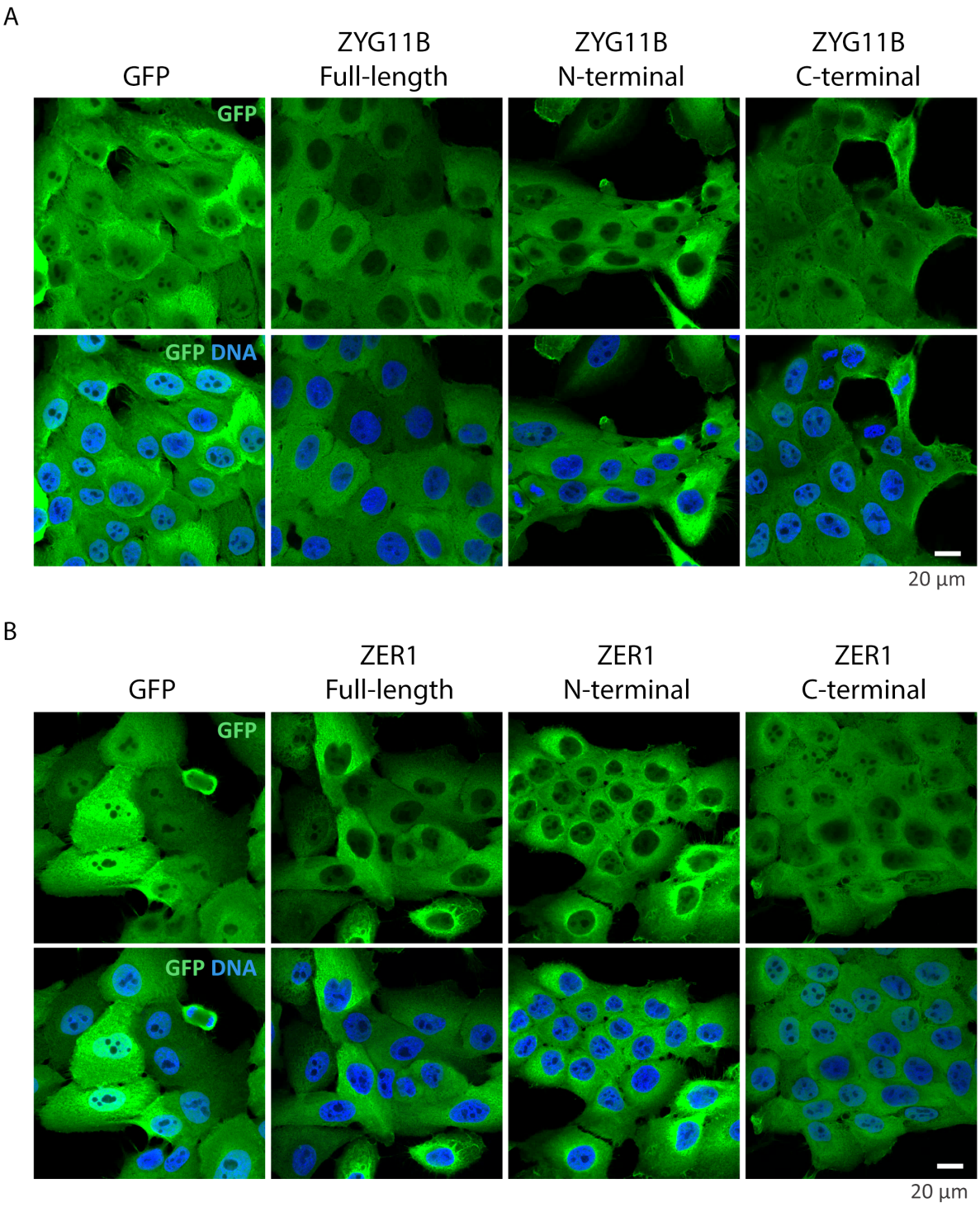



**Figure 3.2 ZYG11B and ZER1 GFP fusion proteins localise similarly to ZYG11A fusions**

Representative confocal images of MCF10A cells stably expressing either a GFP only, or a full length, N-terminal or C-terminal GFP fusion protein for ZYG11B (**A**) or ZER1 (**B**). Fusion proteins are marked by GFP and the DNA is visualised using DAPI. In control cells, GFP is present throughout the cell whereas the full length and N-terminal fusions localise to the cytoplasm. The C-terminal localisation is similar to the GFP only control, present throughout the cell. This is similar to the localisation observed with ZYG11A GFP fusion constructs.

#### 3.2.1.2 Mutations in highly conserved regions of ZYG11A does not alter the localisation of ZYG11A

The exclusion of a protein from the nucleus is typically by virtue of the presence of a nuclear export signal (NES). This localisation motif is responsible for mediating the interaction between the protein and the exportin/CRM1 complex, and from here, the protein is then shuttled into the cytoplasm. This prompted the analysis of the ZYG11A family protein sequence to identify any possible interaction domains by using the ELM (Eukaryotic Linear Motif) prediction resource (Gouw *et al.*, 2018), through which a number of potential binding sites were revealed. Following this, ZYG11 protein sequences from various species were aligned and regions of high conservation were compared to the ELM prediction to identify highly conserved, and therefore likely important, protein motifs. The most robust sequences identified were a nuclear exclusion signal in the N-terminal region, as well as sites for N-glycosylation and PP2A-B62 docking (Figure 3.3A-C). To test whether these sequences played a functional role in the localisation of ZYG11A I performed site mutagenesis on these residues using full length GFP-ZYG11A proteins and stably expressed them in MCF10A cells at equivalent levels using retroviral transduction and FACS sorting based of GFP signal (Figure 3.4A). In all three mutants, ZYG11A remained localised throughout the cytoplasm and failed to significantly increase levels in the nucleus (Figure 3.4B-C). Nuclear export can also be rapidly prevented by targeting the exportin pathway using the inhibitor, Leptomycin B (LMB), confirmed by the accumulation of Myc in the nucleus following LMB (25 nM) treatment for 6 hours (Figure 3.5C). In an effort to delineate whether this mechanism was indeed behind the cytoplasmic localisation of the ZYG11 family fusion proteins, all cell lines (including the mutants) were similarly treated with LMB (Figure 3.5A). Still, the protein localisation appeared the same and quantification of the nuclear to cytoplasmic ratio of GFP fluorescence showed no significant alterations in any of the fusion proteins (Figure 3.5B).

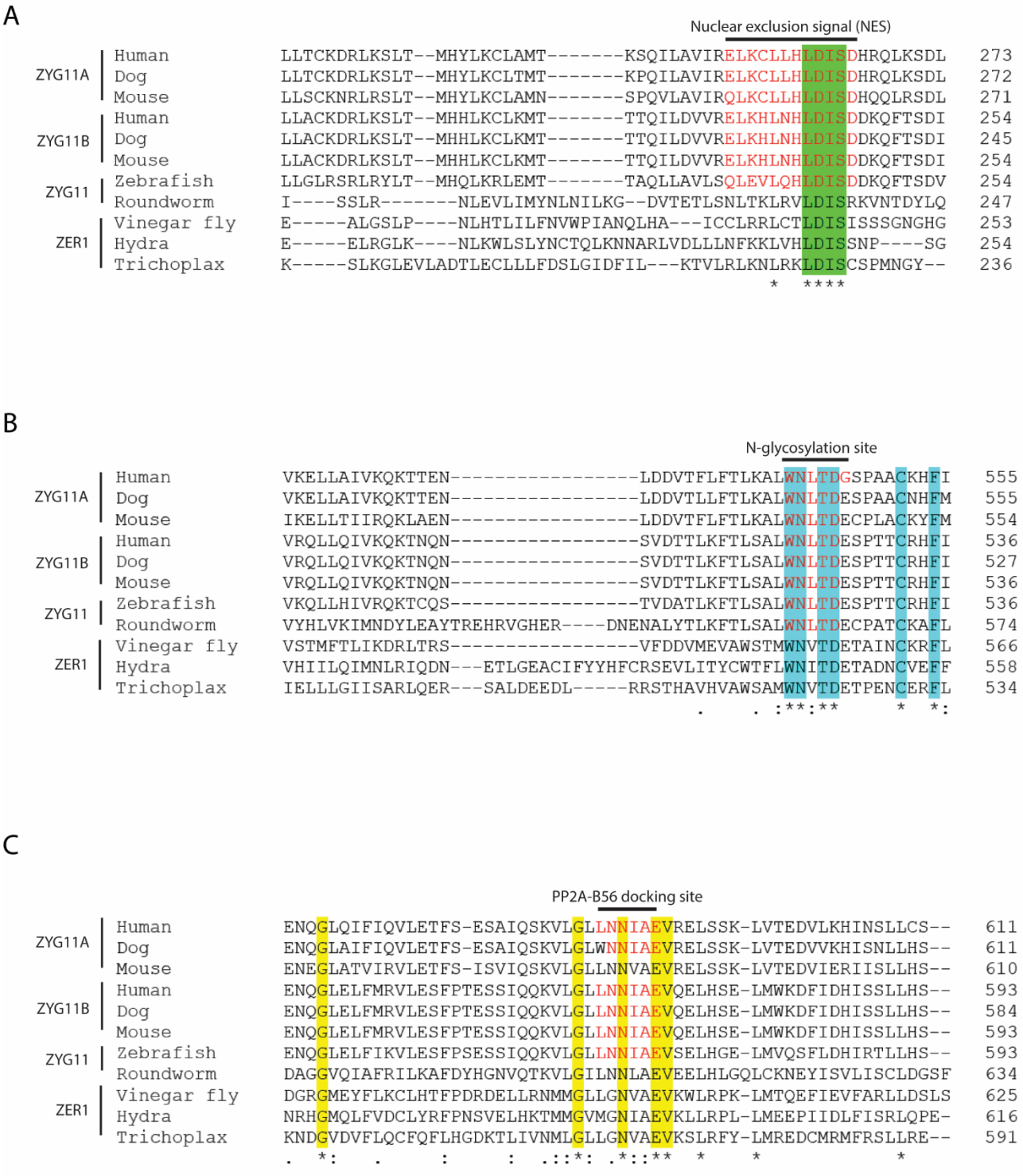



**Figure 3.3 ZYG11A contains highly conserved protein interaction motifs for nuclear export, N-glycosylation and PP2A binding**

Alignment of selected protein regions in all three ZYG11 family members across multiple vertebrate and invertebrate species. Highlighted sequence motifs across each species represent highly conserved regions between each family member as identified by Eukaryotic Linear Motif (ELM) prediction – **A.** Nuclear exclusion/export signal, **B.** N-glycosylation site, **C.** Protein phosphatase 2A-B56 (PP2A-B56) docking site.

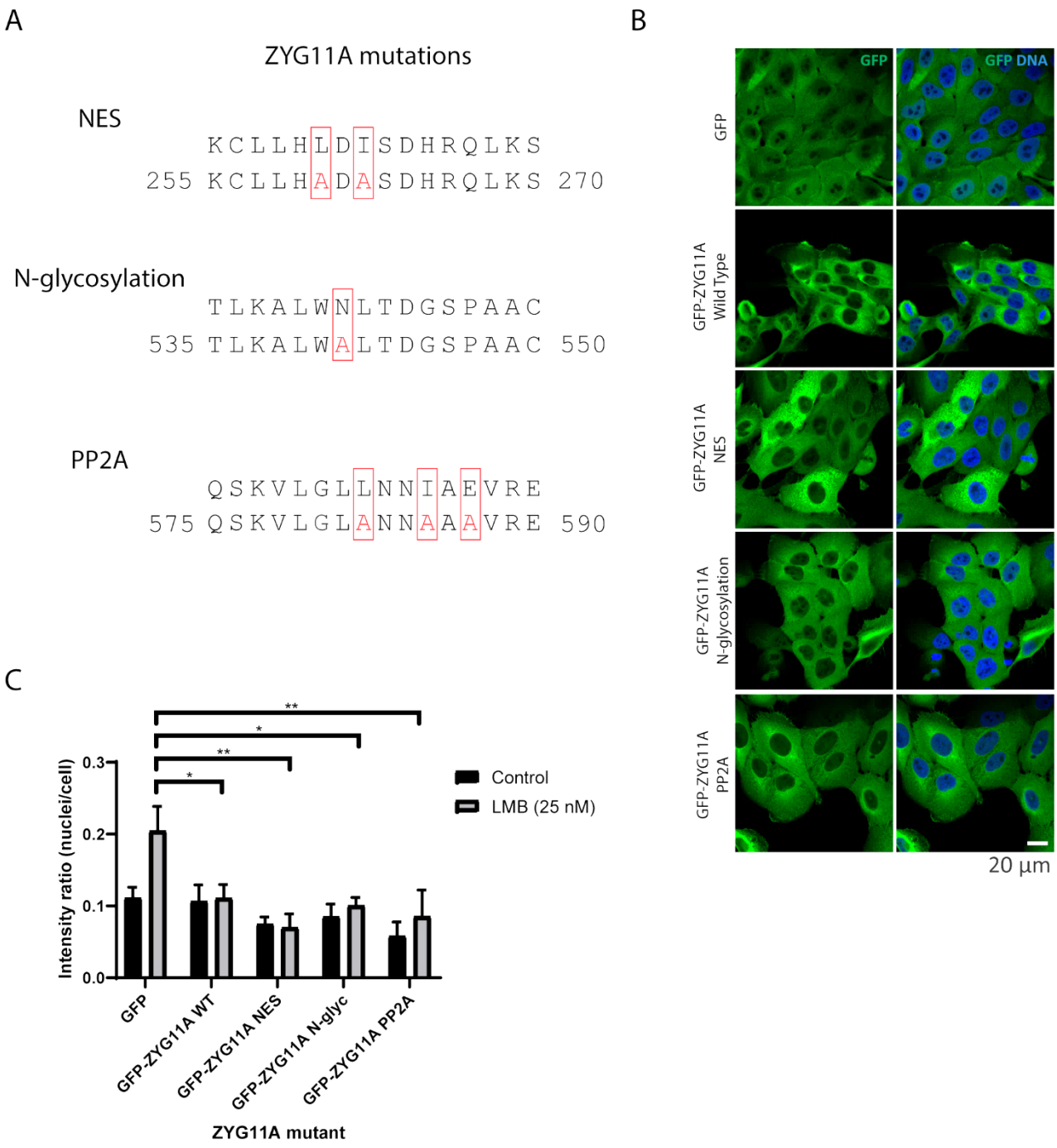



**Figure 3.4 Mutations in highly conserved ZYG11A protein interaction motifs do not alter the localisation of the fusion proteins**

**A.** Alignment of the putative interaction motifs for the following highly conserved sequences: NES signal, N-glycosylation site and PP2A-B56 docking site. Each sequence is accompanied by specific point mutations that have been introduced to the GFP-ZYG11A FL fusion protein. **B.** Representative confocal images of stable cell lines expressing the mutated GFP fusion proteins. **C.** Quantification of the levels of GFP in the nucleus upon LMB treatment measured by the intensity ratio of GFP in the nucleus to the whole cell. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, * = $p \leq 0.05$, ** = $p \leq 0.01$, Two-way ANOVA).

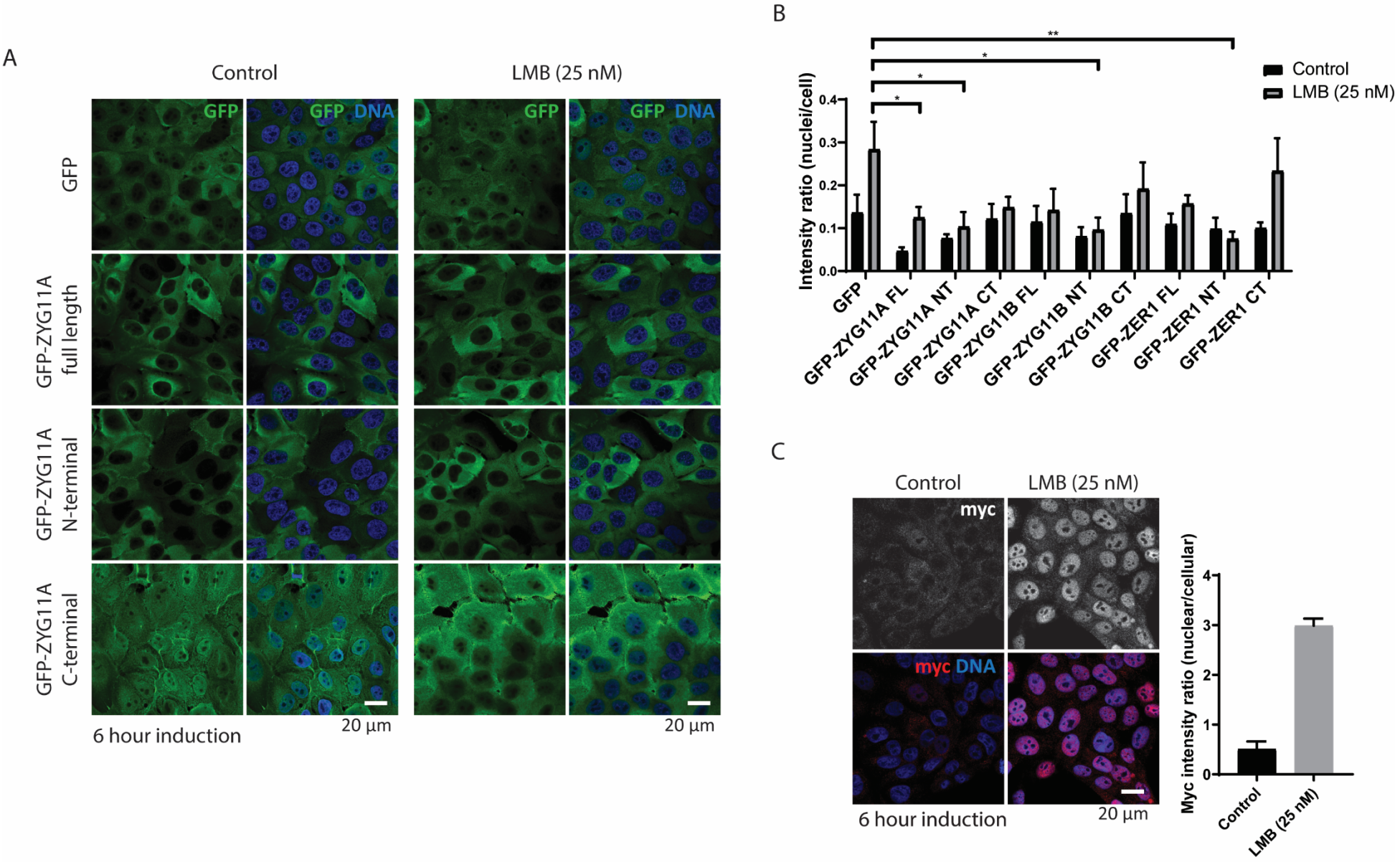

A
Control
LMB (25 nM)
GFP
GFP DNA
GFP
GFP-ZYG11A full length
GFP-ZYG11A N-terminal
GFP-ZYG11A C-terminal
6 hour induction
20 µm
20 µm
B
Control
LMB (25 nM)
Intensity ratio (nuclei/cell)
0.0
0.1
0.2
0.3
0.4
*
*
*
**
GFP
GFP-ZYG11A FL
GFP-ZYG11A NT
GFP-ZYG11A CT
GFP-ZYG11B FL
GFP-ZYG11B NT
GFP-ZYG11B CT
GFP-ZER1 FL
GFP-ZER1 NT
GFP-ZER1 CT
C
Control
LMB (25 nM)
myc
myc DNA
6 hour induction
20 µm
Myc intensity ratio (nuclear/cellular)
0
1
2
3
4
Control
LMB (25 nM)

**Figure 3.5 ZYG11A fusion proteins localise to the cytoplasm and are excluded from the nucleus regardless of nuclear export inhibition**

**A.** Representative confocal images highlighting the normal morphology of MCF10A cells stably expressing either GFP only, GFP-ZYG11A full length (FL), GFP-ZYG11A N-terminal (NT), and GFP-ZYG11A C-terminal (CT) fusion protein constructs. GFP marks the location of all of the GFP fusions and the nucleus is visible by DAPI staining (blue). Note the cytoplasmic localisation of both the FL+NTT ZYG11A fusion proteins compared with the GFP control and CT ZYG11A fusions which are distributed throughout the cells (left panel set). The localisation of these fusion proteins following a 6-hour treatment of the nuclear export inhibitor, Leptomycin B (LMB) (right panel set). **B.** Quantification of the levels of GFP in the nucleus upon LMB treatment measured by the intensity ratio of GFP in the nucleus to the whole cell. Includes the ratios for a similar set of stable cell lines generated using ZYG11B and ZER1 GFP fusion constructs. The lower the ratio, the lower the amount of GFP in the nucleus. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, * = $p \leq 0.05$, ** = $p \leq 0.01$, Two-way ANOVA). **C.** Representative confocal images of MCF10A cells stained for Myc as a readout for nuclear accumulation following a 6-hour treatment with LMB (25 nM). The nuclear to cytoplasmic ratio is increased following LMB treatment in the quantification compared to the untreated control.

Together, this suggests that the exportin complex alone is not the sole reason why these proteins are translocated from the nucleus. One potential explanation for situations where LMB has no effect on nuclear transport could be that the protein requires additional post-translational modifications (e.g. phosphorylation) or protein-protein interactions in order to translocate into the nucleus. The highly conserved nature of the mutated protein domains within ZYG11A do not appear to play a role in its cytoplasmic localisation, nor do the data suggest that these domains directly interplay with the exportin complex in relation to ZYG11A nuclear shuttling.

### 3.2.2 Validation of ZYG11 family knockdown levels and the assessment of genetic compensation between ZYG11 family members

In the initial screen that identified ZYG11A as a novel regulator of mammalian cell polarity and cell morphology (Smith *et al.*, 2016), both ZYG11B and ZER1 were not included in the high-content siRNA library. Therefore, it was important to identify whether the loss of other closely related ZYG11 family members was able to phenocopy the cell morphology alterations seen in *ZYG11A* knockdown ($ZYG11A^{KD}$) cells. In order to confidently assess this, knockdown of mRNA expression by the respective siRNA was validated in the human embryonic kidney cell line, HEK293T cells. HEK293T cells were reverse transfected (i.e.

hit with the siRNA prior to cells adhering to the plate) with a multiplex of commercially available siRNAs alongside an OTP-NT (ON-TARGETplus non-targeting) control. Here, the efficiency and specificity of each siRNA was assessed by quantitative real-time PCR (qRT-PCR) at the assay endpoint of 72 hours post-transfection. Compared with the OTP controls *ZYG11A* mRNA was confirmed to be knocked down by ~60% (Figure 3.6A), and high levels of silencing in *ZYG11B* and *ZER1* knockdown (ZYG11B$^{KD}$ and ZER1$^{KD}$) were observed (Figure 3.6B-C). I then went on to confirm these results in the non-transformed human mammary epithelial cell line, MCF10A grown in 2D cultures. Although I was able to confirm the knockdown of *ZYG11B* and *ZER1* in MCF10A cells, *ZYG11A* showed very low baseline expression in MCF10As and I was unable to confirm its knockdown following RNAi treatment (Figure 3.6D-F). Since genetic compensation is a common occurrence between many closely related gene families, the mRNA levels of each ZYG11 family member were also assessed in all knockdown samples. In the case of HEK293T ZYG11A$^{KD}$ cells, *ZYG11B* levels remained unchanged while *ZER1* levels were significantly decreased in the siZYG11A samples compared to the OTP controls (Figure 3.6B-C). On the other hand, *ZYG11B* knockdown resulted in a significant increase in *ZYG11A* expression only in HEK293T cells (Figure 3.6A), while similar increases in *ZER1* expression were consistently observed in ZYG11B$^{KD}$ cells from both HEK293T and MCF10A cell lines (Figure 3.6C and F). Interestingly, when *ZER1* was knocked down the levels of *ZYG11B* were significantly increased in MCF10A cells with a similar trend emerging in HEK293T cells (Figure 3.6C and F). Whilst I was not able to assess the effects *ZYG11A* knockdown on *ZYG11A* expression in MCF10A due to low baseline mRNA expression levels, the complementary analysis in HEK293Ts confirmed that each siRNA was functioning efficiently and specifically. These data show that the loss of *ZYG11B* can induce the expression of both *ZYG11A* and *ZER1*, and conversely, the loss of *ZER1* can induce the expression of *ZYG11B*. Despite having obtained a number of commercially available antibodies to ZYG11A, ZYG11B and ZER1, none of these antibodies were sensitive enough to detect the endogenous ZYG11 family members by western blot analysis in any of the cell lines tested. Hence, I was unable to confirm that mRNA expression changes led to expected alterations of the protein levels.

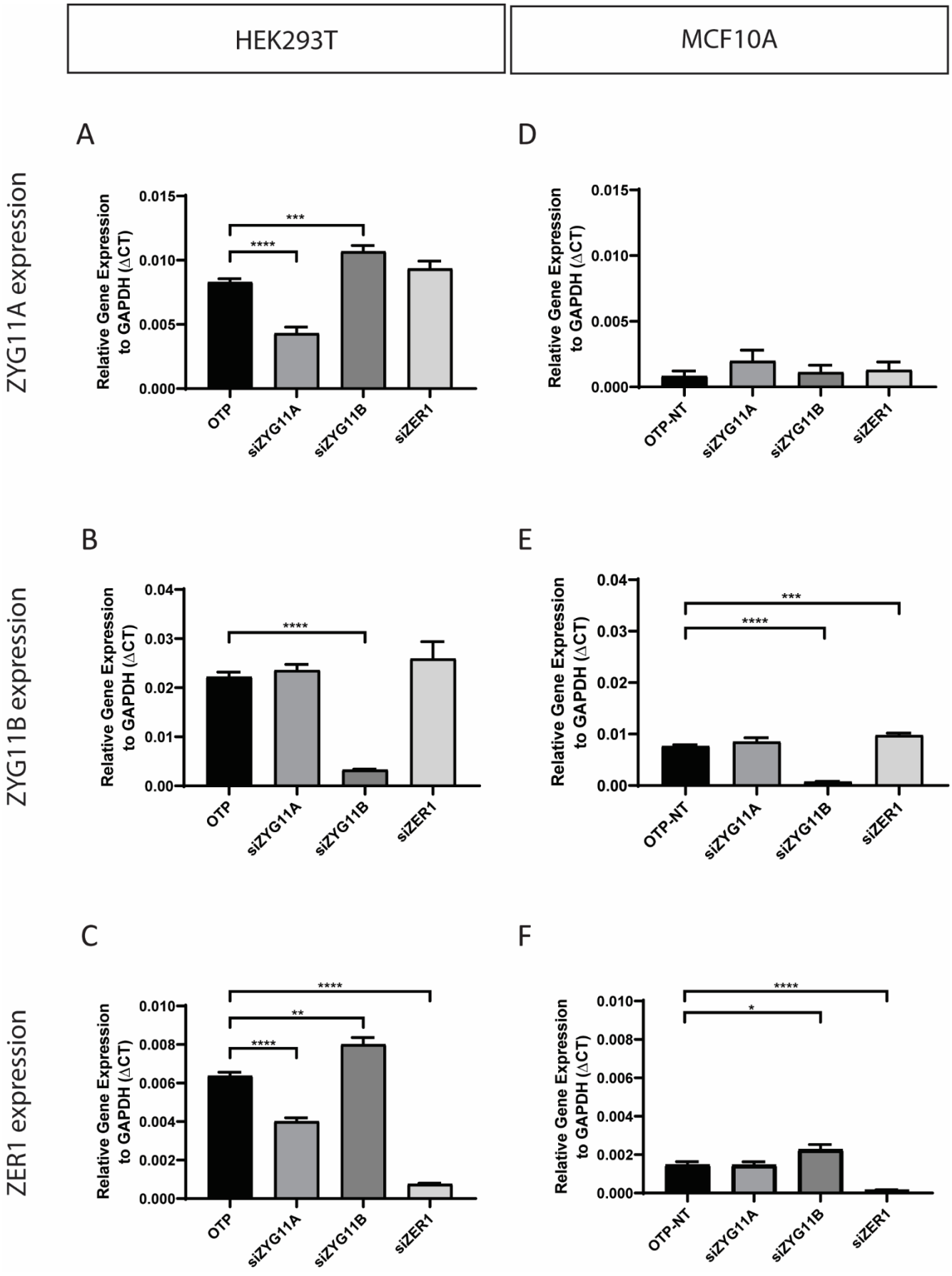

HEK293T
MCF10A
A
D
ZYG11A expression
B
E
ZYG11B expression
C
F
ZER1 expression
Relative Gene Expression to GAPDH (ΔCT)
OTP
OTP-NT
siZYG11A
siZYG11B
siZER1

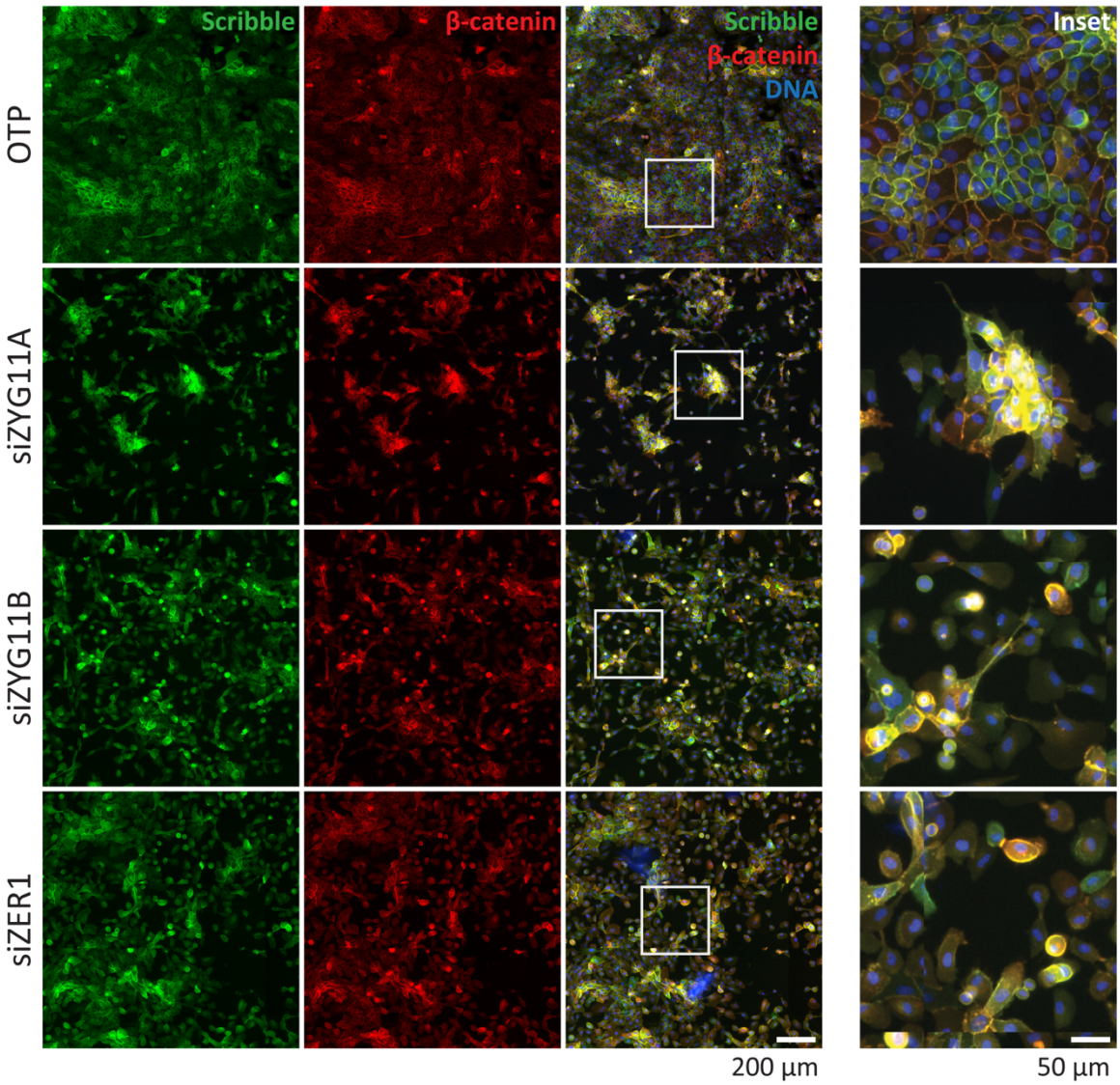
**Figure 3.6 ZYG11A loss does not induce transcriptional compensation by other ZYG11 family members**

The quantitative analysis of mRNA expression following the reverse transfection of HEK293T cells (**A-C**) or MCF10A cells (**D-F**) using either the OTP non-targeting control or siRNAs targeting *ZYG11A*, *ZYG11B* and *ZER1* (40 nM siRNA concentration). Graphs show the expression of *ZYG11A* (**A, D**), *ZYG11B* (**B, E**) and *ZER1* (**C, F**) at 48 hours post-transfection relative to GAPDH. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$, students t-test).

### 3.2.3 ZYG11 family members show distinct alterations in cell shape and cell polarity

Alterations in normal cell morphology, particularly in epithelial cells, are often an indication that apicobasal cell polarity has been compromised. From previously conducted functional screening (Smith *et al.*, 2016), *ZYG11A* was identified as a novel regulator of cell shape, however whether its highly conserved counterparts act similarly was still unknown. Similar to the screening conditions, 2D monolayer cultures of MCF10A normal epithelial cells were used to assess the individual knockdown profiles of each ZYG11 family member. The assay timepoints were optimised such that the OTP control cells would reach confluency at 72 hours post-transfection, exhibiting the characteristic cobblestone pattern of confluent MCF10A cells, emphasised by the presence of Scribble and β-catenin staining at the cell membrane (Figure 3.7A, top row). The knockdown of *ZYG11A* displayed a strikingly different phenotype to the OTP control resulting in cells that formed highly compact and almost clumped cell islands that were able to grow atop one another and maintain extended elongations between the cell clumps (Figure 3.7A-B, second row). In contrast to *ZYG11A* knockdown and to wildtype MCF10A cells that, under normal circumstances grow in small colonies when at a low cell density, both $ZYG11B^{KD}$ and $ZER1^{KD}$ cells showed a tendency to remain as single rounded cells (Figure 3.7A, bottom two rows). Of note, the overexpression of ZYG11 family members did not show any effects on normal cell shape (Figure 3.1B and Figure 3.2).

A


Scribble
β-catenin
Scribble
β-catenin
DNA
Inset
OTP
siZYG11A
siZYG11B
siZER1
200 µm
50 µm

B

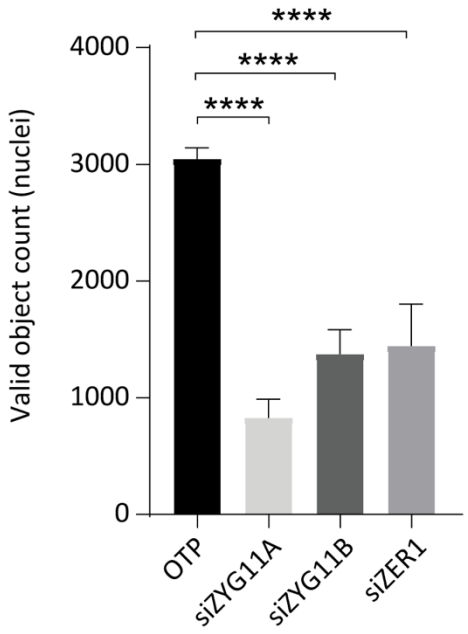

****
****
****
4000
3000
2000
1000
0
Valid object count (nuclei)
OTP
siZYG11A
siZYG11B
siZER1

**Figure 3.7 The ZYG11A knockdown phenotype is distinct from the other two ZYG11 family members**

**A.** Representative confocal images of OTP non-targeting, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. Images display staining for Scribble (green), β-catenin (red) and a merged set of images with the nuclear marker DAPI (blue). **B.** 4X zoomed in images of **A** focusing on regions of high phenotypic variation. **C.** Graph representing the valid object count of remaining adherent cells depicted in **A** based on DAPI staining. Error bars represent ± SEM and are representative of 2 independent experiments each performed in duplicate. (n=4, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$, Students t-test).

Overlapping and clumping of cells in epithelia like those seen in ZYG11A$^{KD}$ is often a sign that the internal regulatory mechanisms of cell polarity have been altered. Although opposite in nature, lack of cell-cell connection displayed by ZYG11B$^{KD}$ and ZER1$^{KD}$ cells is also a potential indicator of polarity defects. Taken together, my data suggest differences as to how each family member affects cell shape and highlights ZYG11A as a distinct member of the ZYG11 family in terms of cell shape regulation.

Since *ZYG11A* knockdown has been previously associated with the mislocalisation of Scribble from the membrane (Lorey Smith, our lab's unpublished results) it was important to further characterise the Scribble (and β-catenin) profile of these cells. In terms of *ZYG11A* knockdown, very few cells remained adhered to the plate at the 72-hour assay end point. Of those left there were no obvious alterations in the location of either proteins with both Scribble and β-catenin present at the cellular membranes (Figure 3.8). Where cells were more confluent (e.g. OTP and siZER1) or where cell clumping was present (siZYG11A), there appeared to be brighter Scrib and β-catenin staining, however this may be due to the concentration of cells in a single area. Additionally, key cytoskeletal proteins (F-actin and α-tubulin) and adhesion proteins that bind to the basement membrane (Integrin α6) were also examined since the cells displayed a reduced capacity to adhere to the coverslips. Although the staining of F-actin appeared to be lower in the ZYG11 family knockdown samples (Figure 3.9), this was also seen in control cells with low plating density (data not shown).

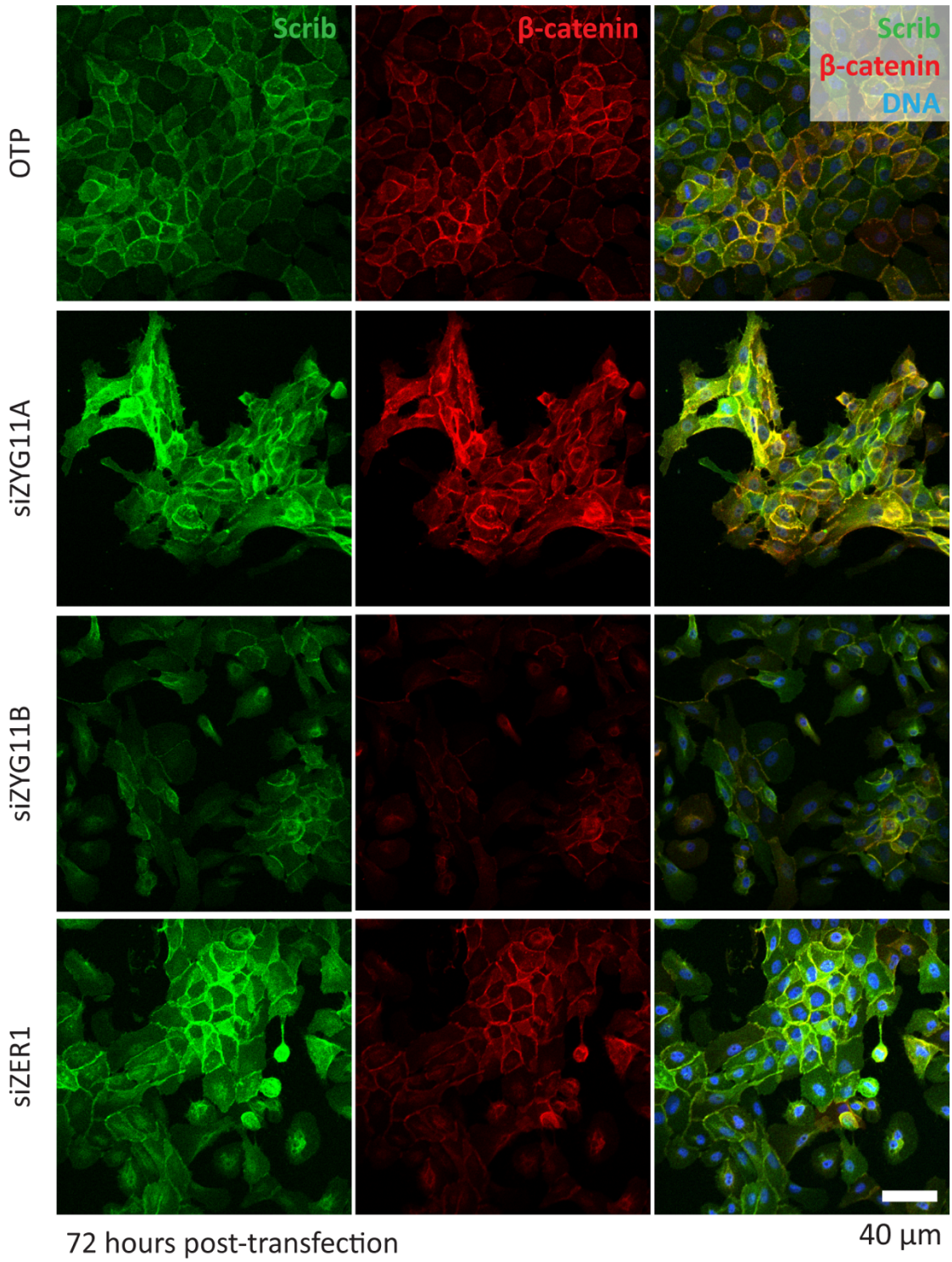



**Figure 3.8 ZYG11 family knockdown does not affect the membrane localisation of Scribble and β-catenin**

Representative confocal images of OTP, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. Individual stains of Scribble (green) and β-catenin (red) are merged in the last column, along with the nuclear stain, DAPI (blue). Note the concentration of Scribble and β-catenin staining at the cell membrane where cells are in contact with one another.

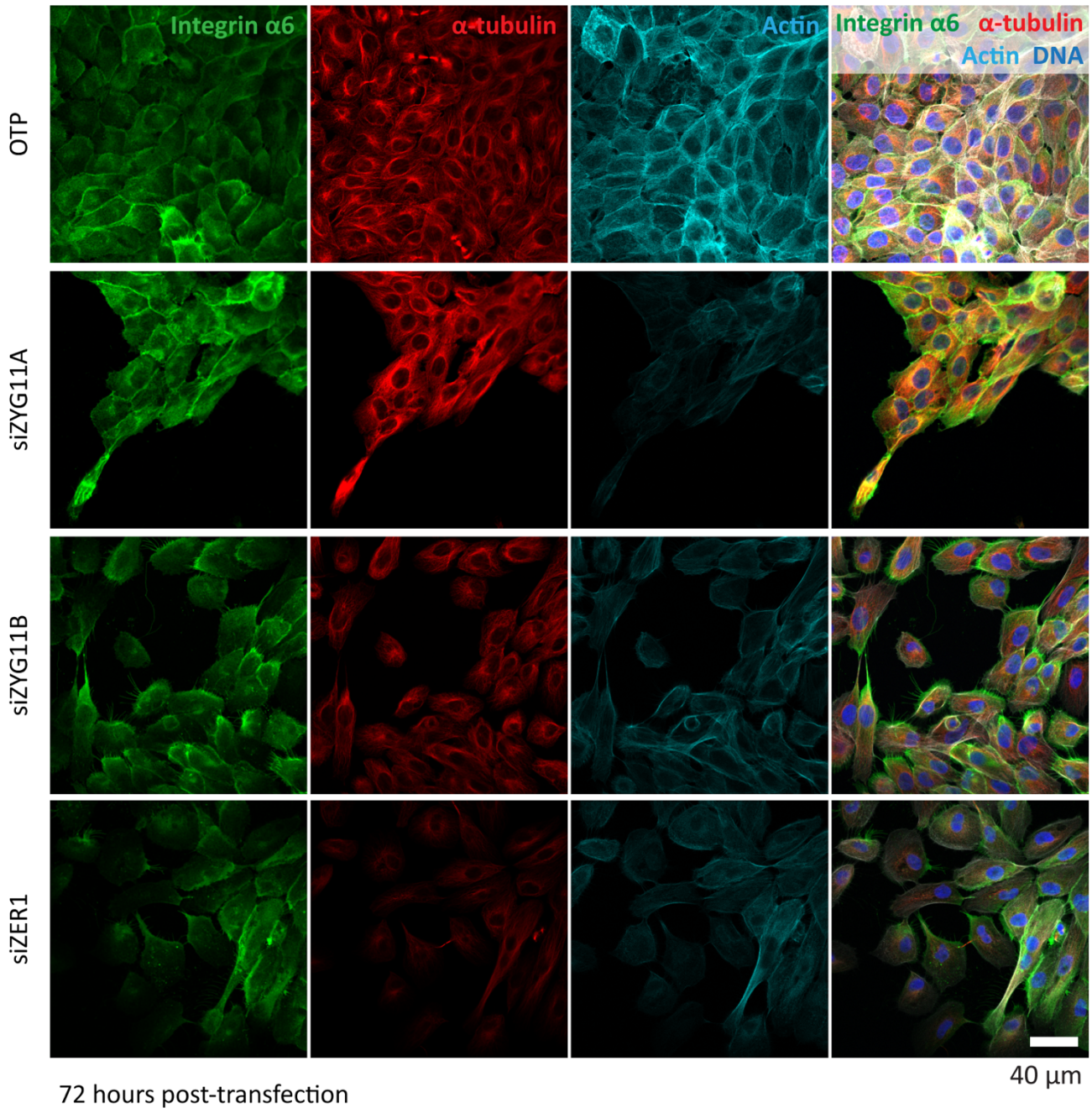
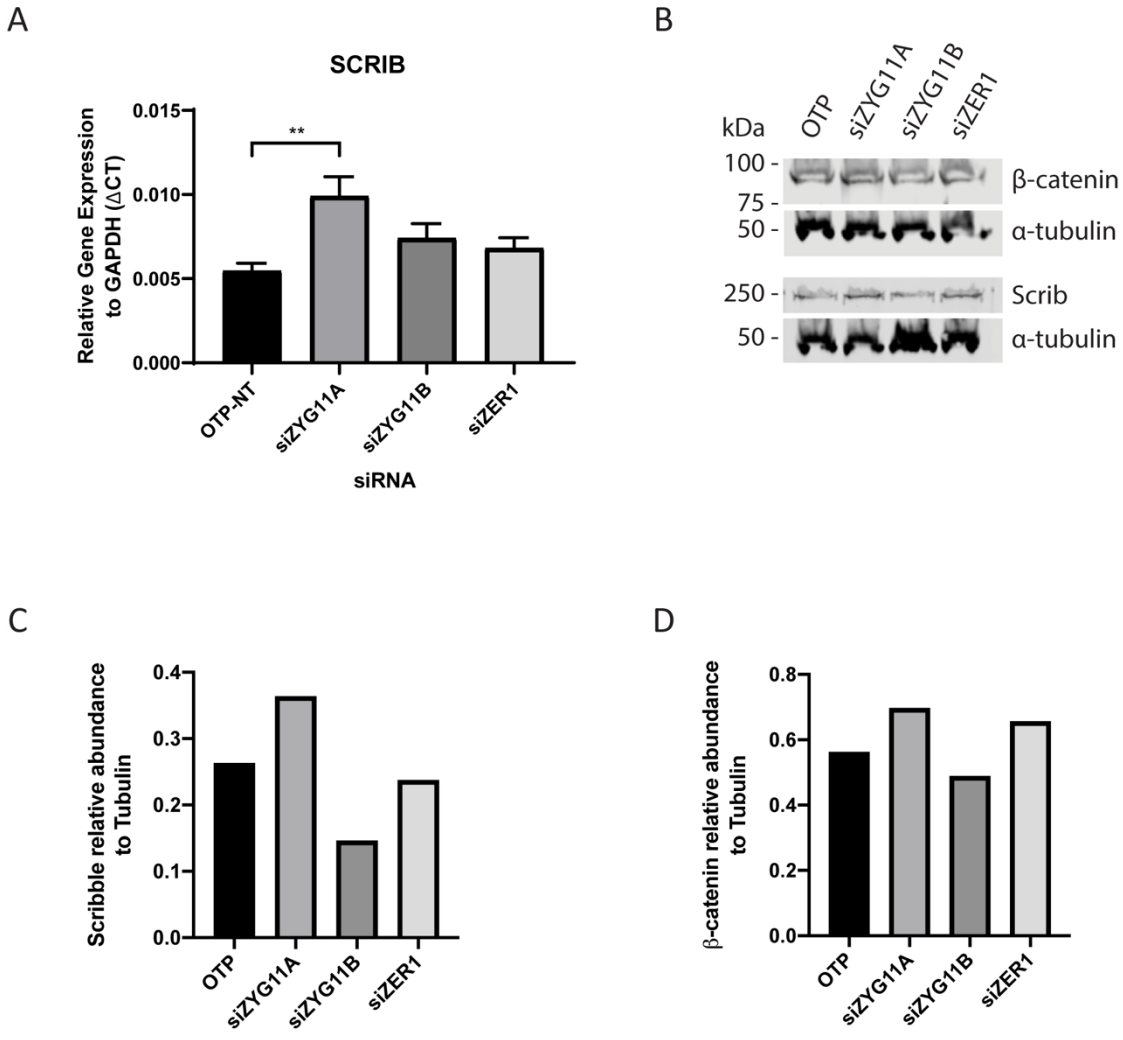



**Figure 3.9 Loss of ZYG11 family members show minimal alterations in proteins related to the cytoskeleton and adhesion**

**A.** Representative confocal images of OTP, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. Individual stains of Integrin α6 (green), α-tubulin (red) and F-actin (cyan) are merged in the last column, along with the nuclear stain, DAPI (blue).

Next, I aimed to determine the protein levels of Scribble and β-catenin within the ZYG11 knockdown cells and found that while there were minor changes in β-catenin, Scribble protein levels appeared to be slightly elevated in *ZYG11A* knockdown cells (Figure 3.10B-C), consistent with my IF studies. To assess whether this effect was due to the direct de-regulation of Scribble proteostasis, qRT-PCR was conducted, revealing significantly increased *SCRIB* mRNA expression levels in *ZYG11A* knockdown cells (Figure 3.10A).

**Figure 3.10 Loss of ZYG11A slightly increases Scribble protein levels and significantly increases SCRIB transcription levels**




**A.** Real-time PCR analysis of *SCRIB* expression in control OTP, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. mRNA levels are quantified relative to GAPDH. **B.** Western blot of the samples in **A** displaying cell adhesion marker, β-catenin (top) and cell polarity marker, Scribble (bottom) protein levels. **C-D.** Relative abundance of Scribble and β-catenin compared to Tubulin. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, ** = p≤0.01, Student's t-test).

Little is known about the regulation of Scribble at the mRNA or protein level and taken together the data present a novel regulatory mode of *SCRIB* expression in MCF10A cells that can be linked to ZYG11A. Whether this is occurring directly via ZYG11A or by the ubiquitination of another protein by ZYG11A will be a key point to address. In addition, as *SCRIB* overexpression in MCF10A cells actually improves adhesion by generating more tight junctions and makes the cells more epithelial-like and cuboidal in shape (Elsum, Martin and Humbert, 2013), it is unlikely that the increased Scribble expression levels alone can account for the altered cell shape of *ZYG11A* depleted MCF10A cells.

### 3.2.4 ZYG11A loss induces cell death

Despite morphology differences, all of the ZYG11 family knockdowns lacked the ability to form a confluent monolayer at the 72-hour timepoint, which from herein will be referred to as a reduction in 'cell plating density'. Consistent with the morphological similarities observed in the *ZYG11B* and *ZER1* knockdown samples, both knockdowns significantly decreased the cell plating density by 50%, and in *ZYG11A* knockdowns showed a 70% decrease (Figure 3.7B). Aberrant changes in a cell's ability to form a confluent monolayer could be explained by either an increase in cell death, a decrease in cell proliferation, or a loss of attachment to the plates.

A cells ability to maintain a homeostatic environment, while certainly effective, is not robust enough to protect a cell in every situation. When overcome with irreparable cellular aberrations, such as the loss of a gene, the cell death cascade is subsequently triggered in order to eliminate the defective cell. To delineate whether cell death was a contributing factor in the reduced plating density, the siRNA transfected cells were analysed for markers of programmed cell death – otherwise known as apoptosis. By 48 hours post transfection, OTP controls showed minimal levels of cell death based on Annexin V and PI staining (Figure 3.11A, top left). Under knockdown conditions, only *ZYG11A* loss showed signs of apoptosis induction via Annexin V binding and PI staining (Figure 3.11A, top right), with both ZYG11B$^{KD}$ and ZER1$^{KD}$ remaining consistent with control apoptosis levels (Figure 3.11A, bottom plots). The increase in apoptotic cells as a result of *ZYG11A* loss was only by ~15% (Figure 3.11B) which is considerably low and proportionally unlikely to be completely accountable for such a large decrease in cell plating density. As alluded to earlier, direct alterations in the cell death pathway are not always the trigger for apoptotic cell events, rather, they are commonly activated following upstream changes that interfere

with normal cellular processes. Therefore, the minimal increase in cell death could be an indication that perhaps functional aberrations elsewhere might be preventing the cells from repopulating the plate, and that cell death might simply be secondary to these changes.

In the cells overexpressing the GFP tagged ZYG11 family constructs there were no changes in cell viability (Figure 3.12A-B), although testing this under cell death sensitisation may reveal effects on cell survival.

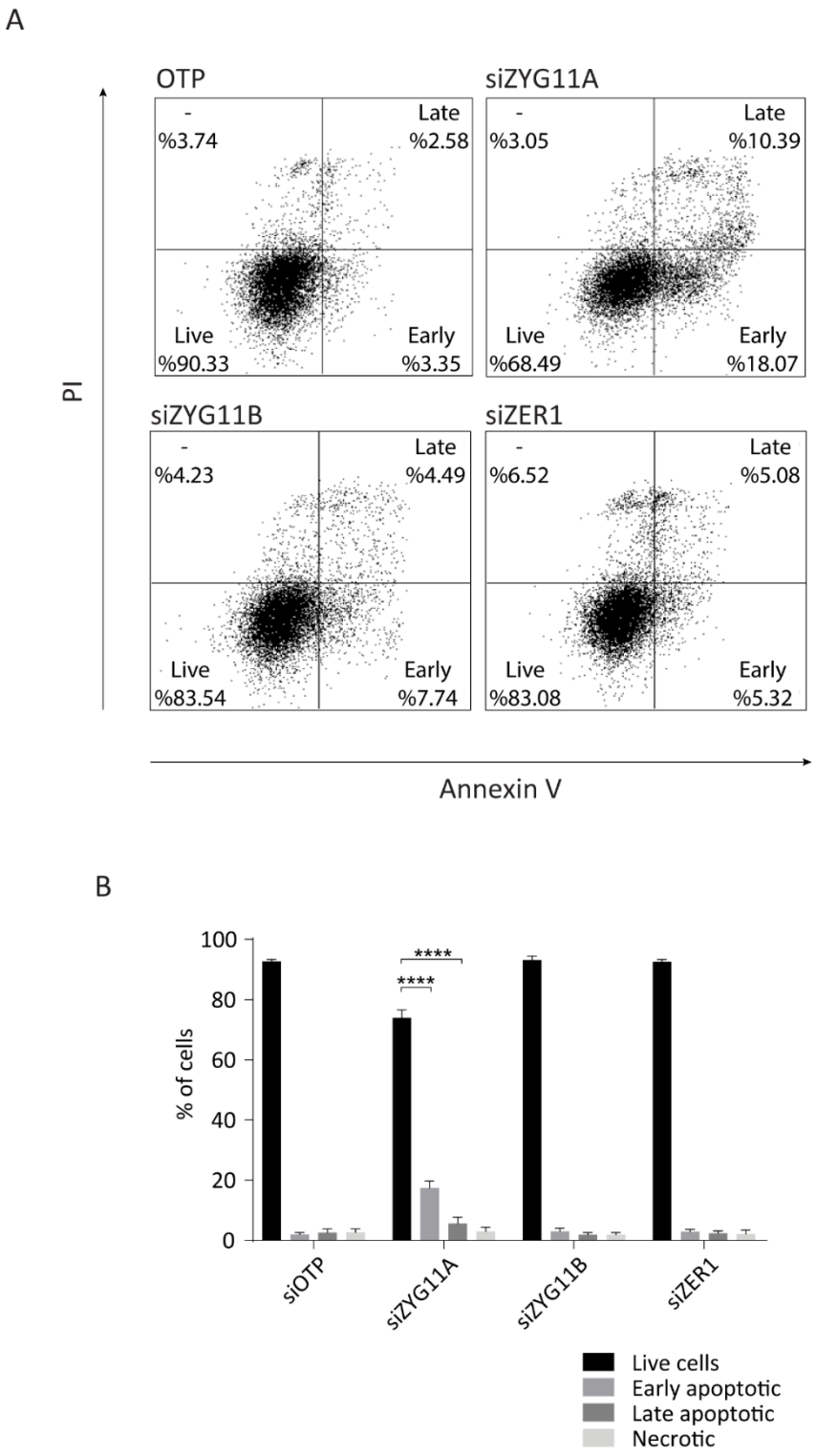



**Figure 3.11 Early apoptotic events are increased in ZYG11A knockdown cells**

**A.** Viability of OTP non-targeting controls, siZYG11A, siZYG11B and siZER1 knockdown MCF10A cells at 48 hours post-transfection. Annexin V and PI staining is used to determine cell death levels and the individual FACS plots represent live (bottom left quadrant), early apoptotic (bottom right quadrant) and late apoptotic (top right quadrant) and necrotic (top left quadrant) cells. **B.** Quantification of the plots in **A** assessed using 10,000 events per sample. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, **** = $p \leq 0.0001$, Two-way ANOVA).

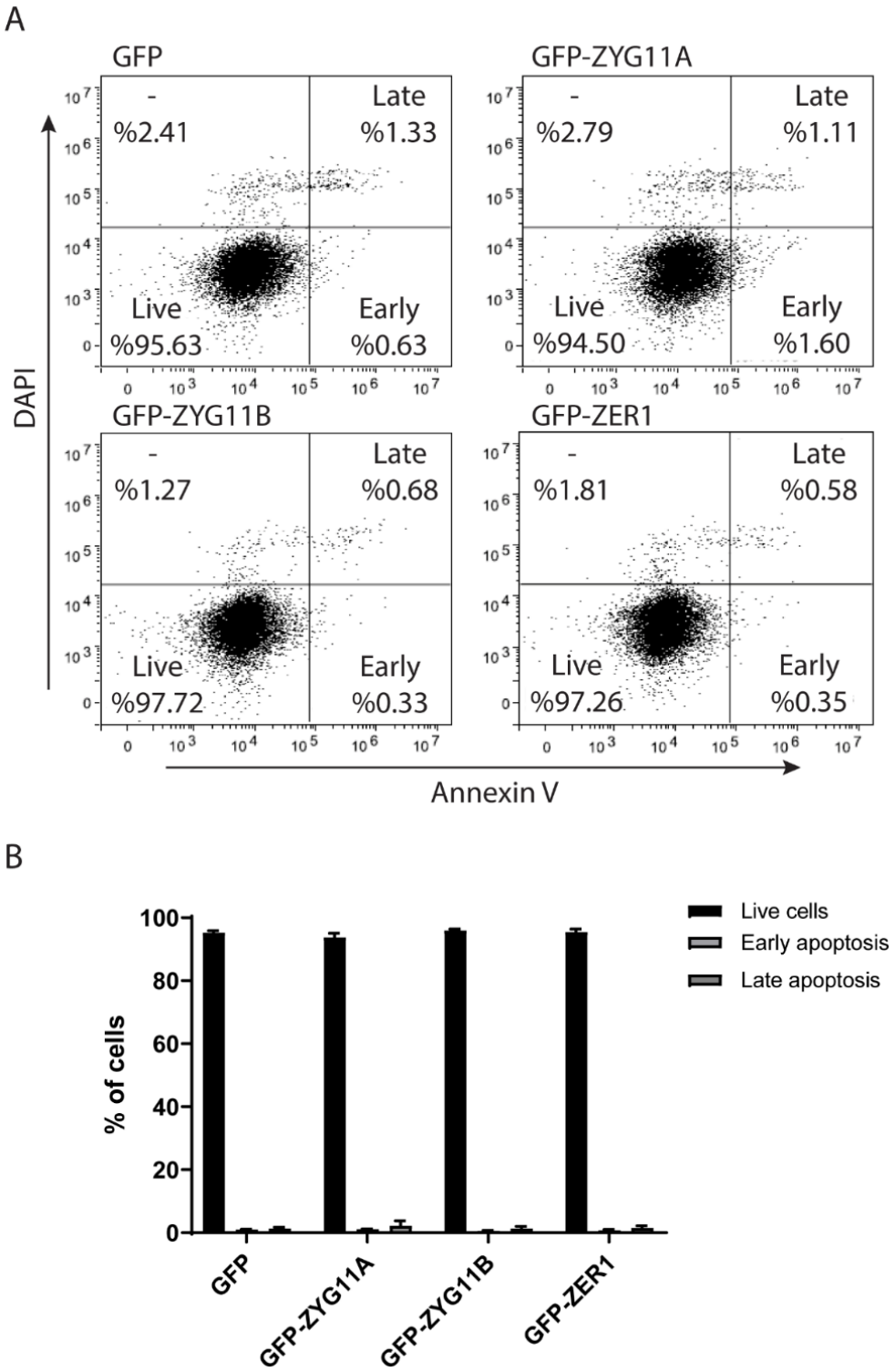



**Figure 3.12 Overexpression of ZYG11 family fusion proteins does not affect cell death in MCF10A cells**

**A.** Viability of MCF10A cell lines stably expressing either GFP only, GFP-ZYG11A full length, GFP-ZYG11B full length, or GFP-ZER1 full length fusion protein constructs at ~70% confluence. Annexin V and DAPI staining marks cell death levels with the individual FACS plots representing live (bottom left quadrant), early apoptotic (bottom right quadrant) and late apoptotic (top right quadrant) and necrotic (top left quadrant) cells. **B.** Quantification of the plots in **A** assessed using 10,000 events per sample. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, Two-way ANOVA).

#### 3.2.4.1 ZYG11A and ZER1 are required for proliferation and the progression through the G0/G1 stage of the cell cycle

The longstanding connection between the ZYG11 family and cell cycle has been extensively studied in the roundworm where *zyg-11* mutants exhibit an extended duration of meiosis II, show increases in cyclin B1, fail to condense chromosomes, and have a delay in mitosis (Liu, Vasudevan and Kipreos, 2004; Sonneville and Gonczy, 2004; Vasudevan, Starostina and Kipreos, 2007; Balachandran *et al.*, 2016). More recently, these functions were found to be conserved in human U2OS cells where ZYG11A and ZYG11B were shown to act redundantly with another E3 ubiquitin ligase, APC/C, to degrade Cyclin B1 at the G2 to M phase transition (Balachandran *et al.*, 2016; Brandeis, 2016). Since ZYG11 family knockdown cells were all unable to form a confluent monolayer, to test whether the loss of these genes was similarly affecting cell cycle and proliferation, transfected cells were pulse labelled with BrdU for 1 hour and stained with 7-AAD to visualise total DNA content. Following the knockdown of either *ZYG11A* (Figure 3.13A, top right) or *ZER1* (Figure 3.13A, bottom right), the cell cycle profiles suggested an arrest at the G0/G1 phase of the cell cycle as evidenced by a significant increase in G0/G1 cells from 50% of the cell population in OTP controls to about 60-65% in *ZYG11A* and *ZER1* knockdown cells. Furthermore, this increase in G0/G1 cells was accompanied by a decrease in S phase cells from 30% in OTP controls to roughly 5% and 10% in siZYG11A and siZER1 samples, respectively (Figure 3.13B). In contrast, *ZYG11B* knockdown showed normal cycling conditions compared to controls (Figure 3.13A, bottom left). Furthermore, consistent with the proposed redundant role of ZYG11 family members in G2 to M phase transition only in the absence of APC/C function, we observed no differences in the proportion of G2/M phase cells 48 hours after siRNA treatment.

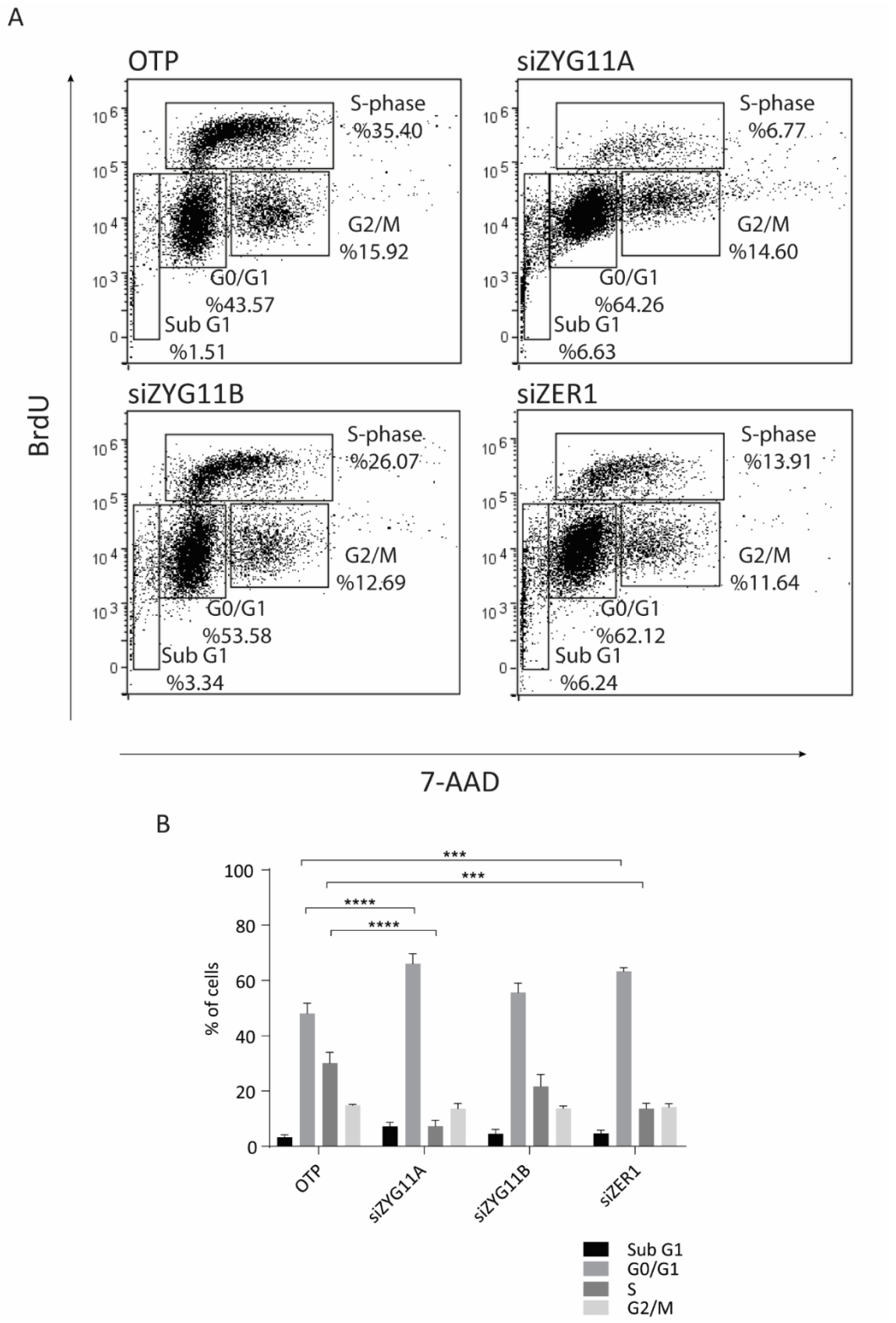



**Figure 3.13 ZYG11A and ZER1 loss affect cell cycle at the G0/G1 to S phase transition**

**A.** Cell cycle analysis of OTP non-targeting controls, siZYG11A, siZYG11B and siZER1 knockdown MCF10A cells at 48 hours post-transfection. Individual FACS plots show each cell cycle phase based on DNA content (7-AAD) and DNA synthesis (BrdU incorporation). **B.** Quantification of the plots in **A** assessed using 10,000 events per sample. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, *** = p≤0.001, **** = p≤0.0001, Two-way ANOVA).

Since Retinoblastoma 1 (RB1), p53 and p21 are often associated with arrests in G0/G1, and ZER-1 has been proposed to regulate Rb expression in a viral context (White *et al.*, 2012), I analysed the mRNA and protein expression of these genes in the knockdown cells. Of these targets, *RB1* mRNA expression was increased in ZYG11B$^{KD}$ cells (Figure 3.14A) while *p53* mRNA expression was only increased in ZER1$^{KD}$ cells (Figure 3.14B). Interestingly, the p53 transcriptional target, *CDKN1A* (p21), was overexpressed in all ZYG11 family knockdown cells with *ZYG11A* showing the greatest increase in mRNA levels (Figure 3.14C). Despite the changes in mRNA expression levels, the protein levels of these genes did not show visible increases in any of the targets in all samples (Figure 3.14D-E). To examine what the effect of overexpression of ZYG11 family members may be on cell cycle regulation, I examined cell cycle phasing in the various stable MCF10A cell lines I had previously generated. Surprisingly, overexpression of the ZYG11 family members did not show any effects on cell cycle exit (Figure 3.15A-B).

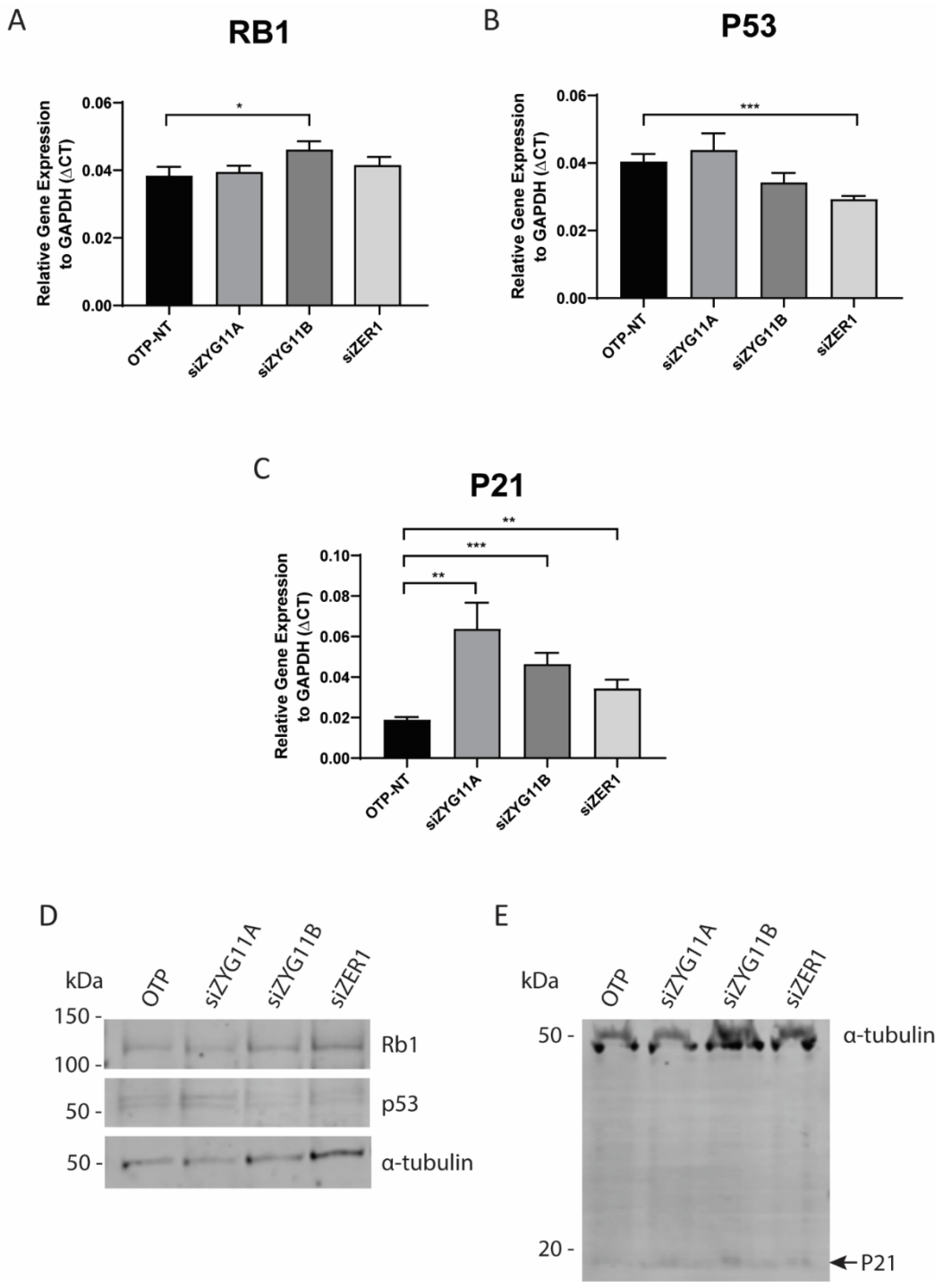



**Figure 3.14 ZYG11A knockdown increases transcription of the G0/G1 to S phase transition marker, P21**

Analysis of OTP control, siZYG11A, siZYG11B, siZER1 mRNA expression levels of G0/G1 phase cell cycle regulators at 72 hours post-transfection, including, *RB1* (**A**), *P53* (**B**) and *P21* (**C**). All samples are analysed relative to *GAPDH* mRNA levels. **D-E.** Protein levels of the cell cycle regulators mentioned in the above samples based on western blot analysis. Rb1 and p53 protein levels (**D**). p21 expression levels (**E**). Protein levels within each sample are compared to α-tubulin controls. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate transfection wells and triplicate wells for RT-PCR. (n=6, * = p≤0.05, ** = p≤0.01, *** = p≤0.001, Student's t-test).

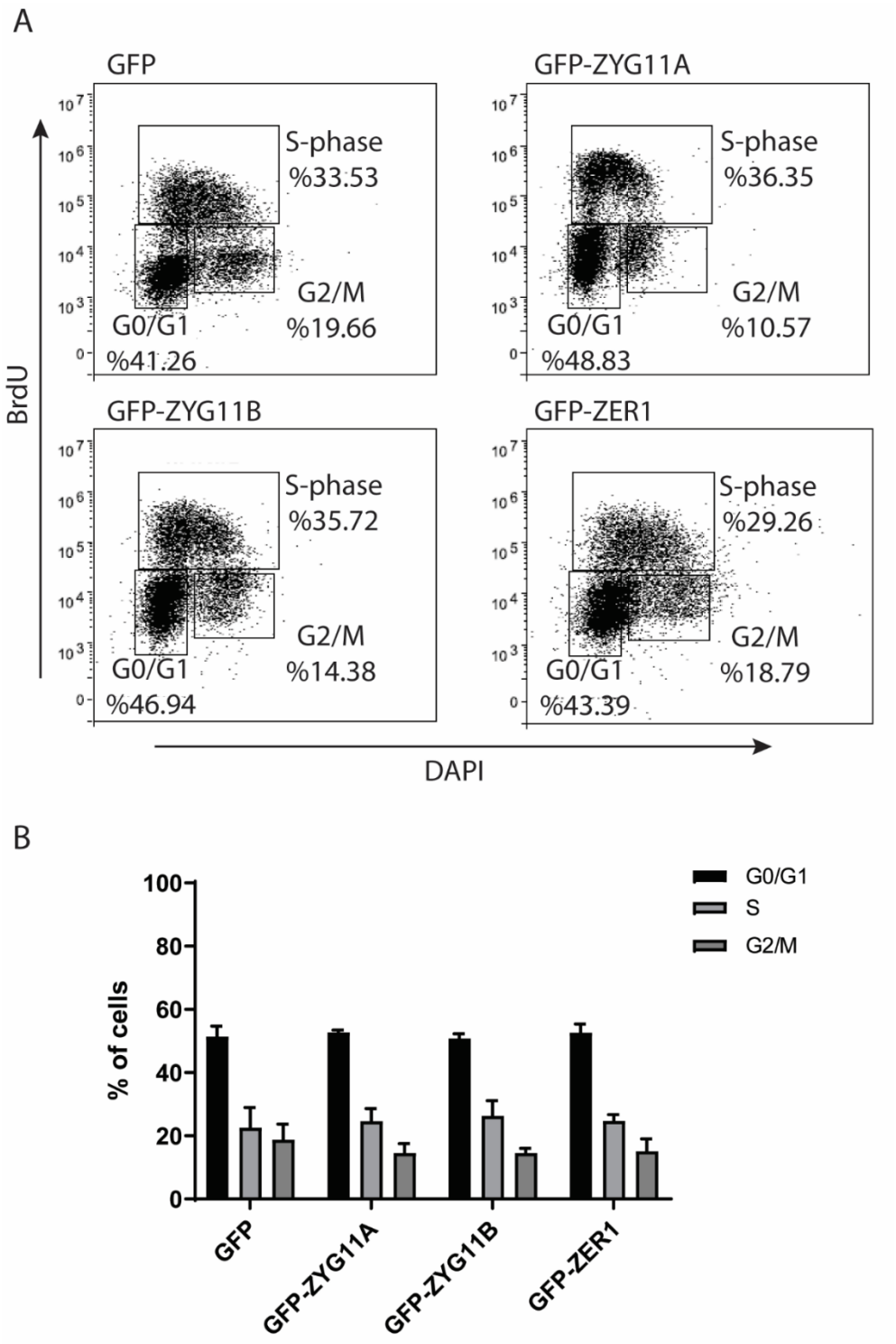



**Figure 3.15 Overexpression of ZYG11 family fusion proteins does not affect cell cycle in MCF10A cells**

**A.** Cell cycle profiles of MCF10A cell lines stably expressing either GFP only, GFP-ZYG11A full length, GFP-ZYG11B full length, or GFP-ZER1 full length fusion protein constructs at ~70% confluence. Each phase is determined by the levels of DNA synthesis (BrdU incorporation) and DNA content (DAPI) by FACS. **B.** Quantification of the plots in **A** assessed using 10,000 events per sample. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, Two-way ANOVA).

### 3.2.5 The ZYG11 family affects directed cell migration and invasion

To examine the effects of loss of ZYG11 family members in more detail, I conducted cell real-time observational studies using live-cell imaging at 24 hours post-transfection, which is the point where the knockdown has come into effect. Leading up to the 72-hour endpoint, consistent with my previous data, while the ZYG11B$^{KD}$ and ZER1$^{KD}$ cells showed a greater capacity to reach confluency (Figure 3.16, bottom two rows), the *ZYG11A* knockdown samples were unable to reach confluency and instead produced an abundance of clumpy islands (Figure 3.16, second row, arrows). Interestingly, these islands appeared to migrate as a mass of cells (Supplementary video 3.1.2).

Due to the inability of these cells to form a confluent monolayer, I was unable to test how they would respond to a different migration stimulus such as scratch induced wound healing. Instead I decided to use the Boyden chamber assay. To directly measure the migration capabilities of these cells, I tested how they responded to a chemotactic stimulus, EGF. With the control EGF starved cells setting a baseline level of migration towards the EGF stimulus, ZYG11A$^{KD}$ cells showed a reduced capacity to migrate. In contrast, there were significant increases in chemotaxis in ZYG11B$^{KD}$ and ZER1$^{KD}$ cells (Figure 3.17A-B). To take these results further, human Matrigel was added to the transwells to test the ability of these cells to invade through a matrix, and although not significant, similar trends were observed in ZYG11A$^{KD}$ and ZER1$^{KD}$ cells (Figure 3.17C-D). While it is clear that the knockdown of *ZYG11A* prevents the effective migration of MCF10A cells, the stark contrast with *ZYG11B* and *ZER1* cells, which increase cell migration, supports the idea that the ZYG11 family have developed divergent functions despite high evolutionary conservation. Consistent with apoptosis and cell cycle studies, overexpression of individual ZYG11 family members did not show any effects on cell migration (Figure 3.18A-B).

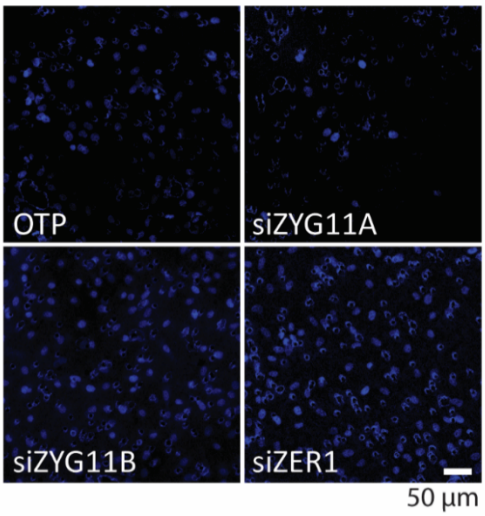
24h
36h
48h
60h
72h
OTP-NT
siZYG11A
siZYG11B
siZER1
200 µm

**Figure 3.16 Knockdown of *ZYG11A* causes cells to move in cell clusters**

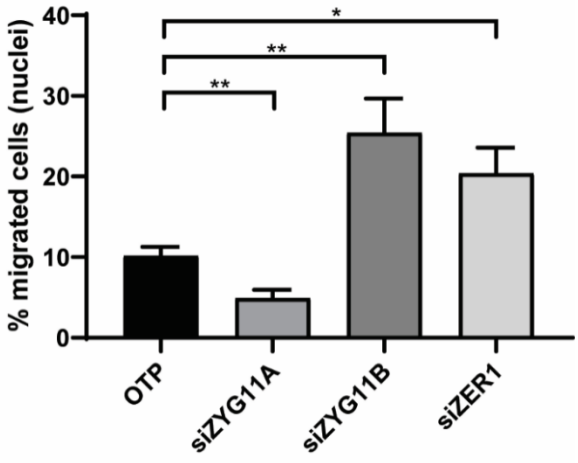
Representative brightfield images of live imaging done on MCF10A cells transfected with either OTP-NT (control), siZYG11A, siZYG11B or siZER1. Image timepoints begin at the 24-hour media change where the siRNAs have come into effect and end at the 72-hour assay endpoint where control cells have reached confluence. Image intervals are in 12-hour blocks. The siZYG11A cells maintain large cell clumps that migrate together over the 48-hour imaging period (black arrows). See Supplementary for full video.

Transwell migration
(- matrigel)

A

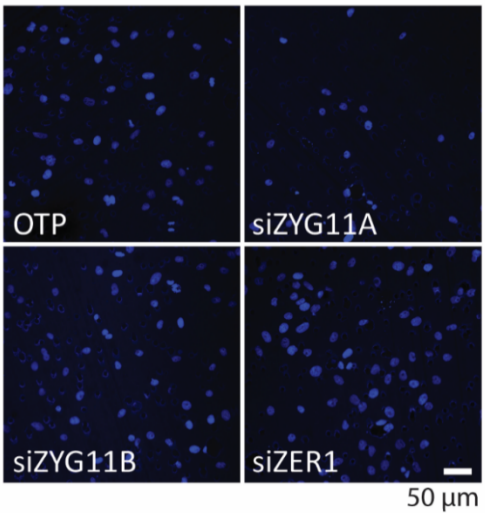



B




Transwell invasion
(+ matrigel)

C




D

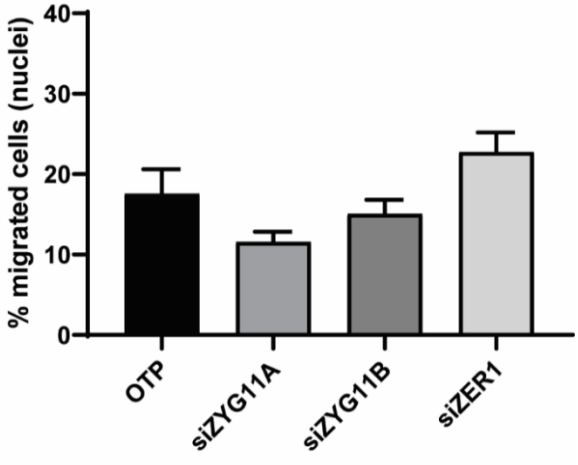


**Figure 3.17 ZYG11 family members can increase or decrease cell migration in response to an EGF stimulus**

**A.** Representative confocal images of the Transwell cell migration of OTP, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. Cells starved of EGF overnight that have migrated towards the EGF stimulus between the 48 to 72-hour period are stained with the DNA stain, DAPI (blue). **B.** Quantification of **A** as a percentage of the total number of seeded cells per well. **C.** Representative confocal images of Transwell invasion through a layer of human Matrigel of OTP, siZYG11A, siZYG11B and siZER1 MCF10A cells at 72 hours post-transfection. Cells that have invaded over the period of time mentioned in **A** are marked with DAPI (blue), quantified as above in **D**. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, * = $p \leq 0.05$, ** = $p \leq 0.01$, Student's t-test).

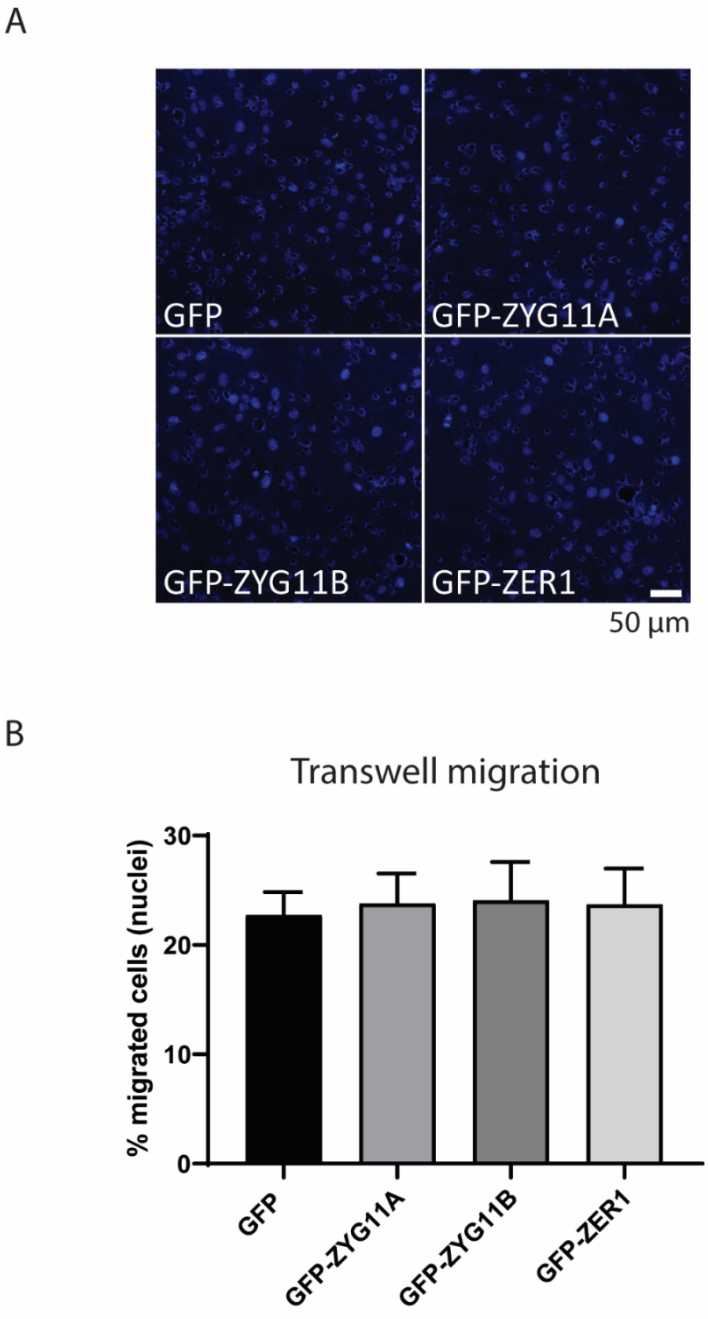


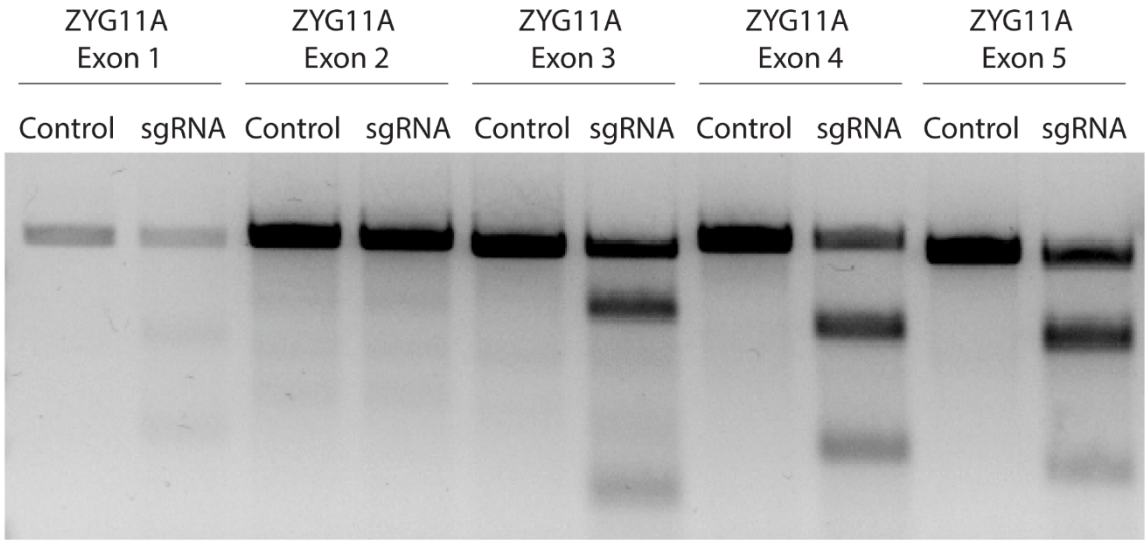

**Figure 3.18 Overexpression of ZYG11 family fusion proteins does not alter migration in response to an EGF stimulus**

**A.** Representative confocal images of MCF10A cell lines stably expressing either GFP only, GFP-ZYG11A full length, GFP-ZYG11B full length, or GFP-ZER1 full length fusion protein constructs. Cells starved of EGF overnight that have migrated towards the EGF stimulus over a 24-hour period are visible through the DNA stain, DAPI (blue). **B.** Quantification of **A** as a percentage of the total number of seeded cells per well. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, Student's t-test).

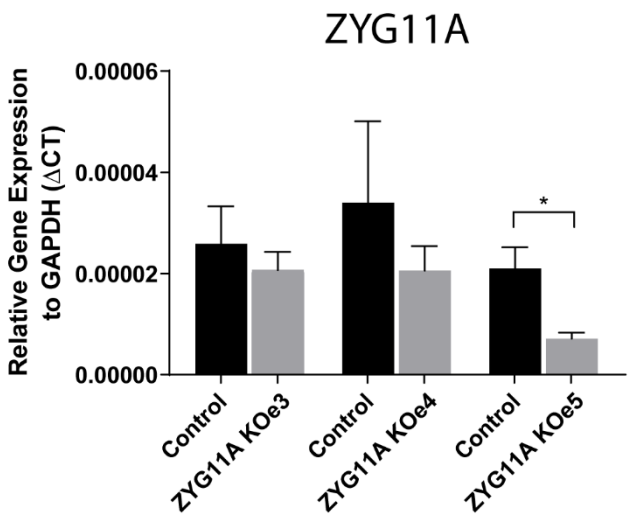
### 3.2.6 Generation and characterisation of *ZYG11A* CRISPR knockout MCF10A cell lines

To further investigate the *ZYG11A* knockdown phenotype obtained using the RNAi approach, I used CRISPR-Cas9 engineering to assess whether the complete knockout of the gene at the DNA level would produce similar or enhanced morphological defects. To achieve this, MCF10A cells harbouring the mCherry-Cas9 construct were transduced with single guide RNA (sgRNA) designed to target the *ZYG11A* sequence across five different exons (1-5). Although I initially tried to establish single cell clones using single cell sorting, control and sgRNA transduced cells immediately became elongated and formed spindle-like colonies in 2D cultures reminiscent of transformed cell morphology (data not shown). Since I was unable to generate single cell clones, the *ZYG11A* CRISPR lines were established as a heterogeneous population, and the CRISPR sgRNA was induced by the addition of doxycycline to the culture media. As mentioned earlier (see section **3.2.2**), *ZYG11A* expression is very low in MCF10A cells with endogenous protein levels undetectable by western blotting or immunofluorescence, thus an alternative T7 endonuclease I (T7EI) mismatch cleavage assay was used. This method identified successful genome editing in exons 3, 4, and 5 (herein referred to as ZYG11A$^{KOe3}$, ZYG11A$^{KOe4,}$ and ZYG11A$^{KOe5}$) as indicated by multiple bands present in the sgRNA lanes compared to the controls following gel electrophoresis (Figure 3.19A). Real-time PCR experiments conducted showed that only ZYG11A$^{KOe5}$ cells were able to reduce the expression of *ZYG11A* (Figure 3.19B), however moving forward, all three cell lines were still assessed based on the T7EI results.

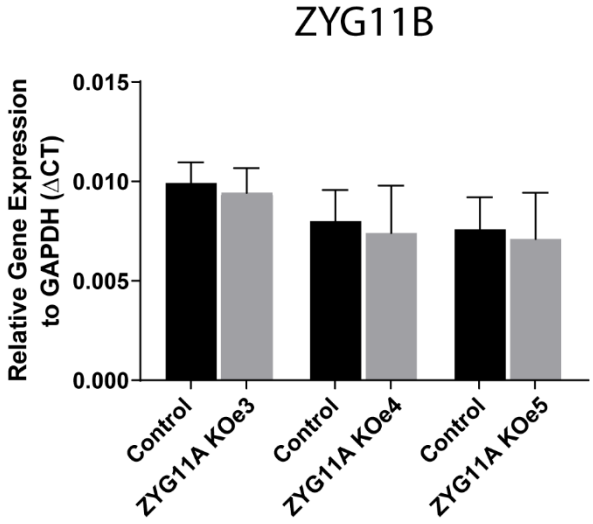
A


ZYG11A Exon 1
ZYG11A Exon 2
ZYG11A Exon 3
ZYG11A Exon 4
ZYG11A Exon 5
Control
sgRNA
Control
sgRNA
Control
sgRNA
Control
sgRNA
Control
sgRNA

B


ZYG11A
Relative Gene Expression to GAPDH (ΔCT)
0.00006
0.00004
0.00002
0.00000
*
Control
ZYG11A KOe3
Control
ZYG11A KOe4
Control
ZYG11A KOe5

C


ZYG11B
Relative Gene Expression to GAPDH (ΔCT)
0.015
0.010
0.005
0.000
Control
ZYG11A KOe3
Control
ZYG11A KOe4
Control
ZYG11A KOe5

D

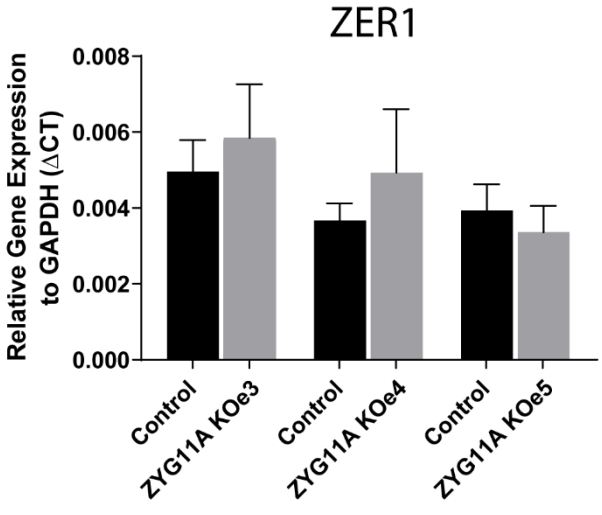

ZER1
Relative Gene Expression to GAPDH (ΔCT)
0.008
0.006
0.004
0.002
0.000
Control
ZYG11A KOe3
Control
ZYG11A KOe4
Control
ZYG11A KOe5

**Figure 3.19 Validation of *ZYG11A* CRISPR-Cas9 knockout lines does not affect ZYG11B or ZER1 levels**

**A.** Validation of CRISPR-Cas9 gene editing by T7E1 endonuclease activity and gel electrophoresis. Each lane represents DNA amplified from the sgRNA exon target following T7E1 mediated recognition of NHEJ mismatch and subsequent cleavage in successfully edited DNA, and no cleavage in unedited DNA. Each pair of lanes displays an unedited empty sgRNA control and an sgRNA targeted to either Exon 1, 2, 3, 4 or 5 of ZYG11A. **B-D.** Real time PCR analysis of *ZYG11A, ZYG11B* and *ZER1* expression levels in control and *ZYG11A* Exon 3, 4 and 5 knockout lines ($ZYG11A^{KOe3}$, $ZYG11A^{KOe4}$, $ZYG11A^{KOe5}$). mRNA levels are determined relative to a GAPDH control. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, * = $p \leq 0.05$, Student's t-test).

### 3.2.6.1 Knockout of *ZYG11A* does not alter cell morphology

An altered cell morphology was by far the most striking change observed in $ZYG11A^{KD}$ cells and with this in mind, I hypothesised that similar defects would be present in the $ZYG11A^{KO}$ cell lines. Surprisingly, the $ZYG11A^{KO}$ cells showed no alterations in cell shape and maintained correctly localised, predominantly membranous Scribble and β-catenin staining as well as typically organised actin filaments (Figure 3.20). These cells did not display any defects in their ability to form a confluent monolayer of cells, nor did they show any signs of a loss of contact inhibition, as seen in $ZYG11A^{KD}$ cells. In the context of a presumptive complete knockout it appears that *ZYG11A* is not required for the maintenance of cell shape or the correct localisation of polarity or cell adhesion proteins.

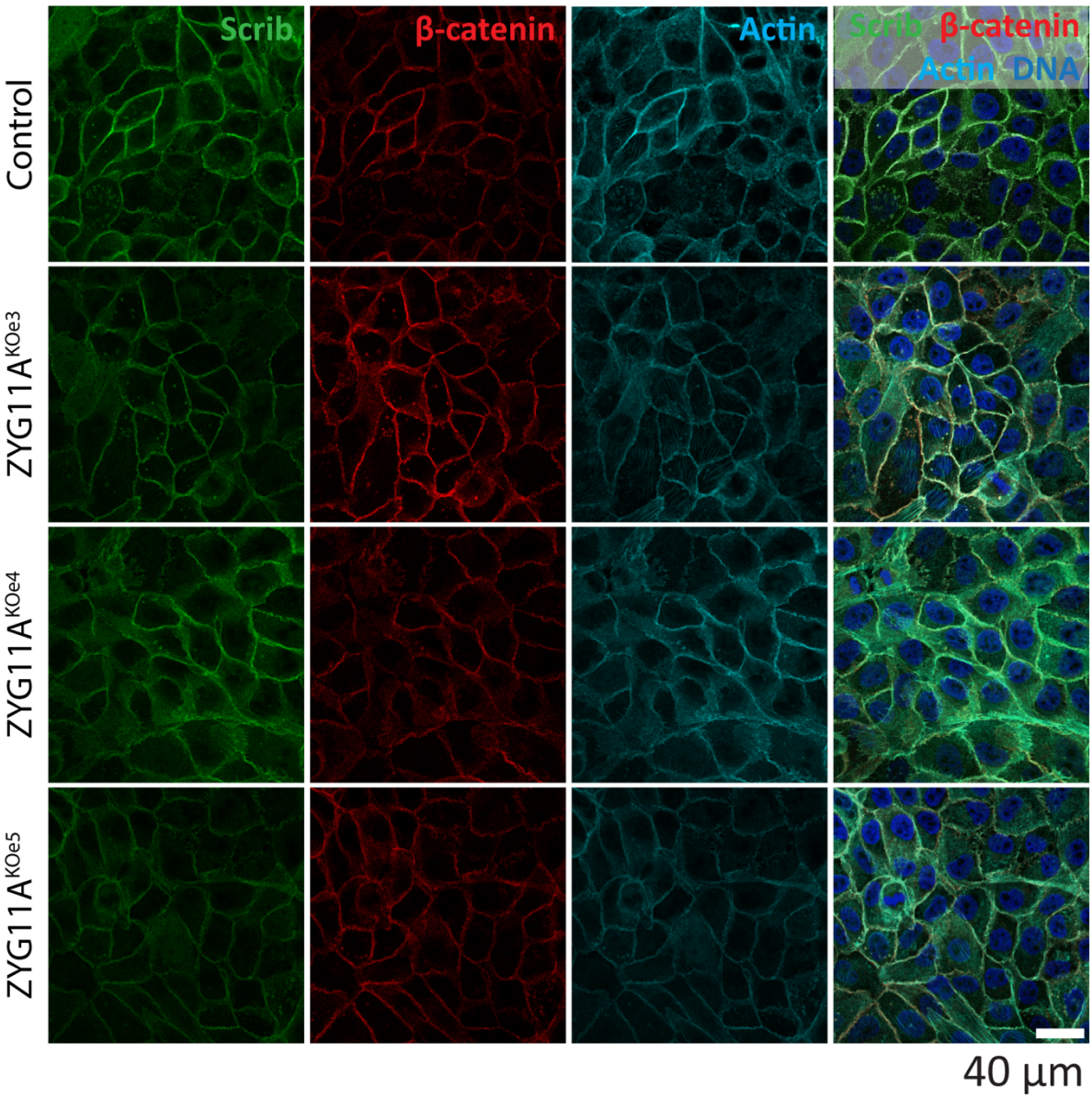



**Figure 3.20 *ZYG11A* knockouts show minimal changes in Scribble, β-catenin and Actin**

**A.** Representative confocal images of control, ZYG11A$^{KOe3}$, ZYG11A$^{KOe4}$ and ZYG11A$^{KOe5}$ MCF10A cells at confluence. Images display the localisation of Scribble (green), β-catenin (red) and F-actin (cyan) followed by a merge of the three channels in the last column along with the nuclear stain, DAPI (blue).

#### 3.2.6.2 *ZYG11A* knockout does not affect cell death, cell cycle, or cell migration

To further compare the RNAi mediated knockdown against the knockout phenotype, the ZYG11A$^{KO}$ cells were subjected to similar cell death, cell cycle, and cell migration analyses as knockdown cells. Unlike ZYG11A$^{KD}$ cells, where there were significant increases in cell death and cell cycle arrests at the G0/G1 phase, the complete loss of *ZYG11A* showed no alterations compared with the controls in any of the tested cellular processes, at least not under the given conditions (Figure 3.21A-D). In addition, the capacity of these cells to migrate in response to an EGF stimulus remained unchanged (Figure 3.22A-B). In stark contrast with *ZYG11A* knockdown, the complete depletion of *ZYG11A* does not increase cell death or prevent proliferation by an arrest in cell cycle, nor does it decrease cell migration.

#### 3.2.6.3 *ZYG11B* and *ZER1* do not compensate for *ZYG11A* knockdown

Compensatory mechanisms that occur in response to complete genetic knockout, but not during genetic knockdown, have been described in multiple contexts including *in vitro* and *in vivo* systems (El-Brolosy and Stainier, 2017). With the idea that perhaps this was the case with *ZYG11A*, the ZYG11A$^{KO}$ lines were analysed for any increases in *ZYG11B* or *ZER1* mRNA expression. Although they were the most likely candidates with respect to genetic compensation, there were no increases in mRNA for either gene, consistent with ZYG11A$^{KD}$ results (Figure 3.19C-D). These data suggest the potential for compensatory mechanisms independent of other ZYG11 family members in a knockout setting.

Despite distinct phenotypic changes in a knockdown setting, ZYG11A$^{KO}$ cells do not trigger any alterations in cell shape or cell function. Unfortunately, these findings indicate that the *ZYG11A* knockout model will not be a useful tool to study cell polarity or morphology changes in the MCF10A cell line.

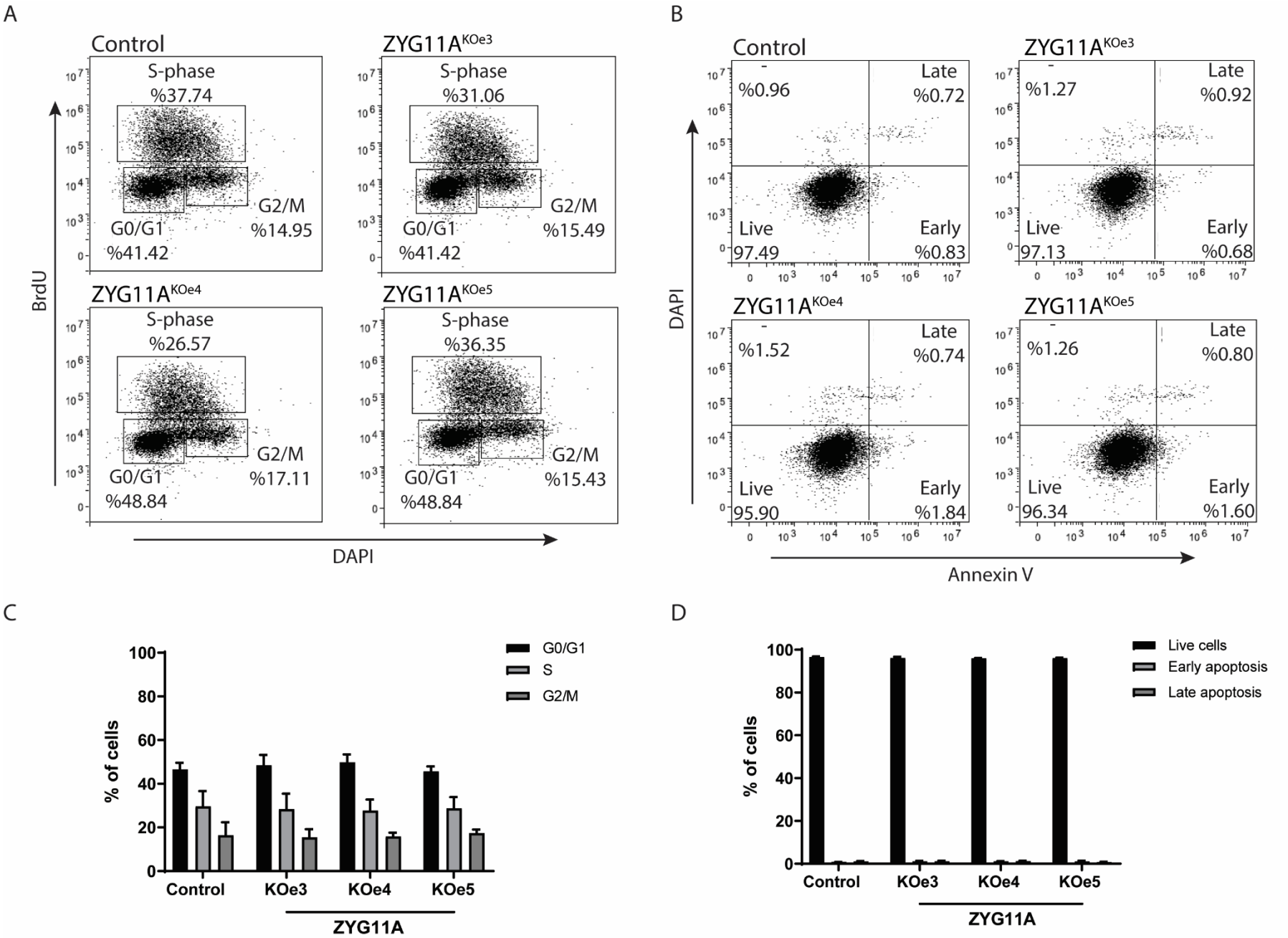

A
Control
S-phase
%37.74
G2/M
%14.95
G0/G1
%41.42
ZYG11A^KOe3
S-phase
%31.06
G2/M
%15.49
G0/G1
%41.42
ZYG11A^KOe4
S-phase
%26.57
G2/M
%17.11
G0/G1
%48.84
ZYG11A^KOe5
S-phase
%36.35
G2/M
%15.43
G0/G1
%48.84
BrdU
DAPI
B
Control
%0.96
Late
%0.72
Live
97.49
Early
%0.83
ZYG11A^KOe3
%1.27
Late
%0.92
Live
97.13
Early
%0.68
ZYG11A^KOe4
%1.52
Late
%0.74
Live
95.90
Early
%1.84
ZYG11A^KOe5
%1.26
Late
%0.80
Live
96.34
Early
%1.60
DAPI
Annexin V
C
G0/G1
S
G2/M
% of cells
Control
KOe3
KOe4
KOe5
ZYG11A
D
Live cells
Early apoptosis
Late apoptosis
% of cells
Control
KOe3
KOe4
KOe5
ZYG11A

**Figure 3.21 Complete loss of *ZYG11A* does not affect cell cycle or cell death**

**A.** Cell cycle profiles of control, ZYG11A$^{KOe3}$, ZYG11A$^{KOe4}$ and ZYG11A$^{KOe5}$ MCF10A cells at ~70% confluence. Each phase determined by the levels of BrdU incorporation (DNA synthesis) and DAPI (DNA levels) staining by FACS. **B.** Quantification of the plots in **A** assessed using 10,000 events per sample. **C.** Viability of the same knockout cell lines described in **A** using Annexin V and PI staining to determine cell death levels with the individual FACS plots representing live (bottom left quadrant), early apoptotic (bottom right quadrant) and late apoptotic (top right quadrant) and necrotic (top left quadrant) cells. **D.** Quantification of the plots in **C** assessed using 10,000 events per sample. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, Two-way ANOVA).

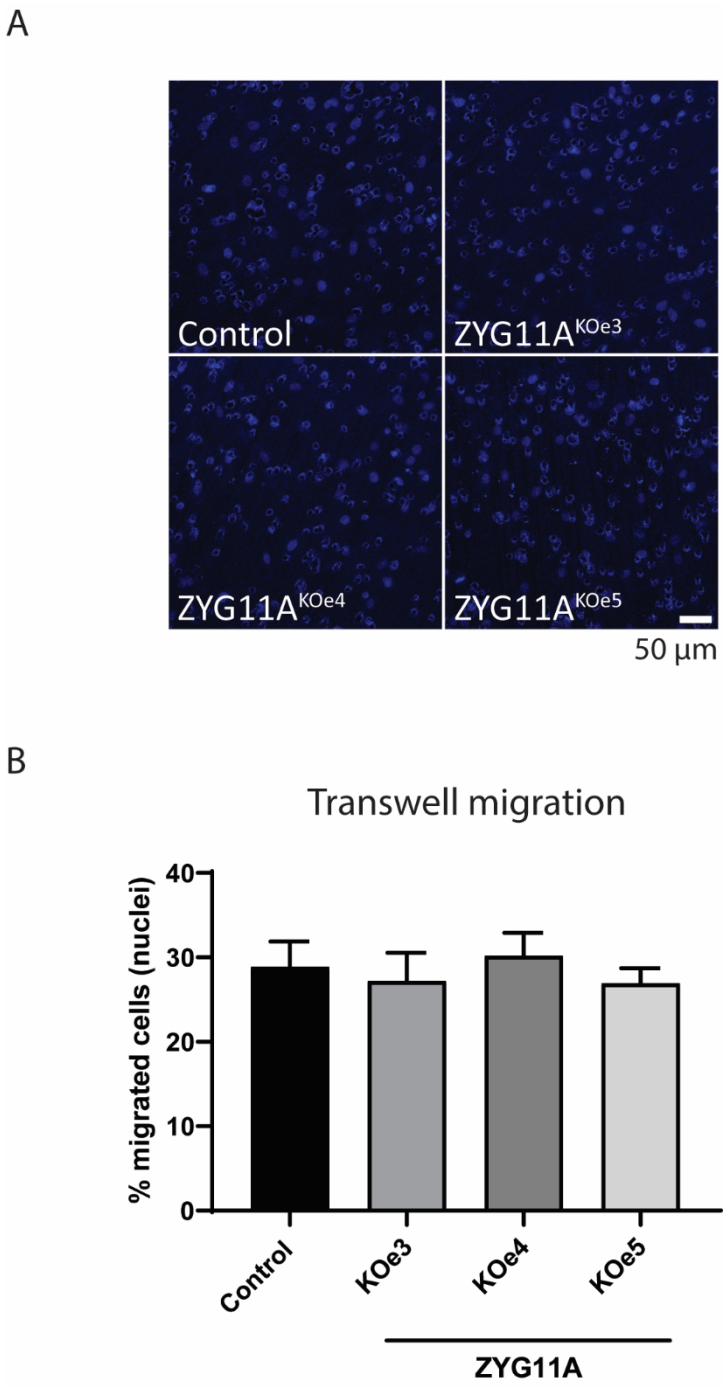



**Figure 3.22 Knockout of *ZYG11A* does not alter cell migration in response to an EGF stimulus**

**A.** Representative confocal images of the Transwell cell migration of control, ZYG11A$^{KOe3}$, ZYG11A$^{KOe4}$ and ZYG11A$^{KOe5}$ MCF10A cells. Cells starved of EGF overnight that have migrated towards the EGF stimulus over a 24-hour period are visible through the DNA stain, DAPI (blue). **B.** Quantification of **A** as a percentage of the total number of seeded cells per well. Error bars represent ± SEM and are representative of 3 independent experiments each performed in duplicate. (n=6, Student's t-test).

### 3.2.7 Generation of *zyg11* and *zer1* CRISPR zebrafish strains

Although *C. elegans* has provided the foundations of ZYG11 function with respect to cell cycle and embryonic polarity establishment, the context of these developmental defects has only been observed in terms of the invertebrate embryo. Like in *C. elegans*, both *zyg11* and *zer1* are present in the zebrafish, yet no studies have investigated the function of these genes in this setting. Genetic knockout (KO) fish for both *zyg11* and *zer1* were generated using CRISPR-Cas9 RNP technologies with single guide (sg) RNAs directed towards the 3$^{rd}$ exon of each gene. These were introduced by microinjection at the 1-cell stage and raised to adulthood where they were then genotyped to identify any genomic alterations (Figure 3.23A). From this, two *zyg11* KO lines and one *zer1* KO line was established. As a reference, the protein lengths of Zyg11 and Zer1 are 746 aa and 774 aa, respectively. In the first *zyg11* strain, a single guanidine base was inserted, resulting in a frame shift mutation at position 386 (+368G), reducing the theoretical amino acid length to 129 through the generation of an early stop codon (Figure 3.23B). The second *zyg11* strain harboured the loss of a thymidine and adenine at the 366/367 positions (-366T, -367A) along with the addition of a guanidine (+366G) (Figure 3.23C). This resulted in a theoretical amino acid length of 141, again due to a frameshift-generated early stop codon. The *zer1* CRISPR knockout led to a 16 bp loss between the 445 → 460 positions (-445[AAGAGCGGCGCCCCGC]) with a frame shift leading to a stop codon at amino acid 183 (Figure 3.23E). Unfortunately, like with the mammalian cell lines there were no antibodies available to check the reduction of protein levels. Nevertheless, sequencing confirmed that the mutations generated in both *zyg11* and *zer1* heterozygous knockouts formed downstream premature stop codons in the 3$^{rd}$ exon, likely to produce proteins ~20% of their original length and lacking most of the LRR and ARM domains (Figure 3.23G).

A

RNPs microinjected into a 1-cell stage embryo

Embryos monitored until ready to raise the fry

Adults sequenced to identify gene editing

B

*Zyg11* Exon 3 sgRNA

```
355  ACCATAGCCGATA-TCCTCCGTGGACTCTCCTCTAATAAA  393
355  ACCATAGCCGATAGTCCTCCGTGGACTCTCCTCTAATAAA  394
     *************************************
```

*Zyg11 (+1 bp)*

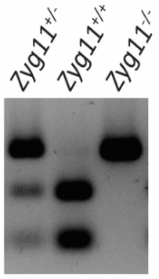
C

*Zyg11* Exon 3 sgRNA

```
355  ACCATAGCCGATATCCTCCGTGGACTCTCCTCTAATAAAA  393
355  ACCATAGCCGAG-TCCTCCGTGGACTCTCCTCTAATAAAA  392
     ***********.*****************************
```

*Zyg11 (-2+1 (-1 bp))*

D



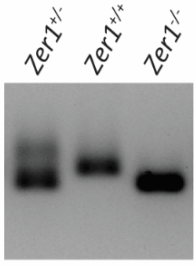

E

*Zer1* Exon 3 sgRNA

```
426  CTCCGGGATCTTCTTCCGCAAGAGCGGCGCCCCGCTGGCG  465
426  CTCCGGGATCTTCTTCC----------------GCTGGCG  449
     ****************                *******
```

*Zer1 (-16 bp)*

F




G

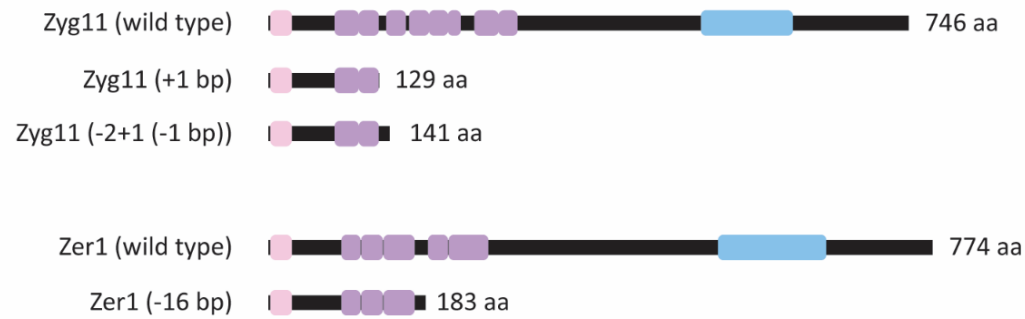


**Figure 3.23 Generation of *zyg11* and *zer1* CRISPR-Cas9 knockout zebrafish lines**

**A.** Schematic of the generation of CRISPR-Cas9 *D. rerio* strains using TU (wild type) 1-day old embryos injected with RNPs. **B-C.** Sequencing of the two *zyg11* knockout strains confirming the DNA alterations in exon 3. **D.** Genotyping of the *zyg11* (+1 bp) and (-2+1(-1 bp)) strains using PCR and gel electrophoresis. Digestion of the native EcoRV cut site reveals the genetic alterations in both strains. Heterozygous fish show 3 bands (203, 129 and 73 bp), WT fish show two bands (129 and 73 bp), and homozygous fish show 1 band (203 bp). **E.** Sequencing of the *zer1* knockout strain showing the exon 3 alteration. **F.** Genotyping of the *zer1* (-16 bp) strain. WT and homozygous fish show 1 band, whereas heterozygous fish show 2 bands. Note that the upper band (WT allele) in heterozygous fish is slightly higher than the band in the WT fish, and this is consistently seen in heterozygous fish that have been confirmed by sequencing. **G.** Predicted protein lengths of the *zyg11* and *zer1* knockout strains, depicting the VHL box, the LRR and ARM domains. Premature stop codons generated by CRISPR-Cas9 editing produce ZYG11 family proteins lacking the ARM domain and shortened LRR domains.

### 3.2.8 The loss of *zyg11* and *zer1* does not affect embryonic development or the viability of adult Zebrafish

If the function of Zyg11 in the zebrafish is similar to that in the roundworm, then the homozygous loss of *zyg11* should result in embryonic lethality. Unlike in the previous roundworm studies, the CRISPR lines generated in this chapter will also assess the effects of *zer1* loss. The three individual CRISPR lines, *zyg11 (+1)*, *zyg11 (-2+1(-1))* and *zer1 (-16),* were in-crossed and each was observed for any aberrant embryonic defects or any increases in embryonic death up to 5 days post-fertilisation (dpf). Unexpectedly, the embryos from each heterozygote in-cross appeared normal with very minimal death over the 5-day period (Mendelian ratios would suggest at least 25% would be lethal if *zyg11* or *zer1* homozygosity is lethal). Considering the early embryonic lethality in *C. elegans* lacking *zyg-11*, it was important to identify the genotypes of these fish to confirm that there were indeed homozygous embryos and fish present in these crosses. To do this, half of the embryos were raised to adulthood while the other half were genotyped. In their respective crosses there were *zyg11 (+1)*, *zyg11 (-2+1(-1))* and *zer1 (-16)* homozygous embryos present that were also identified in the genotyped adults (embryos not shown) (Figure 3.23D and F). To examine whether *zer1* could compensate for the loss of *zyg11* and vice versa, *zyg11 (-2+1(-1))* and *zer1 (-16)* mutant fish were used to generate a double heterozygous strain. Once these fish reached sexual maturity (~5 months onwards), they were in-crossed. Through Mendelian ratios it was expected that 1 in 16 of these fish would harbour a double homozygous knockout, and as these fish would take a few months to reach maturity, only a small population of these fish were genotyped at 5 dpf (Figure 3.24,

Table 1).

**Table 1.** Genotypes of *zyg11 (-2+1(-1 bp))$^{+/-}$; zer1 (-16 bp)$^{+/-}$* inter-cross (embryos)

| Genotype | Observed (O) | Expected (E) | Difference (D) = O - E | $D^2$ | χ2 |
|---|---|---|---|---|---|
| WT/WT; WT/WT | 1 | 1 | 0 | 0.1 | 0.1 |
| WT/WT; zer1/WT | 4 | 3 | 1 | 1.9 | 0.7 |
| WT/WT; zer1/zer1 | 1 | 1 | 0 | 0.1 | 0.1 |
| WT/zyg11; WT/WT | 2 | 3 | -1 | 0.4 | 0.1 |
| WT/zyg11; WT/zer1 | 3 | 5 | -2 | 5.1 | 1.0 |
| WT/zyg11; zer1/zer1 | 2 | 3 | -1 | 0.4 | 0.1 |
| zyg11/zyg11; WT/WT | 2 | 1 | 1 | 0.5 | 0.4 |
| zyg11/zyg11; WT/zer1 | 6 | 3 | 3 | 11.4 | 4.3 |
| zyg11/zyg11; zer1/zer1 | 0 | 1 | -1 | 1.7 | 1.3 |
| Total χ2 | 8.1 | | | | |
| Degrees of freedom | 8 | | | | |
| p-value | 0.420 | | | | |

**Table 2.** Genotypes of *zyg11 (-2+1(-1 bp))$^{+/-}$; zer1 (-16 bp)$^{+/-}$* inter-cross (adults)

| Genotype | Observed (O) | Expected (E) | Difference (D) = O - E | $D^2$ | χ2 |
|---|---|---|---|---|---|
| WT/WT; WT/WT | 5 | 7 | -2 | 3.1 | 0.5 |
| WT/WT; zer1/WT | 22 | 14 | 9 | 72.3 | 5.4 |
| WT/WT; zer1/zer1 | 9 | 7 | 2 | 5.1 | 0.8 |
| WT/zyg11; WT/WT | 14 | 14 | 1 | 0.3 | 0.0 |
| WT/zyg11; WT/zer1 | 29 | 27 | 2 | 4.0 | 0.1 |
| WT/zyg11; zer1/zer1 | 9 | 14 | -5 | 20.3 | 1.5 |
| zyg11/zyg11; WT/WT | 7 | 7 | 0 | 0.1 | 0.0 |
| zyg11/zyg11; WT/zer1 | 7 | 14 | -7 | 42.3 | 3.1 |
| zyg11/zyg11; zer1/zer1 | 6 | 7 | -1 | 0.6 | 0.1 |
| Total χ2 | 11.4 | | | | |
| Degrees of freedom | 8 | | | | |
| p-value | 0.178 | | | | |

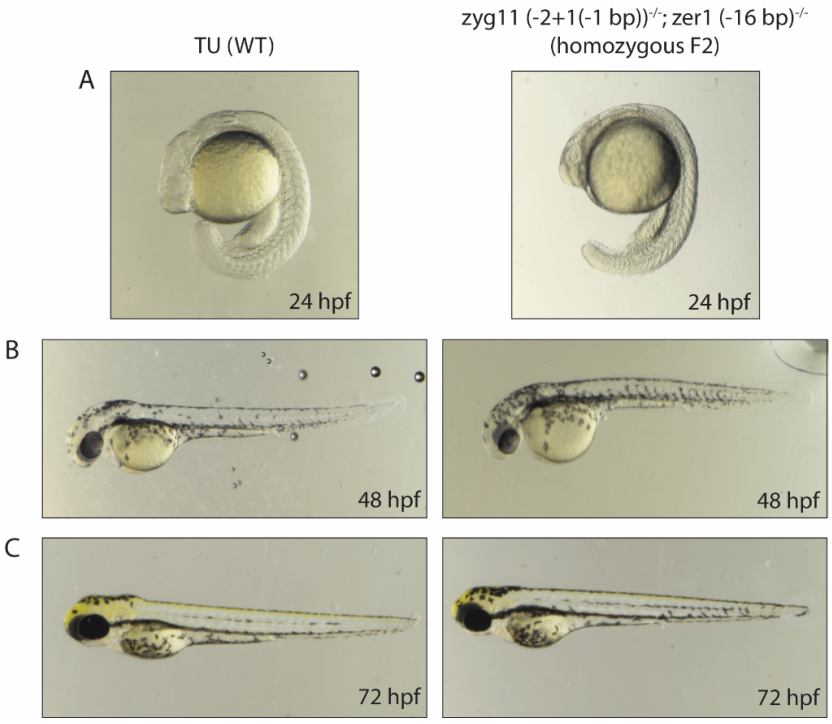


**Figure 3.24 *zyg11* and *zer1* double knockout fish display no obvious phenotypes in embryos or adults**

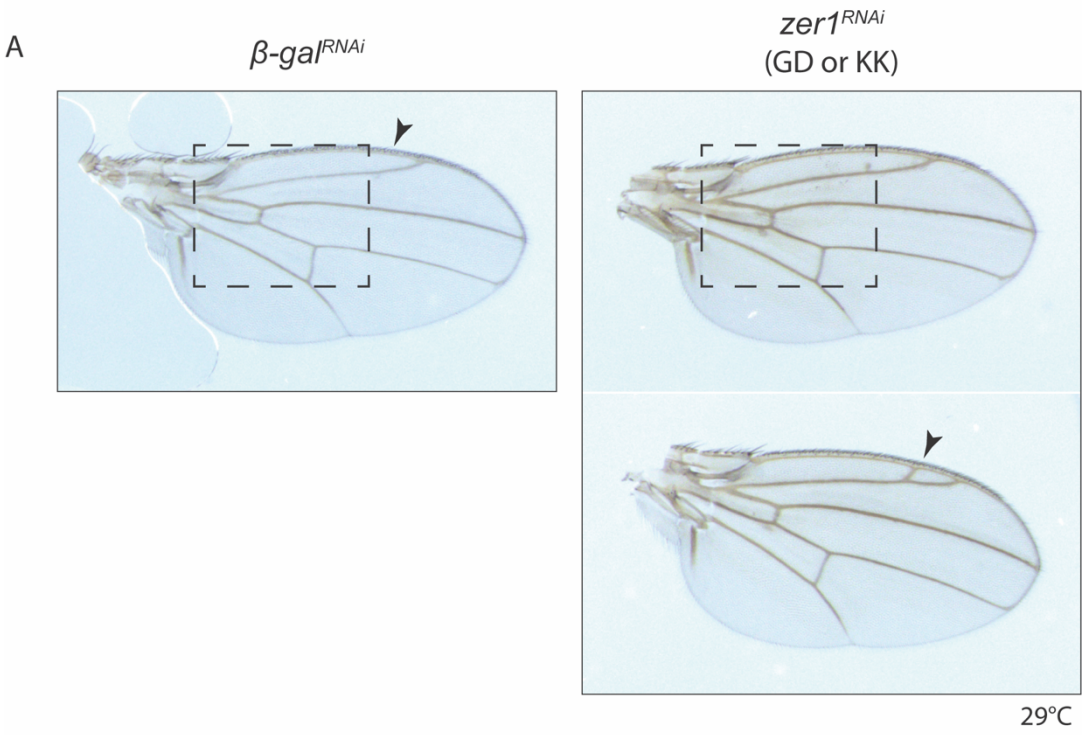
**Table 1-2.** The genotypes of *zyg11 (-2+1(-1 bp))*$^{+/-}$*; zer1 (-16 bp)*$^{+/-}$ in-crossed (INX) embryos and adult fish. The expected (E) number of fish are determined by Mendelian ratios. The 'fit' of the genetic population (displayed as the p-value) is determined by a Chi-squared ($\chi^2$) test, n = 21 (embryos) and 108 (adults). The resulting double homozygous knockout fish were further inter-cross to produce an F2 generation of double knockout fish. **A-C.** Representative images of embryos from the F2 generation of *zyg11 (-2+1(-1 bp))*$^{-/-}$*; zer1 (-16 bp)*$^{-/-}$ inter-cross fish at three stages of embryonic development – **A.** 24, **B.** 48 and **C.** 72 hpf. Note that there are no visible defects in these fish when compared to TW (wild type) embryos and they are able to survive to 5 dpf (fish fry stage).

Although none of the embryos taken for genotyping were double knockouts, the remaining genotypes in the population were present and there were no statistically significant changes seen (p = 0.420). After raising a larger population of embryos to adulthood, genotyping showed that double knockout adult fish were found in the correct Mendelian ratios (p = 0.178) (Figure 3.24, Table 2). Finally, *zyg11* and *zer1* homozygous single and *zyg11; zer1* double knockouts were inter-crossed to determine whether the embryos remain viable without any maternally contributed *zyg11* or *zer1*. These embryos were closely monitored during the early embryonic stages, as well as during the 24 to 72 hours post-fertilisation stages where most organogenesis occurs, and developmental defects can become more obvious. Once again, these second-generation (F2) knockout embryos appeared normal and were viable to the larval stage at 5 dpf, similar to wild type in-cross embryos (Figure 3.24A-C). Overall, these studies are in contrast to the developing *C. elegans* indicating that under the normal developmental conditions used here, the ZYG11 family is not essential for the early developmental stages in the zebrafish *in vivo* model.

### 3.2.9 *zer1* RNAi knockdown produces weak phenotypes in *Drosophila* wings and eyes

As our mammalian RNAi experiments implicated a role for ZYG11 in epithelial cell shape and cellular properties, to test whether the loss of *zer1* affects the development of prominent epithelial compartments in the vinegar fly, *Drosophila* the *zer1*$^{RNAi}$ line was expressed in either the eye or the wing using the GAL4-UAS system. Alongside a *β-gal*$^{RNAi}$ control, the knockdown of *zer1* was targeted to the developing wing through the *UAS-dcr;nub-GAL4* driver using two *zer1*$^{RNAi}$ lines, GD and KK. These two lines represent two differently generated transgenic RNAi lines: 1. A P-element mediated insertion of a short hairpin RNA (shRNA) directly into the fly genome (GD), and 2. A phiC31 integrase

system incorporating a more targeted 2-step integration shRNA (KK). Although there were no observable defects in adult wings when raised at 25˚C (data not shown), a shift to 29˚C that allows the expression of the RNAi to be increased (Duffy, 2002), gave rise to visible wing defects observed in *zer1*$^{RNAi}$ knockdown adults from both lines compared to the controls. Visibly dark and discoloured sections in the wings and/or the formation of an extra cross-vein between the L1 and L2 wing veins were the most penetrant defects observed (Figure 3.25A) These defects were highly penetrant and affected an average of 79.7% of wings in the GD line and 80.5% of wings in the KK lines, compared to 6.7% in the controls (Figure 3.25B).



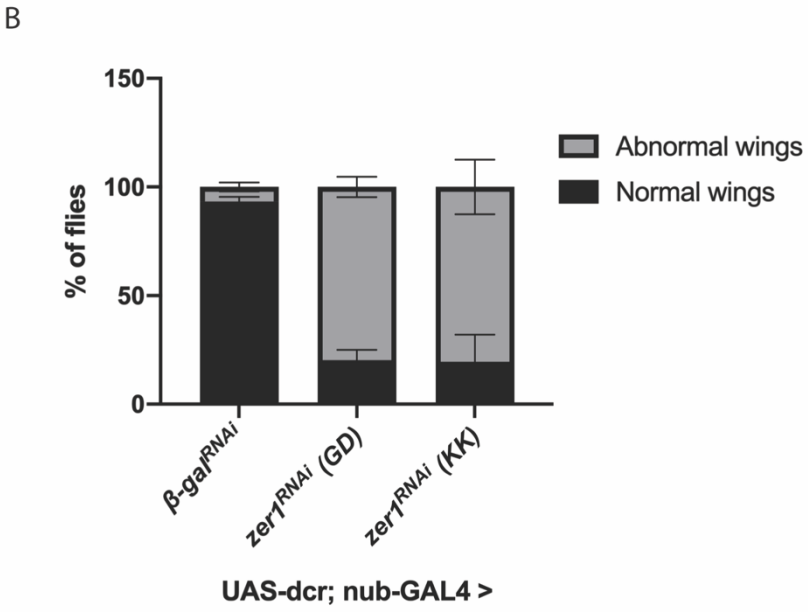



**Figure 3.25 Loss of *zer1* in the developing wing causes wing vein patterning defects in adult flies**

**A.** Representative images of the adult phenotypes observed in $zer1^{RNAi\ (KK)}$ or $zer1^{RNAi\ (GD)}$ knockdown wings using the *UAS-dcr; nub-GAL4* driver, compared to a $\beta\text{-}gal^{RNAi}$ control. The loss of *zer1* affects normal wing patterning, displaying 'abnormal' discolouration and darkening of the wing (square outline, top right image) and/or the formation of an ectopic cross-vein between the L2 and L2 veins (arrowhead, bottom right image). The same areas on control wings are similarly highlighted in the top left image. **B.** The penetrance of wing defects is markedly increased in $zer1^{RNAi\ (GD)}$ (75%) and $zer1^{RNAi\ (KK)}$ (93.1%) knockdown adults compared with controls (4.5%). Scoring is based on the presence or absence of the wing defects described. Number of wings scored is between 239-323 adult wings per cross.

Another epithelial population of cells can be found in the developing eye. Since the KK line produced a more penetrant phenotype in the wing, I focused on using the KK line in the developing eye. When driven through *GMR-GAL4* driver, there were also two subtle changes. In addition to a small but significant decrease in eye size in *zer1* knockdown flies compared with control counterparts, they also showed an enhanced rough eye phenotype (Figure 3.26A-C). The *GMR-GAL4* driver typically presents a mild rough eye phenotype which, under *zer1* loss, was enhanced to an even rougher appearance and a strong yellow undertone to the eyes (Figure 3.26A, arrows). While this was observed in 20% of control eyes, this penetrance was increased to 80% in adult *zer1* knockdowns, suggesting that the loss of *zer1* can enhance *Drosophila* ommatidial disorganisation (Figure 3.26C)

The phenotypes observed in both adult wings and eyes are considerably mild in comparison to the loss of well-known polarity or cell shape regulators like Scrib or β-catenin. However, these data provided early indications that the loss of *zer1* in *Drosophila* epithelial cells can cause defects in tissue organisation and may even induce necrosis (as indicated by the wing darkening) during development.

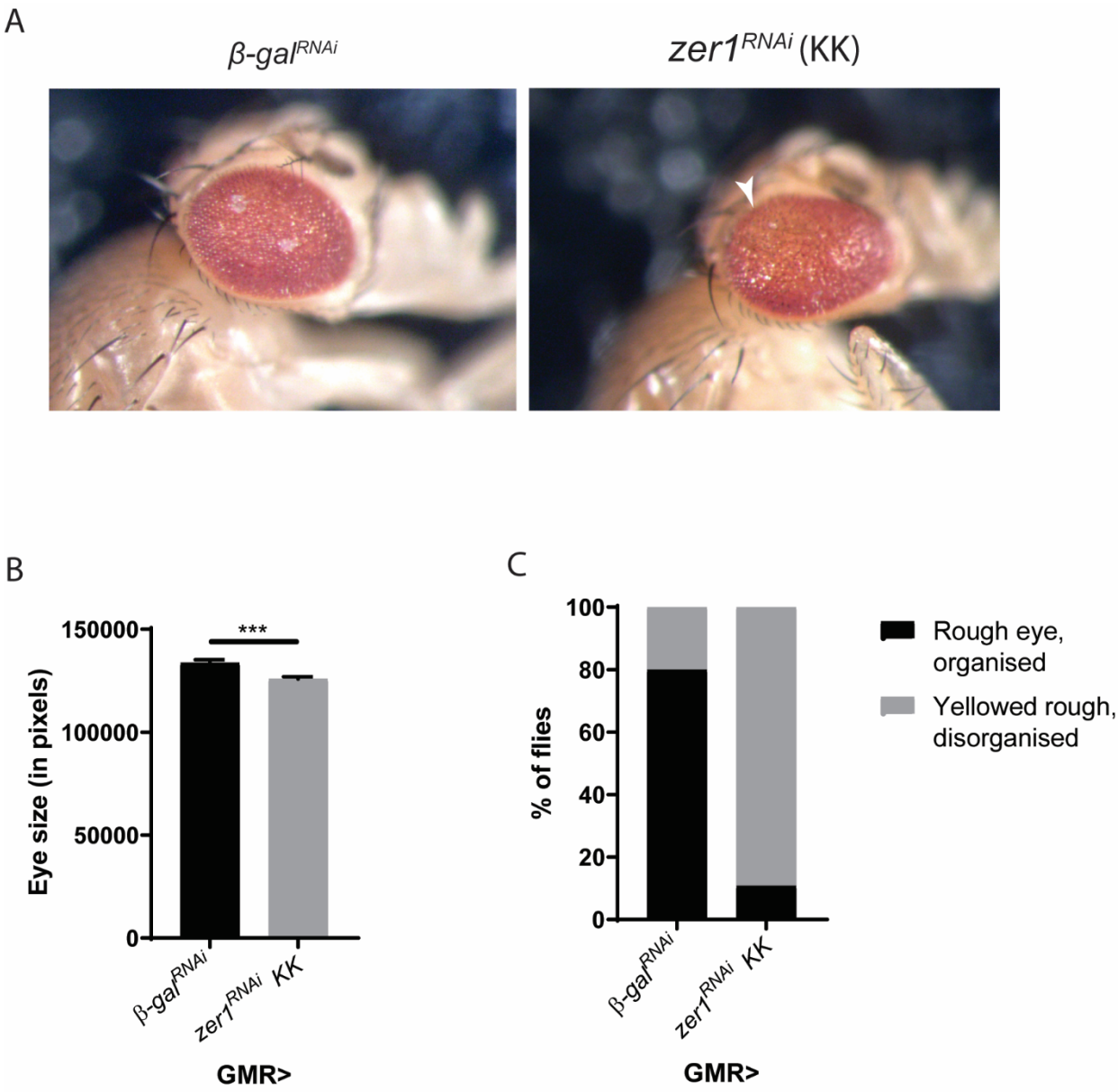



**Figure 3.26 Knockdown of zer1 in the embryonic eye affects the size and patterning of the adult Drosophila eye**

**A.** Representative images of the adult eye phenotypes observed in zer1$^{RNAi\ (KK)}$ knockdown wings using the GMR-GAL4 driver, compared to a β-gal$^{RNAi}$ control. While controls show a rough eye phenotype, the ommatidia remain organised and red in colour. Loss of zer1 displays a rough eye phenotype that has more disorganised ommatidia and visible yellowing compared to the controls (white arrow, bottom image). Eyes also appear more rounded and protrude more than the control eyes. **B.** Quantification of the size of zer1$^{RNAi\ (KK)}$ eyes compared to controls. There is a small but significant decrease in eye size when zer1 is disrupted. **C.** The penetrance of the more disorganised and yellow rough eye phenotype is markedly increased in zer1$^{RNAi\ (KK)}$ (89.2%) flies compared with controls (20%). Number of adult flies is between 11-37 adults. Error bars represent ± SEM.

### 3.3 Discussion

Highly conserved gene families – like the ZYG11 family – exist throughout evolution, maintaining certain functions from species to species. Mounting evidence has highlighted the preservation of the ZYG11 family's role in cell cycle control from *C. elegans* to mammalian cells, and to this, it was predicted that other ZYG11 family functions might be similarly observed in mammalian systems. The establishment of embryonic cell polarity in *C. elegans* has previously linked ZYG-11 to cell polarity *in vivo*, yet whether the mammalian counterparts of ZYG-11 also share this characteristic has been studied for the first time in Chapter 3. In addition, the embryonic lethality linked to ZYG-11 loss in *C. elegans* has also been explored in other *in vivo* models.

Here I undertook a functional characterisation of ZYG11 family members in human epithelial cells. In initial studies, I identified a conserved function of the N-terminal domain for each ZYG11 family member for its nuclear exclusion. The nuclear exclusion appeared to be independent of the exportin pathway with mutations in a predicted NES domain conserved across evolution and between ZYG11 family members, unable to alter its localisation. These findings suggest that there are other regulatory mechanisms, likely shared between the ZYG11 family, that dictate the shuttling of these proteins. Moreover, the overexpression of the wild type or mutant versions of these proteins also did not induce any phenotypic or functional changes. I showed that the loss of *ZYG11A* by RNAi in the human mammary epithelial cell line, MCF10A, caused aberrant alterations in cell morphology, apoptosis and cell cycle phasing while the depletion other ZYG11 family members, *ZYG11B* and *ZER1*, displayed minimal phenotypes. In comparison to siRNA mediated knockdown, the CRISPR mediated knockout of *ZYG11A* did not show any cellular defects in proliferation, apoptosis or migration, and this was in the absence of observable genetic compensation from *ZYG11B* or *ZER1*.

The *in vivo* study of ZYG11 family members has been limited to the loss of *zyg-11* in the developing *C. elegans* embryo, leaving a number of unanswered questions about whether other genetic models of *zyg-11* loss can mimic the phenotypes seen in *C. elegans*. The animal work in this chapter has focused on two widely used genetic models to bridge this gap, using the vertebrate *D. Rerio* zebrafish model and the *D. Melanogaster* invertebrate model. I described the generation of two zebrafish CRISPR lines that harbour premature stop codons in either the *zyg11* or *zer1* coding regions and have demonstrated that fish homozygous for either *zyg11* or *zer1* knockout display no obvious developmental defects

and are able to produce viable and fertile adults. Similarly, double knockout fish also lack any visible phenotypic defects and can grow to adulthood. Together with the CRISPR cell line data, this suggest that the direct alteration of ZYG11 family members at the DNA level may be compensated for by other genes. In the fly where there is only a single ZYG11 family member, the knockdown of *zer1* resulted in mild wing and eye phenotypes that are indicative of a deregulation in the normal tissue architecture. These findings suggest that in the vertebrate model of zebrafish, the ZYG11 family is not essential from a developmental perspective, and in the *Drosophila* model it is minimally required for the formation of normal wing and eye epithelial tissues.

### 3.3.1 The localisation and nuclear to cytoplasmic shuttling of the ZYG11 family

The correct establishment of apicobasal cell polarity hinges on the proper subcellular localisation of the Par, Crumbs and Scribble complexes, specifically at the basolateral membrane of cells in the context of Scribble complex components. In line with the previously reported localisation of ZYG11B (Balachandran *et al.*, 2016), ZYG11B and other ZYG11 fusion proteins are ubiquitously expressed in the cytoplasm but are excluded from the nucleus and do not co-localise with chromosomes during mitosis. While not exclusively at the membrane or co-localised with other polarity proteins, the function of ZYG11 proteins are not limited to polarity and therefore this localisation is unlikely to dictate their contribution to cell polarity regulation. Moreover, one point of interest is the exclusion of each of the proteins from the nucleus which is dictated by an unknown protein element found in the N-terminal half of the proteins. Both the inhibition of nuclear export and the mutation of highly conserved ZYG11A amino acid motifs (including an NES) were unable to force the nuclear accumulation of ZYG11A and other family members. The shutting of ZYG11 proteins into the nucleus may be by virtue of a 'piggybacking' mechanism that is often observed in proteins like ZYG11A that lack a discernible nuclear localisation signal (NLS), therefore preventing the entry of these proteins when without a shuttling partner. I also must consider that aside from active translocation, the tethering of ZYG11A/B and ZER1 by other proteins to the cytoplasmic space may also inhibit entry to the nucleus, particularly in an overexpression scenario where native signalling is likely not active. This was a major caveat as the low expression and lack of suitable antibodies required the use of recombinant proteins. The future development of endogenous knock-in models would help address this issue. Nevertheless, the data presented shows that the localisation of ZYG11 family members cannot be manipulated by candidate targeted

mutations, and hence other regulatory mechanisms are occurring. Identifying the post-translational modifications and protein interactions that dictate the location, as well as other processes of the ZYG11 family will be fundamental to deciphering the mechanistic signalling pathways they are involved in.

### 3.3.2 The ZYG11 family regulates cell cycle and proliferation

The link between ZYG11 and the cell cycle is deeply rooted in *C. elegans* embryonic studies where the *zyg-11* gene was first identified as a zygote-defective mutant (Hirsh and Vanderslice, 1976), and further characterised as a maternal-effect gene that regulated meiosis II divisions (Wood *et al.*, 1980; Kemphues *et al.*, 1986; Carter, Roos and Kemphues, 1990). Here, I demonstrated that *ZYG11A* knockdown in MCF10A cells caused an arrest at G0/G1 and an increase in *p21* mRNA expression. This is consistent with the literature that showed the same cell cycle arrest in the non-small cell lung cancer derived H1299 cells, although their data did not show an upregulation of *p21* mRNA expression (X. Wang *et al.*, 2016). My data, which included the other two family members, showed that ZER1 can also regulate this G0/G1 transition. It is interesting that the knockdown of *ZYG11A* and *ZER1* caused cell cycle arrest with significant increases in *p21* mRNA expression as this was seen not only *ZYG11A* and *ZER1* knockdown cells, but also in the normally cycling *ZYG11B* knockdown cells. Although the cause of this *p21* increase is still unclear – since there was no visible increase in p53 levels – this indicates that *p21* expression alone is not the sole driving factor in *ZYG11A/ZER1* knockdown mediated cell cycle arrest, and that there are divergent roles for each family member in cell cycle regulation. A previous study in the U2OS osteosarcoma cell line has found that the $CUL2^{ZYG11A/B}$ acts redundantly with $SCF^{APC/C}$ at the G2 to M phase transition of the cell cycle (Balachandran *et al.*, 2016). Whether this function is also shown in MCF10A cells will be important to test in the context of APC/C loss.

### 3.3.3 The ZYG11 family regulates cell survival and apoptosis

The induction of apoptosis by the loss of individual ZYG11 family members has not been examined. Here I show that only *ZYG11A* loss is able to promote low levels of apoptosis. Whether this cell death is a direct consequence of *ZYG11A* knockdown, or a side effect of the cell cycle arrest or inappropriate signalling is still unclear. Due to the experimental

timepoints and transient nature of the knockdown I was unable to use the caspase inhibitor QVD to test whether this apoptosis was caspase dependent. It would also be interesting to see whether this death is p53 dependent, by additionally knocking down *P53*, and whether there are signs of DNA damage, through γ-H2AX staining. More recently, ZYG11B and ZER1 have been identified to specifically target N-terminal glycine residues for protein ubiquitination, which often become exposed following caspase cleavage (Timms *et al.*, 2019). Perhaps sensitizing ZYG11 knockdowns to caspase cleavage events will enhance or reveal cell survival defects. Nevertheless, this area of study still holds important questions about the mechanism of *ZYG11A* knockdown mediated cell death.

### 3.3.4 The role of the ZYG11 family in cell polarity and cell shape

In *C. elegans*, the polarity defects of *zyg-11* null embryos appear to function independently from the cell cycle role of ZYG-11, with a recent study also reporting that rescue of the cell cycle by *cyb-1* knockdown does not extend to other mutant phenotypes including membrane ruffling and P-granule distribution (Sonneville and Gonczy, 2004; Balachandran *et al.*, 2016). Thus, there is a separate role for ZYG-11 in *C. elegans* anterior-posterior polarity, yet the question of whether mammalian ZYG11 counterparts also regulate polarity has been largely neglected. However, unlike invertebrate species that only harbour one or two ZYG11 family members, the mammalian ZYG11 family is comprised of ZYG11A, ZYG11B and ZER1. ZYG11A was first implicated in mammalian epithelial cell polarity through high-content screening used to identify genes that when knocked down, were able to disrupt cell shape and/or the localisation of the known polarity and cell shape proteins, Scribble and β-catenin (Smith *et al.*, 2016, our lab's unpublished results). Their localisation is important for adherens junction integrity and cell-cell adhesion, as well as the mutual antagonism of apical determinants from the basolateral cortex. For the first time, in this chapter I have built on this observation and have also extended this work to the other ZYG11 family members. Most striking is the overall contrast between the clumped and overlapping phenotype of the *ZYG11A* knockdown cells compared to the highly separate cells observed in *ZYG11B* and *ZER1* knockdown. The overlapping and clumpy cells observed under *ZYG11A* loss suggests that there may be a loss of contact inhibition and aberrant cell adhesions. From the initial screen it was found that *ZYG11A* downregulation increased the mislocalisation of Scribble from the cell membrane to the cytoplasm. My findings show that these disorganised cells appear to maintain higher levels of Scribble and that this is regulated at least in part through altered mRNA levels.

Therefore, I cannot discount the idea that the previously observed increased levels of Scribble in the cytoplasm may be consequentially linked to its aberrant expression. However, this does not address the question of what exactly is being directly affected by the knockdown of *ZYG11A* since it is unlikely that ZYG11A functions directly as a Scribble transcription factor (although that said little is known about the transcriptional regulation of Scribble to date). These findings are relevant as it is the first time that any mammalian counterpart of ZYG11 has been shown to affect apicobasal cell polarity.

### 3.3.5 The role the ZYG11 family in directed cell migration

Several seminal studies have emphasised the requirement of polarised cells in epithelial cell migration, underscoring the role of Scribble mediated polarity in this process (Qin *et al.*, 2005; Dow *et al.*, 2007, 2008). Indeed, the overexpression and mislocalisation of Scribble has also been implicated in directed migration (our lab's observations). Therefore, it is notable that my experiments show that *ZYG11A* is important for the directed migration of MCF10A cells in response to an EGF stimulus. As I have also shown that ZYG11A is important for survival, more experiments are needed to understand how the alterations in cell and the increase in cell death (described in **3.3.2** and **3.3.3**, respectively) due to *ZYG11A* knockdown may impact on my measurements of cell migration in this assay. Due to the inability of these cells to form a confluent monolayer, I was unable to test how they would respond to a different migration stimulus such as wound healing. In wild type cells, the recruitment of Rac1, β-Pix and F-actin by Scribble is known to play a pivotal role in wound healing. Additionally, cell migration also relies on the formation of lamellipodia and filopodia at the front of a migrating cell, and the concurrent detachment at the rear. Given that these cells display many cell elongations, it will be important to characterise the focal adhesions in these cells and whether other integrins or actin filaments are involved in these migration defects.

One unexplored aspect in this study is the behaviour of these cells in 3D cultures. Due to the transient nature of the siRNA knockdown, I was unable to culture the cells to a mature enough stage of acini formation. Therefore, further analysis of the existing CRISPR cell lines as well as the development of more stable knockdowns (i.e. short-hairpin mediated knockdown) will be crucial in characterising how each ZYG11 family member affects cell polarity and cell shape in 3D.

### 3.3.6 Differences in RNA interference and CRISPR phenotypes in human cells

The knockdown of ZYG11 family members using an RNAi approach has revealed that *ZYG11A* downregulation has the capacity to arrest cells at the G0/G1 to S phase transition, induce mild cell death, inhibit cell migration in response to an EGF stimulus, and affect cell morphology and polarity. In contrast to knockdown data, when using the same cell line, but instead a CRISPR-generated knockout of *ZYG11A*, no effects on any of these functions were observed. Because there were no indications of genetic compensation by *ZYG11B* or *ZER1*, this is relevant as it highlights the potential for other E3 ubiquitin ligases to take over when overlapping protein targets begin to accumulate (discussed further in 3.3.8). It will be also be interesting to utilise the ZYG11A$^{KO}$ cells to see whether these functionally normal lines can be sensitised to cell cycle, cell death or cell migration, and test which signalling pathways, if any, can increase sensitivity to these cues. Finally, although I used multiple independent siRNAs to target the ZYG11 family members in my studies, to exclude the possibility of off-target effects of the various RNAis, it will be important to confirm that human cDNAs for each ZYG11 family member can be engineered to be insensitive to their respective RNAi, and can rescue the individual knockdown phenotypes.

### 3.3.7 Genetic compensation between the ZYG11 family and other ubiquitin ligases

Every essential process of the cell is protected by complex compensatory mechanisms that act in concert to ensure that homeostasis is preserved under various forms of stress. The high conservation between each ZYG11 family member is grounds for the possibility of genetic compensation, particularly between the more closely related ZYG11A and ZYG11B. However, the data in this chapter suggests that ZYG11B and ZER1 are more closely related, at least in terms of knockdown morphology, migration capacity and cell death, where ZYG11A is phenotypically distinct and more severe in comparison. This idea is further supported by the transcriptional upregulation of *ZER1* in *ZYG11B* knockdown cells and vice versa, without any significant changes in *ZYG11A*, nor any sign of compensation by *ZYG11B* and *ZER1* in *ZYG11A* knockdown cells. Perhaps this can be attributed to differences in the mRNA expression patterns of *ZYG11B* and *ZER1*, which are typically expressed in all tissue types, versus *ZYG11A* expression, which appears much less abundant in most tissue types, although is particularly enriched in organs like the

kidney, thyroid and testis (Ardlie *et al.*, 2015). In a recent study by Timms et al. it was found that the ability to target a novel N-degron motif was a characteristic shared by ZYG11B and ZER1 with varying levels of compensation between the two proteins. While it is still possible that ZYG11A is able to target these N-degrons when expressed at higher levels, it is evident that the compensatory mechanisms developed are intricately interwoven between each family member. Moreover, the evolution of ZYG11 poses an interesting perspective as the nature of the gene duplication that gave rise to *ZYG11A* and *ZYG11B* is such that *ZYG11B* sits upstream of *ZYG11A* in the DNA transcript. Perhaps this upstream position of *ZYG11B* is closer to the native promoter for the original *ZYG11* gene making it preferentially transcribed over *ZYG11A* and the reason why *ZYG11A* is less abundant in most tissues. It may also be that *ZYG11B* is more closely related to the ancestral *ZYG11* gene, which could explain some of the shared functions between ZYG11B and ZER1. Moreover, it can be speculated that during the gene duplication of *ZYG11*, the resulting genes were sub-functional (i.e. each gene only inherited partial functionality). To this, the alterations in Scribble expression are very specific to *ZYG11A* knockdown indicating that it may have also acquired neo-functionality throughout evolution and may be important for a completely different subset of functions to ZYG11B and ZER1.

### 3.3.8 Redundancy and Zyg11 family members in the development of *D. rerio*

Information on the ZYG11 family in zebrafish is sparse with no studies that have noted a specific function or phenotype related to *zyg11* or *zer1*. Other than an in-situ hybridisation displaying the ubiquitous expression of *zer1* in the developing embryo, the role of this gene family is essentially a blank canvas. For the first time, genetic knockout fish for both *zyg11* and *zer1* have been generated using CRISPR-Cas9 technologies. Intriguingly, *zyg11* and *zer1* homozygous knockout fish show no developmental defects, nor do they affect the maturation or viability of adult fish. This is strikingly different to *C. elegans* where the loss of *zyg-11* alone is sufficient to cause meiotic defects and embryonic lethality. Indeed, these two systems represent vertebrate and invertebrate models that hold differing levels of embryonic developmental complexity. Aside from this, the disparity between genetic knockdown and knockout phenotypes has long been attributed to genetic compensation. Gene knockdown versus gene knockout can trigger very different outcomes due to a phenomenon that has been largely documented in the *D. rerio* model. In a seminal study by Kok et al using a carefully selected panel of 48 genes, it was found that up to 80% of

gene knockouts showed no or a mild phenotype compared to their knockdown (morphant) counterparts (Kok *et al.*, 2015). The presence of a premature stop codon (PSC), which is a highly favoured genetic alteration in CRISPR-Cas9 knockout models, can trigger the downstream activation of nonsense-mediated-RNA decay (NMD). Activation of this pathway is required for the transcriptional upregulation of genes with homologous sequences, leading to the phenomenon known as genetic compensation (GC) (El-Brolosy *et al.*, 2019; Ma *et al.*, 2019). As discussed in section 3.3.6, this is not the case in *ZYG11A* knockout cells, raising a variety of questions, particularly since these phenotypically normal cells did not show any genetic compensation by other ZYG11 family members. Additionally, the viability of double homozygous knockout *D. rerio* lines for *zyg11* and *zer1* also showed no phenotypic changes, reducing the likelihood of ZYG11 family associated GC. One explanation for the lack of zebrafish phenotypes may lie in the CRISPR-Cas9 sgRNA and subsequent sequencing data that both target the 3$^{rd}$ exon of *zyg11*, as there is one *zyg11* mRNA transcript that begins downstream of the editing site. Indeed, the likelihood of ZYG11 family associated GC is low, however it does not exclude the possibility of the upregulation of other genes with homologous domains. Other than the VHL box, the precise interaction domains for ZYG11 family members are poorly defined. Therefore, I cannot rule out the possibility that this third transcript or possibly another ARM or LRR containing protein may be recruited to prevent any developmental defects.

Aside from genetic compensation, one largely unexplored theory is compensation on a protein level. Since double knockout fish produced embryos that were viable at 5 dpf, the potential masking of a phenotype by maternally deposited *zyg11* or *zer1* is unlikely. During the two meiotic cycles of embryonic development in *C. elegans*, it has been established that while meiosis I requires APC/C to degrade *cyb-1*, at meiosis II this responsibility is transferred to *zyg-11* (Sonneville and Gonczy, 2004). Mitotic cell divisions of *C. elegans* have not yet been studied in the context of *zyg11* loss, however in mammalian somatic cells, ZYG11A/B are known to act redundantly with APC/C at the G2 to M phase checkpoint of the cell cycle (Balachandran *et al.*, 2016). In addition to cell cycle, APC/C has been linked to various forms of cell polarity including asymmetric divisions in *C. elegans* embryos, planar cell polarity in *D. Melanogaster* and ciliary polarity in the *Xenopus* embryo (Ganner *et al.*, 2009; Weber and Mlodzik, 2017). This raises the question of whether APC/C also utilises ZYG11A/B as a backup mechanism in polarity, and more interestingly, entertains the possibility that APC/C (and potentially other E3 ubiquitin

ligases) concomitantly acts as a buffer when ZYG11A/B is lost. Evidently the compensation between not only the ZYG11 family, but other E3 ubiquitin ligases is an interesting concept that requires further investigation.

In the fish, APC/C governs several points in the mitotic cell cycle of the early developing embryo with defects in cell division occurring as early as 3 hours post fertilisation (Wehman, Staub and Baier, 2007). Whether the regulation of meiosis/mitosis in the fish embryo mirror those in *C. elegans* is yet to be determined. However, if this is the case then it may be possible for APC/C to fulfil the role of Zyg11 at meiosis II during vertebrate embryonic development. It would be interesting to see whether the levels of APC/C or Cyclin B are elevated in the zebrafish knockout lines.

All of these possible explanations assume that *zyg11* is as important for embryonic development in the fish as it is in the roundworm, and there is still much work to be done in order to confirm this. As previously mentioned, the ZYG11 family have recently been associated with a novel N-terminal end rule pathway (Timms *et al.*, 2019). These residues are specifically targeted for degradation by $CUL2^{ZYG11B/ZER1}$ and while these motifs are not abundant in animal proteomes, they are the target for the post-translational modification, N-myristoylation, and are also exposed following many caspase-cleavage events. Defects in N-myristoylation, as well as caspase-cleavage by-products likely rely on functional Zyg11 or Zer1 for clearance, although this has not yet been proven in an *in vivo* model. Future studies addressing this question will first need to analyse ZYG11 family null fish for any increase in N-terminal glycine containing proteins. Extending on this in the context of either N-myristoylation inhibition or under increased caspase cleavage will ultimately reveal whether the ZYG11 family fulfils a more protective role by preventing proteotoxicity when normal physiological processes are impaired.

### 3.3.9 Loss of the ZYG11 family in *D. Melanogaster* and epithelial tissue development

Epithelial cell polarity establishment and maintenance is an essential process in the context of both embryonic development and the initiation of cancer, both of which have been extensively modelled in the fly. One example is the polarity protein Scribble, whose disruption can manifest in phenotypes including embryonic lethality and severe defects in tissue morphogenesis. Since there are no studies that have examined the effects of loss of the *D. Melanogaster* ZYG11 orthologue, *zer1*, the phenotypes reported in this chapter are

the first indications that *zer1* loss can affect embryonic viability and normal epithelial tissue patterning in adult flies. Indeed, these provide very rudimentary links to tissue development and due to the lack of fluorescent markers in the genetic drivers used here, there was no definitive way to look at the behaviour of these cells or the specific pathways that were being affected. Perhaps this is related to the specific pathways that these drivers are involved in. For example, the insertion of *zer1*$^{RNAi}$ at the *nubbin* locus may cause an epistatic interaction and affect downstream pathways like Notch signalling (Neumann and Cohen, 1998). This would be interesting to investigate given the link between *zer1* overexpression and Notch in the adult sensory organ (Abdelilah-Seyfried *et al.*, 2000), and the involvement of Notch (and EGF) signalling in wing vein formation (De Celis, 1998; Johannes and Preiss, 2002). In the developing fly eye, previous studies have suggested that reduced eye size caused by *zer1* overexpression can be rescued by *cycE* overexpression (Tseng and Hariharan, 2002). Here I have shown that the knockdown of *zer1* also produces a smaller eye, as well as increased eye roughness and yellowing. Whether manipulating the expression of cell cycle genes like *cycE* can rescue this phenotype will be interesting to test. Still, knockdown using RNAi lines can often produce varying degrees of gene silencing and when tested using other drivers (*en-GAL4, act-GAL4* and *dpp-GAL4*), the RNAi lines did not display any visible phenotypes (data not shown). Together with the mild phenotypes displayed using *GMR-GAL4* and *nub-GAL4* drivers, it is more feasible to instead use the *zer1* mutant for future characterisation. Preliminary observations of the *zer1* mutant line has demonstrated a fundamental requirement for *zer1* prior to the 1$^{st}$ instar larval stage (my unpublished observations). Early lethality of homozygous mutant *zer1* embryos emphasises a need for this protein in embryonic development, similar to that in the developing *C. elegans* embryo. Also reflected in the fly embryo is APC/C mediated degradation of cyclin B, however unlike in *C. elegans* it is required for both meiosis I and II (Swan and Schüpbach, 2007). If both *zer1* and *APC/C* are required for embryonic development, then perhaps in this model the regulation of embryonic polarity is more important than its role in cell cycle. It would be interesting to see whether the lethality observed in either of these mutant embryos can be rescued by the overexpression of the other. Moreover, because of the lethality of *zer1* mutant embryos, I have now generated *zer1* recombinant animals to enable the study of these mutants in a clonal setting. These flies will also more accurately mimic the development of cancer and play a large role in determining how *zer1* loss affects embryonic and adult epithelial tissue development. Unfortunately, due to time constraints these experiments are still ongoing.

## Summary

While it has also been long established that the ZYG11 family is required for the establishment of embryonic cell polarity in *C. elegans*, until now no studies have attempted to progress this work into mammalian systems. The functional characterisation conducted in MCF10A cells in this chapter is the first time that all three mammalian ZYG11 family members have been studied side by side and with a direct focus on epithelial apicobasal cell polarity. This study has shed light on a potential novel mechanism of Scribble regulation at the transcriptional level by only ZYG11A. Few ubiquitination targets for ZYG11A have been identified in regards to cell polarity, so it is now crucial to pinpoint the protein-protein interaction network of ZYG11A in mammalian epithelia as this may also unravel the conundrum of the localisation of ZYG11 proteins. Moreover, the *in vivo* requirement for *zyg-11* in the developing *C. elegans* embryo has emphasised the physiological relevance of this gene family in embryogenesis, and indeed this may also be the case in another invertebrate species, *D. Melanogaster*. I have identified that the knockdown of *zer1* only displays mild alterations in epithelial tissue patterning. Indeed, the use of $zer1^{RNAi}$ lines is the basis for addressing these caveats in future studies, and preliminary observations suggest that *zer1* null embryos display homozygous lethality. Whether the loss of *zer1* affects polarity proteins, the organisation of the actin skeleton, or cell cycle will be an important distinction to make. However, vertebrate models appear to present a different story, suggesting that there is no requirement for *zyg11* or *zer1* in the embryonic and adult development of *D. rerio*. As a well conserved family of genes, the data presented in this chapter questions the fundamental requirement for the ZYG11 family in vertebrate embryonic development. It also raises the question of how many levels of genetic compensation are possible, particularly with an E3 ubiquitin ligase already known to function redundantly with APC/C. Finally, as the sole E3 ligases known to target N-terminal glycine residues *in vitro*, it is possible that the ZYG11 family fulfils an important protective mechanism in *in vivo* vertebrate species. It may be that these are only uncovered through examining the response to variety of environmental challenges and/or in the context of regeneration or pathologies.

# Chapter 4

# Screening and proximity interaction analysis to identify novel ZYG11A genetic and protein interactions in mammalian epithelial cells

## 4.1 Introduction

Protein-protein interactions (PPI) outline the specific functional capabilities of each protein and certainly for ubiquitin ligases like ZYG11A, identifying the nature of these interactions can provide important insights on downstream mechanisms, including those that dictate the phenotypes mentioned in this study. Until now, studies exploring the ubiquitination activity of the ZYG11 branch of the family have focused primarily on cell cycle regulation. Initial investigations in *C. elegans* embryos revealed CYB-1 as a ZYG-11 mediated ubiquitination target during meiosis II (Liu, Vasudevan and Kipreos, 2004). Additionally, the degradation of MEMI ("meiosis to mitosis transition defect") proteins has also been linked to the CUL-$2^{ZYG\text{-}11}$ complex since *zyg11*$^{RNAi}$ embryos maintain MEMI-1, -2 and -3 into mitosis, and the sustained presence of MEMI-1 in mitosis has been found to exhibit similar phenotypes to *zyg-11* and *cul-2* null embryos (Ataeian *et al.*, 2016). In mammalian cell lines, the ZYG11 dependent degradation of Cyclin B1 appears to be evolutionarily conserved with ZYG11A/B providing a back-up mechanism for Cyclin B1 degradation during mitosis following the deregulation of APC/C at the G2 to M phase checkpoint (Balachandran *et al.*, 2016). Furthermore, ZYG11A has been implicated in the regulation of Cyclin E1 levels at the G0/G1 to S phase of the cell cycle (X. Wang *et al.*, 2016). Still, ZYG11A interactions are poorly defined. As ZYG11A is known to act as a substrate recognition subunit, few of the protein-protein interactions that are captured by traditional co-immunoprecipitation (co-IP) and mass spectrometry approaches encapsulate the complete picture. Because co-IP is fundamentally limited to stable protein interactions, a large proportion of ZYG11A binding will be overlooked due to the transient and rapid nature of ubiquitin conjugation, warranting an alternative method of large-scale PPI identification.

In Chapter 3, I described the effects of mammalian ZYG11 family members on cell shape, directed migration, cell cycle and survival. To begin to understand the molecular mechanisms underlying these phenotypes, here I have focused on identifying and functionally validating potential ZYG11A binding partners. I initially identified predicted ZYG11A binding partners in the literature and through database mining and assessed their functional relevance to ZYG11A using a focused mini-RNAi screen. I also assessed the functional relationship between ZYG11/ZYG11A and its known binding partners through rescue screening but did not find any single gene that was able to completely rescue the knockdown phenotype.

Although rescue screening is a reasonable method to rapidly validate ZYG11A candidate substrates, it is intrinsically limited by the scope of the boutique screening library. Without properly characterised protein-protein interaction networks, the loss of a gene like *ZYG11A* leaves limited substrate targets to pursue in terms of identifying the downstream proteins and molecular pathways affected. To complement RNAi rescue screening, I utilised the BioID proximity proteomics method to identify potential new binding partners in the polarised epithelial cell line MCF10As. This is an unbiased proximity biotinylation technique that has already been proven to be able to identify a number of ubiquitination substrates for E3 ubiquitin ligases (Coyaud *et al.*, 2015; Dho *et al.*, 2019). BioID exploits the enzymatic properties of a promiscuous bacterial BirA enzyme which is able to activate free biotin. When this BirA is fused to a protein, biotin activation is localised and results in the biotinylation of lysine resides on proximal proteins which can then be pulled down via streptavidin affinity and analysed by mass spectrometry. This has been achieved here through the generation of a set of myc-BioID-ZYG11A fusion proteins in MCF10A cells. Through BioID, I have revealed a number of ZYG11A proximal proteins, enriched for pathways that have not been previously linked to ZYG11A, as well as cytoskeletal and adhesion proteins that could ultimately explain the *ZYG11A* knockdown phenotype.

## 4.2 Results

### 4.2.1 Identification of potential ZYG11A associated proteins through data mining

The few known ZYG11A substrates identified so far have been associated with the regulation of cell cycle, specifically the regulation of meiosis II in *C. elegans* embryos. Although significant, the minimal cell cycle alterations observed in ZYG11A depleted MCF10A cells is unlikely to explain the other dramatic phenotypes observed such as morphological alterations and decreased survival. Beyond cell cycle machinery related substrates, it is unknown what signalling pathways ZYG11A participates in, and whether these are ubiquitination mediated or not. To identify these pathways, I initially undertook a data mining approach to collate a list of all potential ZYG11A associated proteins. To do this, I datamined for any genes or proteins that had been previously linked to ZYG11A by first searching through the literature and databases for known ZYG11A protein-protein interactions. I used the Biological General Repository for Interaction Datasets (BioGRID),

which hosts a comprehensive list of protein and genetic interactions and retrieved only 6 non-redundant proteins. Since the mammalian branching of ZYG11 gave rise to both ZYG11A and ZYG11B it was also essential to include ZYG11 related proteins from lower organisms and invertebrate models. Aside from the *C. elegans* model, no other genetic models have directly studied ZYG11 functionally, resulting in very few known ZYG11 interacting proteins. However, the *Drosophila* system is particularly interesting as it is known to only have a single homolog that is more closely related to mammalian *ZER1*. *Drosophila* zer1 protein interactions were therefore also included to account for the likely evolutionary conservation of ZYG11A function. These interactions were curated from DroID which is the comprehensive *Drosophila* interaction database. Many of these proteins had multiple human orthologues and therefore I included many of their mammalian counterparts as candidates. In addition, genes that displayed similar morphological and survival phenotypes to *ZYG11A* knockdown in the original screen by Smith et al. (called Cluster Z) were also included (*ZYG11A, FBXO5, FBXO43, VHL, AP2A2* and *BIRC5*) (Smith *et al.*, 2016). Finally, a number of cell cycle regulators, components of the actin cytoskeleton, and other Cullin related E3 ubiquitin ligase associated genes were included as potential candidates for the screen based on relevance to the biology of ZYG11A. This focused list of potential ZYG11A/ZYG11 protein-protein interactions is summarised in Table 4.1.

I then used STRING analysis on the list to identify any prominent pathways (Figure 4.1). This analysis revealed some expected relationships such as the family of Cullin-RING type ligases as well as a number of interesting associated clusters which included the Actin and Tubulin related chaperonin-containing T-complex (TRiC). This protein complex is formed by the oligomerisation of 8 CCT subunits into a ring structure where newly synthesised Actin and Tubulin enter and are correctly folded following a wave of ATP hydrolysis (Sternlicht *et al.*, 1993; Vallin and Grantham, 2019). The ZYG11A associated genes included a number of poorly characterised zinc finger proteins (ZNF169, ZNF264, ZNF483 and ZNF695) that have been predicted to act as transcription factors based on their protein motifs.

**Table 4.1 Summary of potential ZYG11A interacting proteins**

| ZYG11/ZYG11A association | Potential interactors (gene names) |
|---|---|
| **ZYG11 family member**<br>*Vasudevan et al. (2007)* | ZYG11B, ZER1 |
| **Cluster Z**<br>Smith et al. (2016) | ZYG11A, FBXO5, FBXO43, VHL, AP2A2, BIRC5 |
| **Localisation may be affected by ZYG11A**<br>Smith et al. (2016) | SCRIB, CTNNB |
| **Cullin-RING ligase family member**<br>*Petroski and Deshaies (2005) and Willems et al. (2004)* | CUL1, CUL2, CUL3, CUL4A, CUL4B, CUL5, CUL7, CUL9 (PARC), APC2 |
| **CUL2 ligase complex interactions**<br>*Vasudevan et al. (2007)* | RBX1, RBX2, ELONGIN B, ELONGIN C |
| **Bait in Affinity Capture-MS with ZYG11A**<br>*Huttlin et al. (2014)* | FEZ1, ZNF169, ZNF264, ZNF483, ZNF695, C16ORF65 (PDZD9) |
| **FEZ family member**<br>*Alborghetti et al. (2011)* | FEZ2 |
| **Drosophila PPI with zer1 (CG12084)**<br>*DroID database (droidb.org)* | CCT5, CCT3, CCT2, CCT8, CCT6A, CCT6B, CCT7, NUDT21, NRCAM, NFASC, L1CAM, CHL1, SQSTM1, USP42, NUB1, ZNF703, ZNF503, NCAM1, NCAM2, LAMP1, LAMP2, NTN1, NTN3 |
| **TCP family member**<br>*Gene database* | TCP1, CCT4 |
| **Substrates from the literature**<br>*Liu et al. (2004), Sonneville and Gonczy (2004), Wang et al. (2016)* | CYCLIN B1, CYCLIN B3, CYCLIN E1 |
| **Cytoskeleton related**<br>*Gene databse* | ACTB |
| **Proteins that affect Scribble localisation**<br>*Chen et al. (2016)* | ZDHHC7 |

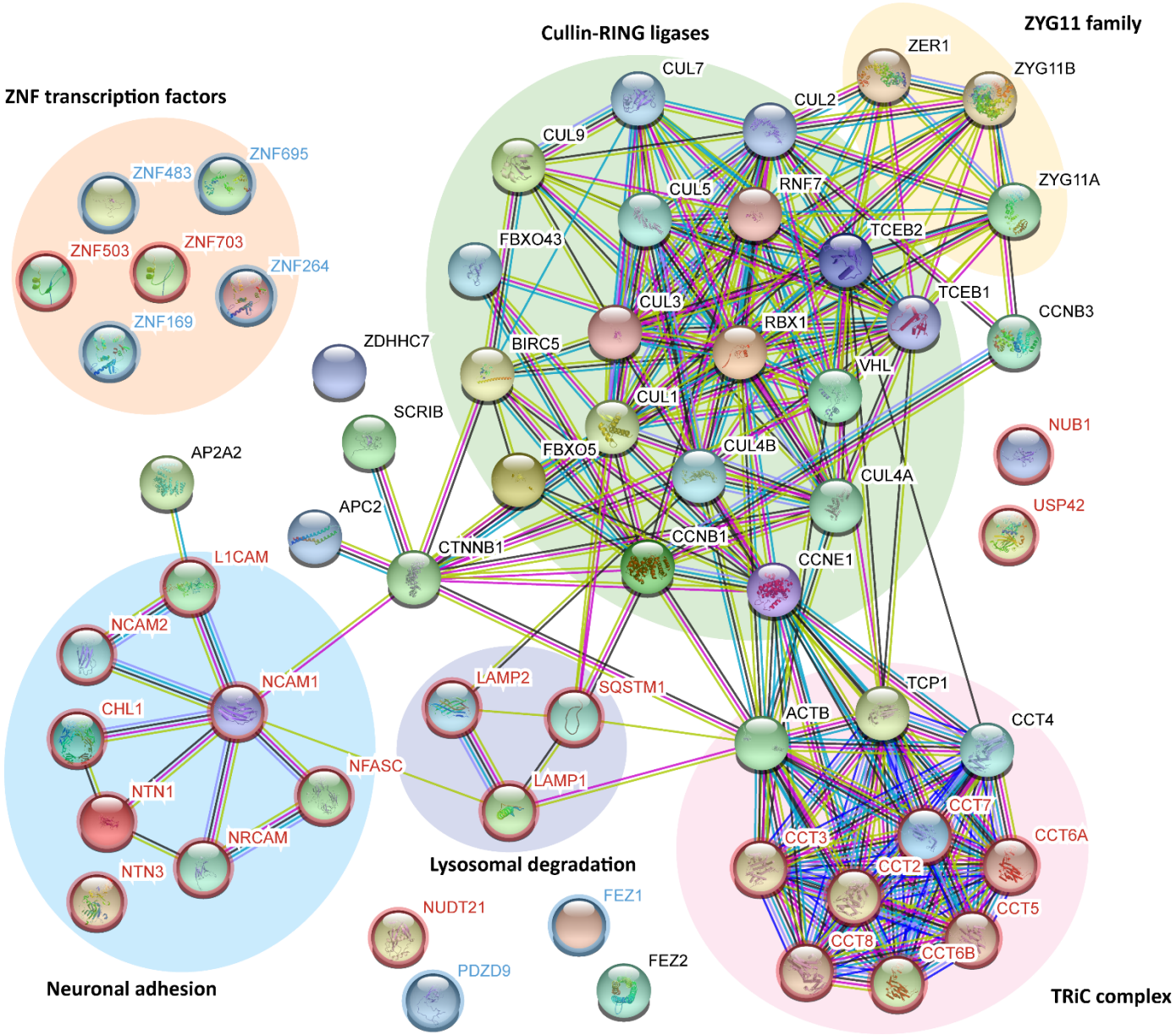

ZNF transcription factors
ZNF695
ZNF483
ZNF503
ZNF703
ZNF264
ZNF169
Cullin-RING ligases
CUL7
CUL2
CUL9
CUL5
RNF7
FBXO43
CUL3
RBX1
BIRC5
CUL1
FBXO5
CUL4B
CUL4A
CCNB1
CCNE1
ZYG11 family
ZER1
ZYG11B
ZYG11A
TCEB2
TCEB1
CCNB3
VHL
ZDHHC7
SCRIB
AP2A2
APC2
CTNNB1
NUB1
USP42
L1CAM
NCAM2
NCAM1
CHL1
NFASC
NTN1
NRCAM
NTN3
Neuronal adhesion
LAMP2
SQSTM1
LAMP1
Lysosomal degradation
NUDT21
FEZ1
PDZD9
FEZ2
TCP1
ACTB
CCT4
CCT7
CCT3
CCT6A
CCT2
CCT5
CCT8
CCT6B
TRiC complex

**Figure 4.1 STRING analysis of the currently known ZYG11/ZYG11A interacting proteins**

Graphical representation of STRING analysis (https://string-db.org/) performed on proteins currently known to physically interact with ZYG11 or ZYG11A. The list is curated from current literature as well as publicly available databases, predominantly BioGRID for ZYG11A binders (blue text and circles), and DroID for *Drosophila* ZYG11 PPIs (red text and circles). Also included are the other Cullin-RING type ligases. Lines between each protein represents a specific association. Clusters of related proteins are distinguished by larger circles, including the ZYG11 family (yellow), the CUL2 complex and other ubiquitin related Cullin-RING proteins (green), the TRiC complex (pink), lysosmal proteins (purple) and neuronal adhesion related molecules (blue). Many zinc-finger proteins are also highlighted in the orange circle.

There were a few proteins that looked promising in view of their high conservation and implication in morphogenesis and ubiquitination. This included Fasciculation and elongation protein zeta-1 (FEZ1), a gene that was first characterised as an important regulator of neuronal axon fasciculation (i.e. the contact and bundling of one axon to others) and axonal guidance in *C. elegans* (unc-76), and in subsequent studies in *Drosophila*, axonal transport via kinesins (Hedgecock *et al.*, 1985; Bloom and Horvitz, 1997; Gindhart *et al.*, 2003). These links with axon elongation and kinesin machinery have been conserved in human cell lines, also describing that FEZ1 is required for neuronal polarisation in the hippocampus through mitochondrial transport (Fujita *et al.*, 2007; Ikuta *et al.*, 2007). Known to be degraded via the 26S proteasome, FEZ1 degradation has been tied to the U-box ligase, E4B as well as the APC/CDC20 ubiquitin ligase, which is particularly interesting since it has been well established that ZYG11A/B can act redundantly with APC/C in human cell lines (Watanabe, Khodosevich and Monyer, 2014; Balachandran *et al.*, 2016). With this list of promising candidate ZYG11A pathways in hand which may be relevant to non-cell cycle related phenotypes, I went on to functionally validate these interactions using an RNAi high content screening approach.

### 4.2.2 Optimisation of high-throughput functional screening

In order to assess the relationship between ZYG11A and my curated list of potential interacting genes, I used a high-content screening method to functionally characterise the knockdown phenotype of each gene individually, then in combination with *ZYG11A* knockdown (Figure 4.2 and Figure 4.5). In the first instance, I wanted to establish whether RNAi depletion of a candidate could phenocopy *ZYG11A* depletion, indicating a role related to ZYG11A. I then used co-depletion of the candidate gene with *ZYG11A* to assess

whether it was an enhancer of ZYG11A (expected stronger phenotype), or a suppressor (expected weaker/rescued phenotype), also potentially indicating it might be a substrate for ZYG11A (Figure 4.4A) To be able to use the original screening method that was used to identify ZYG11A as a regulator of cell shape in MCF10A cells (Smith *et al.*, 2016), I first needed to optimise each parameter of the siRNA transfection conditions and confirm the validity of the system in my own hands as a reliable readout of phenotypic and functional variation.

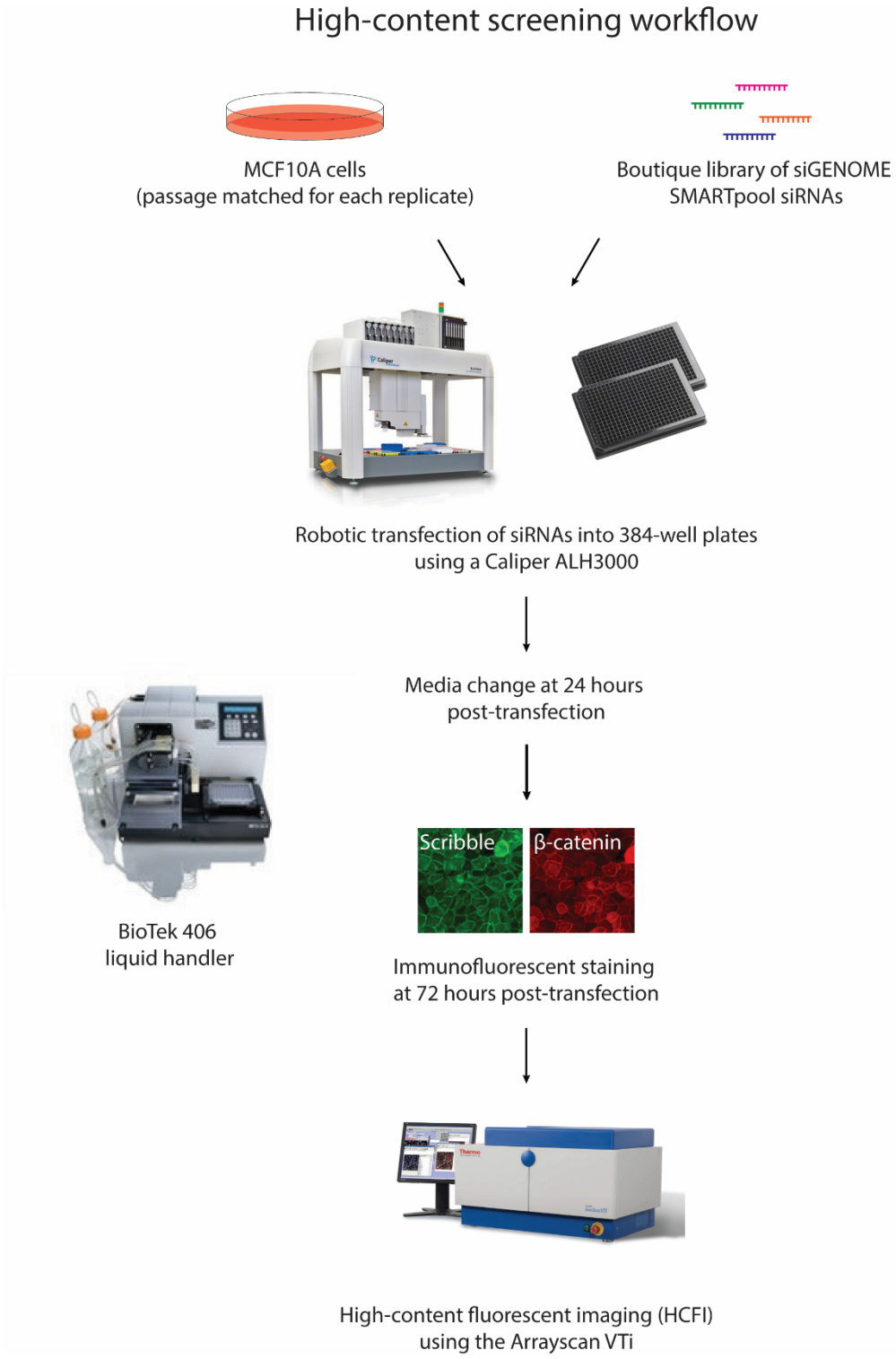


**Figure 4.2 High-content functional genomic screening workflow**

Workflow of the high-content screening pipeline. Passage matched MCF10A cells are reverse transfected with a boutique library of siRNAs towards known ZYG11/ZYG11A binding proteins via robotic transfection in a 384-well plate screening method. At 72 hours post-transfection, cells are immunofluorescently stained with fluorescently labelled antibodies for the known polarity and adhesion markers, Scribble and β-catenin. High-content imaging then allows for the visualisation of any resulting alterations in cell plating density, cell morphology and cell polarity.

To start, MCF10A cells were reverse transfected (i.e. transfected with siRNA prior to cells adhering) in a 384-well screening plate and allowed to settle for 24 hours before replacing the media. I initially optimised the assay for the formation of a confluent polarised monolayer of control MCF10A cells by the assay end point of 72 hours, as well as a high transfection efficiency. To do this, control transfected MCF10A cells (un-transfected and a non-targeting OTP control) were seeded at increasing densities in a 384-well plate to determine the correct number of cells that form a confluent layer using a DNA dye (DAPI) together with the polarised localisation of Scribble and β-catenin at the membrane of cells to count and identify cells. In addition, two lipid concentrations were used to identify the amount which gave the best transfection efficiency without causing lipid toxicity in the target cells. Transfection efficiency was determined by three main transfection controls, two of which were siRNAs that are known to potently induce cell death – siPLK1 (polo-like kinase 1) and siTOX which is a proprietary apoptosis inducing control by Dharmacon. The third control was an siRNA that fluoresces when transfected (siGLO by Dharmacon) that allowed the visual identification of siRNA transfer (Figure 4.3A). To maintain consistency, siRNA concentrations were kept at 40 nM per well. At the 72-hour endpoint, these cells were stained for Scribble, β-catenin and DAPI, and high-content images (25 fields in the centre of each well) were taken using the Arrayscan VTi automated imager. Cellomics software collated all images to create montages of each well and allowed for the manual observation of cell confluence. I identified the optimal transfection conditions to be achieved with 800 cells and 0.06 μl of lipid per well in a 384-well plate. Here the un-transfected, lipid only cells and the OTP controls both reached confluence and showed membranous Scribble and β-catenin. Cell death as a readout of transfection efficiency was similarly observed in both siPLK1 and siTOX where almost 100% of cells were dead at the assay endpoint. Finally, the siGLO transfected cells displayed high levels of fluorescence further confirming high levels of transfection (Figure 4.2B). The optimisation of this assay was a necessary step in implementing a successful screen, and it was now

clear that these conditions would provide a comparable system to determine alterations in cell shape, polarity and overall cell plating density.

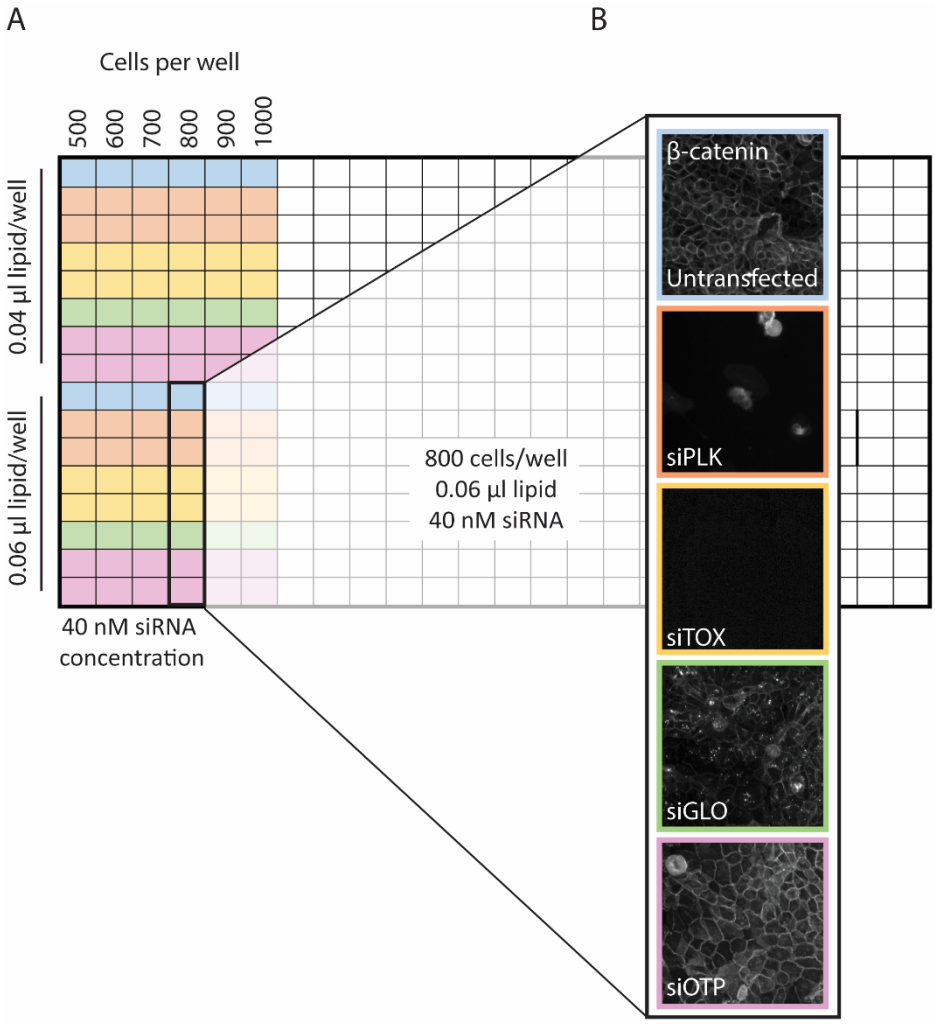



**Figure 4.3 Optimisation of high-content screening**

**A.** Annotated layout of the assessed parameters in a 384-well plate. Cells seeded per well incrementally increase and are noted at the top of each well. Varying lipid concentrations are noted along the left of the plate along with the siRNA screening concentration of 40 nM (final). Each well is colour coded based on the control siRNA used: Untransfected lipid only control (blue), OTP non-targeting control (pink), siPLK1 (orange) and siTOX (yellow) to induce cell death and determine transfection efficiency, as well as siGLO (green) as a method of fluorescently determining transfection efficiency. **B.** Representative immunofluorescent images of the parameters deemed optimal based on 1.) the presence of a confluent monolayer of cells with distinct cell-cell boundaries (marked by β-catenin), 2.) low to no cell viability under cell death induction, and 3.) high levels of fluorescence in each cell based on siGLO transfection. Each transfection was performed in duplicate wells except for siGLO. Images were taken at the 72-hour endpoint.

### 4.2.3 Loss of ubiquitin related proteins CUL2, EMI1 and USP42 phenocopy the loss of ZYG11A

Although it has long been described that the ZYG11 family facilitates Cullin-2 mediated ubiquitination, it is also a possibility that ZYG11A has roles and interactions that are separate from this complex, particularly since the phenotype of ZYG11A is markedly different from ZYG11B and ZER1. Taking the list of potential ZYG11A/ZYG11 associated proteins assembled above, a boutique siRNA library was used to identify any genes, that when knocked down, were able to phenocopy the *ZYG11A* knockdown phenotype. Each knockdown was assessed by manually searching for any knockdowns that showed high cell concentrations (cell clumps) and cell elongations that stretched between cell clumps. The cell counts of any hits was also analysed using Cellomics software by DAPI staining. Although I attempted to use various image analysis softwares and methods to quantify the clumping and cell elongation phenotype, none of these approaches at the time of analysis of this screen could confidently discriminate and quantitate this phenotype. Hence, I had to rely on visual confirmation of the morphological phenotype by eye.

Of the 61 knockdowns observed, careful visual inspection revealed the targets that produced the most similar morphological phenotypes to loss of *ZYG11A* function were Cullin-2 (*CUL2*), ubiquitin specific peptidase 42 (*USP42*) and early mitotic inhibitor 1 (*EMI1*) (Figure 4.4A-B). Of note, the observation that the best characterised ZYG11 interactor, CUL2 appears to behave similarly to ZYG11A in my assay provides confidence of the screen's ability to identify other relevant interactors. Importantly, *CUL2*, *USP42* and *EMI1* loss all showed cell clumping and elongated cell extensions similar to *ZYG11A* loss. It should be noted that in the Smith et al, 2016 screen which tested over 700 individual genes related to cell polarity signalling and cell shape using a similar high content imaging screen method, only 6 candidates ("Cluster Z" that includes ZYG11A and EMI1, see Table 4.1) showed this particular morphological phenotype, i.e. this is not a common alteration in cell morphology seen in MCF10As. In my screen *CUL2*, *USP42* and *EMI1* knockdown also showed the same plating defects as *ZYG11A* knockdown (Figure 4.4A-B). Interestingly, all of these genes have been linked to protein ubiquitination, but also cell cycle regulation (Julie D.R. Reimann *et al.*, 2001; Verschuren *et al.*, 2007; Hock, Vigneron and Vousden, 2014; Cai and Yang, 2016b). It was therefore important to determine whether the morphological and cell plating phenotype of these cells was related to specific cell cycle alterations. Using the Cellomics software on the high content images and DAPI as a readout of DNA levels, the cell cycle profiles of these cells were determined. Although

hits showed aberrations in cell cycle phasing, these were not identical between hits. While the *ZYG11A, CUL2* and *EMI1* knockdown cells were similar to the controls, *USP42* knockdown displayed a significant increase in S phase cells and a decrease in G2 cells (Figure 4.4C). Thus, while there are similarities in phenotypic knockdown, the observation that morphological and survival phenotypes were not associated with any particular cell cycle phasing strongly indicate that cell shape and survival phenotypes may not be cell cycle dependent. These data have highlighted genes whose depletion are able to phenocopy *ZYG11A* knockdown, consistent with the notion that they may be able to work alongside ZYG11A to regulate cell shape and polarity.

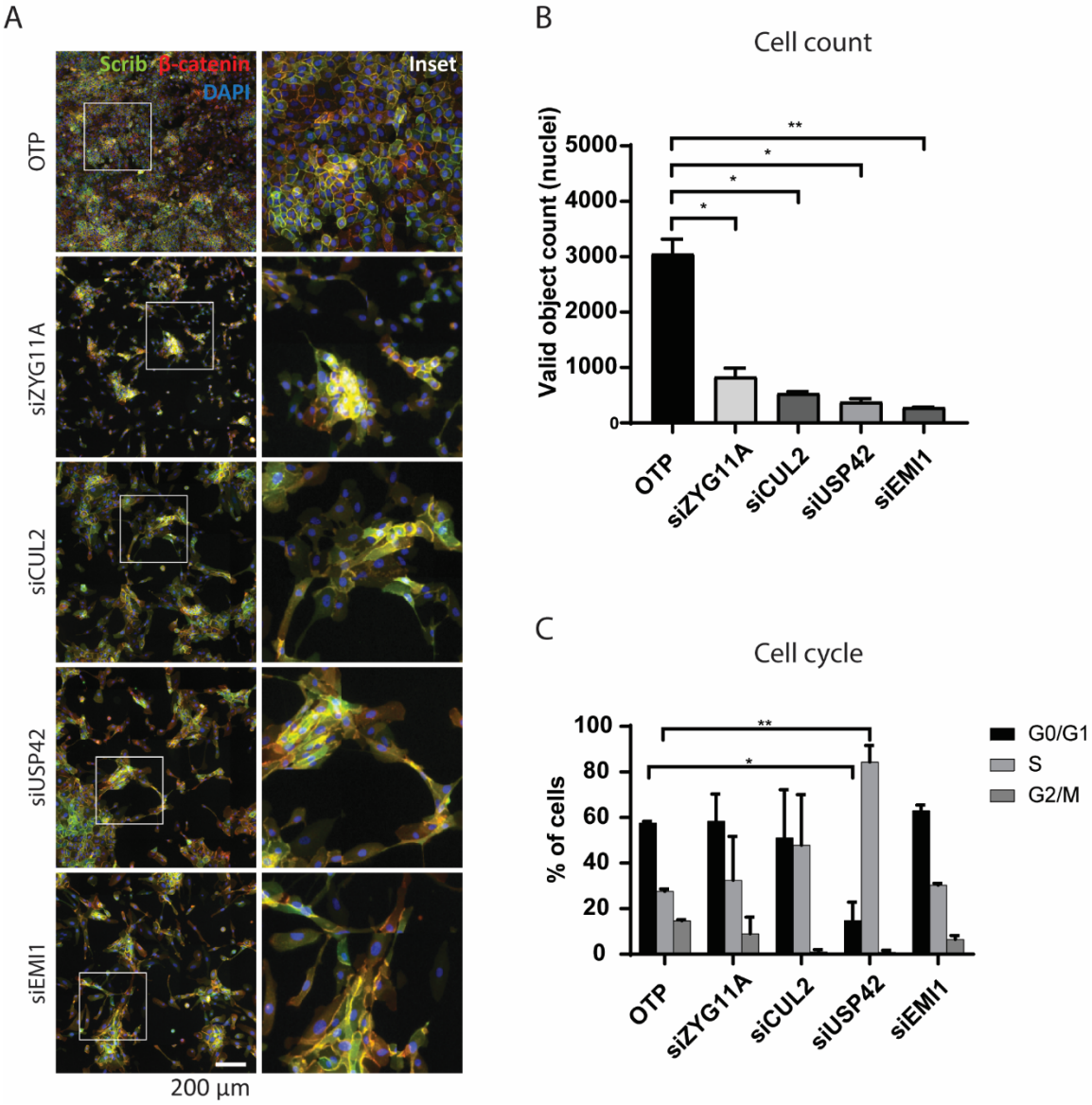



**Figure 4.4 *ZYG11A* loss is phenocopied by other ubiquitin related proteins, CUL2, USP42 and EMI1**

**A.** Representative immunofluorescent images of OTP, siZYG11A, siCUL2, siUSP42 and siEMI1 MCF10A knockdown cells at 72 hours post-transfection. Cell membranes are identified by staining of Scribble (green) and β catenin (red) with DAPI (blue) to mark nuclei in each merged image. White boxes highlight regions of interest that display clumped and elongated cells, and these are enlarged in the left-most inset panels. **B.** Quantification of the valid objects through the number of remaining nuclei from **A**, including the OTP non targeting control. **C.** High-content cell cycle analysis of the cells remaining on the plate at 72 hours post transfection from **A.** Each cell cycle phase is determined by the DAPI channel and is segregated into each cell cycle by Cellomics software based on the fluorescence intensity – 2N = G0/G1, 2N/4N = S and 4N = G2/M. Error bars represent ± SEM and are representative of 2 independent screening experiments each performed in duplicate. (n=4, Student's t-test or Two-way ANOVA).

### 4.2.4 *ZYG11A* knockdown cell plating density is partially rescued by ubiquitin-related proteins

With only 3 genes of the 61 candidates producing a similar phenotype to *ZYG11A* knockdown, I aimed to probe further into another important relationship that these genes may have with ZYG11A. Substrate recognition subunits (SRSs) like ZYG11 family of proteins are extremely important for specifying which proteins will targeted for degradation, playing a large role in ensuring that the selected proteins are within reach of E2 mediated ubiquitin transfer. It is then unsurprising that the loss of a substrate recognition subunit, like in *ZYG11A* knockdown, can lead to the build-up of unwanted proteins that ultimately disrupt proteostasis. With the expectation that the loss of ZYG11A leads to increases in certain proteins that are responsible for the alterations in cell shape and polarity, I tested whether the accumulation of any of the candidate ZYG11A interacting proteins could account for some or all of its loss of function phenotypes. To test this hypothesis, I optimised a ZYG11A loss of function "rescue" screen by first titrating the siRNA concentration down to 25 nM and ensuring that the controls and the ZYG11A knockdown phenotype remained robust but did not cause too much lethality from siRNA multiplexing (data not shown). Once this was verified, rescue screening was done by using 25 nM of the siRNA base (OTP or ZYG11A) and 25 nM of the siRNAs from the boutique library to identify any genes that were able to rescue *ZYG11A* knockdown (Figure 4.5A-B). Taking into consideration the increase in *P21* mRNA expression described in Chapter 3, I also added *TP53* into the library. Unfortunately, none of the candidates were able to rescue the cell morphology defects (see below) however some of these were able to modestly rescue the survival defects as measured by the rescue of the cell plating density phenotype. The siRNA targets that were able to increase the cell plating density were *NYREN18* (*NUB1*), *TCEB2* (Elongin B), *ZNF169*, and surprisingly, *CUL2* and *USP42* (Figure 4.6B). Indeed, as shown previously both loss of function of *CUL2* and *USP42* alone showed similar phenotypes as loss of *ZYG11A* (Figure 4.4). What this means in terms of the functional relationship between *CUL2*, *USP42* and *ZYG11A* remains unclear, but these data demonstrate that the regulatory mechanisms of this ubiquitin ligase are more complex than originally anticipated. It is interesting that some of the partial rescues (*CUL2, TCEB2, USP42* and *ZNF483*) displayed similar cell counts regardless of *ZYG11A* knockdown, suggesting that these may act upstream and dominant to ZYG11A. In addition, two gene knockdowns decreased the plating density, *CCT6B* and *C16ORF65* (*PDZD9*), i.e. were enhancers of the ZYG11A loss of function phenotype (Figure 4.6B). In most cases however, compared with the OTP control, the knockdown of *ZYG11A* in

combination with these partial rescue hits still caused a significant decrease in cell plating density (Figure 4.6A) indicating other potential targets are responsible for this phenotype.

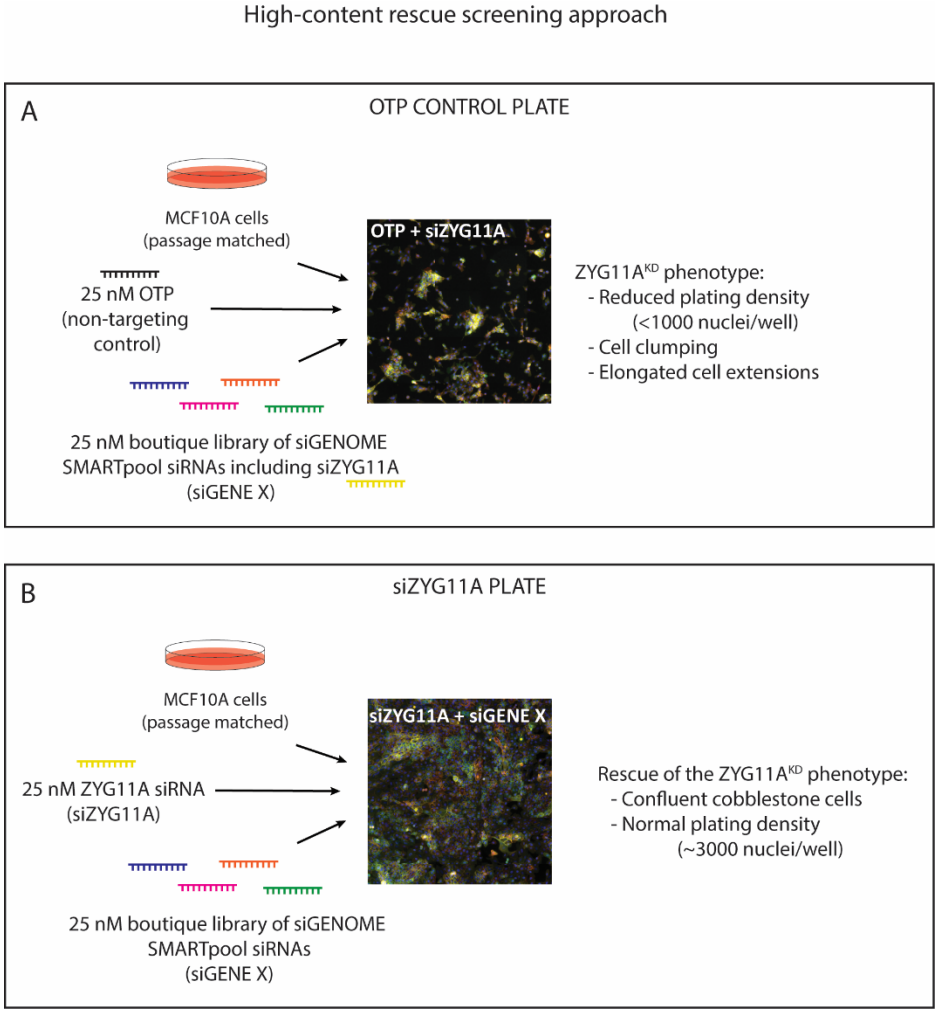



**Figure 4.5 High-content rescue screening workflow**

A simplified workflow of the rescue screening approach beginning with the knockdown of ZYG11A (**A**) or an OTP control (**B**) in MCF10A cells at 25 nM, optimised to robustly produce the ZYG11A knockdown phenotype. On top of this siRNA base is the addition of the boutique screening library (25 nM). If the primary ZYG11A knockdown (siZYG11A) results in the stabilisation of a normally degraded target leading to the aforementioned phenotype, any secondary knockdowns (siGENE X) that restore normal MCF10A morphology are likely to be ZYG11A target substrates.

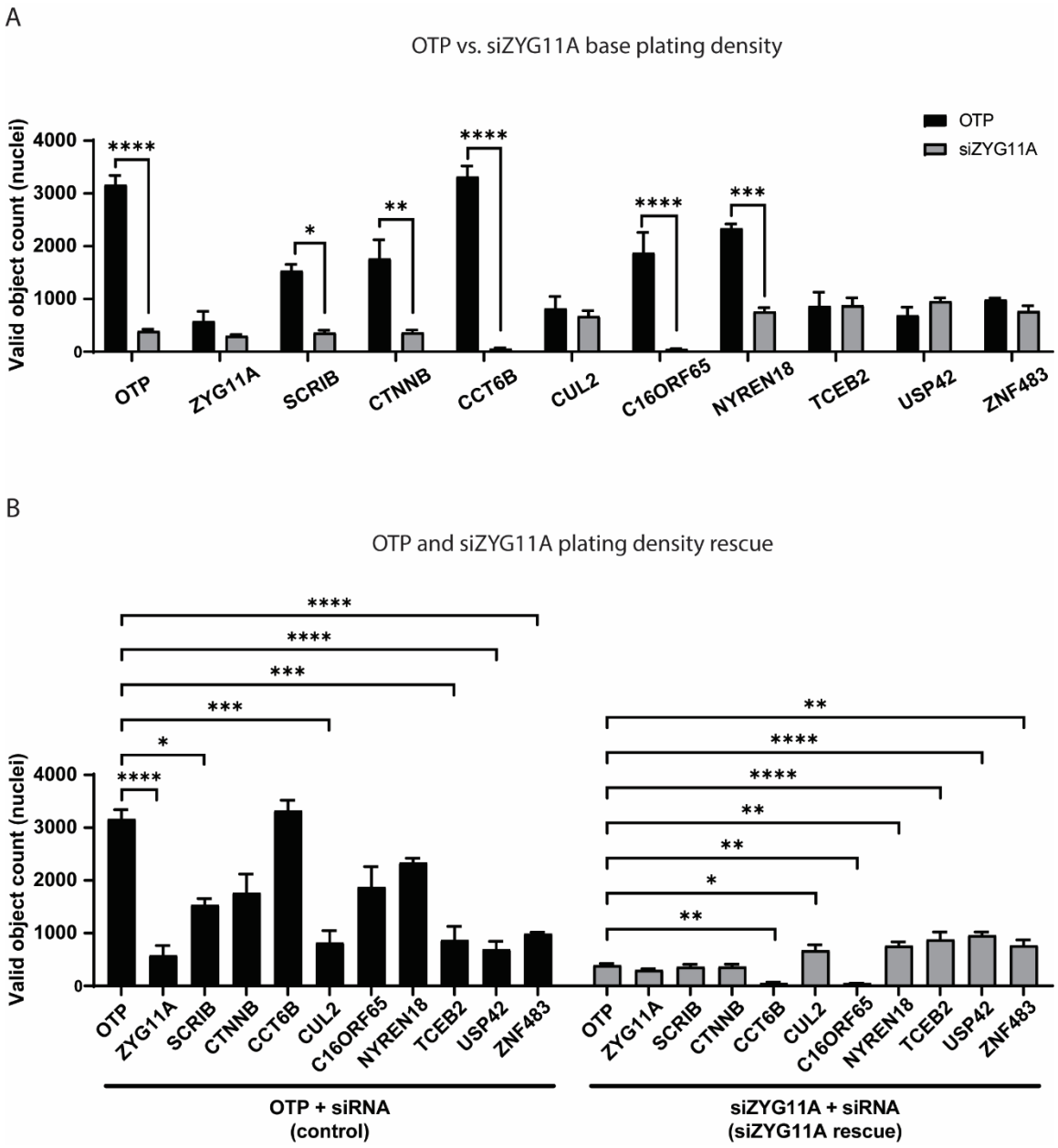



**Figure 4.6 ZYG11A rescue screening reveals novel genetic knockdown interactions that partially rescue the ZYG11A loss phenotype**

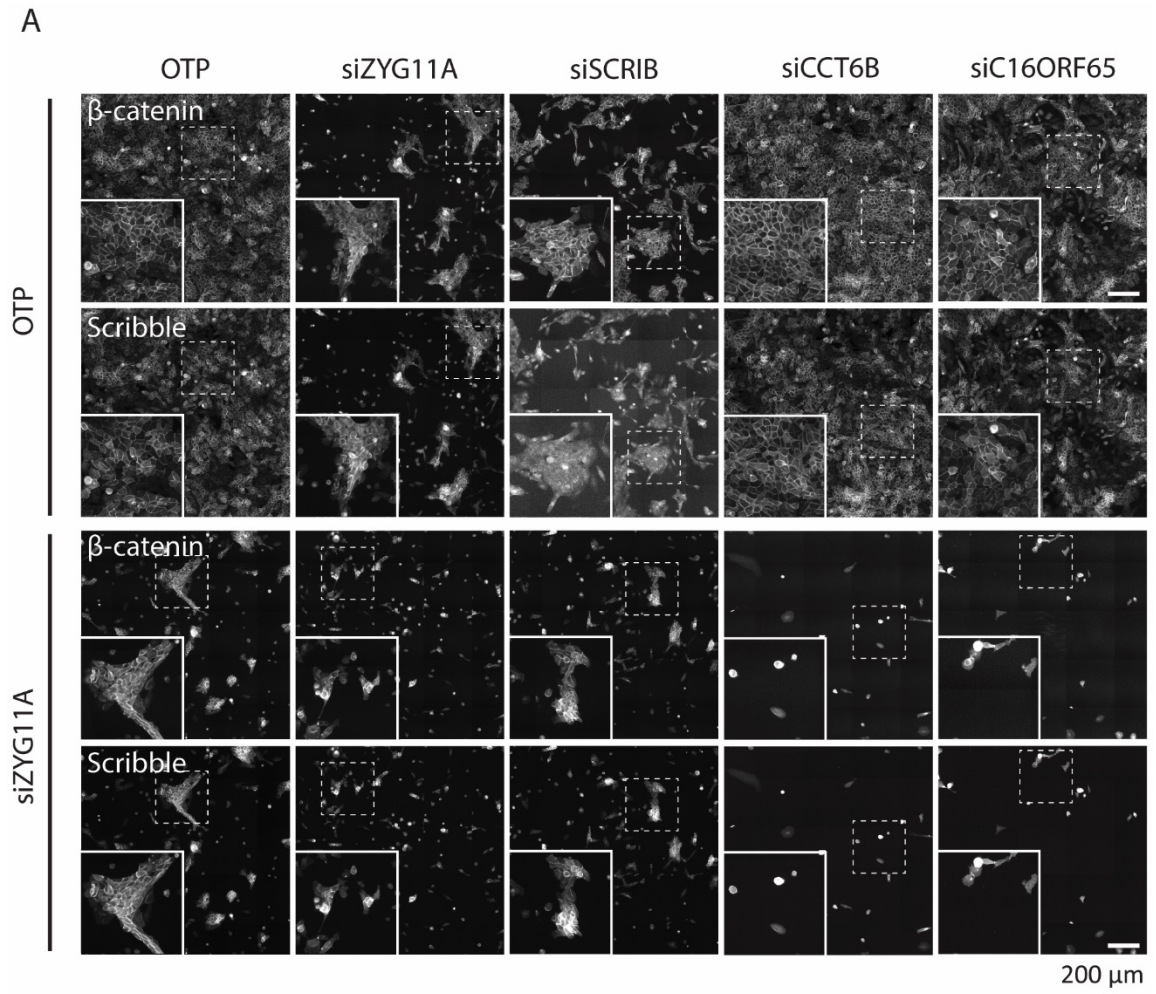
**A.** Quantification of the valid object counts or nuclei remaining at the 72-hour post transfection endpoint following OTP control or siZYG11A double knockdown combinations. Selected to be shown here are the genes specifically able to increase the cell plating density compared to *ZYG11A* knockdown alone. *ZYG11A* knockdown shows a significant decrease in nuclei compared to OTP controls for many library candidates. **B.** The same nuclei counts as in **A,** re-arranged to highlight the rescuers and enhancers of the siZYG11A phenotype. Note that these are only partial rescues as OTP controls typically contain roughly 3000 valid objects. Error bars represent ± SEM and are representative of 2 independent screening experiments each performed in duplicate. (n=4, Student's t-test).

### 4.2.5 ZYG11A regulates polarity and cell shape proteins, Scribble and β-catenin via other ubiquitin-related proteins

Loss of ZYG11A has now been demonstrated to drive alterations in cell cycle, cell death, migration and polarity. The rescue of the cell plating density described above can reflect a number of functional changes, so in order to assess specific changes in cell polarity or cell morphology determinants, I examined the staining of Scribble and β-catenin in more detail. Of note, due to the nature of the automatic image acquisition in my high throughput assay, I routinely observed an automatic overexposure characterised by a white-saturated appearance when Scribble or β-catenin levels are extremely low in the high throughput microscopy (for example, in *SCRIB* or *CTNNB* knockdown cells) (Figure 4.7A and Figure 4.8A, note that the junctional localised signal of Scribble or β-catenin is also absent in these). Neither of the hits that were able to enhance the *ZYG11A* knockdown phenotype showed any alterations in these proteins, albeit there were very few cells left on the plate (Figure 4.7A). Strikingly, 4 of the 5 genes that were able to rescue the plating density due to *ZYG11A* loss displayed a high level of overexposure in the Scribble staining channel and loss of junctional Scribble signals, suggesting a loss of Scribble (Figure 4.7B). These genes were *TCEB2 (*Elongin B*), NYREN18 (NUB1), USP42* and *ZNF483*. The last hit, *CUL2*, did not show any changes in Scribble expression (Figure 4.7B). Moreover, these hits that showed Scribble loss did not harbour any loss of β-catenin, a protein that is typically associated with Scribble at the membrane (Figure 4.7B). Looking outside of the rescue hits, I then sought to identify any double knockdowns that displayed the opposite phenotype – i.e. normal Scribble levels and a loss of β-catenin. I identified the knockdown combinations of *ZYG11A* with either *ZYG11B, ZER1, LAMP2*, and *NTN1* were able to affect β-catenin levels but not Scribble (Figure 4.8 B). None of these showed any rescue or enhancement of the *ZYG11A* knockdown phenotype (Figure 4.8C) Although the mechanism by which polarity and adhesion proteins, Scribble and β-catenin are altered in the above situations remain unclear, it further supports a relationship between ZYG11A and the control of cell polarity and adhesion.

In summary, although the mini screens have identified some interesting but complex genetic epistatic relationships between potential ZYG11A interactors and ZYG11A, it is also evident that none of these postulated ZYG11A associated proteins are able to completely rescue the loss of ZYG11A alone. This indicates other potential targets are responsible for the ZYG11A phenotypes. In the next section, I have focussed on using a

proximity proteomics approach to better understand the signalling pathways that ZYG11A may be regulating in a cellular context.


A
OTP
siZYG11A
siSCRIB
siCCT6B
siC16ORF65
β-catenin
Scribble
OTP
siZYG11A
200 µm

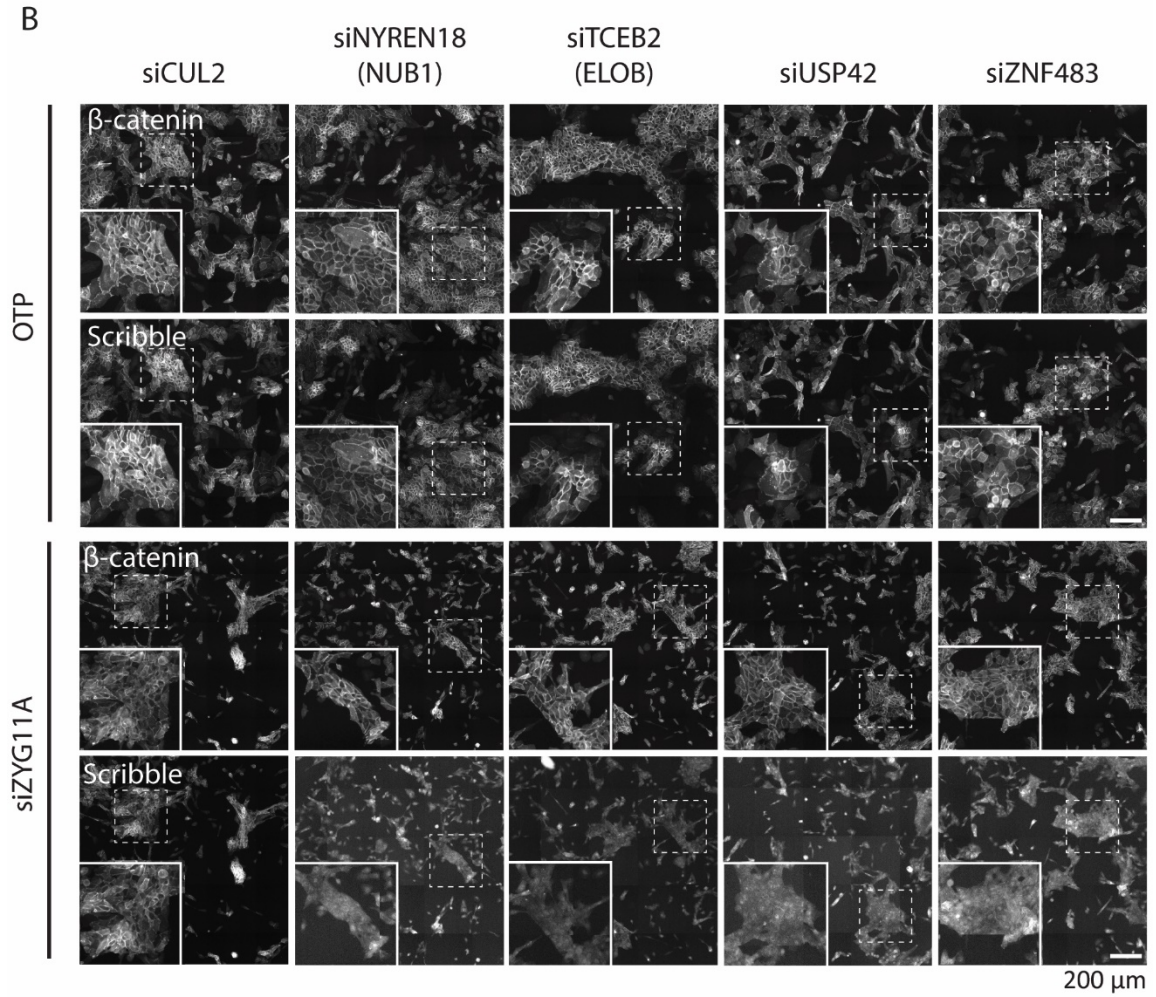
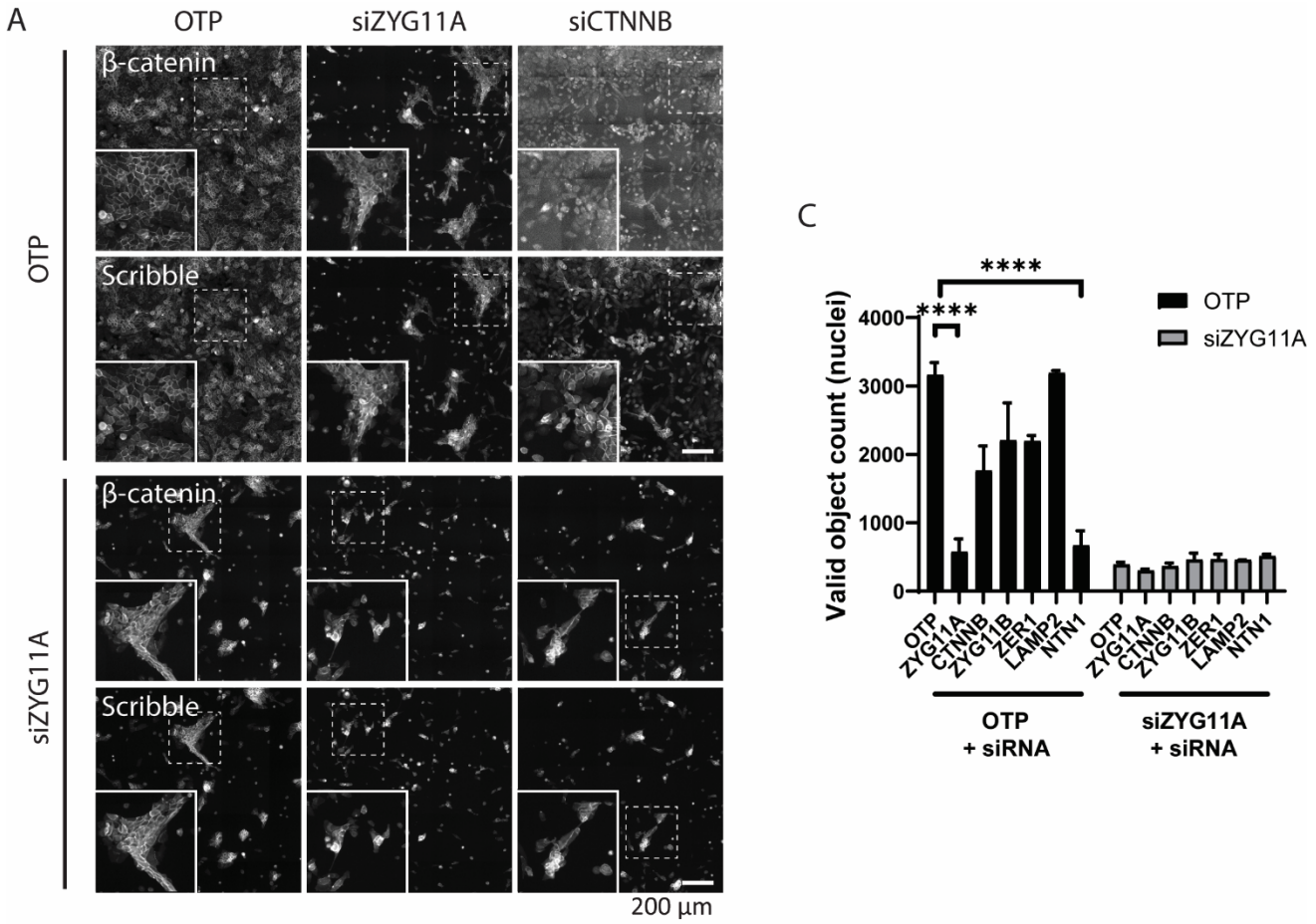

B
siCUL2
siNYREN18
(NUB1)
siTCEB2
(ELOB)
siUSP42
siZNF483
β-catenin
Scribble
OTP
siZYG11A
200 µm

**Figure 4.7 *ZYG11A* knockdown combinations that partially rescue the plating density defects show reductions in Scribble protein levels**

Representative immunofluorescent images of OTP and *ZYG11A* knockdown cells under double knockdown combinations with OTP, siZYG11A, siSCRIB, siCCT6B and siC16ORF65 (**A**) or siCUL2, siTCEB2, siNYREN18, siUSP42 and siZNF483 (**B**). Images display staining of β-catenin (top row) and Scribble (bottom row) at the screening endpoint of 72 hours post transfection in MCF10A cells. On the Arrayscan high content imaging platform, overexposure as seen in the OTP + siSCRIB cells are indicative of the depletion/lack of staining of that protein, i.e. Scribble. Note that β-catenin staining remains normal in these conditions. **B.** 4 of 5 siZYG11A rescue hits (siTCEB2, siNYREN18, siUSP42, siZNF483) show the depletion of Scribble, but not β-catenin. Without *ZYG11A* knockdown, normal Scribble and β-catenin staining are present. All images are representative of 2 independent screening experiments each performed in duplicate. (n=4).


A
OTP
siZYG11A
siCTNNB
β-catenin
Scribble
OTP
siZYG11A
200 µm
C
Valid object count (nuclei)
4000
3000
2000
1000
0
****
****
OTP
siZYG11A
OTP
ZYG11A
CTNNB
ZYG11B
ZER1
LAMP2
NTN1
OTP
+ siRNA
siZYG11A
+ siRNA

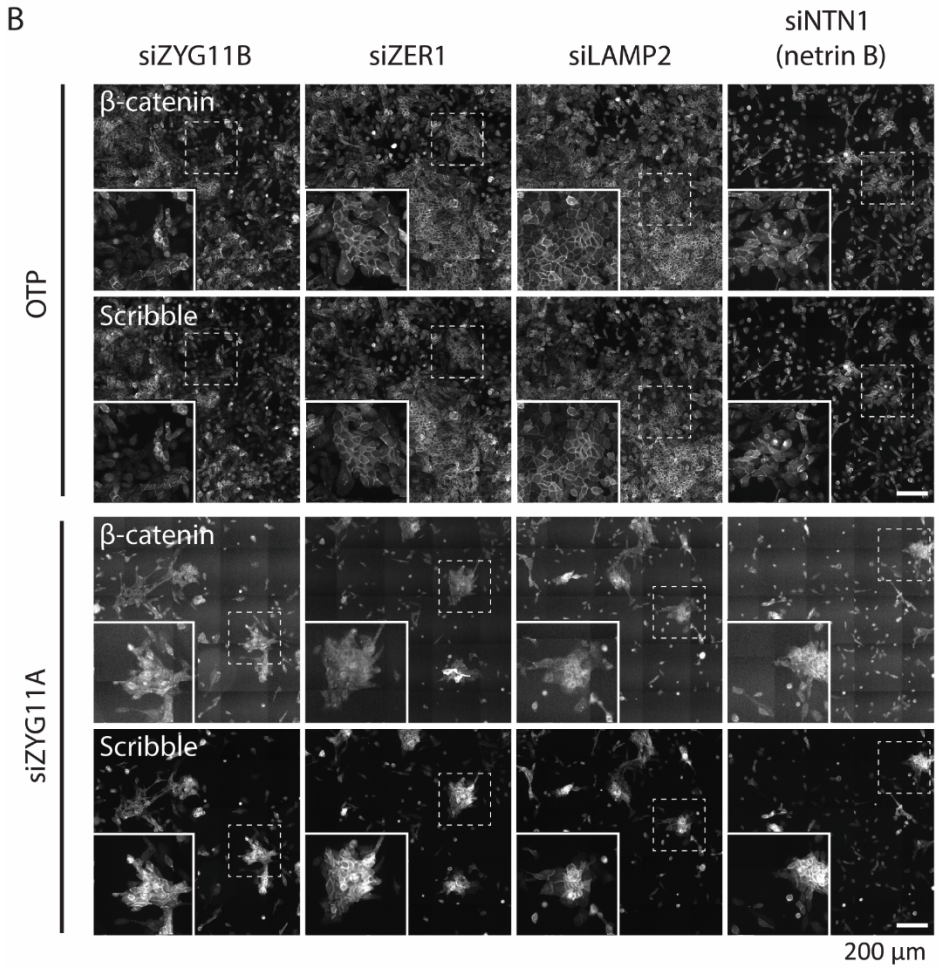

B
siZYG11B
siZER1
siLAMP2
siNTN1
(netrin B)
β-catenin
Scribble
OTP
siZYG11A
200 µm

**Figure 4.8 *ZYG11A* knockdown combinations can also affect β-catenin levels without any effect of Scribble**

**A.** Representative immunofluorescent images of OTP or *ZYG11A* knockdown cells under double knockdown combinations with OTP, siZYG11A or siCTNNB (β-catenin). Images display staining of β-catenin (top row) and Scribble (bottom row) at the screening endpoint of 72 hours post transfection in MCF10A cells. On the Arrayscan high content imaging platform, overexposure as seen in the OTP + siCTNNB cells are indicative of the depletion/lack of staining of that protein, i.e. β-catenin. Note that Scribble staining remains normal in these conditions. **B.** Representative immunofluorescent images of the *ZYG11A* knockdown combinations that show a depletion of β-catenin but not Scribble. These are siRNAs towards *ZYG11B*, *ZER1*, *LAMP2* and *NTN1*. **C.** These do not show any rescue or enhancement of the *ZYG11A* knockdown phenotype. All images are representative of 2 independent screening experiments each performed in duplicate. (n=4).

### 4.2.6 Development of ZYG11A BioID proximity biotinylation fusion constructs

In order to carry out BioID proximity biotinylation, I first needed to generate ZYG11A fusion proteins and validate the biotinylation function. To do this, I fused a myc-BioID tag to the N-terminal of full length ZYG11A. The N-terminal was chosen as it should not interfere with the function of ZYG11A based on previous studies using ZYG11A/B fusion proteins in the degradation of Cyclin B1 (Balachandran *et al.*, 2016). In addition to the tag only control, I also generated two truncated forms of the ZYG11A fusion protein that separated the leucine-rich repeats (N-terminal) and the Armadillo domains (C-terminal). While it is still unclear as to where the substrate recognition activity occurs along the ZYG11A protein, having these two domains separated were likely to reveal more robust interactions that may be specific to N- or C-terminal protein motifs (Figure 4.9A). These cells were stably expressed in MCF10A cells by retroviral transduction and analysed by western blot in order to validate their expression. When expressed alone and under standard RIPA lysis conditions, the BioID tag was detected by western blotting (Figure 4.9B, ~25 kDa in lane 2). However, despite multiple repeats I could not detect the expression of the three ZYG11A fusion proteins (Figure 4.9B, lanes 3-5). I also utilised various other lysis protocols such as NP40 and Laemmli to determine whether this may be the issue, but I was still unable to detect the expression of the three fusion proteins (data not shown).

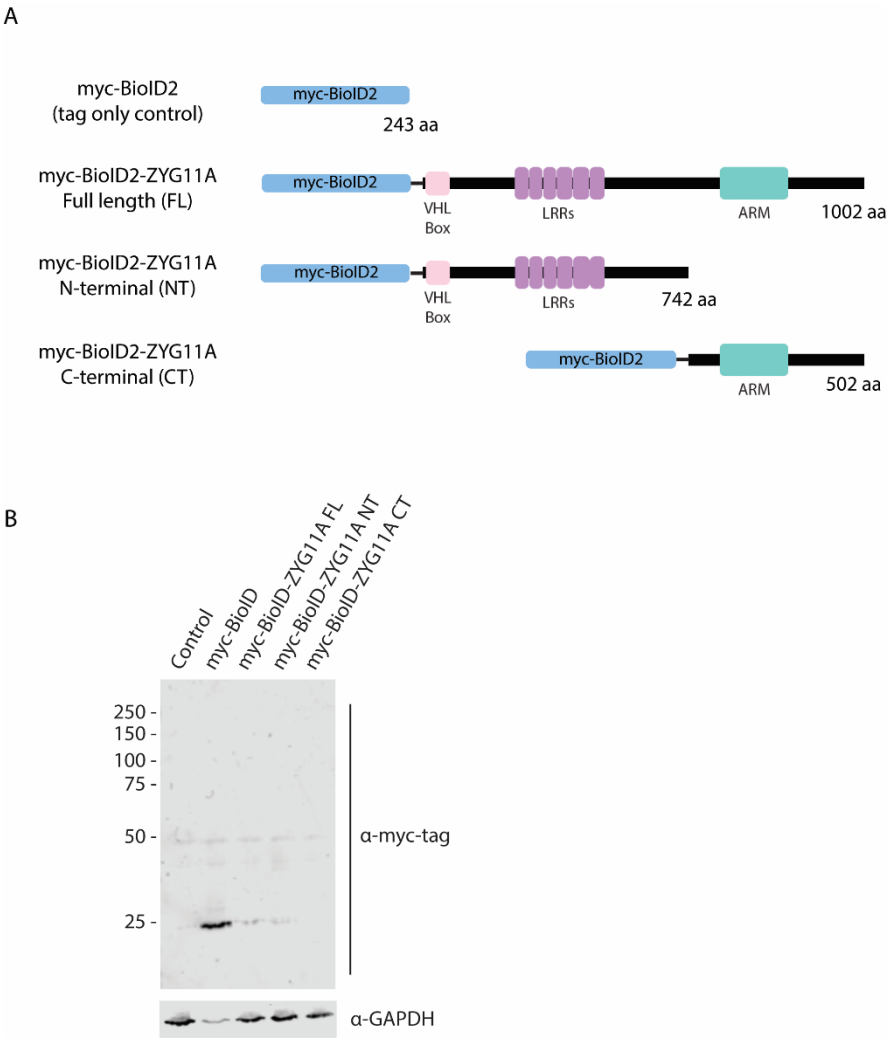



**Figure 4.9 Generation of myc-BioID-ZYG11A fusion proteins**

**A.** Depiction of the myc-BioID-ZYG11A fusion proteins that were stably expressed in MCF10A cells. The full length contains all three known ZYG11 protein motifs. The N-terminal truncation harbours the von Hippel-Lindau (VHL, pink) and leucine rich repeats (LRR, magenta), whereas the C-terminal truncation harbours the armadillo domains (ARM, teal). Each fusion protein is N-terminally tagged with the myc-BioID fusion (blue). **B.** Western blots displaying RIPA soluble protein fractions for empty controls, myc-BioID only, full length, N-terminal and C-terminal myc-BioID fusion proteins for ZYG11A. The myc-BioID tag control is present in lane 2 at ~25 kDa. Fusion proteins in lanes 2-5 show no detectable bands at the predicted molecular weights (ZYG11A (FL = 113 kDa, NT = 83 kDa, CT = 57 kDa).

To determine whether this was a common issue in the GFP-ZYG11A fusion proteins that were previously detected by immunofluorescence (Figure 3.2), protein expression of these fusions was also tested by western blot. Under RIPA lysis conditions, the GFP tag displayed a band at ~25 kDa (Figure 4.10A, lane 1) whereas the full length, N-terminal and C-terminal fusion proteins did not display any bands at the predicted molecular weights (Figure A, lanes 2-4). This was also the case for GFP-ZYG11B and GFP-ZER1 fusion proteins. The detection of these GFP fusion proteins by immunofluorescence (Figure 3.2 and Figure 3.3) suggested that this lack of detectable protein was not due to expression. However, there appeared to be a faint smear running in each of the fusion protein lanes which is characteristic of either protein degradation (i.e. protein lability in solution) or protein insolubility. To distinguish between these two possibilities, I first tested whether protein stability was the issue by treating these cells with the proteasome inhibitor MG132 (also added to the lysis buffer) to prevent protein degradation, however this did not result in detectable protein expression by western blot under RIPA lysis conditions (data not shown). While RIPA buffer is typically used as a whole cell lysis method, it still requires a preclearing step to remove the remaining highly insoluble protein fraction. To determine whether the ZYG11 fusions proteins were present in this normally discarded fraction, RIPA lysed cells were additionally supplemented with 8M urea in order to completely solubilise the remaining proteins. Similar to the RIPA only samples, the addition of urea showed a GFP band at ~25 kDa in lane 1 across all protein sets (Figure B). The fusion protein samples in lanes 2-4 displayed an increase in protein smearing with a visible band present at the end of each smear that corresponded with the expected molecular weights of each fusion protein. This confirmed that the ZYG11A, ZYG11B and ZER1 fusion proteins were indeed present in the RIPA insoluble fraction. Based on the presence of each protein smear despite the addition of urea, it was evident that these fusion proteins were highly insoluble and therefore difficult to manipulate. For simplicity, these three fractions will be herein referred to as the soluble fraction (RIPA soluble supernatant), the insoluble fraction (RIPA insoluble pellet) and the whole cell lysate (WCL) (RIPA + urea soluble fraction).

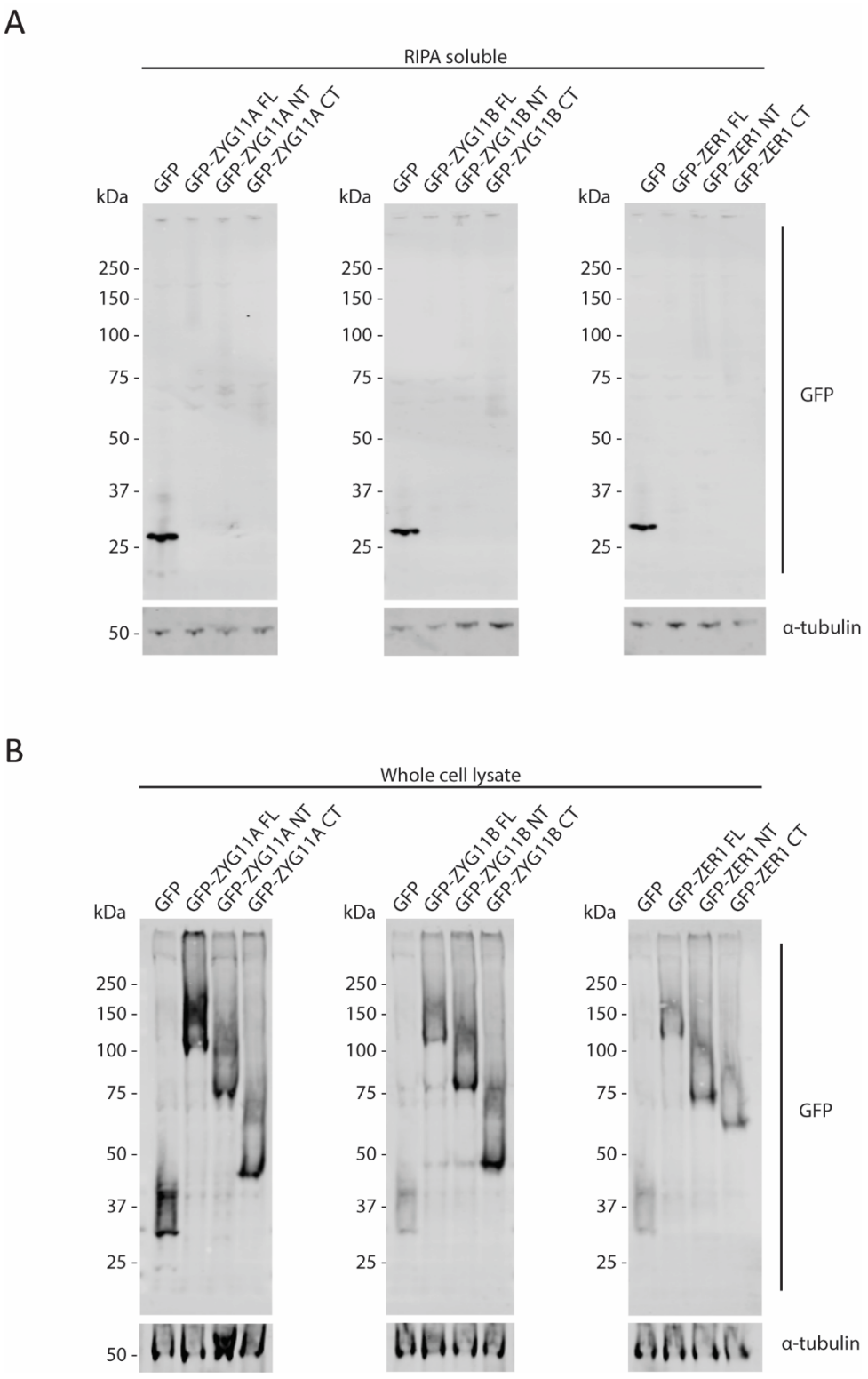

A
RIPA soluble
GFP
GFP-ZYG11A FL
GFP-ZYG11A NT
GFP-ZYG11A CT
GFP-ZYG11B FL
GFP-ZYG11B NT
GFP-ZYG11B CT
GFP-ZER1 FL
GFP-ZER1 NT
GFP-ZER1 CT
kDa
250
150
100
75
50
37
25
GFP
α-tubulin
B
Whole cell lysate

**Figure 4.10 ZYG11 family fusion proteins are largely present in the insoluble protein fraction**

**A.** Western blots displaying RIPA soluble protein fractions for GFP only, full length, N-terminal and C-terminal GFP fusion proteins for ZYG11A, ZYG11B and ZER1. In all three blots, the GFP only control is present in lane 1 at ~25 kDa. Fusion proteins in lanes 2-4 show faint protein smears, but no distinguishable bands **B.** Whole cell lysate western blots using the ZYG11A, ZYG11B and ZER1 GFP fusion proteins described in **A**. The RIPA lysed cells were supplemented with 8M urea lysis buffer to completely solubilise the RIPA insoluble fraction. Protein smears are largely present in each of the fusion protein lanes and display a band at the end each smear at the expected molecular weights: GFP-ZYG11A/GFP-ZYG11B (FL = ~113 kDa, NT = ~83 kDa, CT = ~57 kDa) or ZER1 (FL = ~117 kDa, NT = ~77 kDa, NT, CT = ~68 kDa)).

With these insoluble properties in mind, the same rationale was applied to the MCF10A cells stably expressing the myc-BioID ZYG11A full length, N-terminal and C-terminal fusion proteins. To test whether proteasome inhibition was able to stabilise BioID fusion proteins, they were treated with MG132 (or a DMSO control) and the whole cell lysate was analysed by western blot. In both cases, there was a visible smear ending at the predicted molecular weights in each fusion protein lane, however this was more visible under MG132 treatment (Figure 4.11A, black arrows), suggesting that these ZYG11A BioID fusion proteins are stabilised by proteasome inhibition. Additionally, to test whether the BioID biotinylation activity was functioning, biotin (50 μM) was also added to the media to allow for biotinylation to occur. Under control conditions, there was minimal biotinylation by the fusion proteins compared to the tag only control, however this was increased upon MG132 treatment (Figure 4.11B). This is expected given that the fusion proteins are stabilised following proteasome inhibition, and that many of the ZYG11A interacting proteins are also likely to be stabilised if they are ubiquitination substrates.

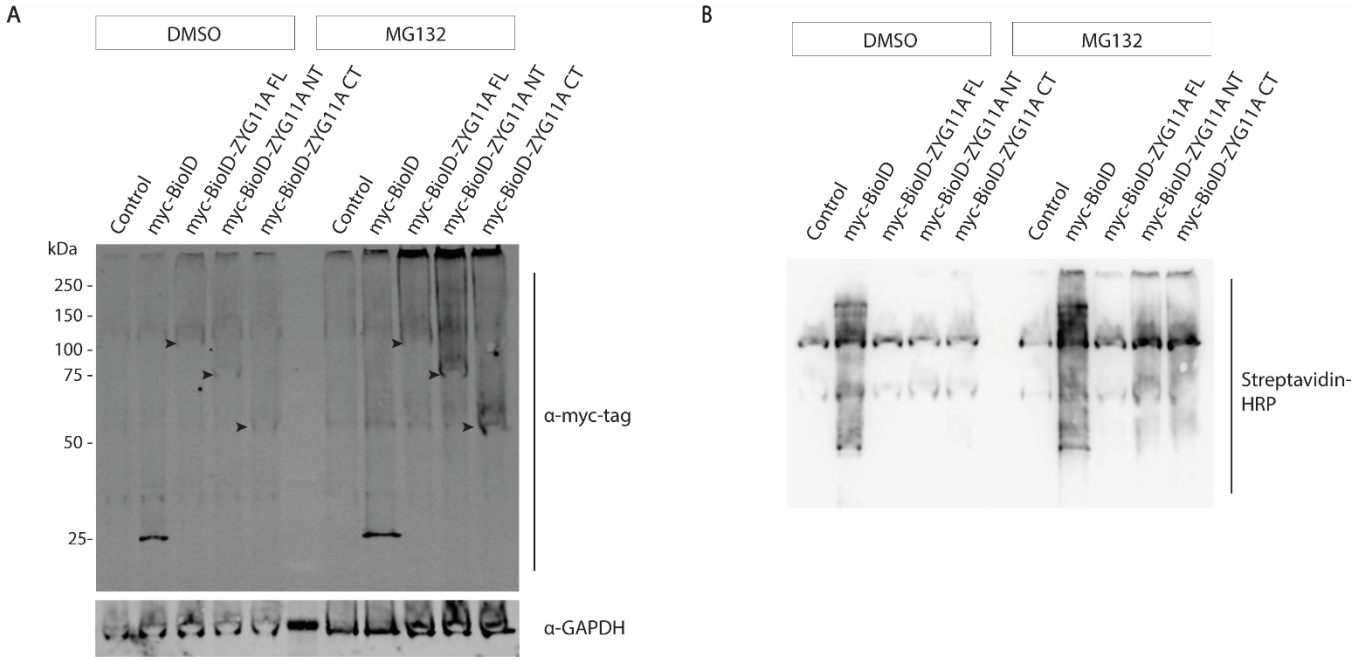



**Figure 4.11 ZYG11A BioID fusion proteins are stabilised by proteasome inhibition**

**A.** Whole cell lysate western blot of myc-BioID-ZYG11A fusion proteins, including am empty vector control, a myc-BioID tag only, and ZYG11A full length, N-terminal and C-terminal fusion proteins under +/- proteasome inhibition (DMSO or 10 μM MG132, 16 hours). Bands are present at the expected sizes for the tag only control (25 kDa), and protein smears run to where each of the ZYG11A fusion proteins are expected to be (FL = 113 kDa, NT = 83 kDa, CT = 57 kDa). Note that protein smearing is darker under MG132 treatment. **B.** Analysis of proximity biotinylation activity in the whole cell lysate western blot in **A**, detected by anti-streptavidin-HRP. The BioID only lane under DMSO and MG132 treatment displays a large number of biotinylated proteins. The biotinylation profiles of each fusion protein appear to vary greatly. The increase in fusion protein smearing in **A** corresponds with visibly darker biotinylation following MG132 treatment.

Under these lysis conditions, protein smearing was still visible in all fusion protein samples. As DNA related viscosity can increase smearing, to better resolve the proteins the whole cell lysates were treated with the endonuclease, Benzonase for 1 hour to digest any DNA and RNA viscosity. This resulted in the improved resolution of the fusion proteins bands (Figure 4.12A). As with the GFP fusion experiments, the BioID fusion proteins were analysed as separate fractions to examine whether these ZYG11A fusion proteins were also found in the insoluble fraction. Following RIPA lysis, the insoluble fraction was separated by centrifugation and completely denatured in 8M urea. Western blots revealed that while the myc-BioID tag was present in both fractions, albeit largely in the soluble fraction, the fusion proteins were almost exclusively found in the insoluble fraction (Figure 4.12B). To again confirm the biotinylation activity of each fusion protein

as well as test the BioID assay conditions, biotin was added to the media at 16 hours prior to harvest. Since many of the proteins that BioID-ZYG11A was expected to capture were likely to be ubiquitination targets with rapid turnover, this was also conducted +/- proteasome inhibition (via MG132 treatment) to enrich for potential substrates. Between these two conditions, biotinylated proteins were mainly detectable under proteasome inhibition (Figure 4.12C), consistent with previous whole cell lysate experiments (Figure 4.11). Importantly, this activity was also largely within the insoluble protein fraction suggesting that ZYG11A interacts with many other proteins associated with the insoluble fraction (Figure 4.12D). Taken together, these results place each ZYG11A fusion protein, along with much of their interaction network, within the insoluble protein space of MCF10A cells.

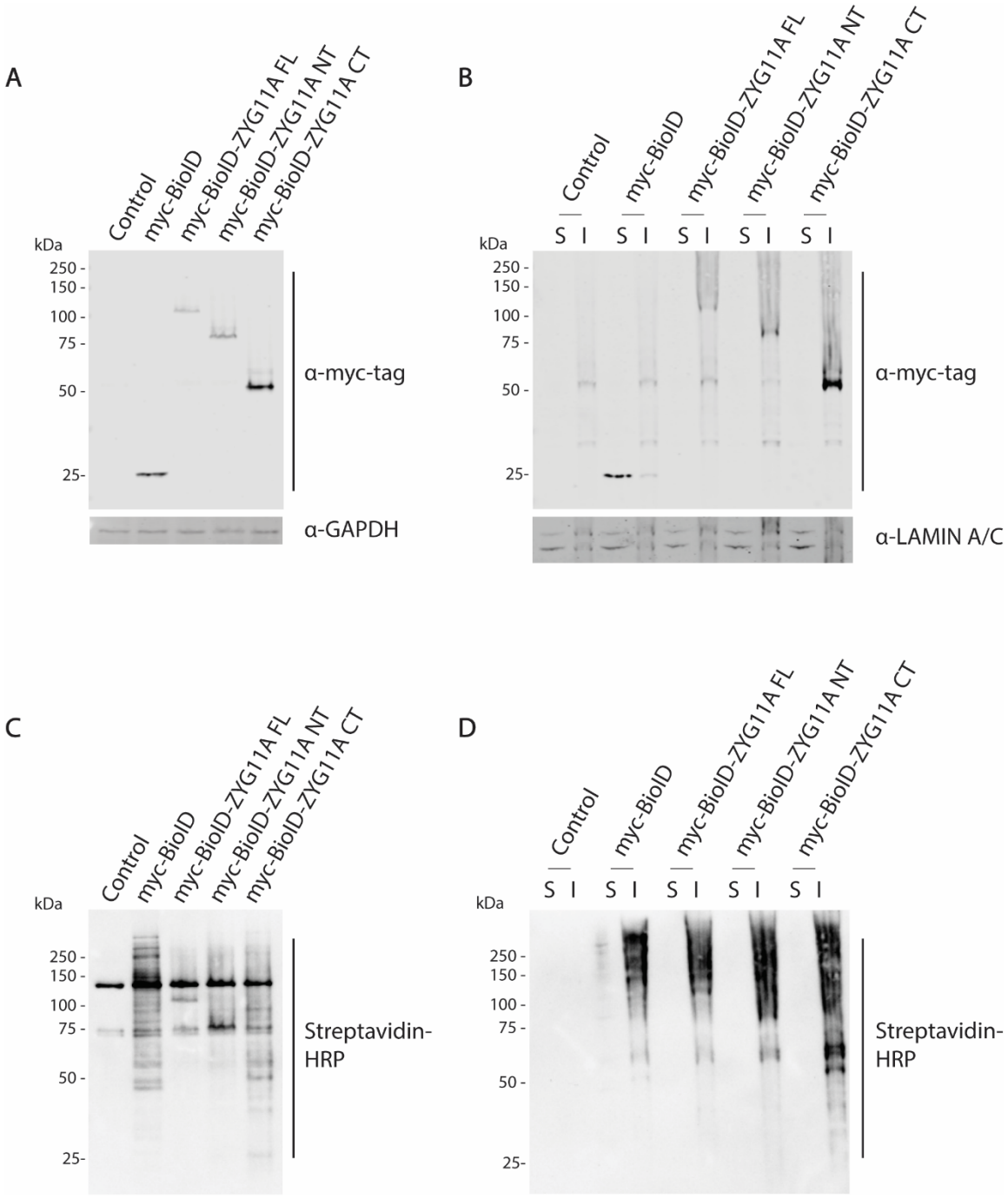



**Figure 4.12 ZYG11A BioID fusion proteins are present in the insoluble protein fraction with biotinylation activity mainly detectible under proteasome inhibition**

**A.** Whole cell lysate western blot of myc-BioID-ZYG11A fusion proteins, including am empty vector control, a myc-BioID tag only, and ZYG11A full length, N-terminal and C-terminal fusion proteins under the BioID proteasome inhibition (10 μM MG132, 16 hours). Bands are present at the expected sizes for the tag only control (25 kDa), and each of the ZYG11A fusion proteins (FL = 113 kDa, NT = 83 kDa, CT = 57 kDa). **B.** Western blot displaying each of the BioID fusion proteins in **A** with lysates separated into the RIPA soluble (S) and RIPA insoluble, urea soluble (I) fractions. The tag only control is mainly present in the soluble fraction with a faint band in the insoluble fraction. Each of the fusion proteins are only visible in the insoluble fractions. **C.** Analysis of proximity biotinylation activity in the whole cell lysate western blot in **A**, detected by anti-streptavidin. The BioID

only lane displays a large number of biotinylated proteins. The biotinylation profiles of each fusion protein appear to vary greatly. **D.** Analysis of the difference in biotinylation activity within the soluble and insoluble protein fractions. The tag only control shows activity in both fractions, prominently in the insoluble fraction. Only insoluble biotinylated proteins are visible in each of the ZYG11A fusion proteins.

### 4.2.7 ZYG11A fusion proteins aggregate under 26S proteasome inhibition

The GFP-ZYG11A fusion proteins previously displayed specific subcellular localisations where the full length (FL) and N-terminal (NT) proteins were excluded from the nucleus and the C-terminal (CT) protein was distributed throughout the cell. To determine whether the myc-BioID fusion proteins showed identical localisation properties to the GFP fusions, the MCF10A stable BioID cell lines were examined by confocal microscopy. Since proteasome inhibition was going to be used as a way to enrich for ubiquitinated proteins, it was also important to determine how these fusion proteins localised after MG132 treatment. As with the GFP lines, the myc-BioID fusion proteins showed nuclear exclusion (FL+NTT) and cell wide distribution (CT) under normal conditions (Figure 4.13A-D, first two columns). Additionally, biotinylated proteins detected by streptavidin immunofluorescence colocalised with the fusion proteins, albeit lowly detectable compared with the myc-BioID tag alone (Figure 4.13A-D, third row of images), which is consistent with levels of detectable biotinylation observed in western blots. However, under MG132 treatment the level of ZYG11A fusion proteins and protein biotinylation increased (Figure 4.13A-D, last two columns). In some cells, this was also accompanied by an enrichment of these proteins in what seemed to be protein aggregates forming at a juxtanuclear position (Figure 4.13B-D, arrows). Control cells did not show GFP aggregates (Figure 4.13A); therefore, they were likely to be caused by the presence of the ZYG11A fusion proteins and this was indeed the case for GFP-ZYG11A fusion proteins as well (Figure 4.14).

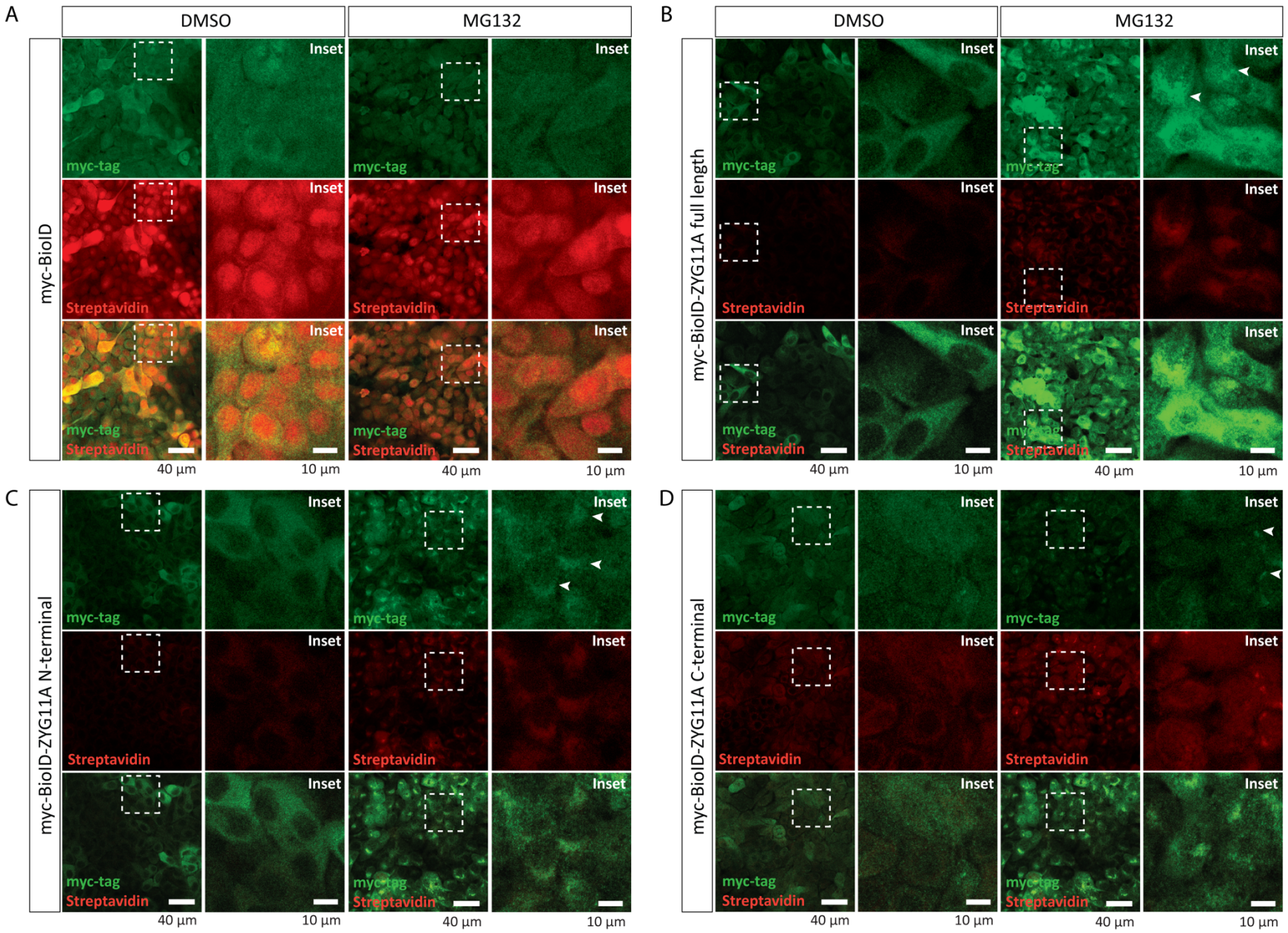

A
DMSO
MG132
myc-BioID
B
myc-BioID-ZYG11A full length
C
myc-BioID-ZYG11A N-terminal
D
myc-BioID-ZYG11A C-terminal
myc-tag
Streptavidin
Inset
40 µm
10 µm

**Figure 4.13 ZYG11A BioID fusion proteins localise similarly to GFP fusion proteins and form aggregates under proteasome inhibition**

Representative confocal images of the localisation of the myc-BioID only, myc-BioID-ZYG11A full length (FL), myc-BioID-ZYG11A N-terminal (NT) and myc-BioID-ZYG11A C-terminal (CT) fusion proteins detected through the myc-tag (green). Note that under 10 μM MG132 treatment, the ZYG11 fusion proteins show altered localisation and form protein aggregates (top row of each fusion protein, arrows). The respective protein biotinylation profile of each fusion protein detected by streptavidin (red), bottom rows of each fusion protein). Note that since there are only low levels of biotinylation in the ZYG11A fusion protein samples, no red is visible when overlayed with the green signal. Although the localisation of the ZYG11A FL+NTT fusion proteins is altered under MG132 treatment, most biotinylated proteins remain in the cytoplasm.

Given that these protein aggregates were not observed under steady state conditions, this suggested that the inhibition of the proteasome was also responsible for protein aggregation. Because the use of MG132 was an essential step required to enrich for protein degradation targets, it was important to identify what these aggregates were. This process was likely to be a form of macroautophagy known as aggrephagy – a mechanism often triggered to alleviate the toxic accumulation of ubiquitinated proteins under proteasomal inhibition. Since each ZYG11A fusion protein was also stabilised by MG132, it was important to test whether this was related to ZYG11A ubiquitination. Confocal analysis of ubiquitin localisation +/- MG132 revealed an increase in ubiquitin staining localised within GFP-ZYG11A protein aggregates, suggesting that this protein aggregation was in response to the aberrant accumulation of the ZYG11A fusion proteins under proteasomal stress (Figure 4.14, right panels). To see whether these aggregates were associated with the juxtanuclear organelle, the Golgi apparatus (GA), the GA marker GM130 was used, however there did not appear to be any co-localisation between the GA and the ZYG11A fusion proteins (Figure 4.15A). A characteristic feature of aggrephagy is the re-routing of ubiquitinated proteins towards the lysosome for degradation. This occurs near the nucleus and can be marked by lysosomal proteins like LAMP2. To observe whether this was the case here, aggregates of the GFP-ZYG11A fusions were stained for LAMP2 and showed an enrichment of LAMP2 surrounding the aggregates (Figure 4.15B). It is important to note that only some cells contained protein aggregates. As the cultures utilised for BioID were at ~90-100% confluence a large proportion of these cells maintained the cytoplasmic localisation of ZYG11A. Likewise, the staining of biotinylated proteins could still be found in the cytoplasm, which can be attributed to a gradual accumulation of each fusion protein, and the concurrent presence of these proteins in the cytoplasm and in the aggregates.

Therefore, despite the shifting localisation of ZYG11A under proteasome inhibition, many of the interactions captured by BioID are still likely to represent its cytoplasmic localisation.

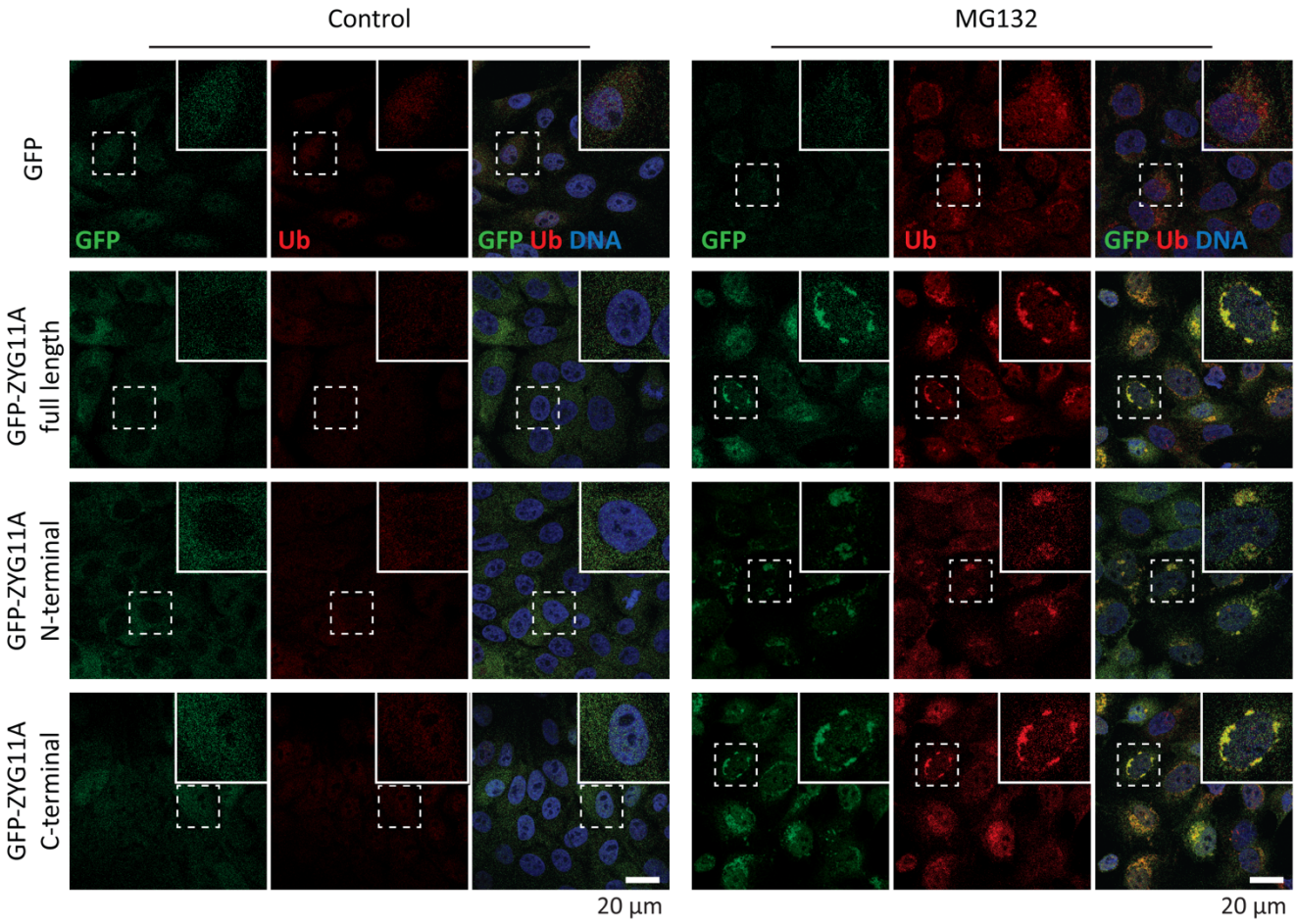



**Figure 4.14 Aggregates of ZYG11A fusion proteins under proteasome inhibition colocalise with ubiquitin**

**A.** Representative confocal images of GFP, GFP-ZYG11A FL, GFP-ZYG11A NT and GFP-ZYG11A CT fusion proteins. Cells are stained for the GFP tag using anti-GFP (green), as well as anti-Ubiquitin (Ub) to mark ubiquitin (red) and DNA using DAPI (blue). Images show the localisation of the fusion proteins under control conditions (DMSO) and proteasome inhibition (10 µM MG132) Note that under proteasome inhibition, ubiquitin levels increase in GFP control cells, however the GFP tag does not accumulate like the ZYG11A fusion proteins.

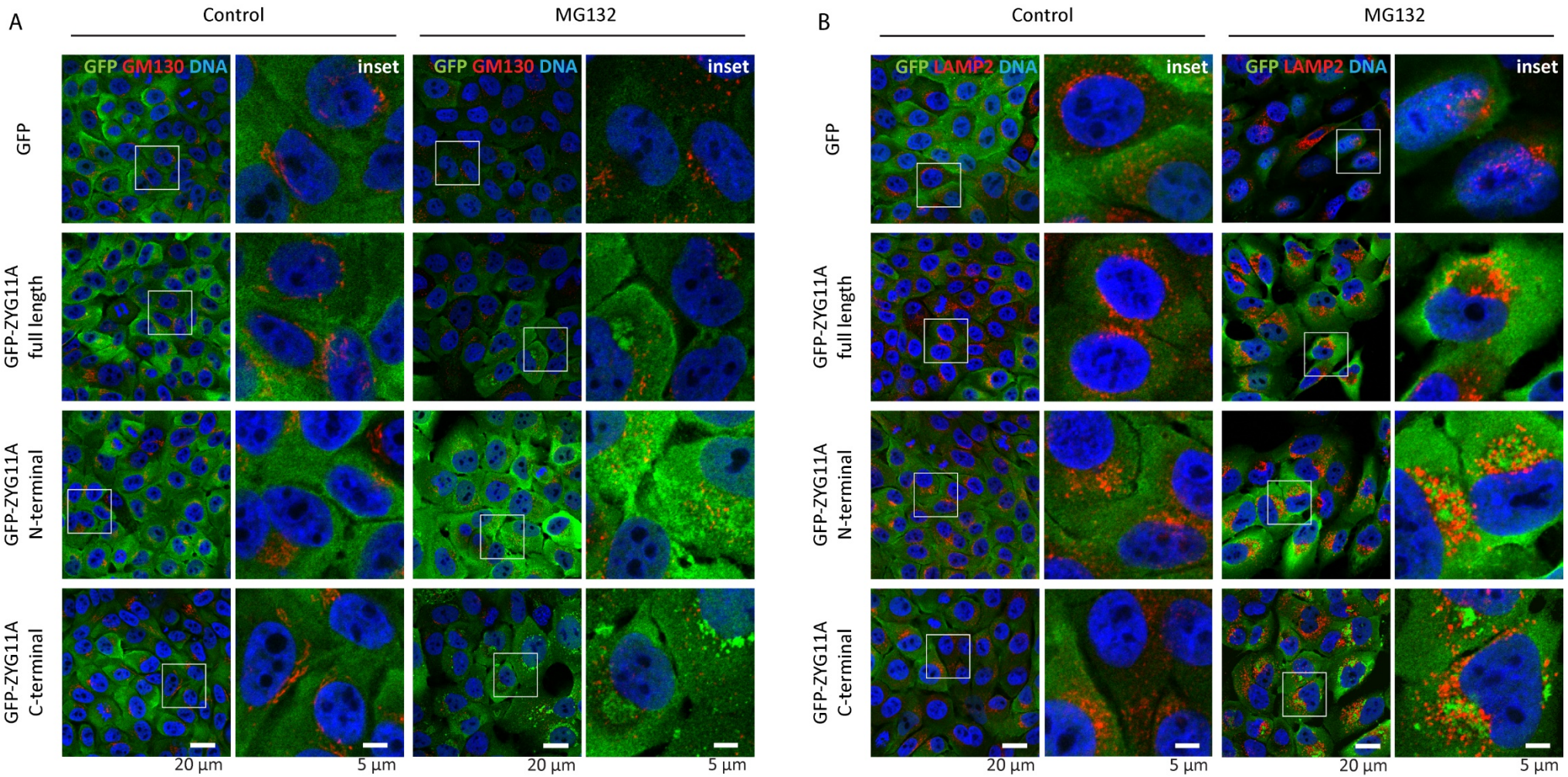

A
Control
MG132
GFP GM130 DNA
inset
GFP
GFP-ZYG11A full length
GFP-ZYG11A N-terminal
GFP-ZYG11A C-terminal
20 µm
5 µm
B
GFP LAMP2 DNA

**Figure 4.15 Aggregates of ZYG11A fusion proteins under proteasome inhibition are surrounded by lysosome proteins**

**A.** Representative confocal images of GFP, GFP-ZYG11A FL, GFP-ZYG11A NT and GFP-ZYG11A CT fusion proteins. Cells are stained for the GFP tag (green), the Golgi apparatus marker, GM130 (red) and DNA (blue). Images show the localisation of the fusion proteins under control conditions (DMSO) and proteasome inhibition (10 μM MG132). Insets highlight the paranuclear localisation of the Golgi apparatus under normal conditions (left panels) and the localisation of the Golgi apparatus relative to MG132 induced protein aggregates. **B.** Fusion proteins described in **A** marked by the GFP tag (green), the lysosomal marker, LAMP2 (red) and the DNA marker DAPI (blue). LAMP2 is localised around the nucleus under normal conditions. GFP aggregates are present under MG132 treatment with high levels of LAMP2 surrounding them. Note that in some cases, the aggregate appears to show lower levels of GFP.

### 4.2.8 Identification of stable and transient ZYG11A interacting proteins by proximity biotinylation and mass spectrometry

Now that I had determined that my ZYG11A constructs were present in the insoluble protein fraction and validated their localisation as well as their ability to biotinylate, I decided to proceed with BioID labelling. To obtain a more complete picture of the ZYG11A interactome, the coupling of BioID with protein stabilisation (via proteasome inhibition) was implemented to capture both stable and transient interactions that where within the ~10 nm labelling radius of the ZYG11A-BioID fusion proteins. Distinct protein interaction domains exist across the ZYG11A protein (N-terminal = VHL box and LRRs, C-terminal = ARM domains) and while it is still unclear which domain is responsible for substrate recognition, the two truncated forms of ZYG11A provided important controls. Theoretically, proteins that interact with the full length as well as either one of the two truncated proteins are likely to be interacting specifically with one of the two domains. Using these fusion proteins as well as the tag alone as a control, BioID labelling and mass spectrometry analysis was conducted with and without MG132 treatment (Figure 4,16A). Each of the constructs (+/- MG132) were run using three biological replicates to allow for statistical analysis. The mass spectrometry data were analysed using MaxQuant with label-free quantitation. The peptides and proteins identified were based on the Uniprot *Homo sapiens* database using Andromeda software to compare the spectra to known peptide signatures. To avoid false positive peptide spectrum matches, protein groups were subjected to a widely used 1% false discovery rate. To ensure the analysis of single protein fragments, the peptide cut-offs were set at the minimum peptide length of 7 residues and a maximum peptide mass of 4600 Da.

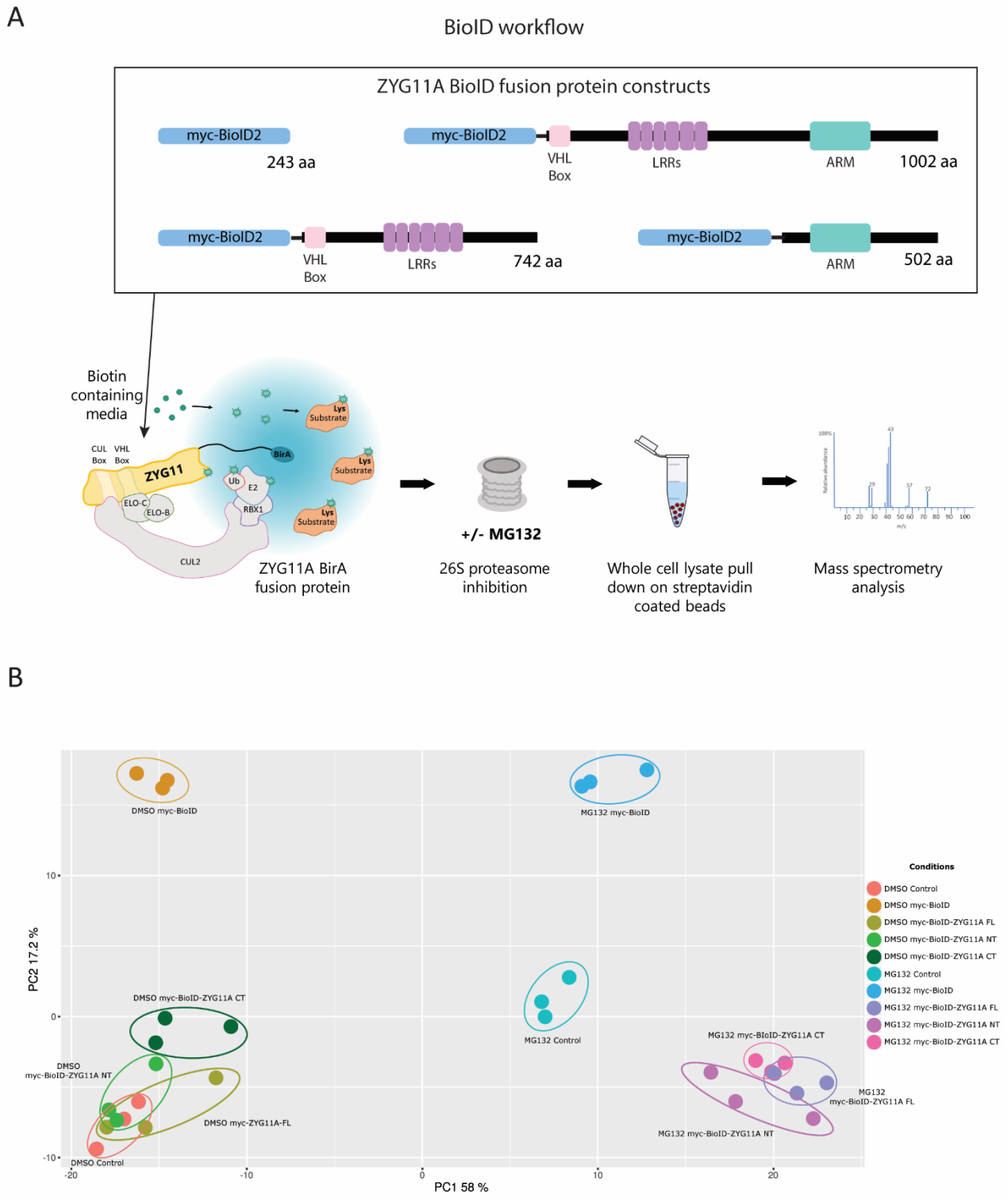



**Figure 4.16 Peptides detected in ZYG11A-BioID experiments are heavily influenced by proteasome inhibition**

**A.** Workflow of the BioID experiment. The BioID fusion proteins are cultured in biotin containing media to allow the biotinylation to occur, ± 10 µM MG132 treatment (or a DMSO control) to inhibit the proteasome and enrich for degradation targets. Whole cells lysates are then harvested for streptavidin pull-downs and analysed by mass spectrometry. **B.** Principal component analysis (PCA) of the BioID experiment displaying a simplified overview of the protein variation between each sample. The PC1 axis represents a greater amount of variation between samples (58%) compared with the PC2 axis (17.2%). Within

the DMSO and MG132 treatment conditions, each fusion protein sample clusters together. The myc-BioID only samples show the most variation from the myc-BioID-ZYG11A samples along the PC2 axis. Each sample performed in biological triplicates.

Prior to statistical analyses, the data were normalised to remove any contaminants and any missing values that would interfere with the statistical analysis. The data were then statistically analysed for differential abundance using the Bayes and the Benjamini-Hochberg methods, and visualised using R. Based on proteomics conventions, comparisons with P-values $< 0.05$ and a logFC (fold change) of $\geq 1.5$ were considered significant hits.

Since there were many different datasets to analyse, a principal component analysis was generated in order to obtain a general overview of the correlations between each sample. (Figure 4.16B). This analysis simplifies the multidimensional data obtained by mass spectrometry to reveal the most robust patterns in the dataset and represents this as a single data point. Here, the principal component (PC) axes, PC1 and PC2 represent the percentage of variability between the proteins that compose each of the datasets. The PC1 axis (58.0%) represents a greater amount of variability compared with the PC2 axis (17.2%). For each condition, the triplicate biological repeats clustered together demonstrating the reproducibility and robustness of the method across independent biological experiments (Figure 4.16B). However, based on the DMSO samples (left of PC1 axis) and the MG132 samples (right of PC1 axis), it is clear that there is a high degree of variability in the proteins within the two treatment datasets. Within each treatment condition, the BioID only samples are largely separated from the fusion proteins along the PC2 axis, which is expected considering the biotinylation previously observed via western blot (Figure 4.11). Between the three fusion proteins, they appear to cluster in the bottom left and right corners for the DMSO and MG132, respectively. This suggests that they are similar in composition and may contain similar protein subsets, perhaps due to the various domains present in the N- and C-terminal fusion proteins. While the empty vector controls do show similar clustering, these are likely to represent ‘sticky’ proteins across the samples since there is no BioID activity. This may also be related to the low levels of biotinylation by the fusion proteins under steady state conditions, making them less distinct from a no-BioID context. Because sticky proteins were going to be widespread across all samples, including the BioID only controls, the BioID only control was used as a reference point as it could account for sticky proteins, as well as non-specific background biotinylation. In order to

look closer at the proteins that were upregulated following proteasomal inhibition for the ZYG11A full length fusion protein, volcano plots were generated.

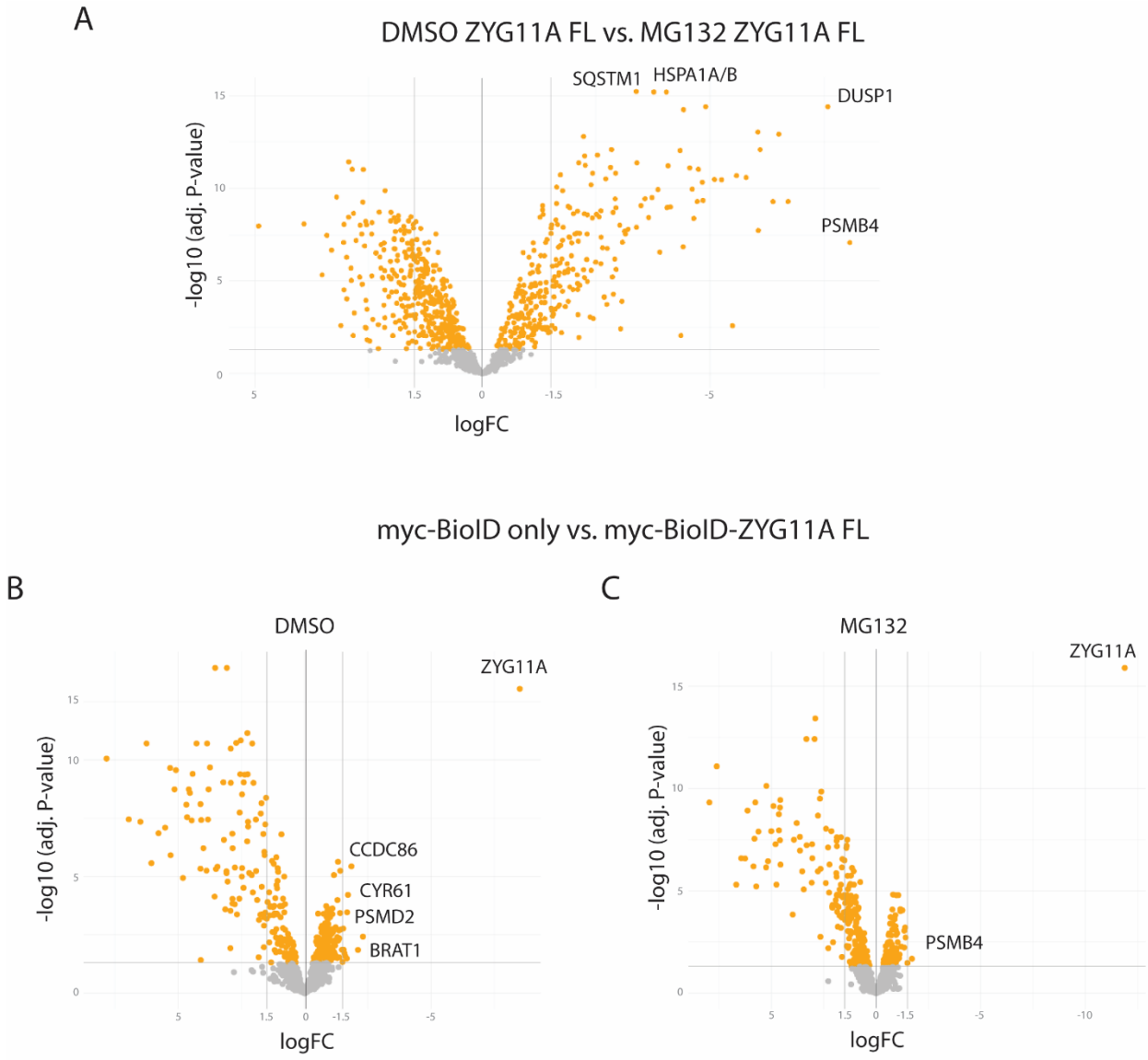



**Figure 4.17 Summary of ZYG11A-BioID hits reveal a novel subset of ZYG11A stable and transient interacting proteins**

**A.** Volcano plots comparing the ZYG11A hits between MG132 and DMSO treatment groups. Proteins enriched in the ZYG11A MG132 samples are displayed in the negative values (left of the plot), and proteins that are enriched in the ZYG11A DMSO treatment are displayed in the positive values (right of the plot). **B-C.** Volcano plots comparing the hits between the myc-BioID only and myc-BioID-ZYG11A FL datasets under normal conditions (DMSO) and proteasome inhibition (MG132), respectively. Note that there are very few statistically significant enriched proteins associated with ZYG11A in either comparison.

The volcano plots showed that there was a much larger subset of significantly biotinylated proteins in the MG132 BioID-ZYG11A full length cell lines compared with their DMSO counterparts (Figure 4.17A). Some of the more significant peptides included proteins like PSMB4, DUSP1, SQSTM1 and HSPA1A. Although there were a number of upregulated proteins in the ZYG11A full length MG132 samples compared to DMSO control, it could be argued that these hits were only upregulated due to an increased abundance of these proteins because of proteasome inhibition, rather than a specific interaction with ZYG11A. In order to more stringently identify specific ZYG11A interactions, the hits from the DMSO and MG132 treatment groups were also compared with the BioID only tag which has also been used as a control for a number of BioID experiments (reviewed by Varnaitė and MacNeill, 2016). With the rationale that the BioID tag is less likely to capture specific interactions due to its widespread localisation, any hits that were significantly upregulated in the ZYG11A fusion proteins were more likely to be due to the specific activity of the ZYG11A protein and its protein domains, rather than random biotinylation. However, when observed within each of the treatment groups (DMSO and MG132) and measured against the BioID only control, the DMSO (Figure 4.17B, right of volcano plot) and MG132 (Figure 4.17C, right of volcano plot) samples had only 11 (e.g. ZYG11A, BRAT1, PSMD2, CYR61 and CCDC86) and 2 (ZYG11A and PSMB4) enriched hits in the BioID-ZYG11A datasets, respectively. Under both conditions, the majority of hits were classified as "depleted" in the ZYG11A full length samples which was expected considering that the BioID only protein is more widely distributed than the ZYG11A fusion, and that previous western blots have shown more biotinylated proteins in this control sample (Figure 4.14C-D, left of volcano plots). While this analysis method was very stringent and generated very few hits, it did not take into account proteins that were only present in the BioID-ZYG11A samples. In order to include proteins of this nature, I also analysed the proteins with 'missing' values. As mentioned earlier, any peptides with missing values were initially removed from the analysis dataset as they would not produce any statistical output, i.e. peptides that are not detected at all in one peptide subset cannot be statistically analysed with subsets where it is detected as their value is zero and hence cannot provide an enrichment value. Therefore, to identify any peptides that were found only in the BioID-ZYG11A fusion samples I performed presence/absence analysis using a program generated by the Comprehensive Proteomics Platform. This compared each set of replicates and returned any genes that were present only in certain groups – in this case, the ZYG11A fusion proteins but not the BioID controls. This increased the number of total DMSO hits to 235, and for the MG132 samples, 226. Through similar data obtained from

the N- and C- terminal truncation BioID samples, the hits from the full length DMSO and MG132 treatments could be further separated by subcategorising hits into either ZYG11A full length only (FL), ZYG11A full length and N-terminal (FL+NT), or ZYG11A full length and C-terminal (FL+CT) (Figure 4.18A-B). Given the different domains present in each of the truncated forms, it was unlikely that a single protein would be able to interact with multiple regions across the ZYG11A protein. Therefore, the hits present in all three fusion protein samples were excluded from further analysis as potentially "sticky" proteins. Only 3 hits were shared between the DMSO and MG132 subsets (VAT1, SLC38A2, UFL1), suggesting differentially enriched subsets due to proteasome inhibition (Figure 4.18C). The complete list of hits from both DMSO and MG132 samples are listed in Table 4.2.

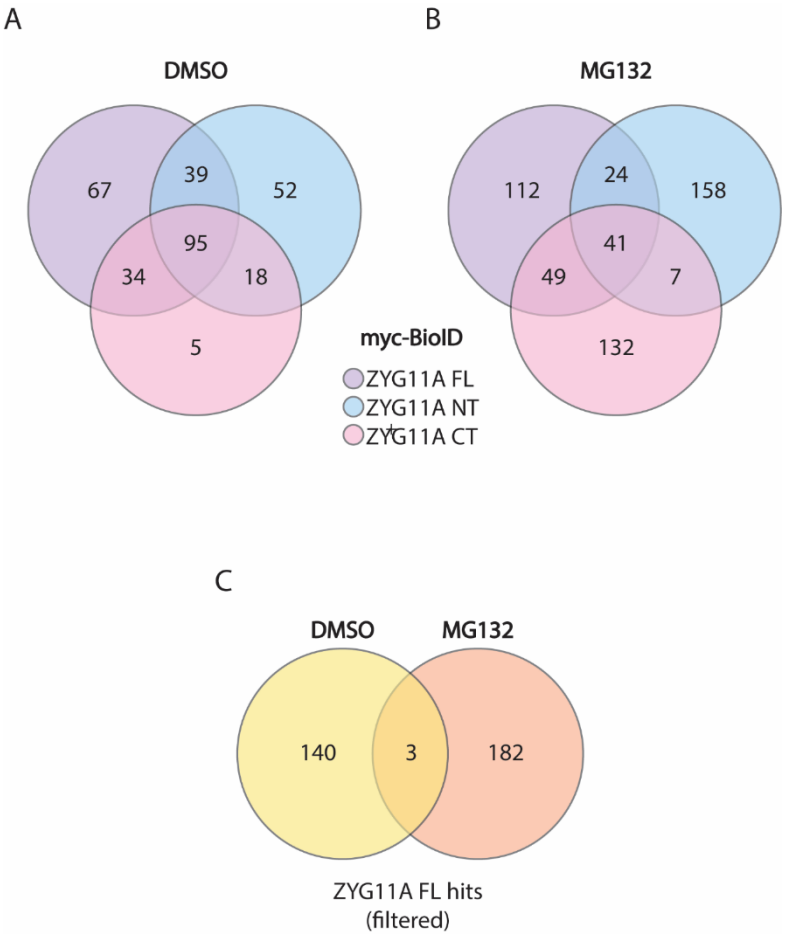



**Figure 4.18 BioID proximity biotinylation shows distinct alterations in interacting proteins under proteasome inhibition**

**A-B.** Venn diagrams summarising the statistically significant and the unique hits (compared with the myc-BioID tag only control) from the FL, NT and CT ZYG11A fusion proteins. Hits from the DMSO controls (**A**) and the MG132 treated samples (**B**) show the distribution of hits between each fusion protein. **C.** Venn diagram representing the ZYG11A FL hits that have been compared to the myc-BioID only control. Hits present in all three ZYG11A fusion samples have also been removed to robustly separate the protein interactions. Note that there are only 3 hits that are present under both DMSO and MG132 treatment following the filtering. Statistically significant hits were based on a logFC of ±1.5 and a P-value ≤ 0.05. Unique peptides were identified by the presence or absence of a peptide between the samples.

**Table 4.2 Summary of BioID hits**

BioID hits separated into DMSO and MG132 conditions and colour coded based on the following distributions: ZYG11A FL only (purple), ZYG11A FL+NTT (blue), ZYG11A FL+CTT (pink). Values are shown for hits that fulfil the P-value ≤ 0.05 and logFC ± 1.5 thresholds. Unique peptides that are only present in the ZYG11A fusion protein samples (and thus cannot be statistically analysed) remain blank. To create a hierarchy of confidence between these unique hits, the presence of each peptide in each of the biological replicates is colour coded: three replicates (blue), two replicates (red), one replicate (green). Genes in each group are then sorted alphabetically. Genes present in both the DMSO and MG132 subsets are marked with a '◆'. Hits possessing an N-terminal glycine are noted in the left column and are subcategorised into: U = followed by an 'unfavoured' residue (D, E, I, N, P, R, S or T), F = followed by a 'favoured' residue (F, G, H L, M or Y), or UC = followed by a residue that is uncategorised (A and V). Proteins known to be N-myristoylated are marked with a 'Y' and those with identified caspase cleavage sites resulting in N-terminal glycines are also subcategorised into U, F and UC as described above. Gene names that are separated by a semicolon represent peptide matches that cannot be distinguished between the listed proteins.

| N-term. glycine | N-myristoylated | Caspase cleaved | DMSO | ZYG11A FL | | ZYG11A NT | ZYG11A CT |
|---|---|---|---|---|---|---|---|
| | | | Gene Name | Adj. P-value | logFC | | |
| | | | BRAT1 | 0.014 | -2.107 | 0.075 | -0.713 |
| | | | ATXN10 | 0.000 | -1.679 | -1.445 | -1.358 |
| | | | RPLP1 | 0.033 | -1.668 | -0.945 | 0.450 |
| | | | DDX39B | 0.026 | -1.606 | -0.640 | -0.882 |
| | | | RPL36AL;RPL36A | 0.020 | -1.572 | -1.127 | -0.729 |
| | | | UAP1 | 0.014 | -1.536 | -0.871 | -1.021 |
| | | | ACOX3 | | | | |
| | | | EXOSC4 | | | | |
| | | | MYL12A;MYL12B | | | | |
| | | | SBDS | | | | |
| | | | TANGO6 | | | | |

| N-term. glycine | N-myristoylated | Caspase cleaved | MG132 | ZYG11A FL | | ZYG11A NT | ZYG11A CT |
|---|---|---|---|---|---|---|---|
| | | | Gene Name | Adj. P-value | logFC | | |
| | | | FBXL5 | | | | |
| | | | FUK | | | | |
| | | | MYL1;MYL3 | | | | |
| | | | NELFCD | | | | |
| | | | TRAF6 | | | | |
| | | | VAT1 ◆ | | | | |
| | | | B4GALT1 | | | | |
| | | | CREB3L2 | | | | |
| | | | MPP1 | | | | |
| | | UC | PSMB10 | | | | |
| | | | RNF135 | | | | |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | ADH5 | | | | | | | | | SHPK | | | | | | |
| | | | ANLN | | | | | | | | | TFAP2D | | | | | | |
| | | | AP1S1 | | | | | | | | | VAV2 | | | | | | |
| | | | ATP5O | | | | | | | | | AAK1 | | | | | | |
| | | F | BCAP31 | | | | | | | | | ACTR10 | | | | | | |
| | | | BCAR3 | | | | | | | | | AGAP3 | | | | | | |
| | | | CDH13 | | | | | | | | | AK4 | | | | | | |
| | | | CDK2;CDK3 | | | | | | | | | ALG2 | | | | | | |
| | | F | CHAF1A | | | | | | | | | ALMS1 | | | | | | |
| | | | CISD1 | | | | | | | | | ANXA3 | | | | | | |
| | | | COLGALT1 | | | | | | | | | AP3S1 | | | | | | |
| | | | COX20 | | | | | | | | | APOBEC3C;APOBEC3F;APOBEC3D | | | | | | |
| | | | ELMO3 | | | | | | | | F | Y | ARF5 | | | | | |
| | | | ELP3 | | | | | | | | | | ATP6AP1 | | | | | |
| | | | ERAL1 | | | | | | | | | | BCAT2 | | | | | |
| | UF | | FBN2 | | | | | | | | | | BRCA2 | | | | | |
| | | | FTL | | | | | | | | | | BTBD1 | | | | | |
| | F | | FXR2 | | | | | | | | | | BUB1 | | | | | |
| | | | HPRT1 | | | | | | | | | | C16orf70 | | | | | |
| | | | ITPRIP | | | | | | | | | | C17orf53 | | | | | |
| | | | JMJD4 | | | | | | | | | | C21orf59 | | | | | |
| | | | JMJD6 | | | | | | | | | | CDC123 | | | | | |
| | | | KIAA1524 | | | | | | | | | | CDR2 | | | | | |
| | | | KIFC3 | | | | | | | | | | CEP57 | | | | | |
| | | | LIMK2 | | | | | | | | | | CEP78 | | | | | |
| | | | MAGT1 | | | | | | | | | | CEP85 | | | | | |
| | | | MAP3K6 | | | | | | | | UF | | COL4A1 | | | | | |
| | | | MRPS9 | | | | | | | | | | COTL1 | | | | | |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| UF | Y | | MTHFD1L | | | | | UC | | | CRY1 | | | | | | |
| | | UC | MTOR | | | | | | | | CSNK2A2 | | | | | | |
| | | | NAA15 | | | | | | | | CUL4A | | | | | | |
| | | | NUB1 | | | | | | | | DENND4C | | | | | | |
| U | | | PDXK | | | | | | | | DFNA5 | | | | | | |
| | | | PPAT | | | | | | | | DMWD | | | | | | |
| | | | PTCD3 | | | | | | | | EMC3 | | | | | | |
| | | | RAB5A | | | | | F | | | FAH | | | | | | |
| | | | RAE1 | | | | | | | | FOCAD | | | | | | |
| | | | SEC14L1 | | | | | | | | FZR1 | | | | | | |
| | | | SERPINA7 | | | | | | | | GLRX2 | | | | | | |
| | | | SH3RF2 | | | | | | | | GNA15 | | | | | | |
| | | | SLC38A2 ◆ | | | | | | | | GNS | | | | | | |
| | | | SNX3 | | | | | | | | GSTCD | | | | | | |
| | | | SPC24 | | | | | F | | | GSTT2B;GSTT2 | | | | | | |
| | | | SREBF2 | | | | | | | | GUF1 | | | | | | |
| | | | SUCLG1 | | | | | | | | ISOC1 | | | | | | |
| | | | TCEB2 (Elongin B) | | | | | | | | ITM2C | | | | | | |
| | | | TINAGL1 | | | | | | | | KCMF1 | | | | | | |
| | | | TRMT1 | | | | | | | | KCTD12 | | | | | | |
| | | | TTK | | | | | | | | KLHL26 | | | | | | |
| | | | TYMS | | | | | | | | KPNA3 | | | | | | |
| | | | UCK2 | | | | | | | | LARS2 | | | | | | |
| | | | UTP18 | | | | | F | | | LRRTM2 | | | | | | |
| | | | VASP | | | | | | | | MALT1 | | | | | | |
| | | | VPS29 | | | | | UC | Y | | MARCKS | | | | | | |
| | | | VTA1 | | | | | | | | MCM8 | | | | | | |
| | | | YME1L1 | | | | | | | | MINA | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | CCDC86 | 3.709E-06 | -1.849E+00 | -1.561E+00 | -1.112E+00 |
| | | | CYR61 | 6.180E-05 | -1.713E+00 | -1.763E+00 | -1.421E+00 |
| | | | PDIA4 | | | | |
| | | | APEX1 | | | | |
| | | UF | EHD4 | | | | |
| | | | GMPS | | | | |
| | | | HSP90AB4P | | | | |
| | | F | PLOD2 | | | | |
| | | F | PSMC3 | | | | |
| | | | SPATS2L | | | | |
| | | | ACAT1 | | | | |
| | | | AK2 | | | | |
| | | | CUL1 | | | | |
| | | | DTYMK | | | | |
| | | | DUSP11 | | | | |
| | | UC | GLRX3 | | | | |
| | | | IKBKB | | | | |
| | | | KLHL21 | | | | |
| | | | METTL13 | | | | |
| | | | MIPEP | | | | |
| | | | MYO1D | | | | |
| | | | NADK2 | | | | |
| | | | NDUFA11 | | | | |
| | | | NDUFB5 | | | | |
| | | | NNMT | | | | |
| | | | POLR1E | | | | |
| | | | PSMC5 | | | | |
| | | | RAB3D | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | MRPL10 | | | | |
| | | | NAA10 | | | | |
| | | | NDUFAF6 | | | | |
| | | | NRBP2 | | | | |
| | | | NT5DC1 | | | | |
| | | | NUDT16;NUDT16L1 | | | | |
| | | | ORC5 | | | | |
| | | | OSBPL9 | | | | |
| | | | OVCA2 | | | | |
| | | | PCBD1 | | | | |
| | | | PCGF1 | | | | |
| | | | PDPK2P;PDPK1 | | | | |
| | | | PIK3C2B | | | | |
| | | | POLH | | | | |
| | | | POLRMT | | | | |
| | | | PPA1 | | | | |
| | | | PPP1R18 | | | | |
| | | | PRKCDBP | | | | |
| | | | PSMD4 | | | | |
| | | | RHOA | | | | |
| | | | RHPN2;RHPN2P1 | | | | |
| | | | ROS1 | | | | |
| | | | RPS19BP1 | | | | |
| | | | SAP18 | | | | |
| | | | SCYL1 | | | | |
| | | | SEC24B | | | | |
| | | | SKA1 | | | | |
| | | | SLC2A12 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | REEP4 | | | | |
| | | | SHC1 | | | | |
| | | | SLFN5 | | | | |
| | | | STXBP3 | | | | |
| | | | STYXL1 | | | | |
| | | | TCEB1 (Elongin C) | | | | |
| | | | TXNL1 | | | | |
| | | | U2AF1;U2AF1L4 | | | | |
| | | | UBTF | | | | |
| UF | | | WDR6 | | | | |
| | | | WDR81 | | | | |
| | | | PSMD2 | 0.002 | -1.526 | -1.187 | -1.968 |
| | | | AKR1D1 | | | | |
| | | | IDH3G | | | | |
| | | | HIGD1A | | | | |
| | | | DNAAF5 | | | | |
| | | | NCAPD2 | | | | |
| | | | RSL24D1 | | | | |
| | | | ADSS | | | | |
| | | | ALDH3A1 | | | | |
| | | | ARPC3 | | | | |
| | | | BPNT1 | | | | |
| | | | CAP1 | | | | |
| | | | CASP1 | | | | |
| | | | CECR5 | | | | |
| | | | DYNC1I2 | | | | |
| | | | GSK3B | | | | |
| | | | GYS1 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | SLC38A2 ◆ | | | | |
| | | | SNX30 | | | | |
| | | | SNX7 | | | | |
| | | | STARD7 | | | | |
| | | | STK38L | | | | |
| | | | STON2 | | | | |
| | | | TIPARP | | | | |
| | | | TMLHE | | | | |
| | | | TNFAIP2 | | | | |
| | | | TRAPPC2L | | | | |
| | | | TSC22D3 | | | | |
| | | | TSPYL1 | | | | |
| UF | | | TSTA3 | | | | |
| | | | TTC12 | | | | |
| | | UC | UBE4B | | | | |
| | | | ZNF268 | | | | |
| | | | ZRANB1 | | | | |
| | | | CDC25C | | | | |
| | | | LRRC14 | | | | |
| | | | NCKAP1 | | | | |
| | | | SESN2 | | | | |
| | | | VPS39 | | | | |
| | | | CARS2 | | | | |
| | | | DERA | | | | |
| | | | MRFAP1 | | | | |
| | | | TBC1D8;TBC1D8B | | | | |
| | | | UFL1 ◆ | | | | |
| | | | USP53 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | INTS7 | | | | |
| | | | KLHL42 | | | | |
| UF | | | LDLR | | | | |
| | | | MTMR6 | | | | |
| | | | NDUFS1 | | | | |
| | | | NKRF | | | | |
| | | | NUP88 | | | | |
| | | | POLD1 | | | | |
| | | | RABGGTA | | | | |
| | | | RTKN | | | | |
| | | | SPATA5L1 | | | | |
| | | UF | TBC1D4 | | | | |
| | | | TUBB4A | | | | |
| | | | UFL1 ◆ | | | | |
| | | | VAT1 ◆ | | | | |
| | | | VPS16 | | | | |
| | | | ZAK | | | | |

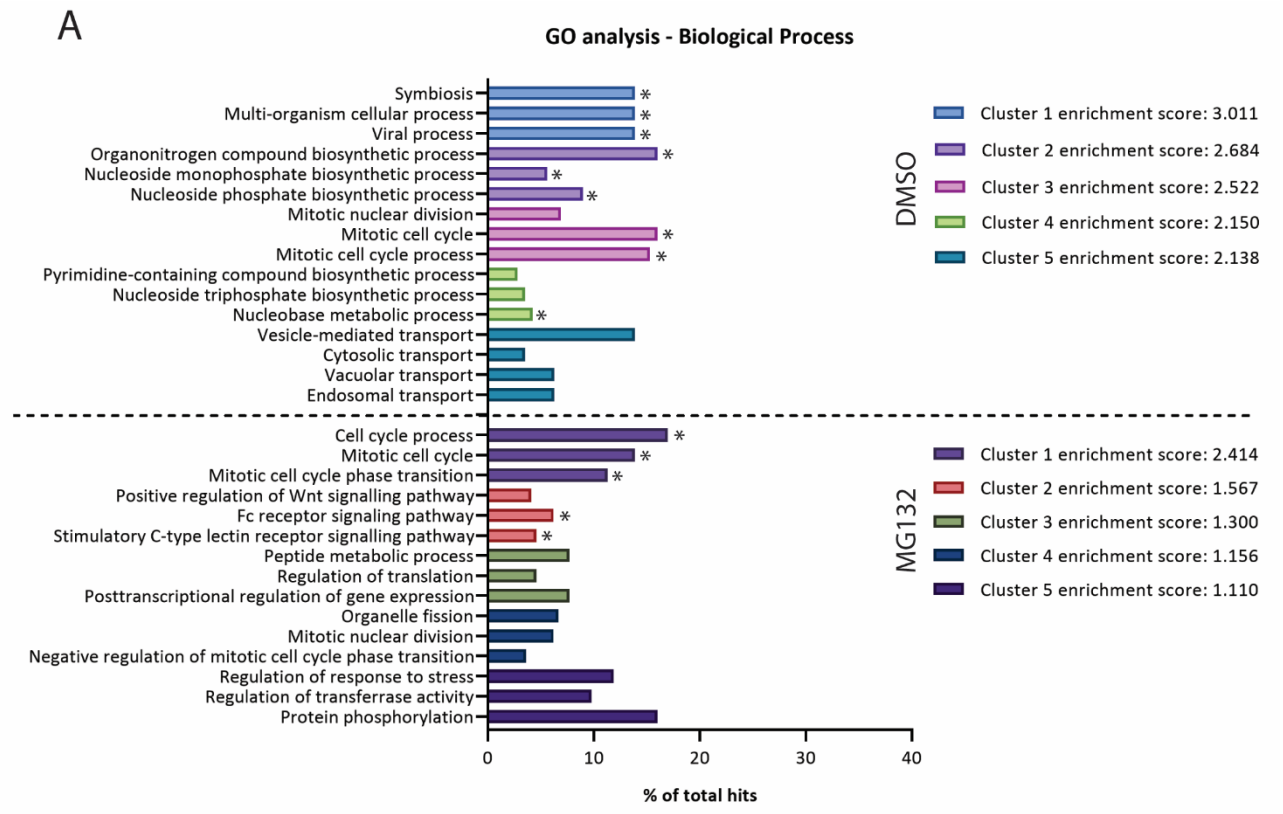
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | UTP23 | | | | |
| | | | CHUK | | | | |
| | | | DSC1 | | | | |
| | | | EBP | | | | |
| | | | F8A1 | | | | |
| | | | FYTTD1 | | | | |
| | | | HIBADH | | | | |
| | | F | MACF1 | | | | |
| | | | METTL1 | | | | |
| | | | PEX5 | | | | |
| | | | PSMD5 | | | | |
| | | | SMG7 | | | | |
| | | | TAF2 | | | | |
| | | | PSMB4 | 0.021 | -1.712 | -1.049 | -1.720 |
| | | | ACAD9 | | | | |
| | | | C16orf62 | | | | |
| | | | CNOT2 | | | | |
| | | | COMMD4 | | | | |
| | | | DLD | | | | |
| UF | | | LCMT2 | | | | |
| | | | NRBP1 | | | | |
| | | | PPIP5K1 | | | | |
| | | | PPP1R3C | | | | |
| | | | QSOX1 | | | | |
| | | | SETDB1 | | | | |
| | | | SF3B5 | | | | |
| | | | THOP1 | | | | |
| | | | APEX2 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | ATG4B | | | | |
| | | | ATP5G3 | | | | |
| | | | CASP2 | | | | |
| | | | CNOT10 | | | | |
| | | | HAUS1 | | | | |
| | | | HAUS7 | | | | |
| | | | HNRNPH2 | | | | |
| | | | KLHL18 | | | | |
| | | | NDUFS8 | | | | |
| | | | OSGIN2 | | | | |
| | | | PAK1IP1 | | | | |
| | | | PPP1R15B | | | | |
| | | | RALGAPB | | | | |
| | | | VRK2 | | | | |
| | | | XPO5 | | | | |
| | | | CLIP4 | | | | |
| | | | CORO7 | | | | |
| UF | | | DLAT | | | | |
| | | | ELMO2 | | | | |
| | | | FGF2 | | | | |
| UF | | | ITGA2 | | | | |
| | | | KDM5A | | | | |
| | | | LAMTOR3 | | | | |
| | | | MIF | | | | |
| | | | NBEAL2 | | | | |
| | | | PCID2 | | | | |
| | | | PRIM1 | | | | |
| | | | SKIV2L2 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| UF | Y | F | SRC | | | | |
| | | | SSH3 | | | | |
| UF | | | TGFBR2 | | | | |
| | | | TRIM7 | | | | |
| | | | TUBE1 | | | | |
| | | | ZBTB33 | | | | |

Considering the known physical association between the ZYG11 family and the Cullin-2 ubiquitin ligase complex (Vasudevan, Starostina and Kipreos, 2007), it was important to determine whether the BioID assay was able to capture any other complex members. While Cullin-2 itself did not appear to be biotinylated by the ZYG11A-BioID fusion, the adaptor proteins Elongin B (TCEB2) and Elongin C (TCEB1) were present in the DMSO ZYG11A full length subset, with Elongin C also present in the N-terminal subset but not in the C-terminal DMSO dataset. This is consistent with the N-terminal containing the VHL-like domain that binds to Elongin C. Although Elongin C is not seen in the MG132 dataset, this may be due to an increase in other proteins which might mask this steady state interaction. Other than Elongin B and C, there was only one other protein that was common between the high-content screening and BioID datasets. NUB1 was previously identified as a ZYG11A genetic interactor, and these BioID data suggest that they may also interact physically or be in close proximity (within ~10 nm of the fusion proteins).

To highlight the presence of any important processes and pathways within the DMSO and MG132 hits, Gene Ontology (GO) analysis was conducted using the DAVID functional annotation database. The processes and pathways chosen here were based on the DAVID enrichment scores, which indicated where the proteins in the ZYG11A datasets were relative to the ranked proteins associated with the specific GO annotations. For each process and pathway, the individual proteins can be assigned to multiple categories based on their enrichment. The higher the enrichment scores, the more enriched the dataset was for high ranked proteins. The biological processes of DMSO proteins enriched for viral processes, nucleoside and nucleobase processes, mitotic cell cycle, and endosomal transport. MG132 hits also enriched for mitotic cell cycle, but also for lectin and cell surface receptor signalling, post-transcriptional and translational regulation, and protein phosphorylation (Figure 4.19A). Alongside this, KEGG pathway analysis revealed that DMSO hits were implicated in pathways in cancer, non-alcoholic fatty liver disease, insulin signalling and focal adhesions. MG132 hits were also linked to focal adhesions and pathways in cancer, as well as actin cytoskeleton regulation, and a number of relevant signalling pathways like PI3K-Atk and MAPK (Figure 4.19B). Other than cell cycle, many of these processes have not been previously linked to the ZYG11 family.


A
GO analysis - Biological Process
Symbiosis
Multi-organism cellular process
Viral process
Organonitrogen compound biosynthetic process
Nucleoside monophosphate biosynthetic process
Nucleoside phosphate biosynthetic process
Mitotic nuclear division
Mitotic cell cycle
Mitotic cell cycle process
Pyrimidine-containing compound biosynthetic process
Nucleoside triphosphate biosynthetic process
Nucleobase metabolic process
Vesicle-mediated transport
Cytosolic transport
Vacuolar transport
Endosomal transport
DMSO
Cluster 1 enrichment score: 3.011
Cluster 2 enrichment score: 2.684
Cluster 3 enrichment score: 2.522
Cluster 4 enrichment score: 2.150
Cluster 5 enrichment score: 2.138
Cell cycle process
Mitotic cell cycle
Mitotic cell cycle phase transition
Positive regulation of Wnt signalling pathway
Fc receptor signaling pathway
Stimulatory C-type lectin receptor signalling pathway
Peptide metabolic process
Regulation of translation
Posttranscriptional regulation of gene expression
Organelle fission
Mitotic nuclear division
Negative regulation of mitotic cell cycle phase transition
Regulation of response to stress
Regulation of transferrase activity
Protein phosphorylation
MG132
Cluster 1 enrichment score: 2.414
Cluster 2 enrichment score: 1.567
Cluster 3 enrichment score: 1.300
Cluster 4 enrichment score: 1.156
Cluster 5 enrichment score: 1.110
0
10
20
30
40
% of total hits

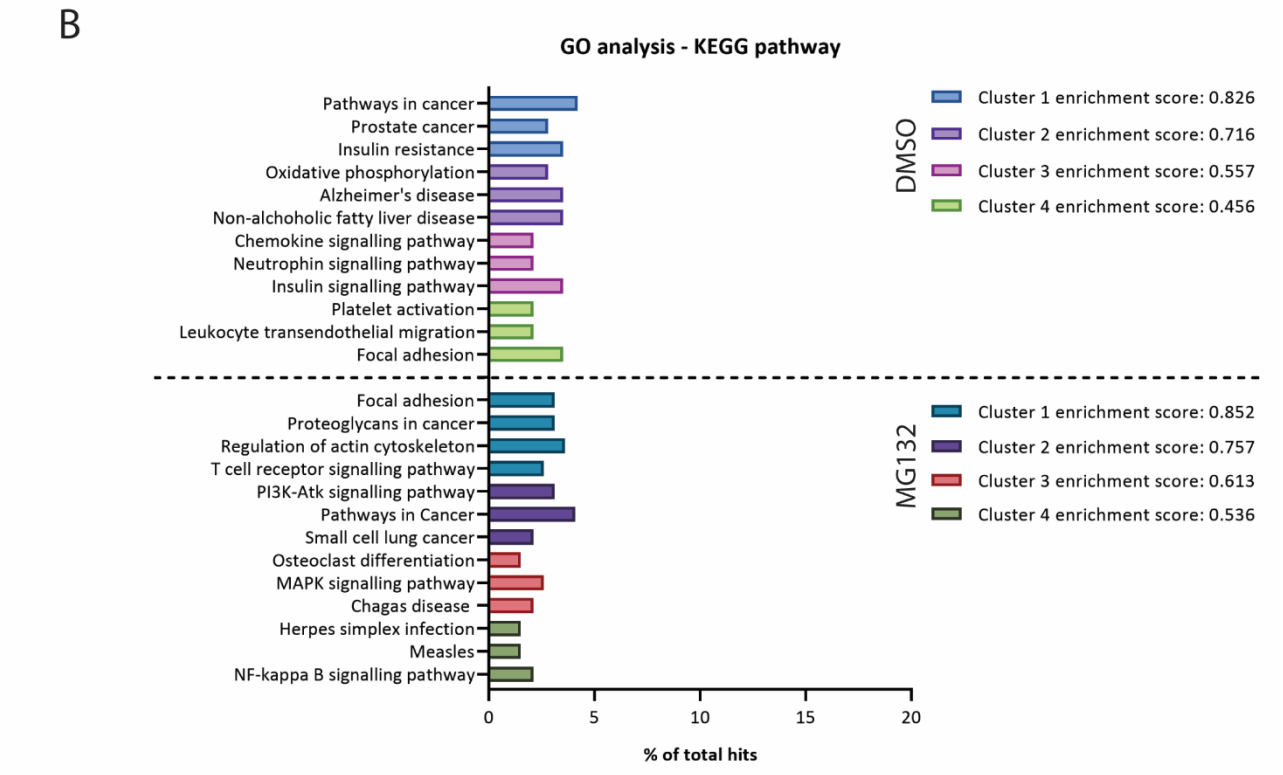

B
GO analysis - KEGG pathway
Pathways in cancer
Prostate cancer
Insulin resistance
Oxidative phosphorylation
Alzheimer's disease
Non-alchoholic fatty liver disease
Chemokine signalling pathway
Neutrophin signalling pathway
Insulin signalling pathway
Platelet activation
Leukocyte transendothelial migration
Focal adhesion
DMSO
Cluster 1 enrichment score: 0.826
Cluster 2 enrichment score: 0.716
Cluster 3 enrichment score: 0.557
Cluster 4 enrichment score: 0.456
Focal adhesion
Proteoglycans in cancer
Regulation of actin cytoskeleton
T cell receptor signalling pathway
PI3K-Atk signalling pathway
Pathways in Cancer
Small cell lung cancer
Osteoclast differentiation
MAPK signalling pathway
Chagas disease
Herpes simplex infection
Measles
NF-kappa B signalling pathway
MG132
Cluster 1 enrichment score: 0.852
Cluster 2 enrichment score: 0.757
Cluster 3 enrichment score: 0.613
Cluster 4 enrichment score: 0.536
0
5
10
15
20
% of total hits

**Figure 4.19 Gene Ontology analysis of the BioID hits**

**A.** A graphical representation of the gene ontology (GO) analysis of the ZYG11A FL BioID hits under DMSO or MG132 treatment based on clustering of the Biological Processes (BP). Each colour marks individual enrichment clusters and the tiers displayed represent the top 3-4 biological processes. **B.** The same BioID datasets as in **A** displaying the top 3-4 pathways from GO KEGG analysis. Bars represent the % of the total analysed hits. For BP analysis, DMSO and MG132 subsets are as a percentage of 137 and 176 candidates, respectively. For KEGG analysis, DMSO and MG132 subsets are both as a percentage of 82 candidates. The processes and pathways with significant adjusted P-values (Benjamini values) are marked by *. Note that no KEGG pathways show significant Benjamini values.

To visualise how all of these datasets linked together based on previously described interactions, STRING analysis was performed on the whole dataset as well as for the FL+NT, and FL+CT datasets for both DMSO and MG132. In the N- and C-terminal containing subsets, some of the top enriched pathways have been highlighted by colour and suggest that the N- and C-terminals are related to distinct processes and pathways (see Figure 4.20 and Figure 4.21). Cell cycle appears to be consistently linked to the FL+CT subset in both treatment datasets which may indicate the relevance of the ARM domain. Metabolic processes are present in both FL+NT, and FL+CT datasets, suggesting that ZYG11A interacts with many metabolism related proteins.

In order to further tease apart these datasets, Gene Set Enrichment Analysis (GSEA) using the Molecular Signatures Database (MSigDB) was also used as it contained a number of independently curated gene sets that were not included in the DAVID analysis. This was done using the DMSO and MG132 datasets which were compared against all of the available datasets in MSigDB. Overall, the top 10 enrichments were quite similar compared to the DAVID analysis, with the DMSO samples enriching for metabolic processes and nucleotide binding (Table 4.3), and the MG132 samples enriching for cell cycle and innate immunity signalling (lectin binding) (Table 4.4). To identify whether the FL+NT or FL+CT datasets showed specific enrichment for any of these signatures, they were also subjected to GSEA and any similarities were also recorded. From this, it appeared that under DMSO conditions, the N-terminal of ZYG11A was largely related to ribonucleotide binding, whereas the C-terminal was more related to the oxidation and reduction process (Table 4.3). In the MG132 dataset, the C-terminal was largely linked to cell cycle processes (Table 4.4). Although only enriched when looking at the entire DMSO dataset, one of the curated databases related to BRCA1 and centrosome dysfunction was enriched (29/140 genes) in the DMSO dataset (Pujana *et al.*, 2007). This included genes

like MTOR, ATP5PO and ARPC3, and was interesting considering that BRCA2 was present in the MG132 ZYG11A FL dataset. Moreover, BRCA2 has been found to be a potential ZYG11A interactor in another study (Malik *et al.*, 2016), providing a promising link between the two genes. How all of these enriched protein datasets may relate to ZYG11A function is elaborated in the discussion of this chapter further below.

**Table 4.3 GSEA analysis using the Molecular Signatures Database (DMSO dataset)**

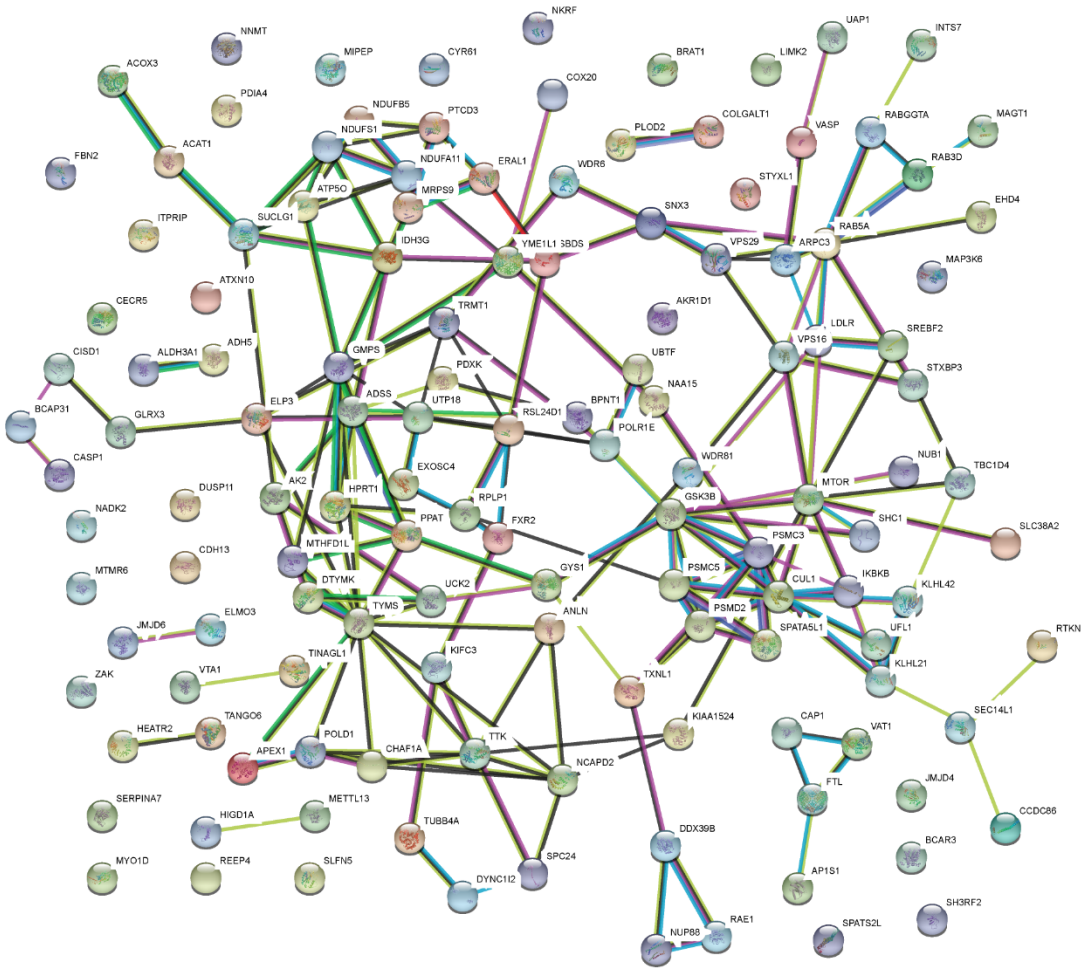
| Gene set name | Genes in overlap | FDR q-value | In FL+N or FL+C datasets? (q-value) |
|---|---|---|---|
| GO_MITOCHONDRION | 31 | 1.25 $e^{-10}$ | |
| GO_OXIDATION_REDUCTION_PROCESS | 25 | 1.72 $e^{-10}$ | C-terminal (8 genes), 7.8 $e^{-3}$ |
| GO_PROTEIN_CONTAINING_COMPLEX_ASSEMBLY | 32 | 1.72 $e^{-10}$ | |
| GO_SMALL_MOLECULE_METABOLIC_PROCESS | 32 | 1.76 $e^{-9}$ | |
| PUJANA_BRCA1_PCC_NETWORK | 32 | 2.58 $e^{-9}$ | |
| GO_RIBONUCLEOTIDE_BINDING | 29 | 2.68 $e^{-9}$ | N-terminal (12 genes), 7.11 $e^{-4}$ |
| GO_NUCLEOBASE_CONTAINING_SMALL_MOLECULE_METABOLIC_PROCESS | 31 | 4 $e^{-9}$ | |
| DODD_NASOPHARYNGEAL_CARCINOMA_DN | 19 | 1.18 $e^{-8}$ | |
| LEE_BMP2_TARGETS_DN | 21 | 2.42 $e^{-8}$ | |
| GO_ORGANOPHOSPHATE_METABOLIC_PROCESS | 22 | 4.12 $e^{-8}$ | |

**Table 4.4 GSEA analysis using the Molecular Signatures Database (MG132 dataset)**

| Gene set name | Genes in overlap | FDR q-value | In FL+N or FL+C datasets? (q-value) |
|---|---|---|---|
| GO_MITOTIC_CELL_CYCLE | 30 | 2.55 $e^{-11}$ | C-terminal (10 genes), 2.05 $e^{-3}$ |
| GO_CELL_CYCLE | 38 | 1.09 $e^{-10}$ | C-terminal (15 genes), 1.57 $e^{-4}$ |
| GO_CELL_CYCLE_PROCESS | 33 | 1.09 $e^{-10}$ | C-terminal (12 genes), 1.07 $e^{-3}$ |
| GO_REGULATION_OF_CELL_CYCLE_PROCESS | 25 | 3.11 $e^{-10}$ | C-terminal (9 genes), 2.05 $e^{-3}$ |
| GO_CELL_CYCLE_PHASE_TRANSITION | 22 | 1.64 $e^{-9}$ | C-terminal (9 genes), 1.07 $e^{-3}$ |
| GO_REGULATION OF_CELL_CYCLE | 29 | 2.41 $e^{-9}$ | C-terminal (11 genes), 1.63 $e^{-3}$ |
| GO_REGULATION_OF_CELL_CYCLE | 22 | 2.66 $e^{-9}$ | |
| REACTOME_INNATE_IMMUNE_SYSTEM | 27 | 7.24 $e^{-9}$ | |
| GO_REGULATION_OF_CELL_CYCLE_PHASE_TRANRANSITION | 18 | 4.07 $e^{-8}$ | C-terminal (8 genes), 1.07 $e^{-3}$ |
| REACTOME_CELL_CYCLE_MITOTIC | 19 | 5.02 $e^{-8}$ | |

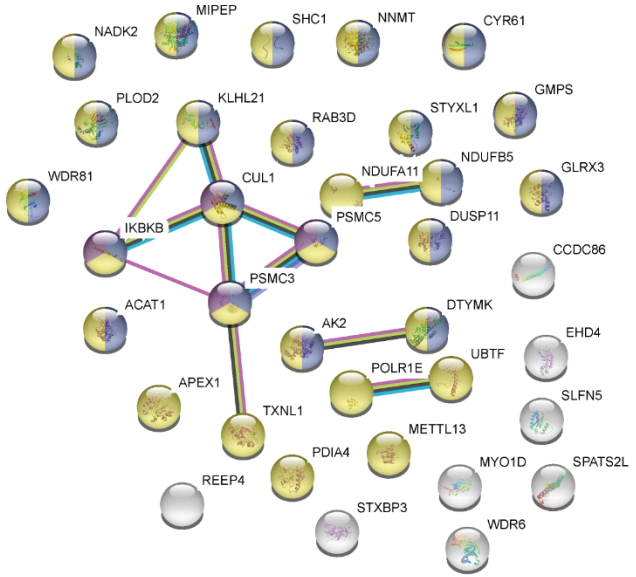
A

Full dataset (FL, FL and NT, FL and CT)




B

FL and NT hits




C

FL and CT hits

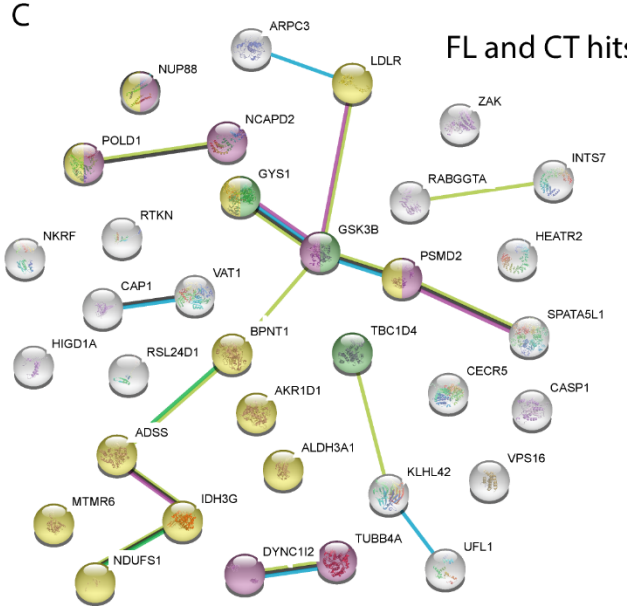
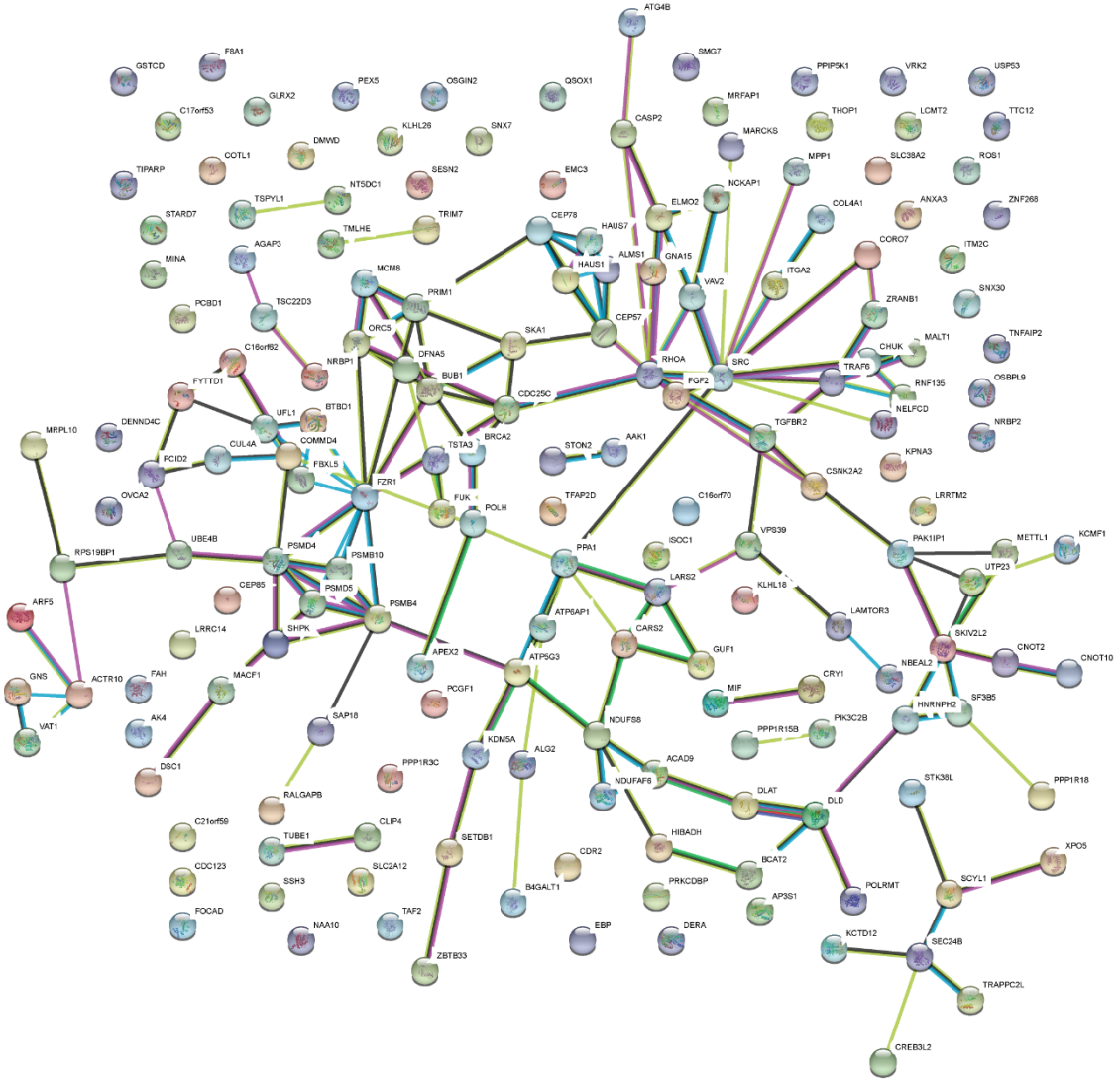


**Figure 4.20 STRING analysis of the DMSO BioID dataset**

**A.** STRING analysis of the hits from the entire DMSO BioID dataset. The STRING protein-protein interaction enrichment value for all hits is 3.72e-07, which indicates significant biological connections compared to what would be expected in a random set of proteins. **B.** STRING analysis of the FL+NT subsets. Although the protein-protein interaction score was not significant, some of the enriched processes that were previously identified by DAVID GO analysis are highlighted. Proteins related to Organonitrogen compound biosynthetic process (dark blue), Metabolic process (yellow) and NF-κB signalling (purple) are enriched in this subset. **C.** STRING analysis of the FL+CT subsets. The protein-protein interaction score was not significant, however some of the enriched processes that were previously identified by DAVID GO analysis are highlighted. Proteins related to Insulin signalling (green), Cell cycle (pink) and Metabolic process (yellow) were enriched in this subset. Note that other than metabolic process, the enriched processes are different between the N- and C-terminal interactions.

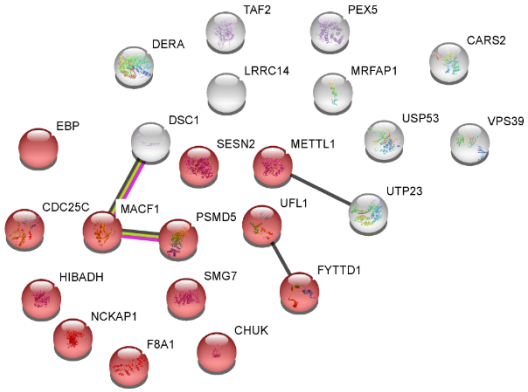
A

Full dataset (FL, FL and NT, FL and CT)




B

FL and NT hits




C

FL and CT hits

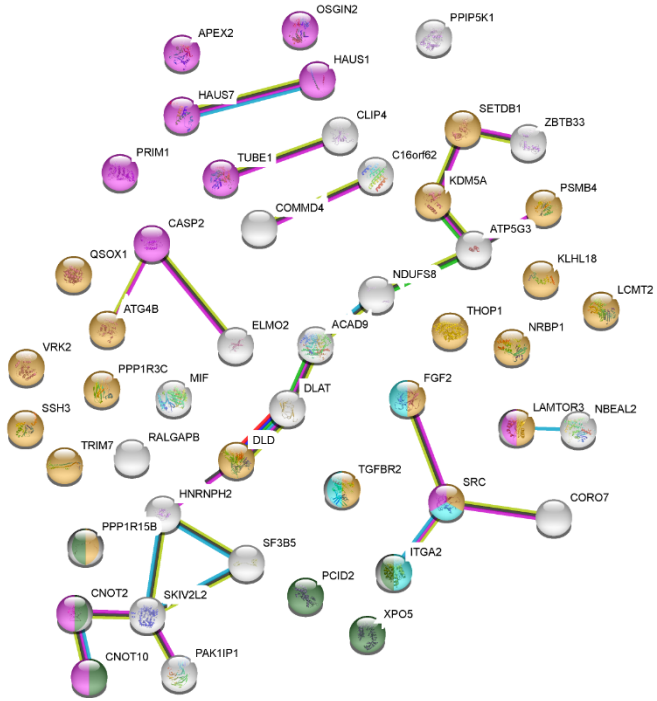


**Figure 4.21 STRING analysis of the MG132 BioID dataset**

**A.** STRING analysis of the hits from the entire MG132 BioID dataset. The STRING protein-protein interaction enrichment value, which indicates significant biological connections compared to what would be expected in a random set of proteins, was not significant. **B.** STRING analysis of the FL+NT subsets. The protein-protein interaction enrichment value was not significant and the only pathway that are enriched in this subset was Acetylation (red). **C.** STRING analysis of the FL+CT subsets. The protein-protein interaction score was not significant, however some of the enriched processes that were previously identified by DAVID GO analysis are highlighted. Proteins related to Macromolecule metabolic process (green), Cell cycle (pink) and Positive regulation of epithelial cell migration (blue) and post-transcriptional regulation of gene expression (green) were enriched in this subset. Note that the enriched processes are different between the N- and C-terminal interactions.

### 4.2.9 ZYG11A interacts with proteins containing N-terminal glycine residues, similar to ZYG11B and ZER1

Recently, it was discovered that ZYG11B and ZER1 specifically target a novel N-end rule pathway through N-terminal glycine residues, however ZYG11A was not found to mediate this pathway, likely due to the low expression of ZYG11A (Timms *et al.*, 2019). This led to the question of whether ZYG11A, when in abundance, can target proteins through the same N-terminal glycine. To see whether this was the case for any of the BioID ZYG11A hits, the N-terminal residues were initially screened for N-terminal glycines. They were then sorted based on the following residue where F, G, H, L, M or Y were 'favoured', and D, E, I, N, P, R, S or T were 'unfavoured' in terms of their susceptibility to be targeted by ZYG11 family proteins for degradation (Timms *et al.*, 2019). From both the DMSO and MG132 datasets, most hits that possessed an N-terminal glycine, were followed by an unfavoured residue in the second position including: PDXK, MTHFD1L, FBN2, WDR6, LDLR, TSTA3, COL4A1, TGFBR2, ITGA2, LCMT2 and SRC. The hits that were 'favourable' included FXR2, PLOD2, GSTT2; GSTT2B (unable to distinguish between these based on mass spectrum) LRRTM2, ARF5 and FAH. Outside of these two categories were proteins like CRY1 and MARCKS, which had N-terminal glycines followed by uncharacterised residues like valine and alanine, respectively. Many of these N-terminal glycine hits were enriched following proteasome inhibition and were typically found in the ZYG11A FL or C-terminal datasets. Since the specific protein interaction domains have not been identified, this suggests that the C-terminal may mediate this function. With only 6% of the total hits containing an N-terminal glycine, it is definitely possible that ZYG11A is indeed able to target N-terminal glycine degrons, although it is also likely that it can also target other degrons. Indeed, N-terminal glycine residues with 'favourable' following residues are lowly abundant in the human genome, so it is not surprising to see low numbers.

Nevertheless, this network of ZYG11A interacting proteins provides a number of new candidates through which the ZYG11A phenotype is propagated. Importantly, the gene ontology implicates ZYG11A in a number of previously unrelated pathways that, along with the identification of N-terminal glycine-bearing proteins, highlight many proteins related to the maintenance of cell shape.

## 4.3 Discussion

The precise molecular mechanisms that govern ZYG11 family activity remain largely elusive with only a handful of bona fide binding partners and ubiquitination substrates that have been validated and linked to a specific function. Certainly, this gap in knowledge is particularly apparent in the case of ZYG11A, which as highlighted by earlier chapters, is the more prominent family member in the maintenance of normal cell morphology and polarity compared to ZER1 and ZYG11B. Therefore, a better understanding of the pathways that ZYG11A functions in and which proteins are de-regulated as a consequence of ZYG11A loss is important to identify. In this chapter I have explored the relevance of the current potential ZYG11A associated genes in the context of functional interaction in a knockdown setting. Approaching this through a boutique high content functional screen, I have identified a small subset of ubiquitin related proteins whose knockdown appear to phenocopy *ZYG11A* loss. While rescue screening failed to identify any complete ZYG11A rescues, it did reveal specific knockdown combinations with the ability to partially rescue the ZYG11A plating density, albeit modestly. In addition, this rescue seemed to be associated with the loss of expression of the known polarity protein, Scribble. As a complementary and an unbiased approach to identifying novel interactions, proximity biotinylation was used to map the spatial interactome of ZYG11A, extending on the identification of novel protein-protein interactions beyond the capabilities of traditionally used co-immunoprecipitation methods. Most surprising was that the majority of this biotinylation activity occurs in the insoluble protein space where the ZYG11A fusion proteins appear to exist. The analysis of the ZYG11A protein interaction datasets highlight a number of interesting candidates, including those that are associated with not only the full length ZYG11A protein, but those that can also can be linked to either the ZYG11A N-terminal, containing the VHL box and LRRs, or the C-terminal, containing the ARM domain. There are distinct interactions that occur under proteasome inhibition which may be linked to the accumulation of these proteins into insoluble protein aggregates under proteasomal stress.

The data in this chapter further links ZYG11A and known cell polarity regulators. It expands on ZYG11A's existing relationship with cell cycle regulation and has also identified a number of novel pathways including focal adhesion regulation and cellular metabolism. These findings outline the proteins surrounding ZYG11A and shed light on a number of candidate substrates for this largely uncharacterised ubiquitin ligase substrate recognition subunit, specifically in a mammalian context. Overall, my studies have

demonstrated that there is still much to uncover surrounding the underlying cause of the *ZYG11A* knockdown phenotype and its protein interactions.

### 4.3.1 Identification of potential ZYG11A associated proteins

The genes and proteins that I have curated that genetically interact or physically bind to ZYG11/ZYG11A were obtained from the databases of generic large-scale genome-wide genetic or proteomic studies. As with most large screening data sets, very few hits are followed up and in the case of ZYG11A, these specific interactions and functional relationships have not been further tested until my study. An important strategy in the validation of these potential interactors was to see if I could initially triage in my functional screen whether these interactors may be 1. Effectors or positive regulators of ZYG11A function, or 2. Direct substrates of ZYG11A. For effectors/regulators of ZYG11A, I would expect their loss to phenocopy loss of function of *ZYG11A* and to perhaps worsen this phenotype in conjunction with loss of *ZYG11A* (i.e. be enhancers of the phenotype) (refer to Figure 4.5). In contrast, degradation substrates of ZYG11A would be expected to be ectopically or overexpressed following *ZYG11A* knockdown, hence although they might not give a loss of function phenotype alone, their depletion should be able to rescue partially or wholly the loss of ZYG11A function phenotypes.

The data in this chapter demonstrate that the loss of a small subset of these known interactors can phenocopy the morphological and cell cycle effects of *ZYG11A* loss, suggesting that these genes may act in a similar pathway to ZYG11A. Importantly, as may be expected one of the hits was *CUL2*, which is the main component of the E3 ubiquitin ligase complex that ZYG11A is known to be part of. *EMI1* is also an interesting hit having been characterised as a negative regulator of the APC/C complex at both the APC/$C^{Cdc20}$ dependent G1 to S phase transition, as well as at the APC/$C^{Cdh1}$ mitotic entry phase (Reimann *et al.*, 2001a; Reimann *et al.*, 2001b). This is relevant as ZYG11A and ZYG11B act redundantly with APC/C at the G2 to M phase transition and may be similarly regulated by EMI1 at this cell cycle transition (Balachandran *et al.*, 2016). Moreover, the loss of *EMI1* is known to induce a reduction of S phase cells, consistent to what I observed in Chapter 3, as well as in other studies, when ZYG11A is lost (Verschuren *et al.*, 2007; X. Wang *et al.*, 2016). Although the analysis of siZYG11A and siEMI1 cell cycle did not display these changes in a high-content screening setting, it is important to note that only the remaining adherent cells were analysed and any cell cycle defects leading to death and

removal from the plate has not been analysed in the high content screening situation. The final hit, *USP42* is a peptidase that acts to de-ubiquitinate the tumour suppressor p53 and histone H2B, which are both independent de-ubiquitination targets (Hock *et al.*, 2011; Hock, Vigneron and Vousden, 2014). The reported stabilisation of p53 by USP42 is interesting in view of the elevated levels of *p21* mRNA expression I observed in *ZYG11A* knockdown cells (Figure 3.14). However, as I already showed in Chapter 3, I observed no evidence of p53 protein stabilisation following ZYG11A depletion. Importantly, the clear differences between each cell cycle profile suggests that the similar morphological and loss of cell survival phenotypes may not be solely due to a common underlying cell cycle aberration, and hence that the alterations in cell shape is likely due to some other as yet to be determined cell cycle independent mechanisms.

Aside from *CUL2*, *EMI1* and *USP42*, the remaining candidates showed very different morphological phenotypes proving to be the major caveat of this screen. With my collaborators at the VCFG, Cellprofiller was used to attempt to separate and cluster each phenotype however it was not possible to create a pipeline that was effective and robust enough to quantitate and separate each of the phenotypes consistently based on morphology. Having said this, the hits presented here were the most robustly similar to the *ZYG11A* knockdown phenotype with the identification of *CUL2* providing a level of confidence in the manual identification of clusters. The findings reported here are relevant because they pinpoint a number of genes that are likely to function in the same pathway as ZYG11A, having already been known to bind or interact with ZYG11/ZYG11A previously. Separating the contribution of cell cycle changes and cell polarity alterations in the phenotype of ZYG11A remains to be a specific point to address in the future.

Using the cell plating density implicated a number of partial rescuers including both Cullin-2 (*CUL2*) and Elongin B (*TCEB2*), which was surprising considering that these are two components of the Cullin-2 E3 ligase complex and that *CUL2* knockdown presented with a similar phenotype as *ZYG11A* knockdown. It could be suggested that the loss of *ZYG11A* triggers a mechanism in another Culln-2 related pathway that ultimately creates the *ZYG11A* knockdown phenotype, although as of yet there is no evidence that *ZYG11A* loss induces compensation by other ZYG11 family members, let alone other Cullin-2 substrate recognition subunits. Of note, *NUB1* was also a hit in the rescue, which as a protein known to function upstream of Cullin E3 ligases by targeting NEDD8 (a Cullin activator) for degradation (Kamitani *et al.*, 2001). However, the loss of NUB1 typically causes an upregulation of all Cullin activity, which appears to not be compatible with the

result that *CUL2* is also able to rescue *ZYG11A* loss in my studies. Another ubiquitin related hit in both screens was *USP42*, and while there is no direct link to ZYG11A, its role in p53 stabilisation and histone de-ubiquitination is still relevant in terms of *p21* and *SCRIB* mRNA upregulation (Hock *et al.*, 2011; Hock, Vigneron and Vousden, 2014).

### 4.3.2 *ZYG11A* knockdown rescue and the regulation of Scribble levels in double knockdown cells

Plating density aside, the important relationship between ZYG11A and polarity is emphasised in this chapter by the loss of Scribble protein staining in most of these double knockdown rescues (apart from *CUL2*). Scribble upregulation following the loss of ZYG11A (see Figure 3.10) marks a specific apicobasal polarity change that poses a novel mode of Scribble regulation. However, the increase in Scribble mRNA raises the question of whether ZYG11A directly targets Scribble. First and foremost, the functional rescue screening presented in this chapter does not suggest that the knockdown of Scribble is enough to rescue ZYG11A loss. Clear separation of the *ZYG11A* knockdown phenotype into cell cycle and cell adhesion/polarity-based changes has proven to be difficult especially since *ZYG11A* knockdown cells overlap and clump together, essentially preventing the establishment of clearly defined cell-cell boundaries by Scribble and β-catenin staining. Little is known about how Scribble turnover is regulated and to date, only the HPV E6 oncoviral protein working alongside the E6AP ubiquitin ligase has been found to destabilise Scribble (Nakagawa and Huibregtse, 2000). Additionally, the proteasome dependent degradation of Scribble can be protected by the presence of the intermediate filament protein, Vimentin, although whether this degradation is also through HPV E6-E6AP activity has not yet been determined (Phua, Humbert and Hunziker, 2009).

Given the increased Scribble levels following ZYG11A loss, it is now important to determine whether these double knockdown hits are affecting Scribble at the protein or transcriptional level. Indeed, *USP42* has been linked to transcriptional regulation and likewise, the final hit *ZNF483* has also been proposed to function as a transcription factor. In addition to the regulation of Scribble, one striking characteristic of these hits is the presence of unchanged β-catenin. As a prominent marker of cell-cell adhesion, the de-coupling of Scribble from β-catenin levels and/or localisation (aside from directly targeting each individual gene) is highly unusual and there are few gene knockdowns that are able to do so. Both myosin heavy chain 9 (*MYH9*) and insulin-like growth factor receptor

(*IGF1R*) are two such proteins, however these individual knockdowns typically de-couple Scribble and β-catenin from their membrane localisations rather than at the protein level seen here in double knockdowns (Smith *et al.*, 2016). Moreover, there were other non-rescuing double knockdown combinations that were able to affect β-catenin expression but not Scribble expression, notably the knockdown of *ZYG11A* with the knockdown of either *ZYG11B* or *ZER1*. This suggests that ZYG11A and at least one of the two other ZYG11 family members are redundantly required for the maintenance of β-catenin, however considering the two other knockdown combinations that were able to achieve this outcome, further work is still required.

The partial rescues identified here supply a novel pool of potential ZYG11A ubiquitination targets that affect cell plating density, polarity and adhesion. Interactions between ligases like ZYG11A and substrates are typically very transient, so these hits are currently being further validated through proximity ligation and ubiquitination studies.

### 4.3.3 ZYG11A in the insoluble protein space and the formation of aggregates under proteasomal stress

While there are only a handful of studies that have focused on mammalian ZYG11 family protein interactions, they have all been in a targeted setting where there was a clear candidate substrate, and in most cases, have been typically met with soluble ZYG11 family proteins. However, here I have shown that in MCF10A cells the ZYG11A fusion proteins are in the insoluble fraction and that the proteins it interacts with are also found in the insoluble space. This unexpected result was specific to the ZYG11 family fusion proteins as a myc-BioID-Scrib fusion protein did not present with any insolubility (data not shown). This suggests two potential situations: 1) There could be a number of conserved residues across the ZYG11A protein, as well as between the ZYG11 family members that cause insolubility, or 2) ZYG11A is bound to proteins or large structures that are insoluble. The most striking property of these fusion proteins was that under proteasomal stress, some cells showed the formation of protein aggregates which was not seen in the GFP or myc-BioID tag alone. The formation of protein aggregates is typically a protective cellular mechanism that works to alleviate proteasomal stress by diverting ubiquitinated proteins to a juxtanuclear position for lysosomal degradation. Two key mediators of this process are p62, which is required for the gathering of polyubiquitinated proteins, and HDAC6, which loads aggregates onto the microtubule network. The proteins are then bound for the

microtubule organising centre where the lysosome facilitates protein degradation. Crosstalk between these two degradation pathways is often triggered by either the presence of misfolded proteins or the accumulation of ubiquitinated proteins. It is unlikely that these proteins are misfolded as the ZYG11A, ZYG11B and ZER1 antibodies that failed to recognise the denatured ZYG11 fusion proteins by western blot, were able to detect the fusion proteins by immunofluorescence, which are typically in their native confirmation (my unpublished observation). Moreover, many studies have shown that the presence of misfolded proteins alone is able to trigger aggresome formation, however there is no sign of ZYG11A protein aggregation unless the proteasome is inhibited. Taken together, it is more likely that the ZYG11A fusions may be constantly turned over via the proteasome in order to maintain a relatively low level of expression. Indeed, this may also be related to the MCF10A cell line where the baseline levels of ZYG11 family proteins are quite low, and thus are not required at the levels of overexpression shown here.

While the capture of proteins by BioID following MG132 treatment may include other aggresome related proteins, there is still much room for the enrichment of transient ZYG11A interactions, given that aggresome formation is not widespread. The integration of the N- and C- terminal controls in the analysis of full-length hits also provide a secondary measure of robustness. It will be interesting to see whether the insolubility and aggresome formation is a specific attribute of the MCF10A cell line. These findings are significant as they present a previously unreported property of the ZYG11 family in the context of mammary epithelial cells and proteasomal inhibition. Why this aggregation occurs remains to be an important question to address in future studies of this protein family as the candidate protein-protein interactions (PPI) described below can no longer be followed up by traditional co-IP methods.

### 4.3.4 A novel ZYG11A protein interaction network

#### 4.3.4.1 Ubiquitin and degradation nodes

The ubiquitin targets of ZYG11A, namely those that have the potential to link ZYG11A to the regulation of cell polarity, are still relatively elusive. This may be due to a number of factors including the two-pronged activity of the ZYG11 family in cell cycle and cell polarity, but also the transient nature of E3 ubiquitin ligase interactions. To add an extra layer of complexity, the protein interaction profile of insoluble proteins is also notoriously difficult to capture given their requirement for harsher lysis conditions, ruling out the use

of co-IP. The use of BioID in this study provided an approach that did not require intact or stable PPIs in order to identify ZYG11A interacting proteins, also lending itself to the transient nature of ubiquitin ligase interactions. The data in this chapter show that both Elongin B and C are within close proximity with the full length ZYG11A and that Elongin C can also be found in the N-terminal subset. Based on the well-studied Cullin-$2^{VHL}$ complex, Elongin C directly interacts with the VHL box on the substrate recognition subunit and mediates the binding of both Elongin B and Cullin-2 (Stebbins *et al.*, 1999). Interactions with the Cullin-2 complex have only been validated in ZYG11B and ZER1 (Vasudevan, Starostina and Kipreos, 2007), therefore this is significant as it is the first time that mammalian ZYG11A has been shown to interact with the components of the Cullin-2 complex, albeit through a proximity dependent manner. While the N-terminal VHL box is the point of contact with the rest of the Cullin-2 machinery, whether the remaining LRRs or ARM domains are responsible for substrate attachment is unknown. Here, I have emphasised the hits present in the FL+CT subsets as more likely to be ZYG11A substrates. However, I cannot discount the possibility that hits in the FL+NT subsets are still reasonable potential substrates for future consideration. Moreover, FL only hits may still provide valid interactions since little is known about the structure of ZYG11A and thus how overall protein confirmation may affect its binding partners. It is also interesting that two other members of the Cullin-RING ligase family were identified, Cullin-1 and Cullin-4A. It is unlikely that they are ZYG11A substrates as they were both included in the high-content rescue screen and did not rescue the *ZYG11A* knockdown phenotype. Perhaps they are responsible for ZYG11A turnover. One other possibility is that they are in the same proximity as ZYG11A due to similar substrates. For example, the SCF complex (SKP1-CUL1-F-box protein) is known to target Cyclin E1 using the substrate recognition subunit, FBXW7, with Cyclin E which has also been reported to be a ZYG11A substrate (Ye *et al.*, 2004; X. Wang *et al.*, 2016). Indeed, there are many ubiquitin related proteins within both subsets. Whether these are shared substrates or part of larger ubiquitin machinery regulation poses an interesting question.

#### 4.3.4.2 Cell cycle nodes

Gene ontology on the BioID-ZYG11A lines treated with DMSO and MG132 both highlighted cell cycle as an enriched biological process, which is unsurprising given the already established roles of the ZYG11 family in this context. What is interesting is that these cell cycle associated proteins were enriched in the FL+CT datasets, indicating that

this terminal, and perhaps the ARM domain is important for cell cycle interactions. The relationship between ZYG11A and the cell cycle can be placed at two points of the cell cycle, the G2/M and G0/G1 to S phase transitions. The most well characterised role of ZYG11A is in relation to APC/C and how it functions redundantly and independently of the APC/$C^{Cdc20}$ complex to degrade Cyclin B1 at the G2 to M phase transition. Therefore, was interesting to find FZR1 (also known as Cdh1) as a hit in the MG132 subset as it is similarly associated with APC/C (APC/$C^{Cdh1}$) complex, but at the G1 to S phase transition. It would be interesting to test whether ZYG11A regulates Cdh1 protein levels, especially considering that the overexpression of Cdh1 can cause G1 arrest, similar to what has been shown in Chapter 3 in *ZYG11A* knockdown cells (Fujita *et al.*, 2009). Although G2 to M phase defects were not observed following ZYG11A loss in Chapter 3, it has already been established that ZYG11A acts during that transition. Other notable hits from the MG132 dataset like CDC25C, which is activated by Cdc2-Cyclin B1 mediated phosphorylation at the G2-M phase transition, further implicates ZYG11A at this stage of the cell cycle (Hoffmann *et al.*, 1993; Takizawa and Morgan, 2000). Perhaps these two processes are tied to ZYG11A's ability to degrade Cyclin B1, especially considering that the activation of Cdc25C occurs in a positive feedback look with Cdc2-Cyclin B1 (Hoffmann *et al.*, 1993). Another interesting node was the BRCA1 related hits enriched in the DMSO dataset. It is interesting that so many of these proteins were interacting with ZYG11A, especially since this was under steady state conditions. While this does not suggest that they are ZYG11A degradation substrates, it is possible that ZYG11A ubiquitinates these proteins for processes other than degradation. Like BRCA proteins, ZYG11A may also have functions unrelated to its catalytic activity which has not yet been described for ZYG11A. Another important point to note is that since the screening was conducted, BRCA2 has also been found as a potential ZYG11A protein-protein interactor on the BioGRID database. BRCA proteins are intimately linked with genome stability including the DNA damage response, cell cycle checkpoints, and centrosome regulation (Pujana *et al.*, 2007; Roy, Chun and Powell, 2012; Malik *et al.*, 2016). If ZYG11A functions in the BRCA pathways, this would be very relevant in terms of epithelial cancer progression, in particular breast cancer, but also prostate cancer.

#### 4.3.4.3 Cell architecture and shape nodes

Biological processes that were identified incorporated various biosynthetic processes, with notable clusters enriching for endosomal transport and protein phosphorylation. While

these can certainly form a bridge to cell polarity and cell shape regulation, of greater interest are the actin cytoskeleton (RHOA, SSH3, NCKP1, MACF1) and focal adhesions (SRC, RHOA, ITGA2) nodes. From the outset, one of the overarching questions surrounding the ZYG11 family was whether it can indeed regulate cell shape by targeting cytoskeletal proteins for degradation. These hits were enriched mainly in the MG132 datasets, lending weight to the possibility that these proteins are indeed ZYG11A ubiquitination targets. Take the focal adhesion proteins for example, the overlapping cells and thin elongations in the *ZYG11A* knockdown phenotype could be due to the deregulation of RhoA, Src and/or Integrin-α2. Indeed, there is much crosstalk between these cellular components (Arthur, Petch and Burridge, 2000; Huveneers and Danen, 2009). How *ZYG11A* knockdown cells interact with the extracellular matrix via the cytoskeletal regulators and adhesion molecules like RhoA, Src and Integrin-α2 will be important to address in the future. It may be that the elongations are caused by a reduced capacity to contract at certain adhesion points, and that the clumps are due to aberrant extracellular matrix interactions. This would also be consistent with the decreased plating efficiencies seen and the increased cell death due to anoikis. Staining for extracellular matrix components in *ZYG11A* knockdown cell clumps could be used to assess if there are altered extracellular matrix interactions. Additionally, these cells could also be plated on to matrix coated surfaces to observe whether matrix interactions can increase plating densities. If these proteins are ubiquitination targets of ZYG11A, this could also provide an explanation for the lack of migration in *ZYG11A* knockdown cells given that both of these pathways can feed into a cell's ability to migrate.

#### 4.3.4.4 ZYG11A and N-terminal glycine interactions

Prior to the recent discovery of N-terminal glycine residues as novel N-end rule degradation pathways and the specific targeting of these by ZYG11B and ZER1, few motifs were known to signal ZYG11 family-mediated degradation. Outside of degradation, N-terminal glycines are the site of N-myristoylation, a type of post-translational modification. This form of lipidation requires the N-myristoyltransferases 1 and 2 (NMT1/2) to catalyse the covalent attachment of a myristic acid, which is then able to target proteins to the plasma membrane. Moreover, many caspase cleavage events can result in the exposure of N-terminal glycines and indeed, there were proteins that harboured caspase cleavage sites with 'favourable' N-terminal glycines. As predicted with both ZYG11B and ZER1, ZYG11A-mediated degradation may also play an important role in

the clearing of un-myristoylated and caspase cleaved proteins. The hits described in this chapter harbour a small handful of N-terminal glycine containing proteins that are also followed by the 'favourable' amino acid residues. However, it is also worth noting that these proteins only make up a small proportion of hits, suggesting that the while N-terminal glycine residues are ZYG11 family targets, they are not the only protein class that ZYG11A is able to target. In fact, none of the other known mammalian ZYG11 family targets contain an N-terminal glycine.

#### 4.3.4.5 Potential models

Individually, there are a handful of interesting and relevant proteins in relation to cell shape and cell polarity. Notably, the kinase GSK3β is known to phosphorylate β-catenin as part of the APC/AXIN complex in order to facilitate its degradation. This is relevant as GSK3β has been shown to be targeted for degradation by NUB1, which from Chapter 3, is able to partially rescue the *ZYG11A* knockdown cell plating density phenotype. If NUB1 is a ZYG11A degradation target, the levels of GSK3β are expected to decrease under ZYG11A loss, which would be an interesting concept to test in the future. Moreover, the downstream effects on β-catenin are important given its established connection to the polarity regulator, Scribble (see Introduction). Indeed, while none of the main components of the apicobasal polarity complexes were hits, a number of proteins with ties to cell polarity, including components of the actin cytoskeleton and the tubulin network were present. In particular, the regulation of cell adhesion (see section 4.3.4.3) through Src, RhoA, and Integrin-α2 would explain many of the cell clumping and elongation defects and may indicate that the reduced plating density may also be due to aberrant cell adhesion rather than cell cycle arrest or apoptosis. Together, these provide a number of novel pathways that may ultimately explain the *ZYG11A* knockdown phenotype and now need to be followed up.

### 4.3.5 Validation of Proximity Proteomics

Proximity biotinylation has proven to be an important tool for identifying novel ZYG11A interactions, however the biological validation of each of these hits is still essential. There are many ways to pursue interesting and relevant hits. To begin with, the knockdown lines can be analysed for protein accumulation of the BioID hits through western blot, immunofluorescence or whole cell proteomic analysis. Since direct binding is difficult in

light of the harsh lysis conditions required for western blotting, proximity ligation assays can be used to further validate proximity protein-protein interactions. The high content rescue screen developed here can be applied to the top hits from the BioID datasets. The rescue of the *ZYG11A* knockdown phenotype will be the gold standard for any ubiquitination target, although if this phenotype is due to the deregulation of many proteins then this may prove more difficult. If this is the case, then rather than using a transient siRNA system, stable short hairpin RNA interference may be more feasible, especially considering that *ZYG11A* CRISPR lines showed no phenotype.

## Summary

Loss of function of ZYG11A leads to altered cellular morphology and decreased survival of mammalian epithelial cells and loss of its single homologue in *Drosophila* is lethal. The underlying mechanisms that govern these changes are unclear. In this chapter I have expanded on this knowledge by validating potential ZYG11/ZYG11A binding partners and revealing a number of proteins that may serve as degradation substrates, or effectors or regulators of ZYG11A function. Furthermore, I have identified a number of potential ZYG11A substrates that further link the polarity protein Scribble to the *ZYG11A* loss phenotype in MCF10A cells. This is the first time that the known protein interaction landscape for ZYG11A specifically has been assessed. However, considering the lack of a complete rescue it is now paramount to expand the knowledge on the protein interaction network of ZYG11A in order to properly decipher which pathways are responsible for these dramatic changes in polarity and cell shape. For the first time, ZYG11A has been shown to interact with components of the Cullin-2 complex in mammalian systems showing that despite the dominance of the $CUL2^{ZYG11B/ZER1}$ complex in the literature, the formation of $CUL2^{ZYG11A}$ is also possible. In light of the recent discovery that ZYG11B and ZER1 target N-terminal glycines, a number of hits also implicate ZYG11A in this function, with favourable hits providing promising links to the actin cytoskeleton and polarity-associated signalling pathways. Most surprisingly, however, is the insolubility of ZYG11A and other ZYG11 family fusion proteins in MCF10A cells, likely due to a low tolerance of high ZYG11 levels. Together, this poses two overarching questions: 1. If ZYG11A is indeed able to target N-terminal degrons, how similar are these substrates to those of ZYG11B and ZER1, and do varying levels of affinity play a role in this?, and 2. Is the family constantly turned over in an attempt to maintain low levels of activity, given

that high levels of N-terminal glycine turnover could compete with the endogenous levels N-myristoylation activity?

With the caveat now being that ZYG11A proteins are linked to the insoluble protein space, the validation of direct interactions will be limited to methods like proximity ligation assays. Therefore, future work will rely heavily on the ability to rescue the cellular defects under *ZYG11A* loss – specifically the polarity and cell shape alterations. This can easily be done using the siRNA screening that was optimised in this chapter. Moreover, the lack of rescue in high-content screening presented here does not reveal the role other potential interactors (like FEZ1), questioning how they are involved in the larger landscape of ZYG11/ZYG11A function. Is it possible that like so many other genes, these connections are context dependent and more relevant in, for example, a neuronal setting? If so, these answers are beyond the scope of this study, but are still important questions to address. Nevertheless, the interactome of ZYG11A has been significantly built upon in this chapter, opening a number of doors between the ZYG11 family and the regulation of cell polarity and morphology.

# Chapter 5

## Discussion and future directions

## 5.1. The ZYG11 family: A conserved family with differing functions and phenotypes?

In this thesis, I have interrogated the functions and potential degradation targets of the highly conserved ZYG11 E3 ubiquitin ligase family in mammalian epithelial cells, with a particular focus on how these proteins may regulate cell polarity and cell shape in this context. I have also further investigated the fundamental requirement for the ZYG11 family in the development of two well-studied vertebrate and invertebrate models. The basis of this study stems from observations in the *C. elegans* embryo where ZYG-11 plays a dual role, regulating both the degradation of Cyclin B (CYB-1) during meiosis II, and the establishment of pnterior-posterior polarity during the transition to the first mitotic division. Using *in vitro* mammalian cell lines, I have identified significant differences in the phenotype and function of the three ZYG11 family members with ZYG11A most divergent from ZYG11B and ZER1. I revealed new functions for ZYG11A in cell survival, adhesion and migration. I have also further investigated the fundamental requirement for the ZYG11 family in the development of two well-studied animal models, *Drosophila* and *D. rerio*. I have also shown that the requirement for *zyg11* and *zer1* in embryonic development is context dependent and appears more important in *Drosophila* likely due to a lack of redundancy. Finally, in Chapter 4 I examined the genetic and physical interactome of ZYG11A, revealing a number of expected associations with the ubiquitination process and the cell cycle, but also a number of potential ubiquitination targets related to cell shape and polarity, and the *ZYG11A* loss of function phenotypes. How these findings contribute to the understanding of development and disease are discussed below.

## 5.2 The differential requirement for ZYG11 family members in invertebrate and vertebrate models

The evolutionary diversity of the ZYG11 family can be traced across a number of organisms, yet since its discovery in *C. elegans* the relevance of these genes in the development of other *in vivo* models has not been assessed until this study. The current dogma in the *C. elegans* field is that *zyg-11* is essential early on in oocyte and embryonic development, and that the maternal contribution of *zyg-11* provides important developmental cell cycle and polarity cues (Kemphues *et al.*, 1986; Carter, Roos and Kemphues, 1990; Liu, Vasudevan and Kipreos, 2004; Sonneville and Gonczy, 2004). Chapter 3, which shifted this focus into the *D. rerio* model, found that neither *zyg11*, nor

its counterpart *zer1*, are required for zebrafish early or late development. This is based on CRISPR-Cas9 knockout models that did not show any developmental defects in single or double knockout fish. Two important differences between the *C. elegans* and *D. rerio* models highlight that: 1. Unlike in *C. elegans*, the loss of *zyg11* alone does not cause embryonic lethality, and 2. Maternal contribution of *zyg11* does not play a role in early zebrafish development. Despite recent compelling evidence that CRISPR-Cas9 knockout can trigger genetic compensation events, the inclusion of *zer1* here has demonstrated that this does not account for the lack of lethality following *zyg11* loss (El-Brolosy *et al.*, 2019). This agrees with CRISPR-Cas9 cell line data which suggest that the knockout of *ZYG11A* does not cause any cellular defects, or compensation by ZYG11B or ZER1. Considering that knockdown of *ZYG11A* (as well as *ZYG11B* and *ZER1*) in MCF10A cells show a number of cellular alterations, it would be interesting to test whether morpholino knockdown of *zyg11* and *zer1*, rather than complete knockout, is able to induce developmental defects in the fish.

There are clearly differing levels of complexity between the vertebrate and invertebrate models, and this is supported by early indications that the loss of *Drosophila zer1* causes embryonic lethality. It has been established that in the *Drosophila* APC/C is responsible for degrading cyclin B at both meiosis I and II (Swan and Schüpbach, 2007). This suggest that there is no direct requirement for zyg11 in cell cycle during meiosis II, which may explain why *zyg11* is not present in the fly genome. What this means in the context of polarity establishment is unclear. By extrapolation it could be proposed that the functions of *zyg11* in embryonic polarity establishment are also fulfilled by other proteins or mechanisms in *Drosophila*. Take the asymmetric distribution of PAR proteins and P granules along the anterior-posterior axis during nematode embryonic polarity. In *Drosophila*, PAR proteins are similarly distributed to form anterior-posterior polarity, however the specific signals preceding their segregation are unclear. The asymmetry is achieved through the reorganisation of either the actin cytoskeleton (for nematodes) or the microtubule network (in flies), yet what triggers this initial polarisation signal remains unknown (reviewed by Raman, Savio and Sonawane, 2018). For *C. elegans*, correct PAR protein polarisation relies on the presence of ZYG-11 and CUL-2, which when lost reverses anterior-posterior polarity (Liu, Vasudevan and Kipreos, 2004; Sonneville and Gonczy, 2004). This presents an important question for the *Drosophila* system: In the fly where there is no *zyg11* ortholog, does *zer1* take on twice the responsibility or is this polarity regulation taken care of by another set of proteins? Because the effects of *zer-1*

loss in *C. elegans* have not been tested, it is difficult to conclude whether the loss of *zer1* in the fly reflects that of *zyg-11* or *zer-1* loss. What is important now is to identify the underlying cause of embryonic lethality in these *zer1* mutants, looking specifically at whether they exhibit altered PAR protein localisation.

With vertebrate and invertebrate models displaying varying requirements for ZYG11 family members, how can we reconcile these differences? The answer is likely to be down to the inherent differences in embryonic polarity establishment and ultimately, the complexity of the organism itself. If the seemingly essential roles for the ZYG11 family in the invertebrate embryo are not reflected in the vertebrate zebrafish model, then what is the reason for their high levels of conservation in mammalian systems? Without a ZYG11 knockout mouse model, it is difficult to test the requirement of this family in mammalian development. Furthermore, here I only observed the general development of these knockout fish, however we can assume that there are likely to be underlying ties to polarity regulation. This is due to evidence in both the mammalian cell line experiments, as well as in the *Drosophila* knockdown studies presented in Chapter 3. It will be important to characterise whether there are any subtle apicobasal or anterior-posterior polarity alterations in the somatic cells of these seemingly normal zebrafish, paying particular attention to the localisation and levels of both the Scribble and PAR complexes.

Given the surprising lack of phenotype in these zebrafish, I can also speculate on the back-up mechanisms and protective roles that ZYG11 plays. The non-canonical role of ZYG11A/B in Cyclin B degradation, and the specific targeting of N-terminal glycines for degradation by ZYG11B and ZER1 are yet to be demonstrated in a vertebrate *in vivo* model. Although it is enticing to delve further into the potential compensatory pathways following Zyg11 loss (e.g. APC/C compensation), the more interesting route would be to test the requirement of these genes under N-myristoylation inhibition or increased caspase cleavage. These fundamental processes are likely to rely on the ZYG11 family to prevent proteotoxicity following their deregulation.

## 5.3 The functional consequences of ZYG11 family loss in cellular homeostasis

### 5.3.1 The regulation of cell cycle progression

For years, the requirement for *zyg-11* in the *C. elegans* meiotic cell cycle has been the cornerstone of ZYG11 family function, yet only recently has this been studied in

mammalian systems. By analysing all three ZYG11 family members, here I show that ZYG11A and ZER1 are required for cell cycle progression at the G0/G1 to S phase transition in normal mammary epithelial cells. This supports a recent study where the loss of ZYG11A resulted in an increase in G0/G1 cells in the H1299 human non-small cell lung carcinoma cell line, and this was associated with an increase in Cyclin E (X. Wang *et al.*, 2016). Moreover, ZER1 has been previously linked to the destabilisation of RB1 (a key player in G0/G1 progression) in HPV infected cells, however this is the first time that it has been shown to impair cell cycle progression in an *in vitro*, non-viral setting. In Figure 3.15 I showed that there were no differences seen in RB1 during ZYG11A and ZER1 loss, suggesting an alternative means of G0/G1 regulation in this context. One important observation is that p21 (*CDKN1A*), which displayed no obvious protein alterations, showed increases in mRNA across all three family members. If not through p53, perhaps p21 mRNA is being regulated post-transcriptionally by Poly(C)-binding proteins (PCBPs) like PCBP4, which has been shown to regulate *p21* mRNA stabilisation (Scoumanne *et al.*, 2011). This upregulation of *p21* mRNA is also intriguing because unlike the other two family members, *ZYG11B* knockdown did not display any cell cycle defects. In Figure 3.7 I showed that ZYG11B and ZER1 are able to genetically compensate transcriptionally when one is knocked down by siRNA, and from recent studies on N-terminal glycine degrons, they are also known to compensate on a functional level (Timms *et al.*, 2019) . While it is likely that ZER1 is compensating for ZYG11B loss, the converse does not appear to be true as there are significant cell cycle alterations in *ZER1* knockdown cells, indicating that the downstream regulation by ZER1 may include unique pathways – at least in the context of G0/G1 exit. Involvement of the ZYG11 family in the mammalian cell cycle has also been explored by Balachandran et al. (2016) where ZYG11A/B appears to function redundantly with APC/C in HEK293T cells. It is unclear whether this is also true in the context of mammary epithelial cells as here I have not observed cell cycle under APC/C loss. Whether ZER1 can also degrade Cyclin B to functionally compensate for APC/C loss still remains untested in mammalian systems. In the *C. elegans* embryo this does not appear to be the case as the loss of *zyg-11* alone results in embryonic lethality. If there were compensation between ZYG-11 and ZER-1, these *zyg-11* mutants would be able to bypass CYB-1 accumulation and survive. Together with the G0/G1 cell cycle arrests discussed above, this highlights that ZYG11 family function can be partitioned, even within the cell cycle, and that each member has an individual capacity to compensate for each other's loss. Building on this, the enrichment of cell cycle regulators in the ZYG11A BioID datasets suggests a greater involvement in cell cycle than initially thought,

particularly since many of these hits were enriched under proteasome inhibition. Of note are important cell cycle regulators like Cdh1 (FZR1) and CDC25C, as well as a number of BRCA1 interacting proteins. Indeed, these potential substrates are across many phases of the cell cycle, suggesting that like so many other cell cycle related proteins, ZYG11A functions widely during cell cycle progression.

### 5.3.2 The regulation of cell shape, cell polarity, and cell migration

In the developing *C. elegans* embryo it has been postulated that ZYG-11 mediated cell cycle and polarity establishment are separable functions of ZYG-11 that are independently regulated. To date, although cell cycle alterations upon ZYG11 family loss have also been found in mammalian studies (Liu, Vasudevan and Kipreos, 2004; Sonneville and Gonczy, 2004; Balachandran *et al.*, 2016; X. Wang *et al.*, 2016; my studies), the role of ZYG11 in cell polarity has so far not been examined outside *C. elegans*. My thesis studies stemmed from previous work from a graduate student in the lab that uncovered ZYG11A as a potential regulator of cell shape and of the localisation of cell polarity regulators Scribble and β-catenin (Smith *et al.*, 2016). I introduced the concept of ZYG11A as a novel regulator of mammalian epithelial cell shape and polarity and tested the idea that during evolution, the three related mammalian homologues, ZYG11A, ZYG11B, and ZER1, acquired differing roles in mammalian species. This is based on differing effects of loss of function of each family members on cell cycle (discussed in 7.1.1.1), and evidence that while *ZYG11B* and *ZER1* knockdowns present with more rounded and separated cells, *ZYG11A* knockdown displays the opposite, tending to form cell clumps with cellular elongations and protrusions (Figure 3.7). These findings are the first to reveal that there are distinct functions for ZYG11A, and ZYG11B and ZER1, in the regulation of cell shape and polarity. Specific to *ZYG11A* loss is an increase in Scribble, however this is unlikely to account for the phenotype given that the overexpression of Scribble in MCF10A cells does not disrupt cell shape or apicobasal polarity in MCF10As, indeed it appears instead to increase the epithelial nature of these cells by enhancing tight junctions and making them more cuboidal (Elsum, Martin and Humbert, 2013). Furthermore, Scribble depletion does not rescue the loss of ZYG11A phenotype (Chapter 5). Other pathways must therefore be involved in the cellular shape changes observed.

One unusual characteristic of the gene knockdowns that partially rescue ZYG11A plating density defects was the de-coupling of Scribble and β-catenin. Work in our lab has

observed an inherent link between the membrane localisation of these proteins, with few individual genes that are able to regulate Scribble or β-catenin independently of each other (Smith *et al.*, 2016). The rescue screening hits presented here are important as it shows that Scribble and β-catenin can also be separated in terms of protein abundance and localisation, albeit how these double knockdown combinations achieve this is still unclear. Findings in *ZYG11A* knockdown cells emphasise a complex relationship where ZYG11A potentially targets a *SCRIB* transcription factor for degradation, rather than Scribble itself. How *SCRIB* expression is transcriptionally activated is still unknown. Analysis of the upstream promoter region of *SCRIB* revealed putative transcription factor binding sites including the BioID hit Kaiso, which is a transcription factor encoded by the *ZBTB33* gene. Moreover, the rescue screening hit *USP42* is known to regulate histone ubiquitination (Hock, Vigneron and Vousden, 2014), and may act to positively regulate Scribble expression, leaving β-catenin unaffected. If these two proteins do in fact regulate *SCRIB* expression, then it will be paramount to interrogate Scribble on a protein and RNA level under individual knockdown and rescue conditions.

Nevertheless, because the ZYG11A plating density defects were not rescued by the loss of Scribble, nor any of the other known ZYG11A interacting proteins, the culprits may lie in the BioID interaction network. One enriched pathway from gene ontology analysis that could account for the reduced plating density, cellular elongations, and overlapping cells is the regulation of focal adhesion proteins (e.g. SRC, RHOA, ITG2A, MYL12A/B and GSK3β). Focal adhesions are involved in how cells interact with the extracellular matrix and can dictate cell migration processes. They are intricately linked with the actin cytoskeleton which can facilitate crosstalk between the integrin-cadherin communication axis, i.e. ECM-cell and cell-cell communication. NUB1, while not traditionally associated with focal adhesions, is known to modulate GSK3β degradation and happens to be one of the only hits in both the rescue screen and BioID (see Chapter 4) (Richet *et al.*, 2012). By extrapolation, ZYG11A may degrade NUB1 and consequently stabilise GSK3β levels affecting a myriad of downstream pathways including β-catenin stabilisation, and the activity of polarity proteins like MARK2 (PAR-1) (Aberle *et al.*, 1997; Cohen *et al.*, 2004; Kosuga *et al.*, 2005). Testing whether GSK3β levels, as well as the other focal adhesion related proteins mentioned above, are deregulated following *ZYG11A* loss will be important to validate these theories.

While these phenotypes focus on apicobasal polarity, I have also underscored a role for the ZYG11 family in cell migration as observed in *ZYG11A* knockdown cells that show a

defect in directed cell migration. Many of the enriched focal adhesion proteins mentioned above are important for front-rear polarisation in migrating cells where the reorganisation of the actin cytoskeleton and cell adhesion is critical for proper cell migration (Arthur, Petch and Burridge, 2000; Huveneers and Danen, 2009). Moreover, additional actin cytoskeleton related proteins in the BioID datasets further underscores how ZYG11A may ubiquitinate and turn over cytoskeletal remodelling proteins, important for cell migration, as well as overall normal cell shape maintenance.

### 5.3.3 Emerging roles of ZYG11 family members in host-pathogen interactions

With the ZYG11A knockdown phenotype aside, gene ontology analysis of the BioID data highlighted many interesting biological processes, including viral processes. ZER1 is known to be required for cell cycle regulation in HPV harbouring cells (White *et al.*, 2012). Very recently ZYG11B has been found to bind with Orf10 of the SARS-CoV-2 protein. It has been suggested that SARS-CoV-2 either hijacks ZYG11B for its ubiquitination activity, or is targeted itself for degradation via N-terminal glycine on Orf10 (Gordon *et al.*, 2020). ZYG11B and ZER1 have both been pinpointed as the only proteins able to target N-terminal glycine residues. However, based on the interaction between ZYG11A and a number of N-terminal glycine containing proteins (Chapter 4) it is likely that the entire family is capable of associating with this specific degron. Due this specific function, perhaps the ZYG11 family has become a common target for viruses, critical for their life cycle in humans.

## 5.4 A model for the function of ZYG11A in mammalian epithelial cells

Under steady state conditions, the ZYG11 family of E3 ligases act to maintain epithelial cell proteostasis by degrading their specific substrates, although in mammalian systems, these targets have largely been limited in number (Balachandran *et al.*, 2016; X. Wang *et al.*, 2016). The knockdown studies I have presented in this thesis highlighted the effects of ZYG11 family loss in a number of cellular processes. In particular the loss of ZYG11A appears to affect a number of functions, particularly the regulation of cell shape and polarity. I have interrogated the overarching relationship between these phenotypes and the function of ZYG11A E3 ubiquitin ligase activity, and I have revealed a number of novel genetic and protein interactions. Moreover, only ZYG11B and ZER1 have been

shown to physically interact with the Cullin-2 complex (Vasudevan, Starostina and Kipreos, 2007; Bennett *et al.*, 2010). The presence of Elongin B (*TCEB2*) and Elongin C (*TCEB1*) in the BioID datasets are relevant as they highlight the first evidence that ZYG11A is able to associate with the substrate recognition subunit-linking components of the CUL2 complex on a protein-protein interaction level. Of the three family members, ZYG11A has the fewest reported protein-protein interactions raising a number of questions on which pathways it can affect. Its link to cell cycle in a mammalian context has already been highlighted previously by Balachandran et al. (2016) and Wang et al. (2016). This relationship is reinforced here by the enrichment of cell cycle related proteins in the BioID hits under both steady state and proteasome inhibition (Figure 4.19). Moreover, the presence of key cytoskeletal and focal adhesion proteins in my BioID datasets sheds light on a previously understudied function of ZYG11A in cell shape and polarity regulation.

Here, I propose a testable model where ZYG11A has developed divergent functions to ZYG11B and ZER1, specifically in terms of cell shape and polarity regulation. During homeostasis, ZYG11A, as part of the Cullin-2 complex is required for the degradation of its own subset of proteins whose turnover by the 26S proteasome is important for normal cell cycle progression at the G0/G1 to S phase transition, as well as the maintenance of epithelial organisation and integrity, and cell adhesion. When ZYG11A is lost, these proteins begin to accumulate within the cell, affecting a number of functions. At the onset of cellular adhesion, these cells have an altered capacity to adhere likely due to increased levels of Scribble, Src, RhoA, GSK3β and Integrin-α2. This also affects their ability to effectively migrate due to alterations in focal adhesion contraction and actin cytoskeleton remodelling, resulting in sustained elongations in migrating cells and aberrant cellular overlapping. Concurrently, the accumulation of cell cycle regulators like Cdh1, and the increased expression of *p21* causes an arrest at G0/G1. While the downregulation of individual proteins like NUB1 and USP42 (which somehow affect Scribble expression, perhaps through transcriptional machinery) may reduce the burden of protein accumulation due to ZYG11A loss, it is likely that there are multiple proteins responsible for these alterations. For this reason, only partial rescues can be achieved.

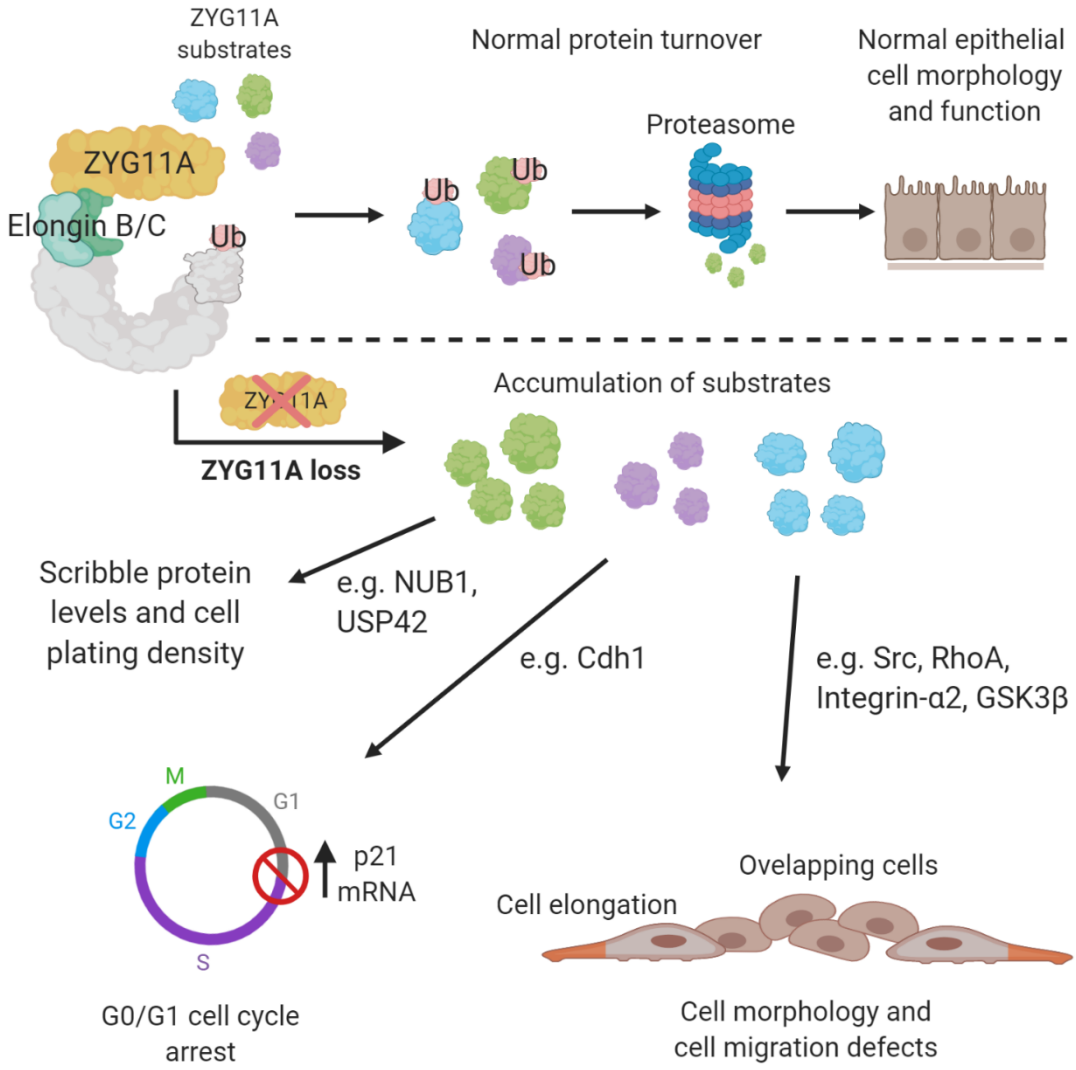



**Figure 5.1 Working model of ZYG11A regulated maintenance of normal cell morphology and function**

ZYG11A interacts with the Elongin B and Elongin C components of the Cullin-2 complex to regulate protein homeostasis under normal conditions. Upon ZYG11A loss, these proteins begin to accumulate deregulating pathways linked to both cell cycle (Cdh1, p21), migration and adhesion (Src, RhoA, Integrin-α2, GSK3β), and cell polarity (Scribble). Partial rescues of the plating density can be achieved through the downregulation of proteins like NUB1 and USP42 which decrease Scribble expression through an unknown mechanism, likely to be linked to transcriptional control machinery.

## Future studies

My study has approached the ZYG11 family from multiple aspects, including *in vitro* and *in vivo* systems. These *in vivo* models in particular provide important systems to identify how these proteins function in a whole organism. Since there is no phenotype in the *D. rerio* models, stressing these systems or sensitizing them to apoptosis or N-myristoylation defects may reveal the protective nature of the ZYG11 family. This concept can be similarly applied to the *ZYG11A* CRISPR cell lines. Comparing *ZYG11B* and *ZER1* knockout models will also be important to properly determine whether a triple knockout of the ZYG11 family is lethal in mammalian systems. The aim of Chapter 4 was to diversify the known interactome of ZYG11A and indeed, I have presented many interesting and testable pathways that ZYG11A may be involved in. Given that ZYG11 family proteins are insoluble in MCF10A cells, we cannot directly test the physical association of these proteins with ZYG11A. By far the best alternatives will require the use of Proximity Ligation Assays (PLAs), traditionally used to detect protein-protein interactions by fluorescence microscopy. The gold standard for determining whether these proteins are the ubiquitin targets responsible for the ZYG11A phenotype is by taking this new library of genes back into siRNA rescue screening. The complete rescue of the ZYG11A phenotype is going to be challenging because the causes are likely to be due to the deregulation of more than one protein. If this is the case, then future work would benefit from a single vector multi-hairpin knockdown system, for example the *Sleeping Beauty* vectors designed by the Stuhmer lab (Fink *et al.*, 2018). Thus far, this is the first study to elaborate on the genetic and protein interactome of ZYG11A, and still it is only the tip of the iceberg. What will be key is to determine exactly where ZYG11A sits amongst all of these novel interactions and interpret these in relation to the larger ZYG11 family.

## Final remarks

Recognising that not all ZYG11 family members were made equal, despite their longstanding cell cycle associations, represents an important step in understanding how this gene family maintains normal cellular function. Here I described novel findings that mammalian ZYG11 family members have varying capacities to regulate cell cycle, cell shape, polarity and migration. I also highlight contextual differences in the fundamental requirement for these genes *in vivo*. Finally, I have presented the first genetic and protein interactome study on ZYG11A, identifying a number of important potential ubiquitination targets. Now more than ever, it is crucial to determine the downstream pathways that

govern ZYG11 family function. These observations provide the necessary framework to further interrogate how the ZYG11 family regulates proteostasis, and how its deregulation can contribute to defects in development, as well as the initiation of epithelial cancer.

# References

Abdelilah-Seyfried, S. *et al.* (2000) 'A gain-of-function screen for genes that affect the development of the drosophila adult external sensory organ', *Genetics*.

Aberle, H. *et al.* (1997) 'B-Catenin Is a Target for the Ubiquitin-Proteasome Pathway', *EMBO Journal*, 16(13), pp. 3797–3804. doi: 10.1093/emboj/16.13.3797.

Akhtar, N. and Streuli, C. H. (2013) 'An integrin-ILK-microtubule network orients cell polarity and lumen formation in glandular epithelium', *Nature Cell Biology*. doi: 10.1038/ncb2646.

Ardlie, K. G. *et al.* (2015) 'The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans', *Science*. doi: 10.1126/science.1262110.

Arthur, W. T., Petch, L. A. and Burridge, K. (2000) 'Integrin engagement suppresses RhoA activity via a c-Src-dependent mechanism', *Current Biology*. doi: 10.1016/S0960-9822(00)00537-6.

Ataeian, M. *et al.* (2016) 'Maternal MEMI Promotes Female Meiosis II in Response to Fertilisation in Caenorhabditis elegans', *Genetics*. 2016/10/13, 204(4), pp. 1461–1477. doi: 10.1534/genetics.116.192997.

Aubrey, B. J. *et al.* (2015) 'An Inducible Lentiviral Guide RNA Platform Enables the Identification of Tumor-Essential Genes and Tumor-Promoting Mutations InVivo', *Cell Reports*. doi: 10.1016/j.celrep.2015.02.002.

Balachandran, R. S. *et al.* (2016) 'The ubiquitin ligase CRL2 ZYG11 targets cyclin B1 for degradation in a conserved pathway that facilitates mitotic slippage', *The Journal of Cell Biology*. 2016/11/05, 215(2), pp. 151–166. doi: 10.1083/jcb.201601083.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society: Series B (Methodological)*. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Bennett, E. J. *et al.* (2010) 'Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics', *Cell*, 143. doi: 10.1016/j.cell.2010.11.017.

Berndsen, C. E. and Wolberger, C. (2014) 'New insights into ubiquitin E3 ligase mechanism', *Nat Struct Mol Biol*, 21. doi: 10.1038/nsmb.2780.

Bloom, L. and Horvitz, H. R. (1997) 'The Caenorhabditis elegans gene unc-76 and its human homologs define a new gene family involved in axonal outgrowth and fasciculation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.94.7.3414.

Brandeis, M. (2016) ‘Slip slidin’ away of mitosis with CRL2Zyg11’, *Journal of Cell Biology*, 215(2), pp. 143–145. doi: 10.1083/jcb.201609086.

Budhidarmo, R., Nakatani, Y. and Day, C. L. (2012) ‘RINGs hold the key to ubiquitin transfer’, *Trends Biochem Sci*. Elsevier, 37(2), pp. 58–65. doi: 10.1016/j.tibs.2011.11.001.

Cai, W. and Yang, H. (2016a) ‘The structure and regulation of Cullin-2 based E3 ubiquitin ligases and their biological functions’, *Cell Division*, 11(1), p. 7. doi: 10.1186/s13008-016-0020-7.

Cai, W. and Yang, H. (2016b) ‘The structure and regulation of Cullin-2 based E3 ubiquitin ligases and their biological functions’, *Cell Division*. BioMed Central, 11(1), pp. 1–11. doi: 10.1186/s13008-016-0020-7.

Campanale, J. P., Sun, T. Y. and Montell, D. J. (2017) ‘Development and dynamics of cell polarity at a glance’, *Journal of Cell Science*, 130(7), pp. 1201–1207. doi: 10.1242/jcs.188599.

Van Campenhout, C. A. *et al.* (2011) ‘Dlg3 Trafficking and apical tight junction formation is regulated by Nedd4 and Nedd4-2 E3 Ubiquitin ligases’, *Developmental Cell*. doi: 10.1016/j.devcel.2011.08.003.

Carter, P. W., Roos, J. M. and Kemphues, K. J. (1990) ‘Molecular analysis of zyg-11, a maternal-effect gene required for early embryogenesis of Caenorhabditis elegans’, *Mol Gen Genet*, 221(1), pp. 72–80. doi: 10.1007/bf00280370.

De Celis, J. F. (1998) ‘Positioning and differentiation of veins in the Drosophila wing’, *International Journal of Developmental Biology*, 42(3), pp. 335–343. doi: 10.1387/ijdb.9654017.

Ciechanover, A. *et al.* (1981) ‘Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system’, *Proceedings of the National Academy of Sciences of the United States of America*, 78(2 II), pp. 761–765. doi: 10.1073/pnas.78.2.761.

Cohen, D. *et al.* (2004) ‘Mammalian PAR-1 determines epithelial lumen polarity by organizing the microtubule cytoskeleton’, *Journal of Cell Biology*, 164(5), pp. 717–727. doi: 10.1083/jcb.200308104.

Cox, J. and Mann, M. (2008) ‘MaxQuant enables high peptide identification rates, individualised p.p.b.-range mass accuracies and proteome-wide protein quantification’, *Nature Biotechnology*. doi: 10.1038/nbt.1511.

Coyaud, E. *et al.* (2015) ‘BioID-based Identification of Skp Cullin F-box (SCF) $^{\beta\text{-TrCP1/2}}$ E3 Ligase

Substrates', *Molecular & Cellular Proteomics*, 14(7), pp. 1781–1795. doi: 10.1074/mcp.M114.045658.

Daulat, A. M. and Borg, J.-P. (2017) 'Wnt/Planar Cell Polarity Signaling: New Opportunities for Cancer Treatment', *Trends in Cancer*, 3(2), pp. 113–125. doi: https://doi.org/10.1016/j.trecan.2017.01.001.

Dho, S. E. *et al.* (2019) 'Proximity interactions of the ubiquitin ligase Mind bomb 1 reveal a role in regulation of epithelial polarity complex proteins', *Scientific Reports*, 9(1), pp. 1–18. doi: 10.1038/s41598-019-48902-x.

Dow, L. E. *et al.* (2007) 'The tumour-suppressor Scribble dictates cell polarity during directed epithelial migration: regulation of Rho GTPase recruitment to the leading edge', *Oncogene*. 2006/10/18, 26(16), pp. 2272–2282. doi: 10.1038/sj.onc.1210016.

Dow, L. E. *et al.* (2008) 'Loss of human Scribble cooperates with H-Ras to promote cell invasion through deregulation of MAPK signalling', *Oncogene*. 2008/07/22, 27(46), pp. 5988–6001. doi: 10.1038/onc.2008.219.

Duffy, J. B. (2002) 'GAL4 system in Drosophila: a fly geneticist's Swiss army knife', *Genesis*. 2002/09/27, 34(1–2), pp. 1–15. doi: 10.1002/gene.10150.

Dyson, N. *et al.* (1989) 'The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product', *Science*, 243. doi: 10.1126/science.2537532.

El-Brolosy, M. A. *et al.* (2019) 'Genetic compensation triggered by mutant mRNA degradation', *Nature*. doi: 10.1038/s41586-019-1064-z.

El-Brolosy, M. A. and Stainier, D. Y. R. (2017) 'Genetic compensation: A phenomenon in search of mechanisms', *PLoS Genetics*, 13(7), pp. 1–17. doi: 10.1371/journal.pgen.1006780.

Elsum, I. A., Martin, C. and Humbert, P. O. (2013) 'Scribble regulates an EMT polarity pathway through modulation of MAPK-ERK signaling to mediate junction formation', *Journal of Cell Science*, 126(17), pp. 3990–3999. doi: 10.1242/jcs.129387.

Eurico-de-Sá, M. *et al.* (2014) 'Slmb antagonises the aPKC/Par-6 complex to control oocyte and epithelial polarity', *Development (Cambridge)*. doi: 10.1242/dev.109827.

Feng, H. *et al.* (1999) 'CUL-2 is required for the G1-to-S phase transition and mitotic chromosome condensation in Caenorhabditis elegans', *Nat Cell Biol*. 1999/12/10, 1(8), pp. 486–492. doi:

10.1038/70272.

Fink, S. *et al.* (2018) 'A simple approach for multi-targeted shRNA-mediated inducible knockdowns using Sleeping Beauty vectors', *PLoS ONE*. doi: 10.1371/journal.pone.0205585.

Fujita, T. *et al.* (2007) 'Axonal guidance protein FEZ1 associates with tubulin and kinesin motor protein to transport mitochondria in neurites of NGF-stimulated PC12 cells', *Biochemical and Biophysical Research Communications*, 361(3), pp. 605–610. doi: 10.1016/j.bbrc.2007.07.050.

Fujita, T. *et al.* (2009) 'An in vivo study of Cdh1/APC in breast cancer formation', *International journal of cancer*, 125(4), pp. 826–836. doi: 10.1002/ijc.24399.

Ganner, A. *et al.* (2009) 'Regulation of ciliary polarity by the APC/C', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0909465106.

Gindhart, J. G. *et al.* (2003) 'The kinesin-associated protein UNC-76 is required for axonal transport in the Drosophila nervous system', *Molecular Biology of the Cell*. doi: 10.1091/mbc.E02-12-0800.

Godde, N. J. *et al.* (2010) 'Cell polarity in motion: Redefining mammary tissue organisation through EMT and cell polarity transitions', *Journal of Mammary Gland Biology and Neoplasia*. doi: 10.1007/s10911-010-9180-2.

Gómez-López, S., Lerner, R. G. and Petritsch, C. (2014) 'Asymmetric cell division of stem and progenitor cells during homeostasis and cancer', *Cellular and Molecular Life Sciences*. doi: 10.1007/s00018-013-1386-1.

Gordon, D. E. *et al.* (2020) 'A SARS-CoV-2 protein interaction map reveals targets for drug repurposing'. doi: 10.1038/s41586-020-2286-9.

Gouw, M. *et al.* (2018) 'The eukaryotic linear motif resource - 2018 update', *Nucleic Acids Research*. doi: 10.1093/nar/gkx1077.

De Groot, R. E. A. *et al.* (2014) 'Huwe1-mediated ubiquitylation of dishevelled defines a negative feedback loop in the Wnt signaling pathway', *Science Signaling*. doi: 10.1126/scisignal.2004985.

Grumati, P. and Dikic, I. (2018) 'Ubiquitin signaling and autophagy', *Journal of Biological Chemistry*. doi: 10.1074/jbc.TM117.000117.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: the next generation', *Cell*.

2011/03/08, 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Harding, K. and White, K. (2018) 'Drosophila as a model for developmental biology: Stem cell-fate decisions in the developing nervous system', *Journal of Developmental Biology*. doi: 10.3390/jdb6040025.

Hedgecock, E. M. *et al.* (1985) 'Axonal guidance mutants of Caenorhabditis elegans identified by filling sensory neurons with fluorescein dyes', *Developmental Biology*. doi: 10.1016/0012-1606(85)90443-9.

Hershko, A. and Ciechanover, A. (1998) 'THE UBIQUITIN SYSTEM', *Annual Review of Biochemistry*. doi: 10.1146/annurev.biochem.67.1.425.

Hirsh, D. and Vanderslice, R. (1976) 'Temperature-sensitive developmental mutants of Caenorhabditis elegans', *Developmental Biology*, 49(1), pp. 220–235. doi: 10.1016/0012-1606(76)90268-2.

Hock, A. K. *et al.* (2011) 'Regulation of p53 stability and function by the deubiquitinating enzyme USP42.', *The EMBO journal*. European Molecular Biology Organisation, 30(24), pp. 4921–30. doi: 10.1038/emboj.2011.419.

Hock, A. K., Vigneron, A. M. and Vousden, K. H. (2014) 'Ubiquitin-specific peptidase 42 (USP42) functions to deubiquitylate histones and regulate transcriptional activity', *Journal of Biological Chemistry*, 289(50), pp. 34862–34870. doi: 10.1074/jbc.M114.589267.

Hoeller, D. *et al.* (2007) 'E3-independent monoubiquitination of ubiquitin-binding proteins', *Mol Cell*. Elsevier, 26(6), pp. 891–898. doi: 10.1016/j.molcel.2007.05.014.

Hoffmann, I. *et al.* (1993) 'Phosphorylation and activation of human cdc25-C by cdc2--cyclin B and its involvement in the self-amplification of MPF at mitosis.', *The EMBO Journal*. doi: 10.1002/j.1460-2075.1993.tb05631.x.

Howlett, A. R. *et al.* (1995) 'Cellular growth and survival are mediated by β1 integrins in normal human breast eqithelium but not in breast carcinoma', *Journal of Cell Science*.

Hu, D. *et al.* (2011) 'The emerging role of APC/CCdh1 in development', *Semin Cell Dev Biol*. 2011/04/19, 22(6), pp. 579–585. doi: 10.1016/j.semcdb.2011.03.012.

Huang, C. *et al.* (2009) 'Talin phosphorylation by Cdk5 regulates Smurf1-mediated talin head ubiquitylation and cell migration', *Nature Cell Biology*, 11(5), pp. 624–630. doi: 10.1038/ncb1868.

Huibregtse, J. M. *et al.* (1995) 'A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase', *Proceedings of the National Academy of Sciences*, 92(7), pp. 2563–2567. doi: 10.1073/pnas.92.7.2563.

Humbert, P. O. *et al.* (2008) 'Control of tumourigenesis by the Scribble/Dlg/Lgl polarity module', *Oncogene*, 27(55), pp. 6888–6907. doi: 10.1038/onc.2008.341.

Humbert, P. O., Russell, S. M. and Richardson, H. E. (2015) 'The Scribble–Dlg–Lgl Module in Cell Polarity Regulation.', in Ebnet K (ed.) *Cell Polarity 1: Biological Role and Basic Mechanisms*. Springer Press.

Huveneers, S. and Danen, E. H. J. (2009) 'Adhesion signaling - Crosstalk between integrins, Src and Rho', *Journal of Cell Science*. doi: 10.1242/jcs.039446.

Hwang, Y. S. *et al.* (2013) 'The Smurf ubiquitin ligases regulate tissue separation via antagonistic interactions with ephrinB1', *Genes and Development*. doi: 10.1101/gad.208355.112.

Ikuta, J. *et al.* (2007) 'Fasciculation and elongation protein zeta-1 (FEZ1) participates in the polarisation of hippocampal neuron by controlling the mitochondrial motility', *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2006.11.142.

Imamura, N. *et al.* (2013) 'Localisation of aPKC Lambda/Iota and its interacting protein, Lgl2, Is significantly associated with lung adenocarcinoma progression', *Tokai Journal of Experimental and Clinical Medicine*.

Johannes, B. and Preiss, A. (2002) 'Wing vein formation in Drosophila melanogaster: Hairless is involved in the cross-talk between Notch and EGF signaling pathways', *Mechanisms of Development*. doi: 10.1016/S0925-4773(02)00083-7.

Kaabeche, K. *et al.* (2005) 'Cbl-mediated ubiquitination of α5 integrin subunit mediates fibronectin-dependent osteoblast detachment and apoptosis induced by FGFR2 activation', *Journal of Cell Science*. doi: 10.1242/jcs.01679.

Kamitani, T. *et al.* (2001) 'Targeting of NEDD8 and Its Conjugates for Proteasomal Degradation by NUB1', 276(49), pp. 46655–46660. doi: 10.1074/jbc.M108636200.

Kamura, T. *et al.* (1998) 'The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families', *Genes Dev*, 12. doi: 10.1101/gad.12.24.3872.

Kamura, T. *et al.* (2004) 'VHL-box and SOCS-box domains determine binding specificity for Cul2 – Rbx1 and Cul5 – Rbx2 modules of ubiquitin ligases', *Genes & Development*, 2, pp. 3055–3065. doi: 10.1101/gad.1252404.

Kemphues, K. J. *et al.* (1986) 'Two loci required for cytoplasmic organisation in early embryos of Caenorhabditis elegans', *Developmental Biology*, 113(2), pp. 449–460. doi: http://dx.doi.org/10.1016/0012-1606(86)90180-6.

Kemphues, K. J. *et al.* (1988) 'Identification of genes required for cytoplasmic localisation in early C. elegans embryos', *Cell*. doi: 10.1016/S0092-8674(88)80024-2.

Kim, D. I. *et al.* (2016) 'An improved smaller biotin ligase for BioID proximity labeling', *Molecular Biology of the Cell*. doi: 10.1091/mbc.E15-12-0844.

Kitagawa, M. *et al.* (1999) 'An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β-catenin', *EMBO Journal*. doi: 10.1093/emboj/18.9.2401.

Kok, F. O. *et al.* (2015) 'Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish', *Developmental Cell*. doi: 10.1016/j.devcel.2014.11.018.

Komander, D. and Rape, M. (2012) 'The Ubiquitin Code', *Annual Review of Biochemistry*, 81(1), pp. 203–229. doi: 10.1146/annurev-biochem-060310-170328.

Kosuga, S. *et al.* (2005) 'GSK-3β directly phosphorylates and activates MARK2/PAR-1', *Journal of Biological Chemistry*, 280(52), pp. 42715–42722. doi: 10.1074/jbc.M507941200.

Li, S. *et al.* (2003) 'The role of laminin in embryonic cell polarisation and tissue organisation', *Developmental Cell*. doi: 10.1016/S1534-5807(03)00128-X.

Li, W. *et al.* (2008) 'Genome-Wide and Functional Annotation of Human E3 Ubiquitin Ligases Identifies MULAN, a Mitochondrial E3 that Regulates the Organelle's Dynamics and Signaling', *PLoS One*. Public Library of Science, 3(1), p. e1487. doi: 10.1371/journal.pone.0001487.

Liu, J., Vasudevan, S. and Kipreos, E. T. (2004) 'CUL-2 and ZYG-11 promote meiotic anaphase II and the proper placement of the anterior-posterior axis in C. elegans', *Development*. 2004/06/25, 131(15), pp. 3513–3525. doi: 10.1242/dev.01245.

Livak, K. J. and Schmittgen, T. D. (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method', *Methods*. doi: 10.1006/meth.2001.1262.

Ma, Z. *et al.* (2019) 'PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components', *Nature*. doi: 10.1038/s41586-019-1057-y.

Macara, I. G. and McCaffrey, L. (2013) 'Cell polarity in morphogenesis and metastasis', *Philos Trans R Soc Lond B Biol Sci*. 2013/09/26, 368(1629), p. 20130012. doi: 10.1098/rstb.2013.0012.

Malik, S. *et al.* (2016) 'BRCA2 mediates centrosome cohesion via an interaction with cytoplasmic dynein', *Cell Cycle*. doi: 10.1080/15384101.2016.1195531.

Manzano-López, J. and Monje-Casas, F. (2020) 'Asymmetric cell division and replicative aging: a new perspective from the spindle poles', *Current Genetics*. doi: 10.1007/s00294-020-01074-y.

Martin-Belmonte, F. and Perez-Moreno, M. (2012) 'Epithelial cell polarity, stem cells and cancer', *Nature Reviews Cancer*. doi: 10.1038/nrc3169.

McCaffrey, L. M. *et al.* (2012) 'Loss of the Par3 Polarity Protein Promotes Breast Tumorigenesis and Metastasis', *Cancer Cell*. doi: 10.1016/j.ccr.2012.10.003.

Mellacheruvu, D. *et al.* (2013) 'The CRAPome: A contaminant repository for affinity purification-mass spectrometry data', *Nature Methods*. doi: 10.1038/nmeth.2557.

Metzger, M. B., Hristova, V. A. and Weissman, A. M. (2012) 'HECT and RING finger families of E3 ubiquitin ligases at a glance', *Journal of Cell Science*. Company of Biologists, 125(3), pp. 531–537. doi: 10.1242/jcs.091777.

Milgrom-Hoffman, M. and Humbert, P. O. (2018) 'Regulation of cellular and PCP signalling by the Scribble polarity module', *Seminars in Cell and Developmental Biology*. doi: 10.1016/j.semcdb.2017.11.021.

Motegi, F. and Seydoux, G. (2013) 'The PAR network: Redundancy and robustness in a symmetry-breaking system', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1629). doi: 10.1098/rstb.2013.0010.

Munger, K. *et al.* (1989) 'Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product', *EMBO J*, 8.

Myllymäki, S. M., Teräväinen, T. P. and Manninen, A. (2011) 'Two distinct integrin-mediated mechanisms contribute to apical lumen formation in Epithelial cells', *PLoS ONE*. doi: 10.1371/journal.pone.0019453.

Nakagawa, S. and Huibregtse, J. M. (2000) 'Human Scribble (Vartul) Is Targeted for Ubiquitin-Mediated Degradation by the High-Risk Papillomavirus E6 Proteins and the E6AP Ubiquitin-Protein Ligase', *Molecular and Cellular Biology*. doi: 10.1128/mcb.20.21.8244-8253.2000.

Nance, J. and Zallen, J. A. (2011) 'Elaborating polarity: PAR proteins and the cytoskeleton', *Development*, 138(5), pp. 799–809. doi: 10.1242/dev.053538.

Narimatsu, M. *et al.* (2009) 'Regulation of Planar Cell Polarity by Smurf Ubiquitin Ligases', *Cell*. doi: 10.1016/j.cell.2009.02.025.

Nejsum, L. N. and Nelson, W. J. (2007) 'A molecular mechanism directly linking E-cadherin adhesion to initiation of epithelial cell surface polarity', *Journal of Cell Biology*. doi: 10.1083/jcb.200705094.

Neumann, C. J. and Cohen, S. M. (1998) 'Boundary formation in Drosophila wing: Notch activity attenuated by the POU protein Nubbin', *Science*. doi: 10.1126/science.281.5375.409.

O'Brien, L. E. *et al.* (2001) 'Rac1 orientates epithelial apical polarity through effects on basolateral laminin assembly', *Nature Cell Biology*. doi: 10.1038/ncb0901-831.

Ozdamar, B. (2005) 'Regulation of the Polarity Protein Par6 by TGF Receptors Controls Epithelial Cell Plasticity', *Science*, 307(5715), pp. 1603–1609. doi: 10.1126/science.1105718.

Pearson, H. B. *et al.* (2011) 'SCRIB expression is deregulated in human prostate cancer, and its deficiency in mice promotes prostate neoplasia', *J Clin Invest*. 2011/10/04, 121(11), pp. 4257–4267. doi: 10.1172/jci58509.

Peña-Rangel, M. T., Rodriguez, I. and Riesgo-Escovar, J. R. (2002) 'A misexpression study examining dorsal thorax formation in Drosophila melanogaster', *Genetics*, 160(3), pp. 1035–1050.

Perez-Mockus, G. *et al.* (2017) 'Neuralised regulates Crumbs endocytosis and epithelium morphogenesis via specific Stardust isoforms', *Journal of Cell Biology*. doi: 10.1083/jcb.201611196.

Phua, D. C. Y., Humbert, P. O. and Hunziker, W. (2009) 'Vimentin Regulates Scribble Activity by Protecting It from Proteasomal Degradation', 20, pp. 2841–2855. doi: 10.1091/mbc.E08.

Pujana, M. A. *et al.* (2007) 'Network modeling links breast cancer susceptibility and centrosome dysfunction', *Nature Genetics*. doi: 10.1038/ng.2007.2.

Qin, Y. *et al.* (2005) 'The mammalian Scribble polarity protein regulates epithelial cell adhesion and migration through E-cadherin', *J Cell Biol*. 2005/12/14, 171(6), pp. 1061–1071. doi: 10.1083/jcb.200506094.

R Core team (2015) 'R Core Team', *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing , Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.*

Raman, R., Savio, C. and Sonawane, M. (2018) 'Polarised Organisation of the Cytoskeleton: Regulation by Cell Polarity Proteins', *Journal of Molecular Biology*. Elsevier Ltd, p. #pagerange#. doi: 10.1016/j.jmb.2018.06.028.

Reimann, Julie D.R. *et al.* (2001) 'Emi1 is a mitotic regulator that interacts with Cdc20 and inhibits the anaphase promoting complex', *Cell*, 105(5), pp. 645–655. doi: 10.1016/S0092-8674(01)00361-0.

Reimann, J. D.R. *et al.* (2001) 'Emi1 regulates the anaphase-promoting complex by a different mechanism than Mad2 proteins', *Genes and Development*, 15(24), pp. 3278–3285. doi: 10.1101/gad.945701.

Richet, E. *et al.* (2012) 'NUB1 modulation of GSK3β reduces tau aggregation', *Human Molecular Genetics*, 21(24), pp. 5254–5267. doi: 10.1093/hmg/dds376.

Ritchie, M. *et al.* (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies | Nucleic Acids Research | Oxford Academic', *limma powers differential expression analyses for RNA-sequencing and microarray studies*.

Ritchie, M. E. *et al.* (2015) 'Limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*. doi: 10.1093/nar/gkv007.

Rodriguez-Boulan, E. and Macara, I. G. (2014) 'Organisation and execution of the epithelial polarity programme', *Nature Reviews Molecular Cell Biology*, 15(4), pp. 225–242. doi: 10.1038/nrm3775.

Roth, S. and Lynch, J. A. (2013) 'Symmetry Breaking During Drosophila Oogenesis', *Oogenesis*, pp. 1–353. doi: 10.1007/978-0-85729-826-3.

Roy, R., Chun, J. and Powell, S. N. (2012) 'BRCA1 and BRCA2: Different roles in a common pathway of genome protection', *Nature Reviews Cancer*. doi: 10.1038/nrc3181.

Royer, C. and Lu, X. (2011) 'Epithelial cell polarity: a major gatekeeper against cancer?', *Cell Death and Differentiation*. Nature Publishing Group, 18(9), pp. 1470–1477. doi: 10.1038/cdd.2011.60.

Rueden, C. T. *et al.* (2017) 'ImageJ2: ImageJ for the next generation of scientific image data', *BMC Bioinformatics*. doi: 10.1186/s12859-017-1934-z.

Saito, Y., Desai, R. R. and Muthuswamy, S. K. (2018) 'Reinterpreting polarity and cancer: The changing landscape from tumor suppression to tumor promotion', *Biochimica et Biophysica Acta - Reviews on Cancer*. doi: 10.1016/j.bbcan.2017.12.001.

Santoro, A. *et al.* (2016) 'Molecular mechanisms of asymmetric divisions in mammary stem cells', *EMBO reports*. doi: 10.15252/embr.201643021.

Sarikas, A., Hartmann, T. and Pan, Z. Q. (2011) 'The cullin protein family', *Genome Biology*. doi: 10.1186/gb-2011-12-4-220.

Savagner, P. (2015) 'Epithelial-mesenchymal transitions: From cell plasticity to concept elasticity', in *Current Topics in Developmental Biology*. doi: 10.1016/bs.ctdb.2014.11.021.

Scheffner, M., Nuber, U. and Huibregtse, J. M. (1995) 'Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade', *Nature*, 373(6509), pp. 81–83. doi: 10.1038/373081a0.

Schimanski, C. C. *et al.* (2005) 'Reduced expression of Hugl-1, the human homologue of Drosophila tumour suppressor gene lgl, contributes to progression of colorectal cancer', *Oncogene*. doi: 10.1038/sj.onc.1208520.

Scoumanne, A. *et al.* (2011) 'The cyclin-dependent kinase inhibitor p21 is regulated by RNA-binding protein PCBP4 via mRNA stability', *Nucleic Acids Research*. doi: 10.1093/nar/gkq778.

Shapiro, I. M. *et al.* (2011) 'An emt-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype', *PLoS Genetics*. doi: 10.1371/journal.pgen.1002218.

Sherr, C. J. and Roberts, J. M. (1995) 'Inhibitors of mammalian G1 cyclin-dependent kinases', *Genes Dev*. 1995/05/15, 9(10), pp. 1149–1163.

Skwarek, L. C. *et al.* (2014) 'The F-box protein Slmb restricts the activity of aPKC to polarize epithelial cells', *Development (Cambridge)*. doi: 10.1242/dev.109694.

Smith, L. K. *et al.* (2016) 'A Phenotypic High-Content Screening Assay to Identify Regulators of

Membrane Protein Localisation', *Assay Drug Dev Technol*. 2016/09/24, 14(8), pp. 478–488. doi: 10.1089/adt.2016.733.

Sonneville, R. and Gonczy, P. (2004) 'Zyg-11 and cul-2 regulate progression through meiosis II and polarity establishment in C. elegans', *Development*. 2004/06/25, 131(15), pp. 3527–3543. doi: 10.1242/dev.01244.

Stamos, J. L. and Weis, W. I. (2013) 'The β-catenin destruction complex', *Cold Spring Harbor Perspectives in Biology*. doi: 10.1101/cshperspect.a007898.

Stebbins, C. E. *et al.* (1999) 'Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function', *Science*. 1999/04/16, 284(5413), pp. 455–461. doi: 10.1126/science.284.5413.455.

Stephens, R. *et al.* (2018) 'The Scribble Cell Polarity Module in the Regulation of Cell Signaling in Tissue Development and Tumorigenesis', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2018.01.011.

Sternlicht, H. *et al.* (1993) 'The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin in vivo', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.90.20.9422.

Sugiarto, S. *et al.* (2011) 'Asymmetry-defective oligodendrocyte progenitors are glioma precursors', *Cancer Cell*. doi: 10.1016/j.ccr.2011.08.011.

Swan, A. and Schüpbach, T. (2007) 'The Cdc20 (Fzy)/Cdh1-related protein, Cort, cooperates with Fzy in cyclin destruction and anaphase progression in meiosis I and II in Drosophila', *Development*. doi: 10.1242/dev.02784.

Takizawa, C. G. and Morgan, D. O. (2000) 'Control of mitosis by changes in the subcellular location of cyclin-B1-Cdk1 and Cdc25C', *Curr Opin Cell Biol*. 2000/11/07, 12(6), pp. 658–665. doi: 10.1016/S0955-0674(00)00149-6.

Thiery, J. P. *et al.* (2009) 'Epithelial-Mesenchymal Transitions in Development and Disease', *Cell*. doi: 10.1016/j.cell.2009.11.007.

Thomas, M. *et al.* (2008) 'Human papillomaviruses, cervical cancer and cell polarity', *Oncogene*. doi: 10.1038/onc.2008.351.

Timms, Richard T *et al.* (2019) 'A glycine-specific N-degron pathway mediates the quality control

of protein <em>N</em>-myristoylation', *Science*, 365(6448), p. eaaw4912. doi: 10.1126/science.aaw4912.

Timms, Richard T. *et al.* (2019) 'A glycine-specific N-degron pathway mediates the quality control of protein *N* -myristoylation', *Science*, 365(6448), p. eaaw4912. doi: 10.1126/science.aaw4912.

Tseng, A. S. K. and Hariharan, I. K. (2002) 'An overexpression screen in Drosophila for genes that restrict growth or cell-cycle progression in the developing eye', *Genetics*, 162(1), pp. 229–243.

Vallin, J. and Grantham, J. (2019) 'The role of the molecular chaperone CCT in protein folding and mediation of cytoskeleton-associated processes: implications for cancer cell biology', *Cell Stress and Chaperones*. doi: 10.1007/s12192-018-0949-3.

Varnaitė, R. and MacNeill, S. A. (2016) 'Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID', *Proteomics*, 16(19), pp. 2503–2518. doi: 10.1002/pmic.201600123.

Vasudevan, S., Starostina, N. G. and Kipreos, E. T. (2007) 'The Caenorhabditis elegans cell-cycle regulator ZYG-11 defines a conserved family of CUL-2 complex components', *EMBO Rep*, 8(3), pp. 279–286. doi: 10.1038/sj.embor.7400895.

Venkei, Z. G. and Yamashita, Y. M. (2018) 'Emerging mechanisms of asymmetric stem cell division', *Journal of Cell Biology*. doi: 10.1083/jcb.201807037.

Verschuren, E. W. *et al.* (2007) 'Loss of Emi1-Dependent Anaphase-Promoting Complex/Cyclosome Inhibition Deregulates E2F Target Expression and Elicits DNA Damage-Induced Senescence', *Molecular and Cellular Biology*, 27(22), pp. 7955–7965. doi: 10.1128/mcb.00908-07.

Wald, J. H. *et al.* (2017) 'Suppression of planar cell polarity signaling and migration in glioblastoma by Nrdp1-mediated Dvl polyubiquitination', *Oncogene*. doi: 10.1038/onc.2017.126.

Wallenfang, M. R. and Seydoux, G. (2000) 'Polarisation of the anterior-posterior axis of C. elegans is a microtubule-directed process', *Nature*. Macmillian Magazines Ltd., 408(6808), pp. 89–92. Available at: http://dx.doi.org/10.1038/35040562.

Wang, A. Z., Ojakian, G. K. and Nelson, W. J. (1990) 'Steps in the morphogenesis of a polarised epithelium: II. Disassembly and assembly of plasma membrane domains during reversal of epithelial cell polarity in multicellular epithelial (MDCK) cysts', *Journal of Cell Science*.

Wang, H.-R. (2003) ‘Regulation of Cell Polarity and Protrusion Formation by Targeting RhoA for Degradation’, *Science*, 302(5651), pp. 1775–1779. doi: 10.1126/science.1090772.

Wang, H. *et al.* (2006) ‘Degradation of RhoA by Smurf1 Ubiquitin Ligase’, in *Methods in Enzymology*, pp. 437–447. doi: 10.1016/S0076-6879(06)06032-0.

Wang, S. *et al.* (2016) ‘Atlas on substrate recognition subunits of CRL2 E3 ligases’, *Oncotarget*. 2016/04/24. doi: 10.18632/oncotarget.8732.

Wang, X. *et al.* (2016) ‘*ZYG11A* serves as an oncogene in non-small cell lung cancer and influences *CCNE1* expression’, *Oncotarget*. 2016/01/16, 7(7), pp. 8029–8042. doi: 10.18632/oncotarget.6904.

Wang, Z. *et al.* (2014) ‘Roles of F-box proteins in cancer’, *Nat Rev Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 14(4), pp. 233–247. doi: 10.1038/nrc3700http://www.nature.com/nrc/journal/v14/n4/abs/nrc3700.html#supplementary-information.

Watanabe, Y., Khodosevich, K. and Monyer, H. (2014) ‘Dendrite Development Regulated by the Schizophrenia-Associated Gene FEZ1 Involves the Ubiquitin Proteasome System’, *Cell Reports*. The Authors, 7(2), pp. 552–564. doi: 10.1016/j.celrep.2014.03.022.

Weber, U. and Mlodzik, M. (2017) ‘APC/CFzr/Cdh1-Dependent Regulation of Planar Cell Polarity Establishment via Nek2 Kinase Acting on Dishevelled’, *Developmental Cell*. doi: 10.1016/j.devcel.2016.12.006.

Wehman, A. M., Staub, W. and Baier, H. (2007) ‘The anaphase-promoting complex is required in both dividing and quiescent cells during zebrafish development’, *Developmental Biology*, 303(1), pp. 144–156. doi: 10.1016/j.ydbio.2006.10.043.

Wei, X. *et al.* (2017) ‘Smurf1 inhibits integrin activation by controlling Kindlin-2 ubiquitination and degradation’, *Journal of Cell Biology*, 216(5), pp. 1455–1471. doi: 10.1083/jcb.201609073.

White, E. a *et al.* (2012) ‘Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses’, *Pnas*, 109(5), pp. E260–E267. doi: 10.1073/pnas.1116776109.

Winston, J. T. *et al.* (1999) ‘The SCF(β-TRCP)-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα and β-catenin and stimulates IκBα ubiquitination in vitro’, *Genes and Development*. doi: 10.1101/gad.13.3.270.

Wood, W. B. *et al.* (1980) ‘Parental effects and phenotypic characterisation of mutations that affect early development in Caenorhabditis elegans’, *Developmental Biology*, 74(2), pp. 446–469. doi: http://dx.doi.org/10.1016/0012-1606(80)90445-5.

World Health Organisation (2017) ‘Global Health Observatory’. World Health Organisation. Available at: http://www.who.int/cancer/en/.

Xue, B. *et al.* (2013) ‘Loss of Par3 promotes breast cancer metastasis by compromising cell-cell cohesion’, *Nature Cell Biology*. doi: 10.1038/ncb2663.

Yang, J.-Y. *et al.* (2006) ‘MDM2 Promotes Cell Motility and Invasiveness by Regulating E-Cadherin Degradation’, *Molecular and Cellular Biology*. doi: 10.1128/mcb.00172-06.

Yang, W. L., Zhang, X. and Lin, H. K. (2010) ‘Emerging role of Lys-63 ubiquitination in protein kinase and phosphatase activation and cancer development’, *Oncogene*. doi: 10.1038/onc.2010.190.

Ye, X. *et al.* (2004) ‘Recognition of phosphodegron motifs in human cyclin E by the SCF Fbw7 ubiquitin ligase’, *Journal of Biological Chemistry*. doi: 10.1074/jbc.M409226200.

Yonemura, S. *et al.* (1995) ‘Cell-to-cell adherens junction formation and actin filament organisation: Similarities and differences between non-polarised fibroblasts and polarised epithelial cells’, *Journal of Cell Science*, 108(1), pp. 127–142.

Yu, W. *et al.* (2005) ‘β1-integrin orients epithelial polarity via Rac1 and laminin’, *Molecular Biology of the Cell*. doi: 10.1091/mbc.E04-05-0435.

Zhang, Y., Wang, H. R. and Wrana, J. L. (2004) ‘Smurf1: A link between cell polarity and ubiquitination’, *Cell Cycle*, 3(4), pp. 389–390. doi: 10.4161/cc.3.4.771.

Zheng, N. and Shabek, N. (2017) ‘Ubiquitin Ligases: Structure, Function, and Regulation’, *Annual Review of Biochemistry*. doi: 10.1146/annurev-biochem-060815-014922.