AN RNAi-MEDIATED GENETIC SCREEN IDENTIFIES GENES THAT PROMOTE TUMOUR PROGRESSION IN A LIVING EPITHELIUM

Thesis submitted for the degree of Doctor of Philosophy

Zoe Elizabeth Cornhill, BSc (Hons)

2016

The University of Nottingham

School of Life Sciences

DECLARATION

I confirm that all the work in this thesis is my own and I take full responsibility for it. Images from other sources or work that has been generated through collaborative projects have been otherwise acknowledged in the text.

Signed:

Date:

ABSTRACT

The complex process by which cancer cells invade local tissue and metastasise is responsible for approximately 90% of cancer related deaths. The cell biological events that underlie this transition to malignancy are driven by invariable alterations within the genome, however relatively little is known about the genetic determinants involved. If identified, novel genes which perturb the rate of tumour progression could become potential targets for future therapeutic intervention.

Using a novel in vivo system, it is possible to characterise the behaviour of transformed cells during the early stages of tumour development and follow these cells in real time, thus improving our understanding of the critical events that initiate cell proliferation, tumour cell invasion and metastasis. Using *Drosophila* as a model organism it is possible to generate neoplastic tumours within the dorsal thorax whereby clones of transformed cells are homozygous mutant for a specific tumour suppressor gene. By specifically labelling these transformed cells with GFP, their behaviour can be observed in high temporal and spatial resolution within the living epithelium. RNAi technology can also be employed to simultaneously knock-down expression of an additional gene specifically within the mutant tissue. This forms the basis of a large-scale screen for novel genes that may promote tumour progression in this epithelium.

The screen is now almost complete and so far we have screened through almost 500 genes, the majority of which have previously been implicated in cancer but remain uncharacterised. We have observed a wide range of phenotypes, with genes affecting cell proliferation, invasion, cell shape, actin organisation, junction integrity and epithelial multilayering. By setting 'thresholds' for particular phenotypes 'hits' have been identified which drastically enhance tumour progression, and these genes are in the process of being fully characterised to further our understanding of their role in tumour progression.

ACKNOWLEDGEMENTS

In loving memory of

my Nan and Grandad

(Gillian and Alan Brunsden)

CONTENTS

FIGURES AND TABLES

ABBREVIATIONS

AJ   Adherens junction

ALT   Alternative telomere lengthening

APC   Adenomatous polyposis coli

APC/C   Anaphase-promoting complex/cyclosome

APF   After puparium formation

aPKC   Atypical protein kinase C

Arp2/3   Actin-related protein 2/3

Bax   Bcl-2 associated X protein

Bcl2   B-cell lymphoma 2

BM   Basement membrane

BRCA1   Breast cancer 1

BSA   Bovine serum albumin

CD8/4   Cluster of differentiation 8/4

CDK   Cyclin-dependent kinases

CIN   Chromosomal instability

CIP   Calf intestinal, alkaline phosphatase

CRISPR   Clustered regularly interspaced short palindromic repeats

CO_2 Carbon Dioxide

DEPC   Diethylpyrocarbonate

Dlg   Discs Large

DMEM   Dulbeccos modified eagle medium

DMSO   Dimethylsulphide

DNA   Deoxyribonucleic Acid

DNA-PK	DNA dependent protein kinase
ECM	Extracellular matrix
EGF	Epidermal growth facto
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
EPP	Epithelial polarity programme
FADD	Fas associated protein with death domain
FBS	Foetal bovine serum
FCS	Foetal calf serum
Flp	Flippase
FRT	Flippase recognition target
G1	gap 1
G2	gap 2
GEF	Guanine-nucleotide exchange factor
GFP	Green Fluorescent Protein
GLUT1	Glucose transporter 1
Grb2	Growth factor receptor bound protein-2
GSKβ	Glycogen synthase kinase 3
HBSS	Hanks balanced salt solution
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia inducible factor
IL6	Interleukin 6
KD	Knock-down
LB	Luna broth
Lgl	Lethal giant larvae
Mad2	Mitotic arrest deficient 2

MAPK	Mitogen-activated protein kinases
MARCM	Mosaic analysis with a repressible cell marker
MDCK	Madin Darby canine kidney cells
Mdm2	Mouse double minute 2 homolog
MET	Mesenchymal-epithelial transition
MMP	Matrix metalloprotease
MPT	Mitochondrial permeability transition pores
M	Mitosis
NF-κB	Nuclear factor kappa B
OE	Over-expression
PBS	Phosphate buffered saline
PBS-T	PBS-tween
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PI3K	Phosphoinositide 3 kinase
PRC2	Polycomb repressive complex 2
Rb	Retinoblastoma
RNA	Ribonucleic acid
RNAi	RNA interference
RT	Room temperature
RT-PCR	Real-Time PCR
Scrib	Scribble
siRNA	small interfering RNA
SH2	Src-homology 2
SJ	Septate Junction
SOS	Son of Sevenless

TAM	Tumour associated macrophage
TGFβ	Transforming growth factor beta
TGFβR	Transforming growth factor beta receptor
TIAM	T-cell lymphoma invasion and metastasis 1
TJ	Tight junction
TNFR	Tumour necrosis factor receptor
UAS	Upstream activating sequence
VEGF	Vascular endothelial growth factor
ZO1	Zonula occludens 1

CHAPTER 1

Introduction

Chapter 1 Introduction

Recent statistics suggest that 1 in 2 people will be affected by cancer during the course of their lifetime, with someone being diagnosed with the disease every 2 minutes ("Cancer Research UK," 2016) . There are more than 200 different types of cancer arising from a multitude of various cell types within the body which acquire the ability to grow in an uncontrolled manner, forming tumour masses. As a result of extensive cancer research using both human and animal models, the diagnosis and treatment of the disease has improved significantly in the last 40 years ("Cancer Research UK," 2016). However at a biological level our knowledge of the genes involved and pathways implicated during tumour progression remain poorly understood.

1.1 Tumour Formation

The focus of my PhD research is on understanding the molecular basis of tumour progression, and more specifically invasion. Therefore in my introduction I will give a brief overview of the molecular basis of tumourigenesis and tumour progression, with a more detailed focus of cancer cell invasion.

The human genome is organised into 23 pairs of chromosomes consisting of large strands of DNA, found in the nucleus of all cells. Genes define the protein-coding segments of chromosomal DNA, and ultimately determine cell fate and maintain tissue homeostasis (Fraser & Bickmore, 2007). Cancer arises through a succession of invariable alterations in the genome, capable of driving the progressive transition of

cells to malignancy. Genetic alterations in 2 distinct classes of genes responsible for tightly regulating progression through the cell cycle can result in uncontrolled proliferation and cellular transformation (Bishop, 1996; Hanahan & Weinberg, 2000; Vogelstein & Kinzler, 2004). In normal cell cycle control, proto-oncogenes and tumour suppressor genes act at checkpoints to respectively stimulate or inhibit cellular growth, or promote differentiation. Dominant gain-of-function mutations in proto-oncogenes therefore drive continuous cell division, whilst the recessive inactivation of tumour suppressor genes prevents differentiation, thus leading to hyper-proliferation and a sustained growth advantage over neighbouring wild-type cells (Hanahan & Weinberg, 2000; Lodish, 2000). Tumour progression in humans is a multi-step process in which malignant growth is dictated by the manifestation of various modifications to cell physiology. As first described by Hanahan and Weinberg (2000), the 'hallmarks of cancer' propose that cancer cells must acquire at least six essential biological capabilities **(Fig. 1.1)** for the efficient conversion of normal cells into malignant derivatives. In 2011, Hanahan and Weinberg proposed an additional four essential hallmarks. Acquisition of these vital hallmarks ensures that incipient cancer cells are not eliminated through the cells natural defence mechanisms.
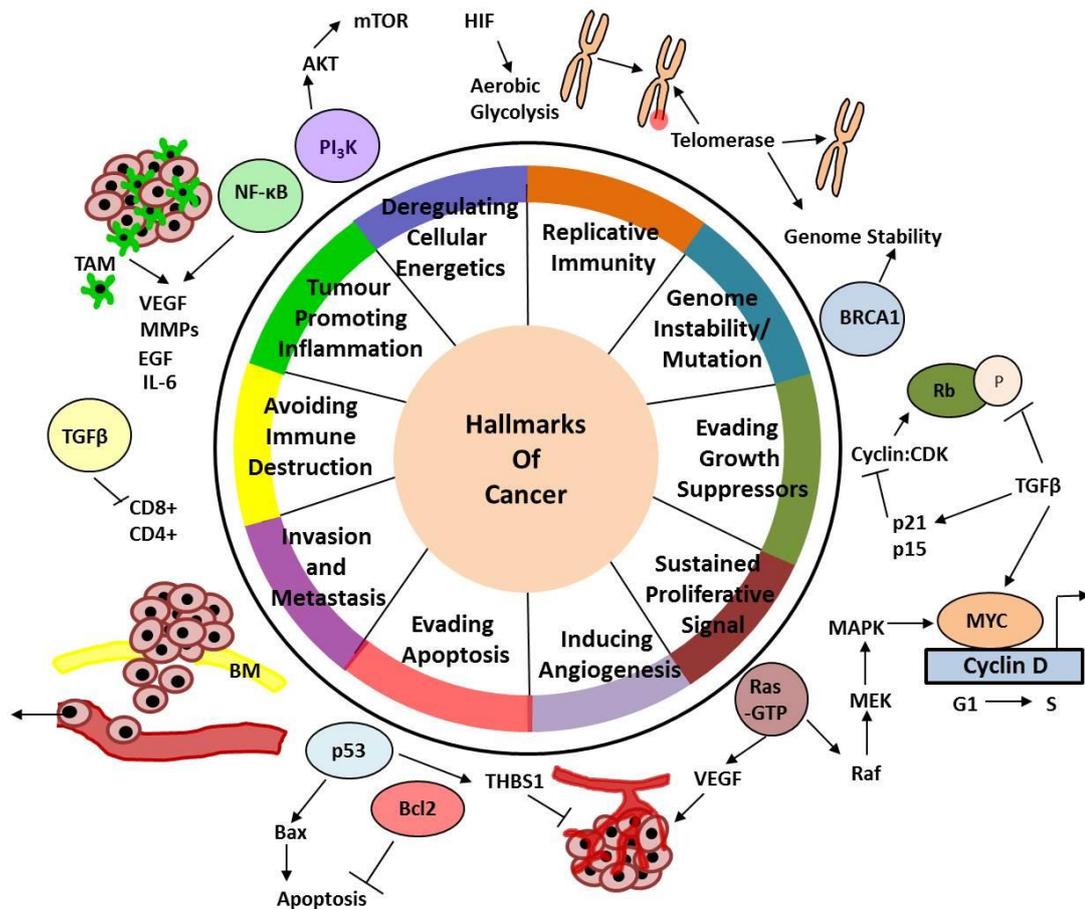
Figure 1.1 Hanahan and Weinberg's 'Hallmarks of Cancer'

The diagram illustrates the essential capabilities that cancer cells must acquire in order to form, survive, grow and progress into highly malignant tumours. The image also depicts some of the key regulators and pathways involved in tumour progression.

1.1.1 Mitogenic signalling

Mitogens convey signals to cells in order for them to progress through the cell cycle and proliferate in a controlled manner (Alberts, 2002). During the intricately regulated cell cycle chromosomal DNA is duplicated and precisely segregated into two daughter cells which are genetically identical **(Fig. 1.2)** (Morgan, 2006). In order to sustain continuous growth, cancer cells can adopt autocrine signalling whereby they are able to synthesise and respond to their own mitogenic growth factors

(Sporn, 1996). This generates a positive feedback loop which enables cancer cells to progress through the cell cycle and grow without extracellular stimulation. Additionally, tyrosine-kinase receptors which are usually responsible for conveying mitogen signals are often overexpressed in cancer cells, such that even with ambient levels of growth factors the cell is hyper-responsive and continues to proliferate (Hanahan & Weinberg, 2000).

One of the main regulators of cell cycle progression is a family of serine/threonine proteins called cyclin-dependent kinases (CDKs) (Malumbres & Barbacid, 2009). As their name suggests the kinase activity of these proteins is controlled by another family of proteins called cyclins. CDKs comprise two subunits: the catalytic core and cyclin-binding site. The catalytic subunit contains an ATP-binding pocket and activating T-loop motif. In its inactivated state the T-loop blocks the catalytic domain preventing substrate activation. However, when the appropriate cyclin binds to the CDK the T-loop becomes displaced which in turn exposes the catalytic site allowing CDK-cyclin complexes to actively regulate the cell cycle (Lim & Kaldis, 2013). The role of CDKs is to phosphorylate and activate specific proteins which are required for various phases in the cell cycle, which, until activated or synthesised prevents the cell from progressing through the cell cycle. While CDKs are present throughout the process of cell replication and division, temporal regulation is determined by cyclins which are synthesised and degraded at different points throughout the cell cycle. During early G1 phase, cells are mitogen-dependent, requiring extracellular growth factors for progression past the 'Restriction point' which defines entry into S phase. During G1 cyclin E must be synthesised in order for cells to be committed to S phase,

for this several biochemical reactions take place which are tightly regulated. The transcription factor E2F binds to the promoter of cyclin E to regulate its transcription. However, retinoblastoma (Rb) in turn determines the activity of E2F (Weinberg, 1995). In its hypophosphorylated state Rb binds to E2F which prevents downstream transcription of cyclin E. The G1-CDK complex, which comprises cyclin D with either CDK4 or CDK6, controls synthesis of cyclin E by regulating the phosphorylation state of Rb. When activated, the G1-CDK complex phosphorylates Rb causing it to dissociate from E2F, enabling the transcription of cyclin E. Once cyclin E is abundant in the cell it can bind to CDK2 forming the G1/S CDK complex and allowing cells to irreversibly pass through the 'Restriction point' (Dyson, 1998; Harbour & Dean, 2000; Ohtani, DeGregori, & Nevins, 1995). In this way CDK/cyclin complexes regulate a cells' progression through the cell cycle and are in turn tightly regulated.
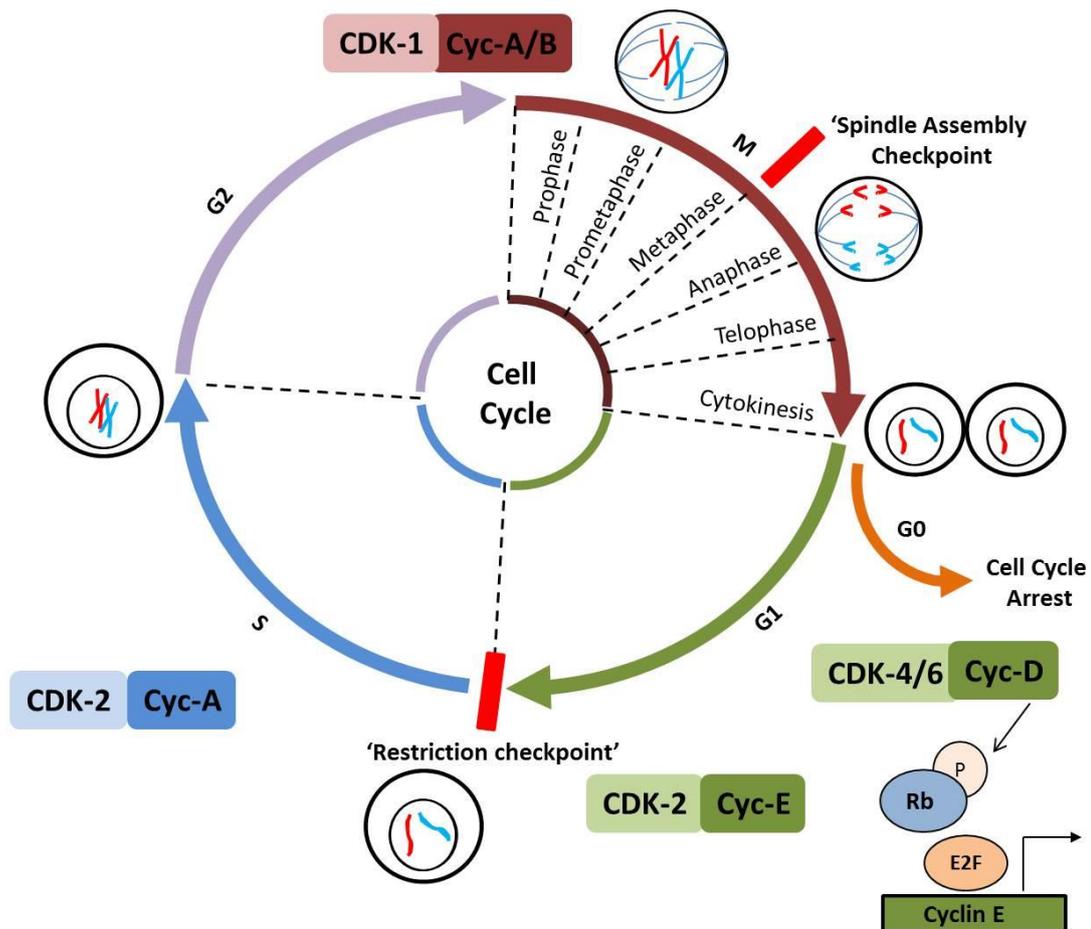
Figure 1.2 Cell Cycle

The diagram illustrates the eukaryotic cell cycle, including stages of mitosis, checkpoints and cell cycle regulators.

Extracellular mitogens can regulate the cell cycle in early G1 phase only; however they can also permit cells to enter the cell cycle from a quiescent state (G0), while anti-mitogenic signals can cause cell cycle arrest and entry into G0 (Alberts, 2002). Epidermal growth factor (EGF) signalling provides a good example of the complexity of signalling pathways and how extracellular mitogenic signals can influence cell cycle progression (Normanno et al., 2006). When EGF binds to cell surface EGF-receptors (EGFR) a cascade of biochemical reactions take place. Initially, the receptor dimerises which results in transphosphorylation of a cytoplasmic tyrosine residue on one of the

receptor subunits. Phosphorylated tyrosines specifically bind to proteins containing src-homology 2 (SH2) domains, such as the growth factor receptor bound protein 2 (Grb2) (Lowenstein et al., 1992). The guanine-nucleotide exchange factor (GEF) Son of Sevenless (SOS) is then recruited to the plasma membrane where it activates Ras (Margolis & Skolnik, 1994). Ras is a small GTPase with central roles in signalling cascades associated with the cell cycle, growth and survival (Coleman, Marshall, & Olson, 2004). Once activated, Ras activates the mitogen-activated-protein-kinases (MAPK) signalling pathway which leads to the transcription of various proteins, including Myc, a transcription factor which regulates the transcription of cyclin D (Zhang & Liu, 2002). Synthesis of cyclin D promotes progression from G1 to S, and therefore cell proliferation (Daksis, Lu, Facchini, Marhin, & Penn, 1994). This signalling pathway demonstrates the complexity of cell cycle regulation and the sheer number of different protein complexes involved.

Equally, as well as sustaining continuous mitogenic growth promoting signals, cancer cells must be able to avoid anti-proliferative signals, thus preventing them from entering a quiescent or post-mitotic state (Hanahan & Weinberg, 2000). Inhibitors of cell cycle transition act primarily to prevent G1 to S transition and typically converge on the Rb pathway. There are several ways that cancer cells elicit their insensitivity to growth signals via the Rb pathway. Firstly, loss of or mutations in TGF-βR, or loss of Smad4 (see below), block TGF-β signalling which in turn promotes the phosphorylation of Rb. This causes Rb to release E2F enabling the transcription of various proteins required for transition through the 'Restriction point' in the cell cycle. Cells are particularly sensitive to a defective Rb pathway and increased E2F

activity. High E2F activity drives expression of genes involved in the apoptotic program, including p14 ARF (an Mdm2 inhibitor) (Weinberg, 1995). Additionally, loss of TGF-β prevents inhibition of the proto-oncogene c-myc and upregulation of CDK inhibitors (Fynan & Reiss, 1993; Moses, Yang, & Pietenpol, 1990). Due to its direct effects on cell proliferation, de-regulation of Rb signalling is a major contributor to human cancer with Rb seen to be downregulated in approximately 80% of human cancers (Weinberg, 1995). Additionally cancer cells can avoid entering post-mitotic states, through the over-expression of Myc. Abundant Myc results in the formation of Myc:Max heterodimers. This dimerisation prevents Max forming its normal complex with Mad which functions to induce differentiation signals (Foley & Eisenman, 1999; Grinberg, Hu, & Kerppola, 2004; Nair & Burley, 2003). To summarise, by sustaining growth signals and evading anti-proliferative signals cancer cells are able to continuously proliferate in an uncontrolled manner.

As their name suggests CDK inhibitors negatively regulate the activity of CDKs. The two main classes of inhibitors are the Ink4 family (p16^{INK4a}, p15^{INK4b}, p18^{INK4c}, p19^{INK4d}) and Cip/Kip family (p21^{CIP1}, p27^{KIP1}, p57^{KIP2}) which inhibit cyclin-CDK complexes (Sherr & Roberts, 1999). Transforming growth factor β (TGF-β) signalling is a prime example of extracellular signalling which activates CDK inhibitors to block the transition from G1 to S phase. Similar to EGF signalling, TGF-β binds to the serine/threonine kinase transmembrane receptor TGF-β-receptor (TGFβR) which phosphorylates cytoplasmic Smad proteins (Heldin, Miyazono, & ten Dijke, 1997). Once phosphorylated Smads are able to translocate into the nucleus where they act as transcription factors. One of the proteins that is transcribed following Smad activation is the CDK inhibitor p21,

resulting in the cell cycle arrest (Pardali, Kowanetz, Heldin, & Moustakas, 2005). Cell cycle arrest can also be initiated when cells contain damaged DNA. Following DNA damage (including UV and ionising radiation) the transcription factor p53 becomes phosphorylated (Siliciano et al., 1997). Normally p53 is present in low levels and quickly turned over in the cell. This is due to its binding with the E3 ubiquitin protein ligase mouse double minute-2 (Mdm2) which sequesters in the transactivation domain of P53 (between Glutamate17 and Proline27) (Meek, 1998). Mdm2 binding results in the degradation of p53. However, under stress p53 becomes phosphorylated by DNA dependent protein kinase (DNA-PK) at sites within the transactivation domain, this prevents Mdm2 binding and thus the abnormal accumulation of p53 in the nucleus (Mayo, Turchi, & Berberich, 1997). Abundant p53 can transcriptionally activate p21 to initiate cell cycle arrest, and when sustained for a longer period of time p53 can activate the Bcl-2 associated protein (Bax) which induces the release of cytochrome c from the mitochondria which initiates the apoptotic cascade (Fridman & Lowe, 2003; Miyashita et al., 1994). **(Section 1.1.2).**

1.1.2 Apoptotic signalling

Programmed cell death (apoptosis) is a homeostatic energy dependent mechanism which is essential for tissue development and cell turnover (Alberts, 2002). It can also be initiated in times of stress, such as DNA damage, hypoxia and prolonged oncogene activation. For this reason, it is essential that cancer cells prevent activation of the apoptotic cascade **(Fig. 1.3)** if they are to survive (Hanahan & Weinberg, 2000). The apoptotic cascade is a complex succession of biochemical reactions resulting in cell shrinkage, pyknosis, cytosolic extrusion, nuclear fragmentation, cell budding and

phagocytosis (Wyllie, Kerr, & Currie, 1980). Both extrinsic and intrinsic signalling pathways can signal apoptotic machinery, both of which converge on the execution pathway and ultimately result in cell death. The extrinsic apoptotic pathway involves a host of proteins belonging to the tumour necrosis factor receptor (TNFR) family which comprises a group of death receptors with homologous 'death domains' including FasR and TNFR-1 (Ashkenazi & Dixit, 1999). Upon binding its ligand (Fas L), FasR recruits the cytoplasmic Fas associated protein with death domain (FADD) an adaptor protein which in turn cleaves pro-caspase 8 also known as a gatekeeper caspase. Caspases are a family of cysteine proteases existing as proenzymes, which, until cleaved remain inactive (Thornberry & Lazebnik, 1998). Unlike the extrinsic pathway, the intrinsic pathway mediates apoptotic signalling via mitochondria. A variety of signals induced by cell stress or lack of sufficient hormones and cytokines impinge on the mitochondria and as a result alter the mitochondrial membrane potential (Brenner & Mak, 2009). This causes the mitochondrial permeability transition pore (MPT) to open, releasing cytochrome c which is a key regulator in the apoptotic cascade. Once released cytochrome c forms the apoptosome, a complex consisting of Apaf-1 and pro-caspase 9. Formation of the apoptosome results in cleavage of caspase 9, another gatekeeper caspase, which in turn initiates the execution pathway (Cain, Bratton, & Cohen, 2002). The apoptotic pathway is stringently regulated by pro-apoptotic (Bcl-2, Bcl-XL, Bcl-W) and anti-apoptotic (Bax, Bak, Bid, Bim) proteins. These B-cell lymphoma 2 (Bcl-2) family proteins largely mediate apoptosis by regulating mitochondrial permeability (Chipuk, Moldoveanu, Llambi, Parsons, & Green, 2010). One of the most notable regulators of apoptosis is p53 which responds to stress signals to induce cell death by upregulating anti-

apoptotic genes including Bax to stimulate the release of cytochrome c. Functionally inactivated p53 is detected in more than 50% of human cancers, rendering cells insensitive to stress signals, and able to evade apoptosis (Levine, 1997). Additionally, overexpression of Bcl-2 promotes cancer cell survival via cooperative effects with c-myc and through the concerted inhibition of cytochrome c release (Vaux, Cory, & Adams, 1988). Anti-apoptotic proteins also include the inhibitor of apoptosis (IAP) family which includes XIAP, C-IAP1, c-IAP2, NAIP and survivin. These proteins preferentially bind to procaspase -3 and -7 which prevent cleavage and subsequent initiation of apoptosis. Moreover, these proteins are commonly overexpressed in human cancer to prolong the life of cancer cells (Tamm et al., 1998). Therefore, by preventing apoptosis cancer cells are able to continue proliferating, bypassing one of the body's natural defence programmes.
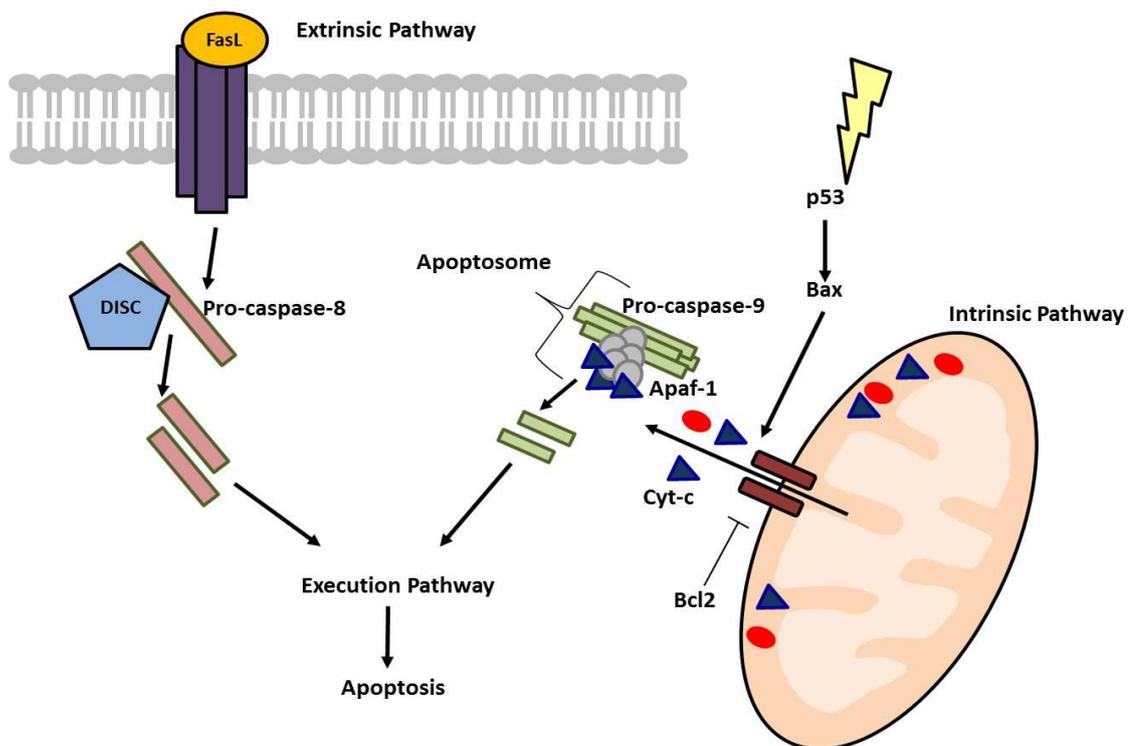
Figure 1.3 Apoptotic cascade

The extrinsic pathway is initiated by extracellular signals leading to cleavage of pro-caspase 8. Alternatively the intrinsic pathway responds to cell stress or other factors and impinges on the mitochondria. Release of cytochrome c ultimately leads to the cleavage of pro-caspase 9. Both caspase 8 and 9 cleave several execution caspases leading to cell death.

1.1.3   Replicative potential

In 1965 Hayflick demonstrated that in culture, normal cells exhibit a limited replicative potential, whereby human fibroblasts ceased growing over many rounds of cell replication. For normal tissue homeostasis this replicative potential is essential in preventing cancer in renewable tissues which undergo many rounds of replication (Shay & Wright, 2000). Telomeric DNA protectively caps the ends of chromosomes and consists of multiple 5' –TTAGGG- 3' sequence repeats (Moyzis et al., 1988). During each succession of the cell cycle 50-100bp are lost as a result of inefficient replication of the 3' end by DNA polymerase. In normal cells, single stranded G-rich tails left at the ends of telomeres can protect itself by generating 'T-loops' in which the telomere folds back on itself, this leads to the single stranded tail invading the double stranded telomeric DNA (Griffith et al., 1999). The 6 subunit shelterin complex then binds to T-loops and protects the DNA from being recognised and treated as 'damaged DNA' by the double –stranded-DNA repair machinery (de Lange, 2005). However, due to the successive loss of telomeric DNA following each round of cell division, eventually telomeric DNA becomes too short and the unprotected chromosomal ends fuse together in karyotypic disarray through breakage-fusion-bridge cycles, resulting in chromosomal instability, senescence and cell death (Counter et al., 1992). Most cancer calls are considered to be 'immortal' due to their

ability to replicate without restriction, and in most cases this is due to increased expression of telomerase. Telomerase, a reverse transcriptase, maintains telomeric DNA by adding hexanucleotides to the chromosomal ends, and in this way cells can continue to proliferate uncontrollably. In a small percentage of cancer cells, unlimited replication is achieved through an alternative telomere lengthening (ALT) mechanism, which involves the homologous recombination of adjacent telomeres (Bryan, Englezou, Gupta, Bacchetti, & Reddel, 1995; Cesare & Reddel, 2010).

1.1.4 Angiogenesis

Another key hallmark that cancer cells must overcome in their transition to malignancy is angiogenesis (Hanahan & Weinberg, 2000). Angiogenesis defines the process in which new blood vessels are formed in order to provide a tissue with essential nutrients and oxygen. In normal tissues, angiogenesis is tightly regulated by maintaining a balance between activating or inhibitory angiogenic signals (Bergers & Benjamin, 2003; Duffy, 1996). A prerequisite to neoplastic tissue growth is the ability of tumours to mediate the 'angiogenic switch', resulting in the upregulation of pro-angiogenic factors and down-regulation of anti-angiogenic factors (Hanahan & Folkman, 1996). For example, loss of p53 in many cancers results in decreased expression of the angiogenic inhibitor thrombospondin-1, while activated oncogenic Ras in tumours leads to the upregulation of vascular endothelial growth factor (VEGF) which promotes growth of neo-vasculature (Dameron, Volpert, Tainsky, & Bouck, 1994; Rak et al., 2000; Volpert, Dameron, & Bouck, 1997). As well as providing the growing tumour with essential nutrients, this novel vasculature also provides a direct

route for cancer cells to spread to other parts of the body (Zetter, 1998) **(Section 1.2).**

1.1.5 Genomic Instability

Hanahan and Weinberg (2011) later proposed additional hallmarks that must be overcome during tumour progression including cancer genomic instability. As discussed, tumour progression is defined by the accumulation of more than one genetic alteration and the concerted co-operation between up-regulated oncogenes and down-regulated tumour suppressor genes. Additionally, genome instability within cancer cells is generated through the down-regulation of so called 'caretaker genes' which detect DNA damage and activate appropriate repair machinery (Ciccia & Elledge, 2010; Kinzler & Vogelstein, 1997; Negrini, Gorgoulis, & Halazonetis, 2010). This renders cancer cells hypersensitive to mutagenic agents and unable to detect and repair genomic alterations. One of the most common forms of genomic instability: chromosomal instability (CIN) sees the rate of accumulated chromosomal alterations increase over time. A common example in human cancer involves mutations in the breast cancer (BRCA1) gene which regulates DNA homologous recombination repair. Mutations in this gene force cells to adopt non-homologous end joining repair machinery which is highly error prone, resulting in the accumulation of more mutations (Powell & Kachnic, 2003). Genomic instability is also a key factor in hereditary cancers, such that all cells within a patient present germline mutations for a gene – such as BRCA1 – making the patient more susceptible to additional mutations, where a single event could drive tumour progression (Deng, 2006; Narod et al., 1993). Additionally, as telomerase protects cancer cells from

limited replication, it is also considered a regulator of genome instability (Hanahan & Weinberg, 2011). **Figure 1.4** shows a karyotype taken from normal cells and cancer cells where genomic instability- aneuploidy and hybrid chromosomes can be easily visualised.
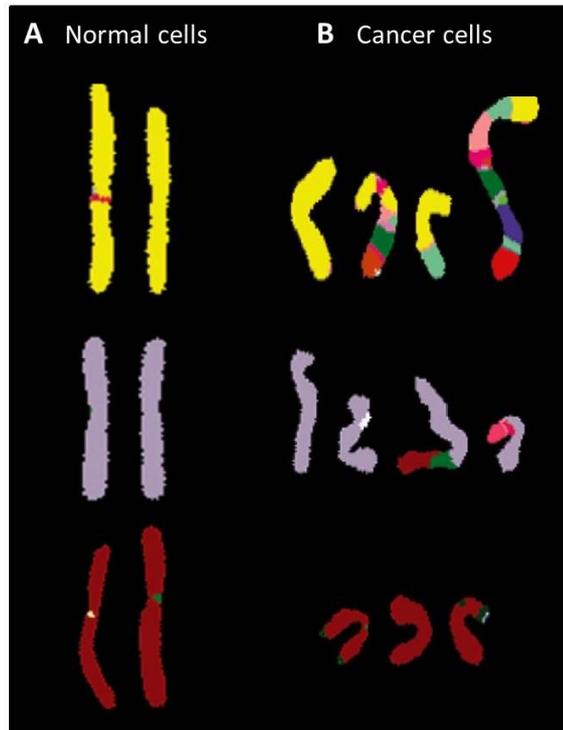
Figure 1.4 Chromosomes from a karyotype of normal and cancer cells
Comparison of a few chromosomes taken from a karyotype of **(A)** normal cells and **(B)** cancer cells. These images (taken from (Weinberg, 2014)) were generated using multi-coloured FISH which labels chromosomes with different fluorescent dyes. Cancer cells (B) are often aneuploidy with chromosomal aberrations, and chromosomal translocations (multi-coloured). Adapted from (Weinberg, 2014).

1.1.6 Tumour associated inflammation

The identification of immune cells within tumour lesions adds an additional layer of complexity to tumour progression and the tumour microenvironment. Tumour-associated inflammation and the innate immune response have been shown to enhance tumour progression by promoting many of the other hallmarks of cancer

(Dvorak, 1986). Chronic inflammation is a result of a variety of factors including infection and environmental toxins. Inflammation due to environmental factors such as smoking, obesity and alcohol therefore link these to higher risks of cancer. Examples of inflammatory induced tumours include bladder carcinomas which can be caused by chronic cystitis, Hepatocellular carcinoma caused by Hepatitis B/C and colorectal carcinoma which can be caused by inflammatory bowel disease (Weinberg, 2014).

Up to 50% of tumour masses can comprise immune cells including tumour-associated macrophages (TAMs). Immune cells within a tumour promote tumour progression through the secretion of mitogens, survival and pro-angiogenic factors as well as matrix metalloproteases (MMPs) which facilitate epithelial-mesenchymal transition (EMT) **(Section 1.2)** (DeNardo, Andreu, & Coussens, 2010; Grivennikov, Greten, & Karin, 2010). In many cancers the protein nuclear factor kappa B (NF-κB) is constitutively activated or activated by inflammatory cells secreted by TNF-α. This transcription factor orchestrates a tumour-promoting inflammatory response by activating cytokines such as interleukin 6 (IL6) (Hoesel & Schmid, 2013). Additionally cancer cells may present immune-evasion whereby regulators of immune surveillance are suppressed. As demonstrated in transgenic mice, tumours developed more readily in mice with deficiencies in CD8+ cytotoxic T lymphocytes and CD4+ helper cells compared to control immune-competent mice. In human cancers, secretion of immune-suppressors such as TGF-β eliminates CD8+ cells (M. Y. Kim et al., 2009; Teng, Swann, Koebel, Schreiber, & Smyth, 2008).

Additionally, the tumour microenvironment adds layers of complexity to the process of tumour progression (Hanahan & Weinberg, 2011). As mentioned, various other molecules have been shown to elicit tumour promoting effects and are recruited to the tumour environment, forming the tumour microenvironment. These non-transformed cells include members of the immune system, lymphatic system and surrounding stromal cells, all of which can aid tumour growth. The complex tumour microenvironment therefore presents a problem with regards to cancer therapy across patients (Junttila & de Sauvage, 2013; Singh & Settleman, 2010).

1.1.7 Tumour energy metabolism

One of the more recently defined hallmarks of cancer indicates that developing tumours are capable of reprogramming their energy metabolism, initiating a 'metabolic switch' (Hanahan & Weinberg, 2011; Hsu & Sabatini, 2008). Tumour cells favour the less efficient aerobic glycolysis over normal respiration in order to produce energy at a faster rate and generate a host of biosynthetic molecules for tumour growth. Glycolytic fuelling in cancer cells is typically associated with the upregulation of the glucose transporter (GLUT1) which encourages aerobic glycolysis. Oncogenes such as Ras and Myc, along with p53 have also been linked to increased glycolysis which can act pleiotropically along with hypoxic tumour conditions (DeBerardinis, Lum, Hatzivassiliou, & Thompson, 2008; Jones & Thompson, 2009). Additionally, upregulation of phosphionostide-3-kinase in cancer cells leads to increased hypoxia inducible factors (HIF) which promote aerobic glycolysis, fuelling growing tumours (Semenza, 2010).

1.2 Invasion and Metastasis

The final 'hallmark of cancer' depicts the progression of solid tumours into aggressive metastatic tumours (Hanahan & Weinberg, 2000). Despite the increasing trend in cancer survival rates, patients with locally advanced cancers have a reduced prognosis compared to those with primary tumours alone; with the formation of secondary metastases in distant anatomical sites being accountable for more than 90% of deaths in cancer patients (Gupta & Massague, 2006; Sporn, 1996). Sometimes certain primary tumours have a tendency to form secondary metastases in particular organs/tissues. For example, cancers originating in the ovaries tend to metastasise to the peritoneum (abdomen lining) whilst breast and colorectal cancers commonly spread to the liver (Hess et al., 2006). The invasion-metastasis cascade is an elaborate multi-step process comprising a succession of complex cell-biological changes which facilitate the dissemination of cancer cells from a primary tumour to distant sites (Valastyan & Weinberg, 2011).

Initial stages of the cascade include the local invasion and migration of cancer cells away from the primary tumour mass. Should migrating cancer cells encounter the vasculature, their subsequent intravasation into the blood and/or lymphatic vessels would allow their transportation and dissemination around the body. Surviving cancer cells extravasate at secondary sites to form small tumorous nodules, known as micrometastases, which have the potential to colonise into fatal secondary tumours (Fidler, 2003). However these micrometastases usually remain dormant for many years due to the inefficiency of colonisation and formation of macroscopic metastatic tumours (Aguirre-Ghiso, 2007). This metastatic cascade of events results in dramatic

changes in cell shape, cell polarity, cytoskeletal organisation, cell-cell junctions and focal contacts. The complex metastatic cascade can be broadly divided into three events: invasion, intravasation and extravasation (Hanahan & Weinberg, 2000).

1.2.1 Invasion and Epithelial – Mesenchymal Transition (EMT)

The vast majority of cancers originate from epithelial cells, which are specialised cells organised into adherent sheets. These epithelial sheets form a closed epithelium found lining most body organs, as well as the skin, and provide protection from the external environment. There are various types of epithelial cells, classified primarily by their site of origin, and cancers arising from these cells are known as carcinomas (Valastyan & Weinberg, 2011). These specialised cells are highly polarised with a distinct planar polarity and apicobasal polarity, which defines separate domains along the apical-basal axis. This polarity and organisation of epithelial sheets is maintained by various signalling networks involving polarity complexes, adhesion molecules and the actin cytoskeleton (Fristrom, 1988). The physiological function of epithelial cells is partly regulated by various adhesive junctions connecting neighbouring epithelial cells (Ebnet, 2008). In vertebrates, tight junctions (TJs) are found on the apical membrane, above the adherens junction (AJ) and are composed of transmembrane claudins which regulate the permeability of the cell and govern the passage of various molecules between adjacent cells (Anderson & Van Itallie, 2009). Instead of TJs, invertebrates possess septate junctions, which are functionally similar but molecularly distinct and are located below the AJ. AJs, located apically on the lateral membrane, are an essential feature of the epithelium in both vertebrates and invertebrates, defining the apical-basal axis and connecting neighbouring epithelial

cells **(Fig. 1.5)** (W. Meng & Takeichi, 2009). Desmosomes, which are composed of glycoproteins and intermediate filaments, are located beneath the AJs and function to maintain adhesion between cells (Alberts, 2002).

Mature AJs are formed by an E-cadherin-catenin complex that binds to cytoskeletal components along with various regulatory and signalling molecules (Gumbiner, 2005) **(Fig. 1.5)**. E-cadherin is a calcium-dependent glycoprotein which is characterised by three distinct domains: a long extracellular domain with five cadherin repeats (EC1-EC5) which establishes trans- and cis-homophillic interactions, a single transmembrane domain and a cytoplasmic tail, which associates with various intracellular proteins (Hartsock & Nelson, 2008; Patel et al., 2006). The extracellular domain forms trans-cadherin bonds with adjacent cells via histidine-alanine-valine (HAV) domains located on EC1 and requires calcium for correct cadherin conformation (Pokutta, Herrenknecht, Kemler, & Engel, 1994). The cytoplasmic tail contains a catenin binding domain and juxtamembrane domain which enables the important interactions with p120-catenin and β-catenin which is pivitol to the adhesive properties of the AJs (Yap, Niessen, & Gumbiner, 1998). While β-catenin binds to the cytoplasmic domain of E-cadherin with high affinity, it in turns binds to α-catenin which regulates actin filaments (Aberle et al., 1994). α-catenin has been described as an allosteric protein that exists either as a monomer or homodimer. The different conformation states of α-catenin define its affinity to bind to either β-catenin or actin filaments. High concentrations of α-catenin which tend to form clustered pools at the plasma membrane around E-cadherin force its monomeric state and high affinity binding to β-catenin. Meanwhile dimeric α-catenin is found in

cytoplasmic pools allowing it to interact with actin filaments (Drees, Pokutta, Yamada, Nelson, & Weis, 2005; Huber, Stewart, Laurents, Nelson, & Weis, 2001; Pokutta & Weis, 2000). This 'catenin molecular switch' indicates the dynamic nature of junctional adhesion and its interactions with the actin cytoskeleton. Additionally, α-catenin dimers drive the formation of bundle actin filaments while repressing the actin-related protein 2/3 (Arp2/3) complex and subsequent branching actin networks; Arp2/3 is an actin nucleator complex which initiates actin polymerisation leading to branched actin filaments (Drees et al., 2005; Pollard & Beltzner, 2002). Branched actin networks are associated with lamellipodia and leading edge extension seen commonly in mesenchymal migratory cells (Pollard & Borisy, 2003). As shown, the vast network of cadherins, catenins and associated signalling pathways link cellular adhesion, epithelial polarity and the actin cytoskeleton to one another. Additionally, the binding of various proteins to the E-cadherin cytoplasmic tail regulates the processes of recycling and endocytosis for essential turnover of E-cadherin and junctional material (Georgiou, Marinari, Burden, & Baum, 2008; Le, Yap, & Stow, 1999).
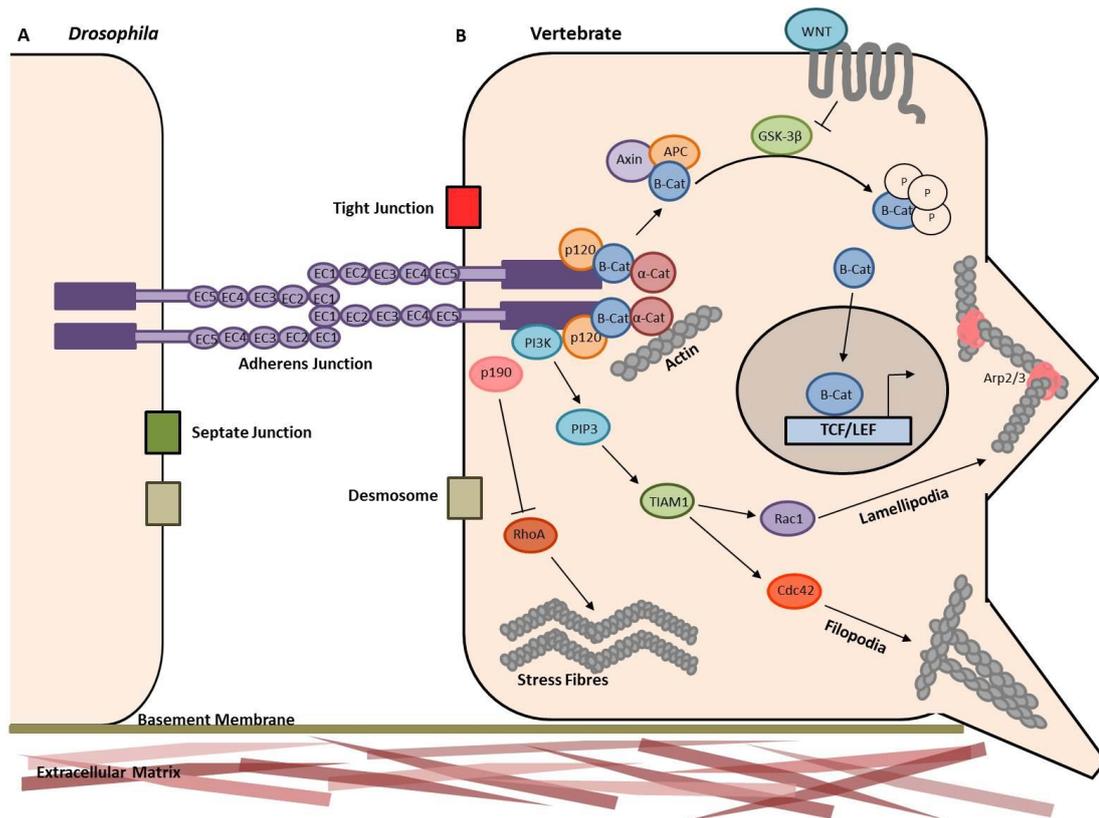
Figure 1.5 E-cadherin and EMT signalling

Diagram shows E-cadherin at the AJ of epithelial cells and the associated signalling networks linking E-cadherin to the actin cytoskeleton and transcription. E-cadherin binds β-catenin and p120, which interacts with the microtubule cytoskeleton and α-catenin which in turn interacts with actin filaments. Loss of E-cadherin results in abundant cytoplasmic β-catenin which translocates to the nucleus and induces transcription of cancer-associated genes. Additionally WNT signalling prevents degradation of β-catenin by inhibiting GSK-3β/axin/APC complex which normally targets β-catenin for ubiquitination and degradation. Signalling pathways associated with E-cadherin, such as PI3K regulates Rho family GTPases which mediate membrane actin protrusions which when mis-regulated can contribute to EMT and mesenchymal cell morphology. **(A)** Septate junction and adherens junction in *Drosophila* epithelial cells compared to **(B)** adherens junction and tight junction in vertebrate.

The process of invasion is typically defined by the loss of adhesion between cells,

thus allowing cells to dissociate from the epithelium and spread. In particular, the

loss of E-cadherin and adhesion junction regulators is associated with invasion and metastasis of cancer cells (Cavallaro & Christofori, 2004). There are various mechanisms by which E-cadherin-mediated adhesion is abolished in invasive tumours, for example, at a transcriptional level, E-cadherin expression can be repressed by the transcription factors Snail and Twist which are known repressors of the E-cadherin promoter (Batlle et al., 2000). Additionally, the E-cadherin protein can undergo proteolytic degradation by secreted MMPs (Werb, 1997). The loss of E-cadherin disrupts cell-cell adhesion, cell polarity and cell shape, thereby facilitating the detachment of tumour cells from the epithelium. Moreover, downregulation of E-cadherin in cancer triggers various downstream signalling pathways which contribute to tumour progression (Christofori & Semb, 1999). For example, in the absence of E-cadherin, β-catenin is liberated from the adhesive junctional Cadherin/Catenin complex. Normally, non-sequestered β-catenin would be phosphorylated by the APC/axin/GSK3β complex which subsequently results in its degradation. However, in many cancers (notably colon carcinomas) accumulation of β-catenin in the cytoplasm occurs due to inactivated mutant APC or increased WNT signalling which prevents the targeting of β-catenin for degradation (Brabletz et al., 2001; Kishida et al., 1998). Translocation of β-catenin to the nucleus and its subsequent binding to Tcf and Lef transcription factors results in the expression of specific genes implicated in hyper-proliferation and tumour progression (Christofori & Semb, 1999). Loss of E-cadherin function is also associated with Rho signalling leading to reorganisation of the actin cytoskeleton (Christofori, 2006). The Rho-GTPases Rac1, RhoA and Cdc42 function to regulate the formation of lamellipodia, contractile forces in migratory cells and formation of actin-rich filopodia respectively (Etienne-Manneville & Hall, 2002; Nobes

& Hall, 1995). In normal cells E-cadherin is able to suppress RhoA and subsequent migratory behaviour by recruiting p190 Rho-GAP, as well as regulating actin assembly by recruiting phosphoinositide 3 kinase (PI3K). PI3K signalling in turn recruits the protein T-cell lymphoma invasion and metastasis 1 (TIAM1) which results in the activation of Rac1 and Cdc42 which regulates the actin cytoskeleton (Arthur & Burridge, 2001; Christofori, 2006).

In order for cancer cells to invade and spread they must lose their specialised characteristics including apical-basal polarity and cell-cell adhesion. EMT is the process in which highly differentiated epithelial cells become motile and invasive by adopting mesenchymal characteristics (Klymkowsky & Savagner, 2009; Polyak & Weinberg, 2009). Normally EMT acts as an essential regulatory program implicated during gastrulation and wound healing, however tumour cells have been shown to 'hijack' this cellular program during tumour progression (Thiery, Acloque, Huang, & Nieto, 2009) (Hanahan & Weinberg, 2011). Unlike epithelial cells, mesenchymal cells are solitary cells with a front-rear polarity making them capable of migrating. The process of EMT is accompanied by a host of physiological changes from the dissolution of adhesive junctional complexes, the switch between apical-basal and front-rear polarity, reorganisation of the actin cytoskeleton and subsequent change in cell shape (Nelson, 2009; Ridley et al., 2003; Yilmaz & Christofori, 2009). Furthermore, at a transcriptional level EMT is associated with the downregulation and upregulation of epithelial and mesenchymal genes respectively. As mentioned previously, the loss of AJ integrity is defined by the dissolution of E-cadherin-mediated adhesion between neighbouring cells (Cavallaro & Christofori, 2004;

Gumbiner, 1996). Additionally, decreases in occludin and claudin along with the loss of zonula occludens 1 (ZO1) at cell contacts impair TJ integrity (R. Y. Huang, Guilford, & Thiery, 2012). **Section 1.3** describes the various polarity protein complexes which localise along the apical-basal axis defining specific apical, lateral and basal domains. De-regulation or mis-localisation of these polarity regulators results in the loss of epithelial apical-basal polarity. Additionally, re-organisation of the actin cytoskeleton facilitates the switch between apical-basal polarity and front-rear polarity which enables directional movement of mesenchymal cells (Christofori, 2006; Nelson, 2009). Rho-GTPases regulate reorganisation of the cortical actin cytoskeleton producing membrane extensions at the leading edge, while Rho activated stress fibres drive cell contractility for motility and directed mesenchymal-like migration **(Fig. 1.5)** (Nobes & Hall, 1995).

The most recognised switch in gene expression during EMT is the 'cadherin switch' where E-cadherin is down-regulated and expression of N-cadherin is upregulated (J. P. Johnson, 1991). This switch in cadherin expression along with the upregulation of EMT transcription factors evokes the loss of AJs, allowing tumour cells to subvert their differentiated polygonal characteristics and adopt a fibroblastic morphology for enhanced cell motility (Nelson, 2009; Polyak & Weinberg, 2009). These gene expression changes are often regulated by various transcription factors associated with early EMT, such as Snail and Twist **(Fig. 1.6)**. Snail regulates transcription by specifically binding to E-box domains in target gene promoters (such as E-cadherin) (Batlle et al., 2000). Snail then recruits polycomb repressive complex 2 (PRC2) which drives a host of histone modifications which are capable of activating or repressing

41

transcription of the target gene (Herranz et al., 2008). In this way Snail can drive

transcription of mesenchymal cell components whilst suppressing epithelial markers.

Similarly, Twist drives histone modifications on target genes by binding to DNA

recognition sites and recruiting SET8 (Yang et al., 2012). Various signalling pathways,

including TGF-β, WNT and Notch, can regulate the expression of these transcription

factors and ultimately promote EMT. For example, under hypoxic conditions HIF1α

signalling activates Twist. In addition to Snail and Twist, Zeb, Fox and Sox

transcription factors are also associated with EMT (Lamouille, Xu, & Derynck, 2014).
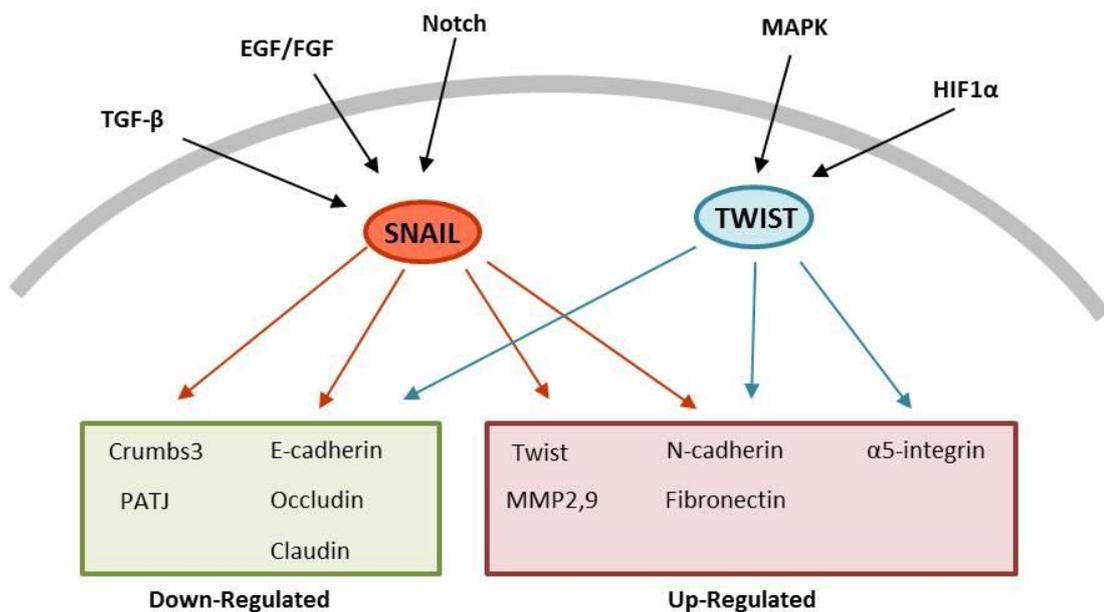
Figure 1.6 EMT transcription factors
Activators and transcriptional targets of the transcription factors Snail and Twist,
which regulate EMT by initiating the 'cadherin switch'.

To summarise, by breaking adhesive interactions between neighbouring cells, altering

the actin cytoskeleton and initiating EMT, tumour cells are able to break away from

each other and exchange their epithelial characteristics for a mesenchymal-like

morphology, allowing them to invade and metastasise. It is important to note however, that whilst EMT serves as the predominant method enabling epithelial cells to invade and metastasise, there are several additional discrete modes of invasion that have been implicated in metastatic cancers. For example, 'collective invasion' describes the advancement of cancer cells into adjacent tissue en- masse, whilst 'amoeboid invasion' sees individual cancer cells escaping through interstices pre-existing in the extracellular matrix (ECM), by acquiring a morphological plasticity phenotype (Hanahan & Weinberg, 2011; Sabeh, Shimizu-Hirota, & Weiss, 2009). During amoeboid invasion, increased Rho signalling leads to actomyosin contractility and subsequent cell cortex remodelling (Sahai & Marshall, 2003). In some cases, inflammatory cells assembling at tumour boundaries have also been shown to aid invasion by secreting enzymes to degrade the ECM (Joyce & Pollard, 2009).

1.2.2 Intravasation and Extravasation

For cancer cells to spread from the primary tumour to distant sites they must first break through the basement membrane (BM) and enter the blood and lymphatic system **(Fig. 1.7)**. Both tumour cells and neighbouring stromal cells release MMPs which are zinc-dependent endopeptidases released to facilitate degradation of the basement membrane and remodelling of collagen fibres within the ECM. By releasing extracellular MMPs, which are normally required for tissue morphogenesis, growth and repair, cancer cells are able to invade through the basement membrane (Kleiner & Stetler-Stevenson, 1999; Werb, 1997).

A prerequisite for metastatic dissemination is the formation of immature neo-vasculature, which supplies oxygen and nutrients to the growing tumour (Bergers & Benjamin, 2003; Duffy, 1996). The release of pro-angiogenic and pro-inflammatory cytokines such as VEGF from tumour cells and stromal cells in the surrounding microenvironment stimulate tumour angiogenesis (Bergers & Benjamin, 2003). Cytokines, MMPs and hypoxic conditions at the tumour site initiate migration of endothelial cells up a chemoattractant gradient towards the tumour to establish the novel, albeit leaky vascular network (Carmeliet & Jain, 2000). To metastasise, locally invaded cancer cells must enter this novel neo-vasculature along with the bodies' blood and lymphatic vessels (Hanahan & Weinberg, 2011). This process of intravasation is considered to be the rate limiting step of invasion and metastasis as circulating cancer cells are often unable to overcome the challenges they face in the circulatory system. For example, circulatory cancer cells often have to rely on anti-apoptotic proteins such as bcl-2 to protect them from anoikis, which occurs due to the loss of matrix and stromal interactions (Guo & Giancotti, 2004). Additionally, platelets are often recruited to shield cancer cells from immune cells and shearing forces present in the circulatory vessels (Gay & Felding-Habermann, 2011; Joyce & Pollard, 2009). Furthermore, cancer cells typically get lodged in the first capillary they enter due to geometric constraints, unless they can enter the larger venous system first, enabling them to travel further (Gupta & Massague, 2006). Once cancer cells have become stuck in a capillary they form adhesive interactions with the endothelial wall via interactions with various integrins and chemokine receptors. These interactions facilitate the extravasation of cancer cells into secondary tissues. Cancer cells achieve trans-endothelial migration by either breaking through endothelial

junctions or mechanically bursting through the endothelial wall (Valastyan & Weinberg, 2011). Successfully extravasated cancer cells undergo mesenchymal-epithelial transition (MET) and form micro-metastases in the new location. MET simply defines the process of metastasised cancer cells regaining their epithelial characteristics so they can colonise in the new tissue, and is achieved by reversing the reactions performed during EMT. These cells usually enter G0 phase of the cell cycle remaining quiescent and dormant for many years. Eventually, macromastatic growth and colonisation is achieved with interplay from the microenvironment which releases growth factors such as TGF-β and IL-6 (Tsai & Yang, 2013). Additionally, it has been hypothesised that a pre-metastatic niche contributes to successful colonisation in which migrating cancer stem cells and bone marrow derived cells facilitate growth of tumours at the new site (Qian & Pollard, 2010). Colonisation and remodelling of the secondary tumour site is achieved through complex heterotypic signalling between cells in the tumour microenvironment. Paracrine signalling, secretion of cytokines and increased tumour associated fibroblasts contribute to the premetastatic niche making it susceptible to colonisation. For example, breast cancer cells have been shown to induce secretion of periostin which promotes WNT signalling and cell growth at the secondary site (Malanchi et al., 2012).
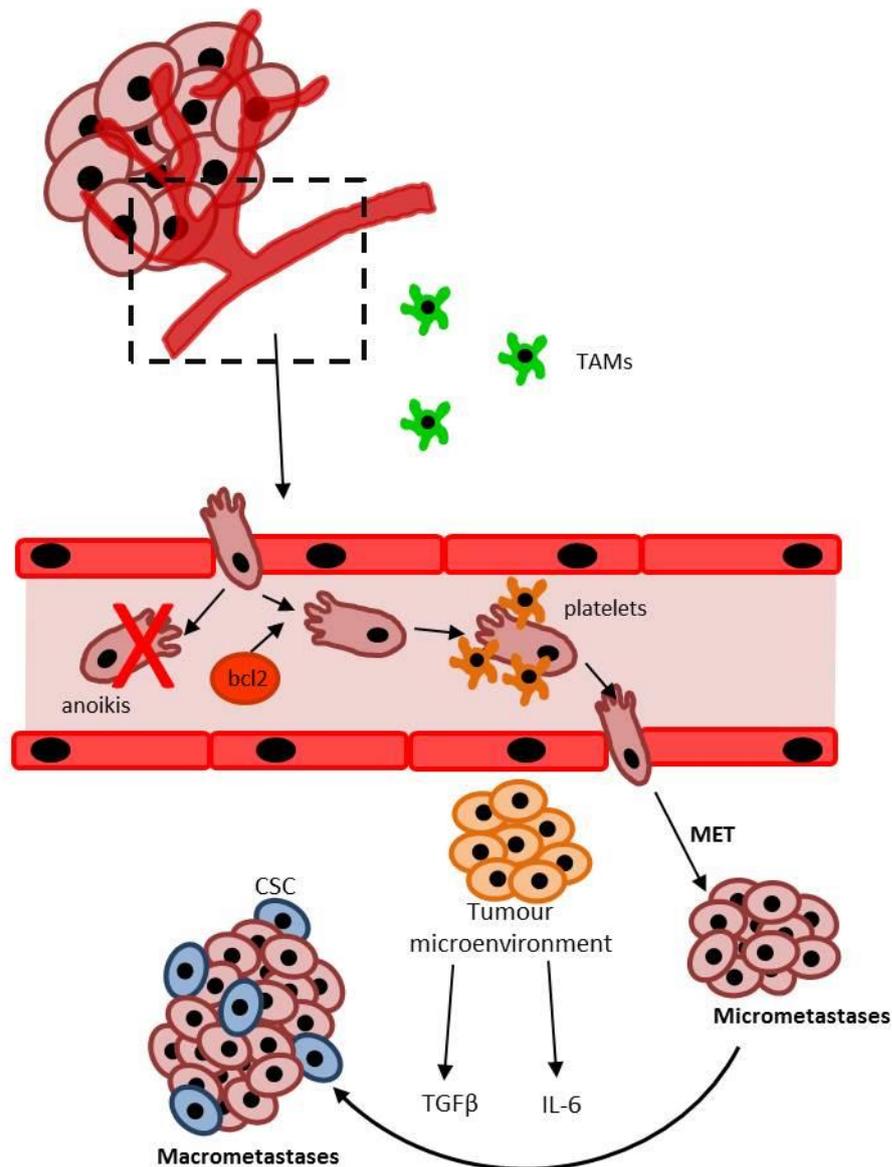
Figure 1.7 Intravasation and Extravasation

The process in which motile cancer cells metastasise from the primary tumour and form macro-metastases in new locations, and some of the dangers they must overcome.

1.3 Epithelial Cell Polarity

Cell polarity is a fundamental characteristic of all epithelial cells, defining distinct

apical, lateral and basal domains for the asymmetric organisation of cell components,

essential for normal cell physiology and tissue homeostasis. Tumour progression is

largely associated with the loss of this epithelial cell polarity. A distinct apical-basal axis of polarity enables the exchange of ions and solutes across the epithelium and is required for regulating asymmetric cell division as well as maintaining the apical junctional complex (Fristrom, 1988; Januschke & Gonzalez, 2008). The apical junctional complex consists of TJs (SJs in *Drosophila*) and AJs, which are structurally dependent on the integrity of several polarity complexes **(Fig. 1.8)**. These polarity complexes which make up the epithelial polarity programme (EPP) localise separately along the apical-basal axis for tight coordination of cell polarity, cell-cell junctions and cytoskeletal organisation (Januschke & Gonzalez, 2008; Rodriguez-Boulan & Macara, 2014).

Epithelial cell polarity is established through the generation of apical and basolateral domains which defines asymmetry along the apical-basal axis. These domains are generated through mutually exclusive interactions formed between distinct polarity proteins. The apical Par and Crumbs complex and basolateral Scribble and Yurt/Coracle complex interact with one another generating zones of mutual exclusion surrounding the mature epithelial junctions **(Fig. 1.8)** (Assemat, Bazellieres, Pallesi-Pocachard, Le Bivic, & Massey-Harroche, 2008).
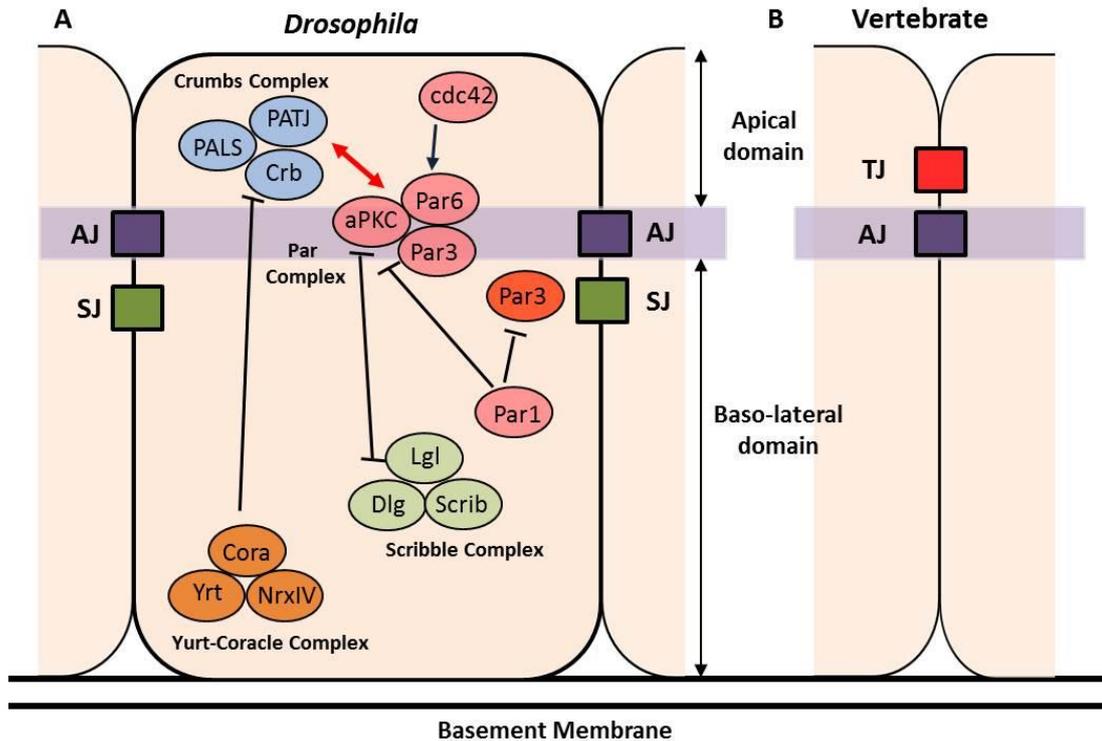
Figure 1.8 Epithelial polarity complexes

Interactions between polarity complexes which are mutually antagonistic define the apical and basolateral domains. The Crumbs and Par complex are restricted to the apical domain whilst the Scribble complex localises to the basolateral domain to maintain apical-basal polarity. **(A)** Adherens junction and septate junction in *Drosophila* compared to **(B)** Tight junctions and adherens junctions in vertebrate epithelial cells.

1.3.1 Apical Par and Crumbs complexes

The apical Par complex was first characterised in 1988 in *C.elegans* and was closely linked to asymmetric cell division (Nance, Munro, & Priess, 2003). There is a host of Par family proteins including the Par kinases (Par1, Par4) and scaffold PDZ-containing proteins (Par3/bazooka, Par6). The Par complex itself is formed by the interactions between Par3 and Par6 along with atypical protein kinase C (aPKC) and the small GTPase cdc42 (Bilder, 2003; Joberty, Petersen, Gao, & Macara, 2000; D. I. Johnson, 1999; Lin et al., 2000). The Par complex plays an important role in regulating spatial

restriction of the actin cytoskeleton, spindle orientation during asymmetric cell division and driving epithelial polarity through its adverse effects at epithelial TJs (Suzuki & Ohno, 2006). Par3 binds to TJs (AJs in *Drosophila*) through adhesion proteins associated with mature junctions; this anchorage to the apical-lateral border allows Par3 to recruit the remaining members of the Par complex. Par6 initially binds to Par3 and subsequently recruits aPKC to the complex. Cdc42 is the last member of the Par complex to join, and binds directly to Par6 (Joberty et al., 2000). GTP bound cdc42 interacts directly with Par6 which changes the conformation of the Par complex, leading to aPKC activation and subsequent phosphorylation of target proteins (Yamanaka et al., 2001). aPKC is pivotal in the maintenance of polarity by generating feedback loops and mutually exclusive interactions. The first feedback loop initiated is between aPKC and Par3 (Baz). As shown in *Drosophila,* the Baz/aPKC interaction is inhibited when Baz is phosphorylated on serine 980 by aPKC. In addition, Crumbs prevents the interaction between Baz with Par6, maintaining its localisation to the apical lateral border. Localisation of Baz is also restricted basally by the Scribble complex which regulates Par1 localisation which subsequently results in phosphorylation of Baz at serine 151 or 1085. (Morais-de-Sa, Mirouse, & St Johnston, 2010; Suzuki & Ohno, 2006).

The Crumbs complex (Crumbs, PALS/Stardust and PATJ) is also found in the apical domain and is important in the establishment of AJs and apical-basal polarity. Crumbs was first identified in *Drosophila* (Tepass, Theres, & Knust, 1990) but has since been widely studied in other organisms. The crumbs complex regulates formation of the apical membrane and forms feedback loops with the Par complex to

restrict apical proteins to the apical-lateral domain whilst restricting the third polarity complex (Scribble) to the basolateral domain. Apical restriction is achieved through interactions between crumbs and Par6 which subsequently upregulates aPKC (Suzuki & Ohno, 2006; Tanentzapf & Tepass, 2003).

1.3.2 The Scribble complex

The scribble complex consists of Scribble, discs large (dlg) and lethal giant larvae (lgl) and localises to the basolateral domain. These polarity proteins are all capable of recruiting a variety of different protein networks due to their highly conserved guanylate kinase, PDZ and SH3 domains (Bilder, 2003). Lgl was the first recognised member of the scribble complex, being identified in *Drosophila* in 1930, although it wasn't until the 1980's that it was fully sequenced and characterised (Mechler, McGinnis, & Gehring, 1985). Dlg was also first discovered in *Drosophila* in 1972 and was quickly identified as an essential regulator of epithelial structure and development. Dlg belongs to the membrane-associated guanylate kinase (MAGUK) family and localises at septate junctions in flies. There are 5 Dlg human homologues all of which consist of conserved PDZ and SH3 domains (Woods & Bryant, 1991; Woods, Hough, Peel, Callaini, & Bryant, 1996). Finally, Scribble was identified in *Drosophila* in 2000 (Bilder & Perrimon, 2000) as a key regulator of epithelial morphogenesis. Scrib specifically regulates positioning of AJs and restriction of apical determinants from the basolateral domain which defines epithelial polarisation.

The Scribble complex establishes the basolateral domain through a number of antagonistic interactions which inhibit lateral localisation of the Par complex. Apical

localisation of Lgl is inhibited by a phosphorylation event mediated by aPKC (Betschinger, Mechtler, & Knoblich, 2003). Additionally, lgl prevents Par6 from localising to the cell cortex. Moreover the scribble complex regulates Par1 localisation which inhibits basolateral localisation of Par3 (Assemat et al., 2008). Importantly, components of the Scribble complex have been linked to cell migration events. Functional alterations to members of the protein networks responsible for orchestrating cell polarity is thought to be crucial to the pathology of cancer (Royer & Lu, 2011).

The genes coding these polarity proteins, *lethal(2)giant larvae (lgl), discs large (dlg)* and *scribble (scrib)* also act as neoplastic tumour suppressors which were first identified in *Drosophila melanogaster,* and along with other functionally related genes have been shown to play a role in human cancers (Bilder, 2004; Gateff, 1978; Schneiderman & Gateff, 1967). In the imaginal discs loss of these genes results in neoplastic overgrowth, however, when clones are introduced these mutant cells are outcompeted through cell competition **(Section 3.1)** (Brumby & Richardson, 2003; C. L. Chen, Schroeder, Kango-Singh, Tao, & Halder, 2012; Pagliarini & Xu, 2003). It has also been shown that these tumour suppressors act together through a common pathway to regulate epithelial cell polarity and growth (Bilder & Perrimon, 2000). Additionally, direct interactions have been observed between the mammalian homologues of *Drosophila* Scrib and Lgl (Lg1-2) (Kallay, McNickle, Brennwald, Hubbard, & Braiterman, 2006). Cortical localisation of Lgl, a membrane-associated protein, has also been shown to facilitate formation of the cytoskeletal network. All three members of the scribble complex have been shown to interact genetically and

the concerted activity of these genes regulates polarity and cytoskeletal organisation (Bilder & Perrimon, 2000).

Mutations resulting in the loss of function in these polarity complexes results in apical-basal polarity defects, however these defect phenotypes can be compensated by reducing expression of the unaffected complex. This demonstrates that the apical-basal axis is defined by the mutual exclusion of these polarity complexes to their apical/basolateral domains (Bilder, 2003; Tanentzapf & Tepass, 2003).

1.4 *Drosophila* as a genetic model organism for tumour progression

For over a century, *Drosophila melanogaster* has been used as a genetic tool for developing our understanding of genetics, development, ageing and various signalling cascades. The American geneticist Thomas Hunt Morgan is undoubtedly the most recognised scientist in *Drosophila* history, popularising the fruit fly as a model organism for genetic studies and developing 'classical' *Drosophila* genetics. He studied white eyed mutant flies and their pattern of inheritance. He was awarded the Nobel Prize in 1933 for his earlier work on the chromosome theory of inheritance and evidence that some genes are linked such that they are inherited together. Additionally, Morgan's students Muller and Bridges produced ground-breaking research in genetics. In 1918, Muller developed the use of balancer chromosomes to stably maintain homozygous lethal mutations, a genetic tool which is still extremely important in *Drosophila* genetics today. Furthermore, Bridges produced an array of

incredibly accurate polytene and cytogenetic maps between 1935 and 1938. *Drosophila* has since become a very popular model to carry out large genetic screens for various diseases and processes (Rubin & Lewis, 2000). One of the first and most famous screens was carried out by Volhard and Wieschaus and was awarded the Nobel Prize in 1980. They carried out a mutagenesis screen studying developmental and patterning defects in the embryo. From this they identified a host of 'classic' genes required during embryogenesis, including hedgehog and Krüppel (Nusslein-Volhard & Wieschaus, 1980).

Additionally, *Drosophila* has become an increasingly popular model organism for studying tumour progression, allowing research on many cancer pathways, due to high conservation of the human genome and preservation of more than 60% of cancer proteins (Potter, Turenchalk, & Xu, 2000). Due to the vast complexity of cancer and the number of molecules and pathways implicated, *Drosophila* has become an ideal tool for studying the mechanisms and genes involved, as many of the same pathways and genes found in *Drosophila* are found in humans (Bernards & Hariharan, 2001; Dahman, 2008; Potter et al., 2000). **Chapter 3** outlines some of the important cancer research and loss of function screens that have been carried out in *Drosophila* over the years.

There are many reasons why *Drosophila* has become such a popular model organism. For example, compared to other model organisms *Drosophila* is relatively cheap and easy to use. With a short life cycle **(Fig. 1.9)** and simple maintenance, experiments can be performed quickly and cheaply. Additionally, with the developing and

excelling fly community a wide range of genetic tools are available, including large RNA interference (RNAi) stock centres and online resources such as flybase (Attrill et al., 2016). Genetically, *Drosophila* is easy to manipulate, with reduced redundancy and a lack of recombination in males. Meanwhile the use of balancer chromosomes prevents random meiotic recombination in females and allows stocks to be kept which carry homozygous lethal mutations (Bernards & Hariharan, 2001; Dahman, 2008; Potter et al., 2000).
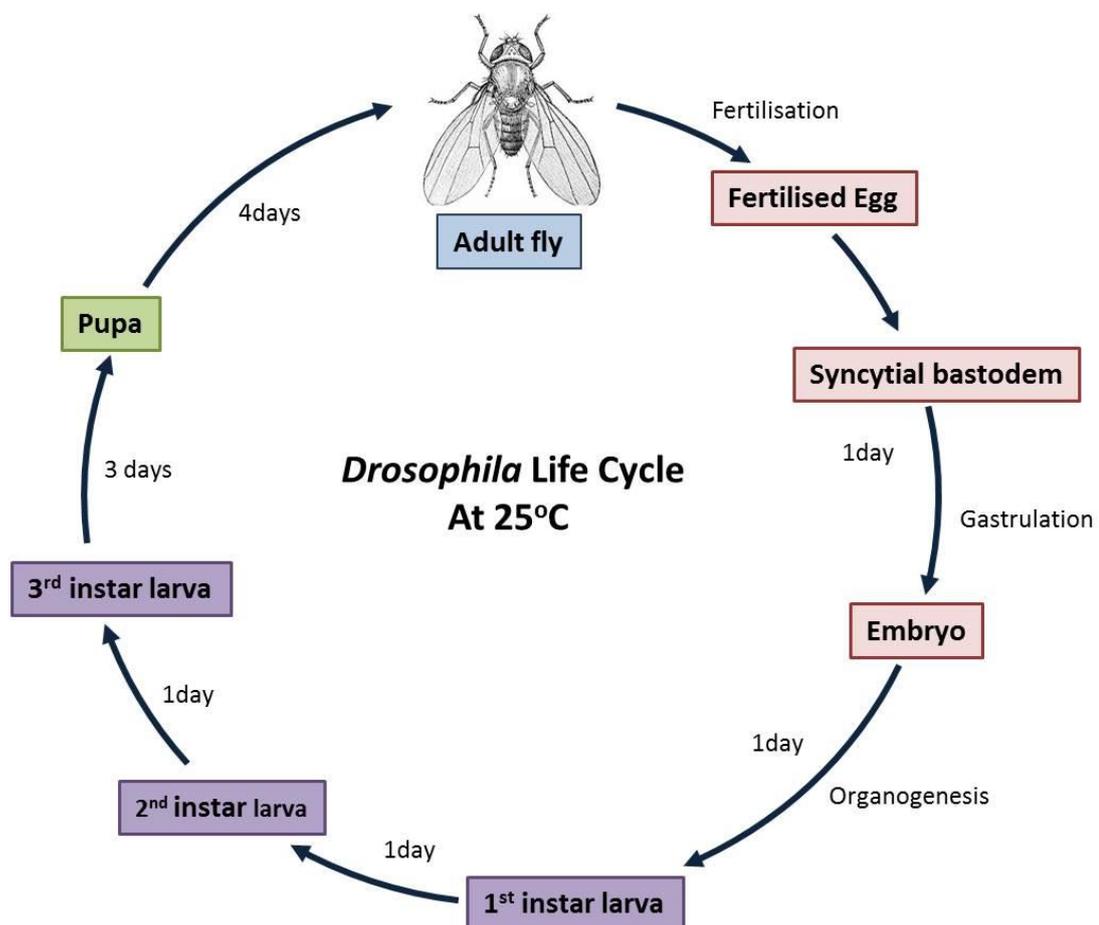
Figure 1.9 *Drosophila* Life Cycle

The different stages of the *Drosophila* life cycle when kept at 25°C and the time scale of development. The *Drosophila* life cycle is approximately doubled in time when flies are kept at 18°C.

1.5 Research Aims

To characterise the critical cell biological events underlying tumour cell invasion, the notum of *Drosophila* was used as a model system in a genetic screen, combining sophisticated *Drosophila* genetics with cell biology. The notum (the dorsal thorax of the fly) is a well-organised, columnar, monolayered epithelium, which is formed during pupal development. Through a succession of morphogenic events at early pupal stages, lateral epithelial sheets approach and meet at the midline where they fuse together to form the dorsal thorax (Zeitlinger & Bohmann, 1999).

Clones of transformed cells labelled with GFP can be visualised within the living epithelium of *Drosophila*, where the shape, dynamics and behaviour of transformed cells can be followed in high resolution over time as the tissue becomes neoplastic. Significantly this novel genetic system is one of the few able to characterise the precise molecular and cell biological events during early tumourigenesis (Cohen, Georgiou, Stevenson, Miodownik, & Baum, 2010).

By inducing small mutant clones, which are homozygous mutant for a specific tumour suppressor gene, it is possible to generate neoplastic tumours in the notum of the fly

These tumours show overgrowth, loss of apical-basal polarity and cells that undergo EMT which can survive beneath the epithelial sheet (Georgiou & Baum, 2010; Georgiou et al., 2008).

Using this system it is possible to generate clones of transformed cells homozygous mutant for *lgl* **(see Chapter 3)**, allowing an enhancer-suppressor screen to be carried out for genes that either promote or inhibit tumour progression. For the screen, transformed cells, homozygous mutant for the tumour suppressor *lgl,* will simultaneously express a GFP reporter (allowing us to observe tumour size, shape and behaviour) and an RNAi transgene that knocks down the expression of an additional gene in the *Drosophila* genome. Importantly, using the MARCM technique **(see Section 2.2.2.3)** the expression of both the GFP reporter and the RNAi transgene will be specific to the homozygous mutant tissue, so only the developing tumour will be labelled in the notum and only the developing tumour will have knockdown of the target gene expression. In this way, genes will be identified that promote or inhibit tumour progression by affecting tumour size, cell morphology, cell behaviour and/or the number of invading cells.

The overall aim of my PhD was to optimise and set up a large-scale genetic screen using RNAi to identify novel genes involved in tumour progression. From this, I hoped to characterise these genes in more detail to try to further our understanding of their biological role in cancer.

One of the initial aims of my research was to carry out a preliminary pilot screen which would be used to ensure the system was working efficiently in the notum, determine ideal conditions and identify which mutant background was most suitable for the main screen. The aim of the main candidate screen was to study approximately 700 RNAi transgenes targeting 500 different genes over 3 years. Then once specific genes that perturb tumour progression had been identified, I would be able to study in more detail the role of these genes in tumour progression.

Targeted cancer therapy has become of particular interest to cancer researchers, whereby drugs are strategically targeted to specific genes directly linked to cell growth and proliferation. Current targeted therapeutics aim to specifically attack cancer cells with minimal damage to normal tissue in the tumour microenvironment. With this in mind, high relapse rates associated with current therapies such as chemotherapy are hoped to be significantly reduced. The anti-cancer drug avastin for example specifically prevents endothelial cell recruitment during angiogenesis, an essential step in the metastasis cascade (Cappuzzo et al., 2003; Miller et al., 2007; Savage & Antman, 2002). It is hoped that novel genes implicated during the invasion-metastatic cascade will be identified through the *in vivo* genetic screen and by investigating how oncogenes and tumour suppressor genes cooperate to affect epithelial cell morphology and invasion; the results obtained will be of great significance to human health. Therefore the overall aim of my research was to identify genes that affect tumour progression from over-proliferation to invasion and metastasis.

CHAPTER 2

Materials and Methods

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 *Drosophila* stocks

Table 2.1 contains a list of all the *Drosophila* stocks used and where they originated. The majority of stocks were made in the Georgiou Lab, while the rest are available publicly. A full list of all the RNAi lines used can be found in **Appendix B-C**.

Table 2.1 *Drosophila* Stocks

Stock	Genotype	Source
M383	w,P[mw,Ubx-flp]; IF/P[CyO-GFP]; P[mw,GAL4][Pnr], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
8217	;P[FRT][40A], P[FRT][G13]/CyO	Bloomington
M413	w, P[mw, Ubx-flp]; P[FRT][40A],P[FRT][G13]/P[CyO-GFP]; P[mw,GAL4][Pnr], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
M380	w;P[mw,Tub-GAL80], P[FRT][40A]; MKRS/TM6b	Georgiou Lab
M442	w, P[mw, Ubx-FLP];*lgl*[4], P[FRT][40A]/P[CyO-GFP];P[mw, GAL4][Pnr], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
M471	P[mw,Ubx-flp]; P[mw,GAL4][Act], P[mw,UAS-Moe-GFP]/ P[CyO-GFP]; P[FRT][82b], *scrib*/TM6b	Georgiou Lab
M485	IF/ P[CyO-GFP];; P[mw,Tub-GAL80], P[FRT][82b]/TM6b	Georgiou Lab
M371	P[mw,Ubx-flp], *dlg*, P[FRT][19A]/FM7, P[Kr-GFP];; P[mw,GAL4][Pnr], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
M382	P[mw,Tub-GAL80], P[FRT][19A]; MKRS/TM6b	Georgiou Lab
M59	;; P[GAL4][Neu], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
M147	w;;;	Bloomington
M520	P[mw,Ubx-flp]; IF/ P[CyO-GFP];MKRS/TM6b	Georgiou Lab
5626	W; P[FRT][42D], P[Uni-GFP.nls]/CyO	Bloomington
M530	P[mw,Tub-GAL80], P[FRT][42D]/ P[CyO-GFP]; P[GAL4][Neu]/TM6b	Georgiou Lab
M288	w;; P[mw,GAL4][Pnr], P[mw,UAS-Moe-GFP]/TM6b	Georgiou Lab
32123	w;; P[UAS-Apoliner]/TM3, sb	Bloomington
	w; P[FRT][42D], *dscam*$^{21/23}$/CyO	Gift from Dr. Larry Zipursky
V51247	Lgl-RNAi (III)	VDRC
58742	w; P[FRT][42D], shg/CyO; P[Ubi-p63E-shg.GFP]	Bloomington
35496	y, w; P[FRT][42D], P[Ubi-mRFP.nls]	Bloomington
6298	P[UAS-p35];;	Bloomington

3703	w; P[CyO-GFP]/Sco; MKRS/TM6B	Bloomington
54591	y1 P(*nos-cas9, w+*) M(*3xP3-RFP.attP*)ZH-2A w*	Bloomington

2.1.2 Chromosome balancers, markers and genetic constructs

Chromosome balancers (homozygous lethal alleles with a dominant visible phenotype) were used to suppress meiotic recombination that naturally occurs in females and to track progeny genotype through multiple genetic cross schemes. **Table 2.2** contains all the balancers and genetic markers used.

Table 2.2 Balancers and genetic markers

Name	Chromosome	Type	Phenotype
CyO (Curley of Oster)	2	Balancer	Curly Wings
FM7 (In(1)FM7)	1	Balancer	Kidney-shaped eyes
TM3 (In(3LR)TM3)	3	Balancer	Stubble (shortened bristles) or serrate wings
TM6b (In(3LR)TM6)	3	Balancer	Tubby (short, fat body), Humeral (extra macrochaetes), and stubble
IF (Irregular facets)	2	Marker	Small 'rough' eyes with irregular ommatidia.
MKRS	3	Marker	Stubble
Sco (Scutoid)	2	Marker	Loss of bristles

Table 2.3 Genetic constructs
Table outlining the function of genetic constructs incorporated into the *Drosophila* genome.

Genetic construct	Function
Pnr-GAL4	GAL4 expression is driven by the pannier promoter which restricts expression to the central region of the notum.
Act-GAL4	GAL4 is expressed ubiquitously under control of Actin5C promoter.
Neu-GAL4	GAL4 is expressed specifically in the sensory organ precursor cells and their progeny.
Tub-GAL80	The tubulin promoter ubiquitously drives expression of GAL80 which represses GAL4 activity.
FRT	Flippase Recognition Target (FRT) sites are located at various

	positions along chromosomes (40A, 19A, G13, 82b, 42D) and are the site of flp induced post-mitotic recombination.
UAS-Moe-GFP	The actin-binding domain of Moesin (Moe) is tagged to GFP which is under the control of UAS. GFP therefore localises to actin filaments.
Kr-nls-GFP	GFP expression is driven by the Kruppel (kr) promoter and the nuclear localisation sequence (nls) restricts GFP localisation to the nucleus.
Ubx-flp	The Ultrabithorax (Ubx) promoter drives expression of flippase (flp) in the nucleus.
UAS-RNAi	Various RNAi constructs under the control of UAS.
CyO-GFP	GFP expression in the abdomen is under the control of the Curley of Oster promoter.
Ubi-nls-GFP	The ubiquitous (Ubi) promoter expresses GFP in the nucleus.
UAS-Apoliner	Genetic construct consisting of GFP and RFP separated by a caspase recognition site. Used as a marker for apoptosis.
UAS-p35	Expression of p35 under UAS control. P35 blocks apoptosis.
Ubi-nls-RFP	The ubiquitous (Ubi) promoter expresses RFP in the nucleus.
Ubi-shg.GFP	The ubi promoter drives expression of E-cadherin (shg) tagged at the C-terminus with GFP (live cadherin)

2.1.3 Mutagenesis and cloning oligonucleotides

For mutagenesis using CRISPR/Cas technology, oligos **(Table 2.4)** were designed using

CRISPR optimal target finder and ordered (desalted, dried) from Sigma-Aldrich®

Table 2.4 CRISPR/Cas oligonucleotides

Name	Primer sequence 5'-3'
CG10600 Target 2 sense	**GTCGGTCGCGCTCCCGCTCCCTAT**
CG10600 Target 2 anti-sense	**AAACATAGGGAGCGGGAGCGCGAC**

Table 2.5 Plasmids

Plasmid	Source
pCFD3-dU6:3gRNA	gift from Simon Bullock (Addgene plasmid # 49410)

2.1.4 Primers for PCR and Sanger Sequencing

Sanger sequencing was used to confirm DNA plasmid sequences and verify CRISPR mutants. Sequencing was carried out by the DNA sequencing facility in the School of Life Sciences, at the University of Nottingham. Non-universal primers were designed using primer 3.

Table 2.6 Primers

Name	Primer Sequence 5'-3'	Supplier
T3	**CAATTAACCCTCACTAAAGG**	**Universal primer supplied by DNA Seq facility, University of Nottingham.**
CG1	**AAGAAGAGAGACAGGGAGCG**	**Sigma-Aldrich®**
CG2	**CGTTCTGCACTTTCGAAACG**	**Sigma-Aldrich®**
CG3	**CGATAAGGTGAACGCGAGAC**	**Sigma-Aldrich®**
RSB1	**GCAGCTTTCTACGTGATGGG**	**Eurofins**
RSB2	**ATTGGGCCAGCTCTCTGTAAGT**	**Eurofins**
ING1	**TACTGTCTGTGCAACCAGGT**	**Eurofins**
ING2	**TTCTCCAGGGCTTTGTCCAT**	**Eurofins**
GAP F	**ATGTTCGTCATGGGTGTGAA**	**Thermofisher Scientific**
GAP R	**GTCTTCTGGGTGGCAGTGAT**	**Thermofisher Scientific**

2.1.5 siRNAs

Table 2.6 lists all the ON-TARGET plus siRNAs used for targeting human genes as well as scrambled siRNA control and siGLO for optimisation experiments. SMARTpool (a mix of 4 different siRNAs targeting the same gene) were ordered to improve specificity.

Table 2.7 siRNA

siRNA target	Oligo Sequence	Supplier
ING1	**AGAGAGGGCUUACAACAGG** **CGAGAAGACCAUGGACAAA** **CCACGUACUGUCUGUGCAA** **GCGUGGGGCUCAAUCAUAA**	**Dharmacon**
RSBN1L	**GAUAUAGAGACAACGACUA** **UAGAAGUAGUUCAACGAAU** **GUGUGGUAUUGUAGUGCAA**	**Dharmacon**

UCUAGUUGAUACAAGGCAA		
Non-targeting control		**Dharmacon**
SiGLO Cyclophilin B control	**GGAAAGACUGUUCCAAAAA**	**Dharmacon**

2.1.6 Antibodies

Table 2.8 Primary Antibodies

List of primary antibodies and corresponding supplier. Target epitopes and dilution for immunofluorescence (IF) shown.

Primary Antibody	Dilution for IF	Supplier
Mouse anti-GFP	**1:300**	**Roche**
Rabbit anti-GFP	**1:1000**	**Molecular Probes**
Rabbit anti-PKC zeta	**1:50**	**Santa Cruz (California, USA)**
Rat anti-DE cadherin	**1:100**	**Developmental studies hybridoma bank (DSHB) (Iowa, USA)**
Guinea Pig anti-Bazooka	**1:1000**	**Prof. Andreas Wodarz**
Mouse anti-Fasciclin III	**1:400**	**DSHB (Iowa, USA)**
Mouse anti-Dscam	**1:250**	**Gift from Larry Zipursky**
Mouse anti-Armadillo	**1:100**	**DSHB (Iowa, USA)**

Table 2.9 Secondary Antibodies

List of secondary antibodies with supplier and dilution for IF.

Secondary Antibody	Dilution (IF)	Supplier
Alexa Fluor ® 488 goat anti-mouse	**1:300**	**Molecular Probes**
Alexa Flour ® 546 goat anti-rabbit	**1:300**	**Molecular Probes**
Alexa Flour ® 546 goat anti-rat	**1:300**	**Molecular Probes**
Alexa Flour ® 488 goat anti-rabbit	**1:300**	**Molecular Probes**
Alexa Flour ® 546 goat anti-mouse	**1:300**	**Molecular Probes**
Alexa Flour ® 546 goat anti-guinea pig	**1:300**	**Molecular Probes**
Alexa Flour ® 546 goat anti-Rat	**1:300**	**Molecular Probes**
DAPI	**25mg/ml**	**Thermofisher Scientific**

Table 2.10 General Reagents

Reagent	Supplier
Agarose	Life Technologies
Alkaline Phosphatase, Calf Intestinal (CIP)	NEB
Ampicillin	Sigma
BbsI	NEB
BSA	Sigma
Chloroform	Sigma
CIP	NEB
DEPC	Thermofisher Scientific
Dharmafect Reagent	Dharmacon
DMEM	Thermofisher Scientific
DMSO	Thermofisher Scientific
DNase 1	Invitrogen
10X DNase 1 Buffer	Invitrogen
dNTP Mix	Thermofisher Scientific
EDTA	Invitrogen
Ethanol	Sigma
Foetal bovine serum (FBS)	Thermofisher Scientific
Fetal calf serum (FCS)	Thermofisher Scientific
Glycerol	Sigma
Glycogen	Sigma
Hanks balanced salt solution (HBSS)	Thermofisher Scientific
HEPES	Thermofisher Scientific
Isopropyl alcohol	Sigma
iQ SYBR Master Mix	BioRad, UK
LB agar	Sigma
LB broth	Sigma
L-Glutamine	Thermofisher Scientific
NEBuffer 2.1	NEB
Non-essential amino acids	Thermofisher Scientific
N-Propyl galate (anti-fade)	Sigma
Nuclease-free water	Thermofisher Scientific
Oligo(dT)18 Primer	Thermofisher Scientific
Paraformaldehyde (PFA)	Thermofisher Scientific
PBS	Sigma
Penicillin/Streptomycin (P/S)	Thermofisher Scientific
Proteinase K	Thermofisher Scientific
5X Reaction Buffer	Thermofisher Scientific
RevertAid Reverse Transcriptase	Thermofisher Scientific
RiboLock RNase Inhibitor	Thermofisher Scientific
5x siRNA buffer	Dharmacon
sybersafe	Invitrogen
50x TAE	GeneFlow

TRIzol Reagent	**Ambion, Life Technologies**
Trypsin	**Thermofisher Scientific**
T4 DNA ligase	**New England Biolabs (NEB)**
T4 DNA Ligase Reaction Buffer (10X)	**NEB**
T4 Polynucleotide Kinase	**NEB**
Tween 20	**Sigma**

Table 2.11 Cells

Cells	Source
U87 (adult human glioblastoma)	Gift from Dr Ruman Rahman
MDA-MB-231 (human breast adenocarcinoma)	Gift from Dr Sally Wheatley
Chemically competent DH5alpha cells	Gift from Dr Jen Patel

2.2 Methods

2.2.1 *Drosophila* Husbandry

Drosophila stocks were maintained at $25^{o}C$ and $18^{o}C$ (where the life cycle of the fly is 10 and 21 days respectively (Dahman, 2008)) and raised on a standard cornmeal medium made as follows (8L batch): 7340mlH20, 565ml golden syrup, 127g yeast, 73g soya flour, 536g cornmeal, 42g agar, 35.4ml propanoic acid. Stocks were flipped every 2-3 days at $25^{o}C$ or once a month when stored at $18^{o}C$.

Flies were anaesthetised using a CO_2 porous polyeythylene gas diffuser and viewed under a Leica M60 or fluorescence microscope (Leica MZ10F) for virgin collections and analysis. Flies were sexed based on the identification of ventrally located claspers on the posterior abdomen of males and virgin females were collected every 8 hours when stored at $25^{o}C$ or 16 hours when stored at $18^{o}C$ (as females will not mate within the first 8 hours of adulthood (Ashburner, 1989). Genetic crosses were set up with approximately 20 females and 5-10 males and kept at $25^{o}C$. F1 flies and pupae

were then screened against appropriate genetic markers and balancers to ensure the inheritance of desired genetic constructs.

2.2.2 *Drosophila* Genetic Manipulation

The majority of transgenic *Drosophila* stocks (RNAi lines) used in this research were generated by insertional mutagenesis by various *Drosophila* stock centres (VDRC, NIG and Bloomington). To manipulate the *Drosophila* genome P-elements (a class of transposon) carrying specific constructs are micro- injected into the *Drosophila* embryo. Once P-elements have integrated into the genome transformant flies are identified by markers (typically eye colour) present in the P-element vector. *Drosophila* have 1 pair of sex chromosomes and 3 pairs of autosomes. The X, 2nd and 3rd chromosomes can all be genetically modified, whilst the 4th chromosome is rarely used due to its smaller size and heterochromatic regions. Across these 4 chromosomes there are approximately 14,000 genes which have been fully mapped and annotated (Bernards & Hariharan, 2001; Dahman, 2008; Potter et al., 2000).

Most experiments in this research required genetic manipulation using mosaic techniques which enabled the generation of subsets of mutant cells (clones) within a wild-type environment. This process comprises 3 sophisticated genetic techniques:

2.2.2.1 Upstream activating sequence (UAS)/Gal4 system

The Gal4/UAS system enables the expression of transgenes in a specific cell type at a particular time point in the animals' development. The Gal4 gene encodes a yeast

transcription activator protein that recognises and binds to upstream activation sequences (UAS) located alongside a gene of interest, such as a GFP reporter or RNAi hairpin. In the presence of the Gal4 protein, the UAS-transgene is subsequently expressed. Gal4 can be engineered under the control of different promoters **(Table 2.3)** (Brand, Manoukian, & Perrimon, 1994).

2.2.2.2 Flp/FRT system

The genetic background of the fly can be easily manipulated using flippase (Flp)/ flippase recognition target (FRT) which is a site-directed recombination system. Induced by the recombination enzyme flippase, post-mitotic recombination can occur between identical FRT sites located on homologous chromosomes during G2 phase of the cell cycle. In this system the exchange of chromosome arms generates daughter cells which are either homozygous mutant or homozygous wild-type. Here flippase is under the control of the *Ultrabithorax (Ubx)* promoter, which restricts expression to the notum (Xu & Rubin, 1993).

2.2.2.3 Mosaic Analysis with a Repressible Marker (MARCM)

To enhance the degree of UAS responder regulation, the mosaic analysis with a repressible marker (MARCM) technique combines these systems with the addition of a Gal4 repressor, Gal80 **(Fig. 2.1).** When present, Gal80 represses gene activation initiated by Gal4 in heterozygous animals. However, following flp induced recombination; daughter cells that are homozygous mutant will have lost the Gal80 repressor, thus allowing expression of UAS-transgenes specifically in the homozygous mutant tissue (T. Lee & Luo, 2001). For example, if the transgene is a GFP-reporter

then mutant cells will be GFP positive whilst wild-type cells will be GFP negative. Multiple UAS-constructs can be inserted into the genome, allowing homozygous mutant cells to not only be specifically labelled, but also to simultaneously express an additional transgene.
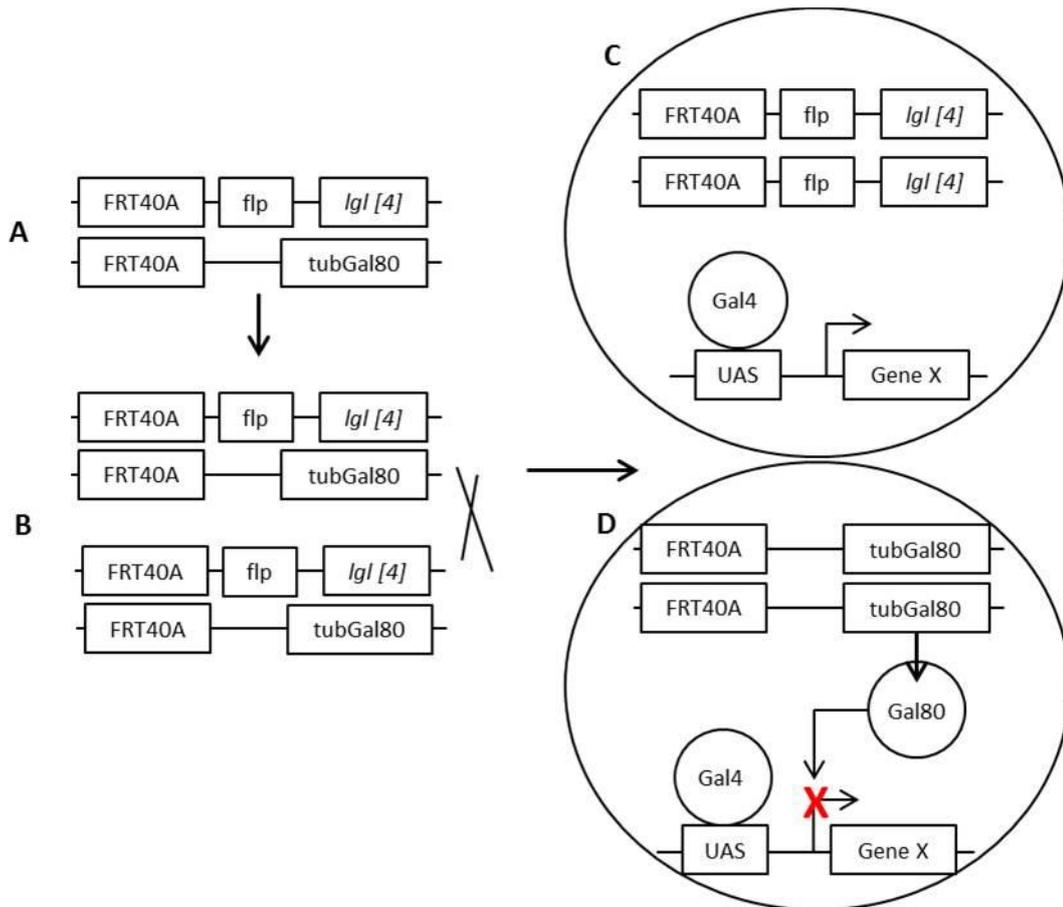
Figure 2.1 MARCM system

One chromosome arm of the parental cell contains the mutant *lgl[4]* allele, flp and FRT site whilst the other contains an identical FRT site and GAL80 under the tubulin promoter **(A)**. Following DNA replication, flp drives recombination between the two FRT sites **(B)**. Daughter cells are either homozygous mutant **(C)** in which tubGAL80 is lost, allowing Gal4 to drive transgene expression, or homozygous wild-type **(D)** where tubGAL80 represses Gal4 activity.

2.2.2.4 Labelling Individual Cells

As well as generating GFP positive homozygous mutant clones individual well-spaced mutant cells can also be labelled on the back of the fly for a more detailed analysis of cell morphology. This is achieved by using a GAL4 driver under the control of the neuralised promoter, which restricts transgene expression to the sensory organ precursor cells and their progeny, which at the time of analysis are not yet fully committed to a precursor cell fate and so are representative of all cells within the epithelium (Cohen et al., 2010).

2.2.2.5 RNA interference (RNAi)

For loss of function studies P-elements can be inserted into the *Drosophila* genome which have a DNA inverted repeat, which, when transcribed forms an RNA hairpin. . Expression of a transgenic RNAi hairpin results in reduced expression of the target gene. dsRNA hairpins are cleaved by dicer into small interfering RNA (siRNA) of 20-22nt, which are transported to the RNA induced silencing complex (RISC). RISC unwinds the siRNA to produce small single stranded RNA, of which the anti-sense strand then binds with high specificity to target messenger RNA (mRNA) resulting in its degradation or loss of function. **Appendix B-C** lists all the RNAi lines used in this research (Dietzl et al., 2007).
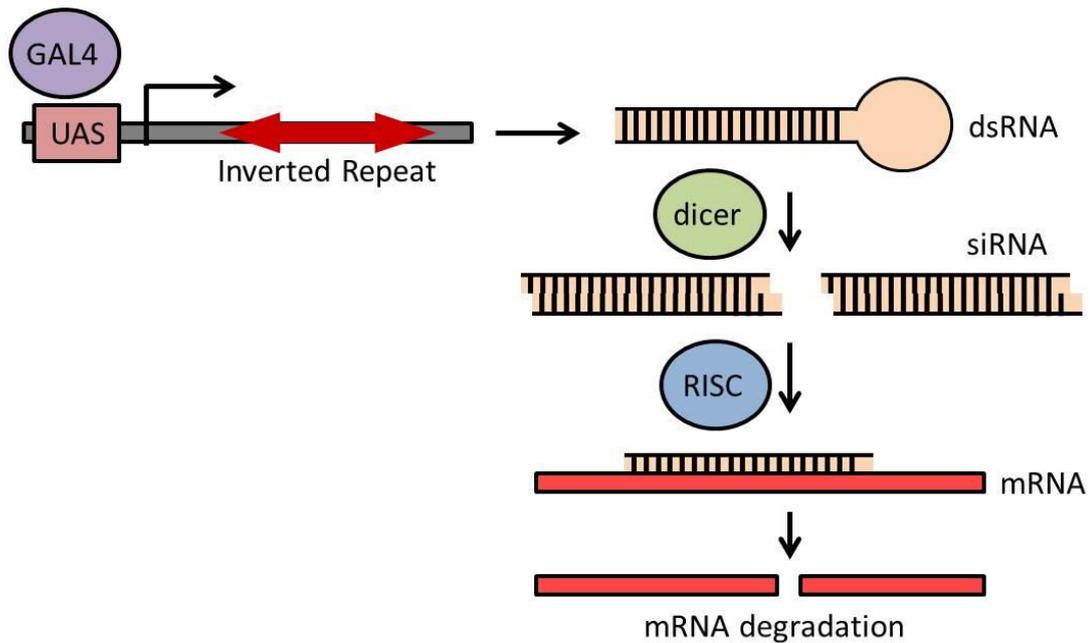
Figure 2.2 Expression of a UAS-RNAi transgene

UAS-RNAi transgenes can be expressed in a tissue specific manner using the Gal4/UAS system. Cleavage of dsRNA hairpins by dicer generates siRNA complementary to target endogenous mRNA. Binding of the complimentary strands results in the degradation of mRNA and subsequent gene silencing.

2.2.3 Fly genotypes for genetic experiments

2.2.3.1 Using MARCM to generate wild-type GFP labelled clones

Genotype imaged was: Ubx-flp; FRT40A, FRTG13/Tub-GAL80, FRT40A; Pnr-GAL4,UAS-Moe:GFP/MKRS

2.2.3.2 Using MARCM to generate homozygous mutant clones for scribble complex polarity genes

Genotypes imaged for *lgl-/-* mutant: Ubx-flp; *lgl4*, FRT40A/Tub-GAL80, FRT40A; Pnr-GAL4, UAS-Moe-GFP/MKRS. *dlg -/-* mutant: Ubx-flp, *dlg*, FRT19A/Tub-GAL80,

FRT19A;;Pnr-GAL4, UAS-Moe-GFP/MKRS. And *scrib-/-* mutant: Ubx-flp; Act-GAL4, UAS-Moe-GFP/IF; FRT82b, *scrib*/FRT82b, Tub-GAL80.

2.2.3.3 Using MARCM to specifically express UAS-transgene on the 3rd chromosome in otherwise WT clones.

Note UAS-transgene includes RNAi, Apoliner and P35. Genotype imaged: Ubx-flp/+; Tub-GAL80, FRT40A/FRT40A, FRTG13; Pnr-GAL4, UAS-Moe-GFP/UAS-transgene. Ubx-flp;*lgl[4]*, FRT40a/Tub-GAL80, FRT40A; Pnr-GAL4, UAS-Moe-GFP/UAS-transgene.

2.2.3.4 Using MARCM to specifically express UAS-transgene on the 3rd chromosome in *lgl* homozygous mutant clones

Genotype imaged: Ubx-flp;*lgl[4]*, FRT40a/Tub-GAL80, FRT40A; Pnr-GAL4, UAS-Moe-GFP/UAS-transgene.

2.2.3.5 Using MARCM to specifically express UAS-RNAi on the X in *lgl* homozygous mutant clones

Genotype imaged: UAS-RNAi/Ubx-flp; Tub-GAL80, FRT40A/ *lgl[4]*, FRT40a; Pnr-GAL4, UAS-Moe-GFP/+. Only female flies were used.

2.2.3.6 Labelling individual well-spaced cells in the notum

Genotype imaged: Neu-GAL4, UAS-moe:GFP/+

2.2.3.7 Generating negatively marked Dscam mutant clones

Genotype imaged: Ubx-flp/+ ; FRT42, *Dscam$^{21/23}$* /FRT42, Ubi-nls-GFP

2.2.3.8 Generating positively marked Dscam mutant clones

Genotype imaged: Ubx-flp/+ ; FRT42, *Dscam$^{21/23}$* /FRT42, Tub-GAL80 ;Pnr-GAL4,UAS-moe:GFP/+

2.2.3.9 Generating positively marked Dscam mutant clones with lgl RNAi

Genotype imaged: Ubx-flp/+ ; FRT42, *Dscam$^{21/23}$* /FRT42, Tub-GAL80; Pnr-GAL4,UAS-moe:GFP/*lgl*-RNAi

2.2.3.10 Expressing live Cadherin in Dscam mutant animals

Genotype imaged: Ubx-flp; FRT42, *dscam^{21}*/FRT42, Ubi-mRFP-nls; Neu-GAL4, UAS-Moe-GFP/Ubi-cad-GFP.

2.2.3.11 Blocking apoptosis in *lgl* mutant clones

Genotype imaged: Ubx-flp/UAS-p35; *lgl[4]*, FRT40a/Tub-GAL80, FRT40A; Pnr-GAL4, UAS-Moe-GFP/+. Only females imaged.

2.2.4 Mounting for live imaging

For live imaging, white pre-pupae (0 hours after puparium formation (APF)) were collected and incubated at 29°C. Once aged accordingly, pupae were immobilised ventral side down onto a microscopic slide and the notum was exposed by removing a region of the pupal case with tungsten forceps. Glass coverslips (0.13x0.17mm) were fixed to the anterior and posterior ends of the microscope slide generating a stage on which a long coverslip (24x50mm) coated with 10S Voltalef injection oil was placed creating an interface between the coverslip and exposed notum.

2.2.5 Dissecting *Drosophila* nota

Pupae were pinned dorsal side down in a silicon dish in cold PBS and the pupal casing was removed. The head and ventral regions were then excised using dissecting scissors and the animal flushed with PBS. The lateral regions of the remaining tissue were then removed leaving the central region of the notum where Pnr-GAL4 drives transgene expression (Simpson, 2007).

2.2.6 Molecular Biology Techniques

2.2.6.1 Immunostaining

Dissected nota **(Section 2.2.5)** were fixed in 4% para-formaldehyde (PFA) for 25mins at room temperature (RT) and washed in PBS-T. Tissue samples were incubated in a humidified chamber overnight at 4°C with primary antibody **(Table 2.8)**, then washed again in PBS-T and incubated for another hr in secondary antibody. Samples were left

overnight in PBS(50%):Glycerol(50%) at 4°C and mounted in mounting media (20% PBS, 80% Glycerol, 2% N-Propyl galate, 25mg/ml DAPI) the following day for imaging.

2.2.6.2 Mutagenesis and Cloning

pCFD3 plasmid **(see plasmid map in Appendix A)** was digested for 3 hrs at 37°C (1µg pCFD3, 5µl 2.1 RE buffer, 1µl BbsI made up to a final volume of 50µl with ddH$_2$0) and dephosphorylated by adding 0.5µl CIP for an additional 30mins. Digested pCFD3 was then run on a 1% agarose gel for 1hr at 120V. The band was visualised using a transilluminater and excised with a scalpel. The DNA was then purified using Qiagen gel extraction kit (Qiagen, 28704) following manufacturers protocol.

Oligos required for mutagenesis and cloning were designed using Primer3 software **(Table 2.4)**. Sense and anti-sense oligos were annealed and phosphorylated by setting up the following reaction: 1µl each oligo, 1µl 10x T4 ligation buffer, 0.5µl T4 PNK, and 6.5µl ddH$_2$0. The following thermocycler program was then used: 37^{o}C 30min, 95^{o}C 5min, ramped down to 20^{o}C at a rate of 0.1^{o}C per second. Once annealed and phosphorylated the oligos were then diluted 1:200 with ddH$_2$0.

For ligation, 50ng of previously digested pCFD3, 1µl annealed oligos (diluted 1:200), 1.5µl 10x T4 buffer, 1µl T4 ligase made up to a final volume of 15µl, was incubated at RT for 2hrs. The reaction was then heat inactivated at 65°C for 15 mins. 4µl of the ligation mix was then added to 50µl of chemically competent *E-coli* DH5α cells and kept on ice for 30mins. The cells were then heat-shocked for 45 seconds at 42°C and immediately placed back on ice for a further 2 mins. 250µl of SOC media (0.5% Yeast

Extract, 2% Tryptone, 10mM NaCl, 2.5mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM Glucose) was added to the mix and placed for 1 hr in a 37°C shaking incubator. 100μl of transformed cells were spread on LB agar plates containing 100μg/μl ampicillin and incubated overnight at 37°C.

Single colonies were used to inoculate 3ml of LB broth (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) containing 100μg/μl ampicillin and grown overnight at 37°C in an orbital shaker at 225rpm.

1.5ml of bacterial culture was centrifuged to generate a pellet and plasmid DNA was purified using Qiagen miniprep kit (Qiagen, 27104) following manufacturer's protocol. 100ng/μl of purified plasmid was finally sent for Sanger sequencing (DNA sequencing facility, University of Nottingham) using the T3 primer.

Qiagen PCR purification kit (Qiagen, 28104) was then used according to manufacturer's protocol to further purify samples before sending to the Cambridge fly facility, UK, for microinjection of DNA into flies where Cas9 on the X chromosome is under the nanos promoter for germline specific expression (Bloomington 54591).

2.2.6.3 PCR to check CRISPR mutagenesis

Surviving micro-injected flies received from the Cambridge fly facility were screened for vermilion (present in the pCFD3 backbone) and individually crossed to a balancer line containing CyO-GFP. Heterozygous CRISPR*/CyO-GFP flies were crossed back to

the balancer line and used to make a stock. Parent CRISPR*/CyO-GFP were simultaneously used to verify desired mutations. Genomic DNA was extracted from the flies by crushing individual whole flies in 50µl squishing buffer (10mM Tris-HCl pH8, 1mM EDTA, 25mM NaCl) containing 200µg/ml Proteinase K (added fresh). The following thermocycler conditions were then used:

37^oC/ 30min, 95^oC/ 2min and hold at 4^oC.

To amplify the target DNA, 5µl of DNA, 25µl OneTag 2x MasterMix, 1µl of each 10µM primer made up to a final volume of 50µl with dH_20 was prepared and run in the PCR machine with the following program:

94^oC/30secs, then 35 cycles of 94^oC/30secs, 65^oC/1min, 68^oC/50secs, finally 68^oC/5mins and hols at 4^oC.

Samples were run on a 2% agarose gel for 1hr at 120V to check for a single band at the appropriate PCR product size. Samples were then purified using the Qiagen PCR purification kit (Qiagen, 28104) following manufacturers protocol and eluted in ddH_2O before being sent for sequencing at the DNA sequencing facility, University of Nottingham, using the forward CG1 primer **(Table 2.6).**

2.2.6.4 RNA extraction

To extract RNA from cell lines, cells were pelleted **(Section 2.3.3)** and lysed by pipetting in 0.75mL Trizol reagent. Samples were then incubated for 5mins at RT then shaken vigorously for 15 seconds following addition of 0.2mL chloroform. To separate the RNA from the phenol-chloroform phase, samples were centrifuged for 15mins at

12,000 rpm, 4°C. The aqueous layer containing RNA was then isolated. 0.5mL 100% Isopropanol was then added to the aqueous layer and incubated for 10mins at RT before centrifuging a further 10mins at 12,000 rpm, 4°C. The supernatant was removed leaving the RNA pellet which was washed with 1mL 75% ethanol and centrifuged for 5mins at 7,500rpm, 4°C. Excess wash was removed and allowed to air-dry for 10mins. The pellet was re-suspended in 50µl RNase-free water and incubated for 15mins at 55°C. RNA concentration was measured using a Nanodrop2000.

2.2.6.5 cDNA synthesis

cDNA synthesis of RNA extracted from cell lines **(section 2.2.6.4)** was carried out in order to perform RT-PCR . For DNase treatment 1µg RNA, 1µl 10x buffer and 1µl DNase1 made to a final volume of 10µl was incubated at 37°C for 30min. The reaction was then stopped by adding 1µl EDTA and incubated at 65°C for 10mins. For cDNA synthesis 1µl OligoDT, 1.5µl dH20, 4µl 5x reaction buffer, 0.5µl RNase inhibitor, 2µl 10mM dNTP and 1µl RevertAid were added to the mix and incubated at 42°C for 60mins then 70°C for a further 10mins.

2.2.6.6 RT-PCR

Real-time RT-PCR was used in order to determine the relative mRNA levels of target genes in each cell line following transfection using SYBER-green detection. For each sample 1µl cDNA, 12.5µl IQ SYBR master mix, 2.5µl each of each primer, was made up to a final volume of 25µl in a 96-well plate. A C1000 thermal cycler CFX96 RT system was used with the following conditions:

95°C/10min then 40cycles of 95°C/15secs, 60°C/1min, 72°C/30secs then 72°C/10min.

2.3 Cell Culture

2.3.1 Cell maintenance

U87 and MDA-MB-231 cells were cultured in DMEM supplemented with 10% (v/v) foetal bovine serum (FBS), and 1% penicillin/streptomycin (P/S). Cells were grown in T75 culture flasks at 37°C in a 5% CO_2 atmosphere. Cells were routinely passaged when 80% confluent. Growth medium was removed and cells washed in Hanks Balanced Salt Solution (HBSS). Cells were incubated for 5mins at 37°C with 0.05% Trypsin in HBSS to detach cells. Pre-warmed growth medium was then added to deactivate the trypsin. Desired numbers of cells were seeded into new T75 flasks containing appropriate amount of media.

2.3.2 Cell counting

To count cells, 10µl of cell suspension was applied to a haemocytometer BS 748 (Hawksley, UK). Cells were counted in all 4 squares in the corners of the haemocytometer. As the number of cells in each square equates to the number of cells x 10^4/ml the following equation was used to count and seed appropriate number of cells.

$$Cells\ per\ ml = \left(\frac{Total\ number\ of\ cells\ from\ 4\ corner\ squares}{4}\right) x\ 10^4$$

2.3.3 Cell storage and recovery

Growth medium was removed from a confluent T75 flask of cells and washed in HBSS. Cells were trypsinised **(Section 2.3.1)** and re-suspended in growth medium.

Cells were then pelleted by centrifuging for 5mins at 800rpm and re-suspended in freezing medium consisting of 90% fetal calf serum (FCS) and 10% dimethylsulfoxide (DMSO). Cells were counted as previously described (section 2.3.2) and diluted to 1×10^6 cells/ml in freezing medium. 1ml of cell suspension was transferred to cryovials and stored at -80°C in a Mr Frosty freezing container for at least 24hr. For long term storage tubes were transferred to liquid nitrogen.

2.3.4 Transfections

5×10^4 cells were seeded in single wells of a 6-well plate in antibiotic-free medium and incubated at 37°C with 5% CO_2 for 24hrs until 60-80% confluent. 5µM siRNA stock solution was diluted 1:20 with serum-free medium and added to DharmaFECT diluted 1:100 with serum-free medium and incubated 20mins at RT. Transfection medium was then diluted with antibiotic-free medium to a final siRNA concentration of 25nM and total transfection volume of 2ml per well. Transfection medium was added drop-wise to each well after removing the culture medium and incubated at 37°C with 5% CO_2 for 48hrs. Culture medium was removed from wells and cells were washed with HBSS. Cells were trypsinised and re-suspended in growth medium as previously described **(Section 2.3.1)** and centrifuged for 5mins at 800rcf to harvest a cell pellet for downstream application (RT-PCR – **Section2.2.6.4**, invasion assay- **Section 2.3.5**)

Percentage knockdown was calculated using the Δ ΔCq method:

ΔCq1 = Cq(Target) – Cq (Ref), ΔCq2 = Cq(Scram) – Cq (Ref)

ΔCq expression = $2^{-\Delta Cq}$

Δ ΔCq = ΔCq1exp/ ΔCq2exp

$$\%KD = (1- \Delta\ \Delta Cq)*100$$

(Target represents the gene being knocked down, Ref represents the reference gene (GAPDH) and Scram is the non-targeting control)

2.3.5 Invasion Assay

Invasion assays were carried out in invasion chambers containing inserts with an 8μm pore size polycarbonate membrane coated in an ECL cell attachment matrix (Merk Millipore) in single wells of a 24-well plate. Prior to performing the assay the matrix was diluted to 20μg/mL with serum-free medium and added to the inserts to a final concentration of 4μg and incubated for an hr at $37^{o}C$. 48 hours post-transfection cells were washed with HBSS and incubated for 10mins at $37^{o}C$ with HBSS with 0.05% trypsin and 25mM HEPES. Serum-free medium containing 5% BSA (quenching medium) was then added to each well and cells were centrifuged at 1500rpm for 10mins to produce a pellet which was re-suspended in 1ml quenching medium. Cells were counted (section 2.3.2) and bought to a final volume of $0.5x10^{6}$ cells/mL. 250μl of cell suspension was added to each insert and 500μl of culture medium with 10% FBS was added to the lower chamber. Plates were incubated for a further 48hrs at $37^{o}C$ with 5% CO_2. Following incubation the cell suspension was removed from the upper chamber and the insert was placed into a new well containing detachment solution (ECM554, Merk Millipore). Cells that had invaded through the matrix were allowed to detach from the membrane for 30mins at $37^{o}C$. CyQuant dye was diluted 1:75 with lysis buffer (ECM554, Merk Millipore) and added to wells containing the detached invading cells and incubated for 15mins at RT. Samples were transferred to a 96-well plate and read on a fluorescent plate reader using 480/520 nm filters and

optimal gain. To image invading cells, instead of placing inserts in detachment solution, the inserts were washed with ddH_2O and the upper-side of the membrane was blotted with cotton buds to remove non-invasive cells. The membrane was then stained with 0.1% crystal violet solution for 2 mins and washed again with ddH_2O prior to imaging.

2.4 Microscopy and Image Analysis

For screening, imaging was performed using an inverted Leica SP2 confocal laser scanning microscope with 40x oil immersion objective and Ar/HeNE (488nm) laser. Clones were imaged by generating z-stacks of 1µm planes from the cuticle to basal lamina at a resolution of 1024x1024. For other experiments requiring multi-channel acquisitions a Zeiss LSM880 or LSM exciter confocal laser scanning microscope was used with 40x oil immersion objective, optimal resolution and lasers 405nm, 488nm, 543nm and 633nm. Fiji (ImageJ) software was used to analyse images and Prism (GraphPad) was used to generate graphs and carry out statistical analysis. Where student t-tests were performed, significance was shown as: $p<0.05^*$, $p<0.01^{**}$, $p<0.001^{***}$, $p<0.0001^{****}$. Additionally ApoE software was used to analyse chromatograms, and R Studio for hierarchical clustering.

CHAPTER 3

Pilot Screen

Chapter 3 Pilot Screen

3.1 Introduction

As discussed in **Section 1.3** functional alterations to epithelial polarity protein networks are crucial to the pathology of cancer. The polarity proteins Scribble (Scrib), Discs-Large (Dlg) and Lethal giant larvae (Lgl) which form the basolateral Scribble complex, genetically interact to cooperatively regulate cell polarity, junction integrity and cell proliferation. The Scribble polarity module maintains epithelial apicobasal polarity, defining distinct cellular domains and positioning of adherens junctions (AJ) and septate junctions (tight junctions in mammalian cells) (Bilder & Perrimon, 2000).

First identified in *Drosophila*, members of the Scribble complex contain highly conserved protein-protein interaction domains, including SH3 and PDZ domains, which enable them to recruit various protein networks to cooperatively regulate polarity (Bilder, 2003; Schneiderman & Gateff, 1967). Additionally, the correct localisation of these proteins is vital for their function in maintaining apicobasal polarity **(Section 1.4)**. In mutants for Scrib/Dlg/Lgl AJs become disorganised through the mislocalisation of apical determinants, resulting in the loss of communication between cells and integrity of the epithelium (Bilder, 2003; Tanentzapf & Tepass, 2003).

Members of the Scribble complex have long been classified as neoplastic tumour suppressors and are linked to the progression of mammalian tumours (Bilder & Perrimon, 2000). There are various human orthologues of these polarity genes which

play a role in the formation of epithelial malignancies where expression is either downregulated or lost. Gene expression studies have shown that human Dlg1 binds to known tumour suppressor genes APC and pTEN, and is implicated in HPV-induced malignancies (S. S. Lee, Weiss, & Javier, 1997; Matsumine et al., 1996; Valiente et al., 2005). Similarly, the mammalian orthologue of Scribble is associated with the APC/β-catenin pathway. While human Lgl1 has been linked to a variety of cancers, including breast and colon adenocarcinomas, glioblastomas, and melanomas (Grifoni et al., 2004; Kuphal et al., 2006; Schimanski et al., 2005). The cancer phenotypes of *lgl*/*dlg*/*scrib* mutants can be rescued by their mammalian homologues, suggestive of their tumour suppressive function in human cells, and implicating them as potential targets of anti-cancer therapies. In both transgenic mouse models and human cancers these genes commonly show a reduced expression, mislocalisation or depletion of protein products. Significantly, the human homologue of Lgl is downregulated or lost in a variety of cancers including breast cancer (76%), colon carcinomas (75%) and lung cancer (63%).

Scrib/Dlg/Lgl are widely used in loss-of-functions (LOF) studies when investigating the transformation of epithelial cells into their malignant derivatives. Studies in *Drosophila* larvae have demonstrated that mutations causing LOF phenotypes in any one of the polarity proteins forming the Scribble complex results in polarity defects, aberrant cell architecture, cellular hyper-proliferation and invasion, all of which are characteristics of mammalian carcinomas (Bilder & Perrimon, 2000).

Loss of *lgl* in the *Drosophila* eye results in mislocalisation of the hippo kinase, an integral component of the highly conserved Salvador/Warts/Hippo (Hpo) pathway which prevents phosphorylation of the transcriptional activator Yorkie (Yki). This leads to upregulation of the pathway which is a known regulator of cell proliferation and survival. Targets of the Hpo pathway, including Cyclin E are upregulated in *lgl* mutant tissue, directly linking defects in cell polarity, hyperproliferation and evasion of cell death in the neoplastic tissue **(Fig. 3.1)** (Grzeschik, Amin, Secombe, Brumby, & Richardson, 2007; Grzeschik, Parsons, Allott, Harvey, & Richardson, 2010). Additionally, when neoplastic *lgl* mutant brain tumours from *Drosophila*, labelled with lacZ encoded β-galactosidase, were grafted onto the abdomen of a wild-type fly, lacZ positive cells were later detected in the ovaries. As *Drosophila* ovaries are enclosed within a non-porous epithelial sheet with two basement membranes, these results conclude that *lgl* mutant cells are also capable of invading and metastasising (Beaucher et al., 2007).
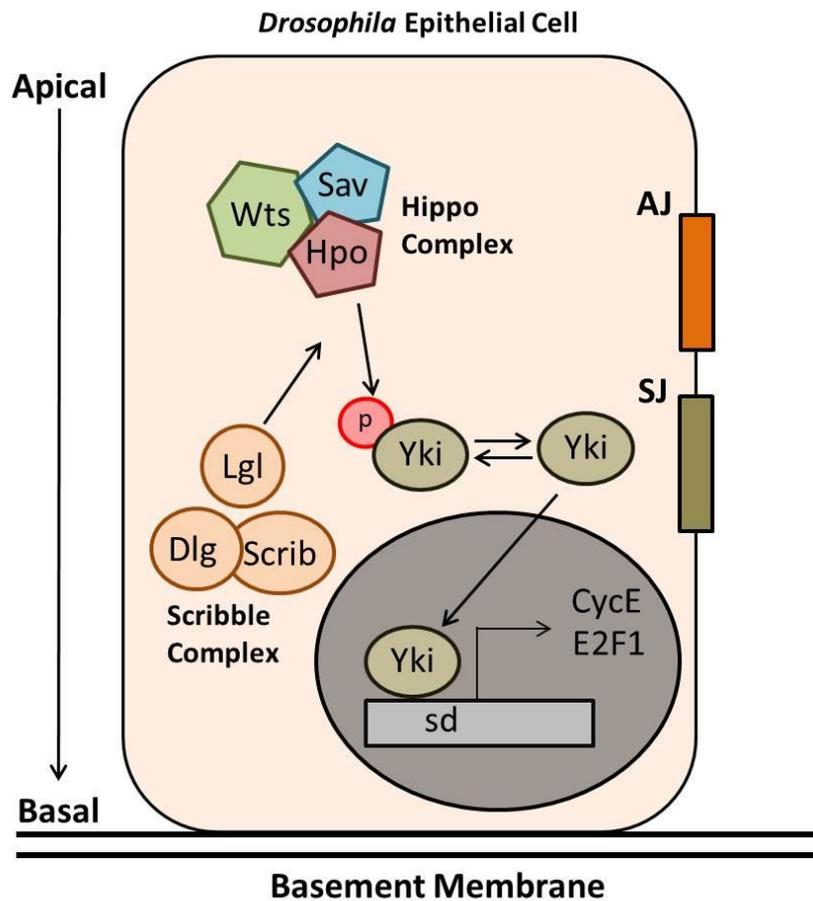
Figure 3.1 Lgl and the Hippo pathway

Lgl facilitates the interaction in which Hippo kinase phosphorylates the transcriptional activator Yorkie (Yki) preventing transcription of genes involved in cell growth. Loss of Lgl therefore prevents inactivation of Yki allowing it to translocate to the nucleus and transcriptionally activate target genes such as cyclin E.

The formation of epithelial tumours is largely dependent on the cooperation between mutations in tumour suppressor genes and oncogenes. In 2003, a model of tumour progression in the larval eye antennal imaginal disc was generated in the Xu and Richardson labs to study the development of metastatic tumours (Pagliarini & Xu, 2003). They showed that over-expression of activated Ras (RasV12) in the eye imaginal disc generates large non-invasive tumours which, when combined with a homozygous *scrib* mutation (*scrib-/-*) resulted in the metastasis of GFP positive

RasV12; *scrib-/-* cells to the ventral nerve cord and other tissues. Moreover, this metastatic behaviour could be inhibited by over-expressing Scrib. As *scrib-/-* flies alone produce non-invasive tumours, Pagliarini and Xu demonstrated that both mutations in scrib and over-expression of Ras were required for invasion. Similarly, during a clonal analysis of the eye disc (using MARCM), *scrib-/-* clones were also shown to require cooperative interactions with oncogenic Ras or Notch in order to exhibit clonal overgrowth (Brumby & Richardson, 2003). Additionally, in the absence of oncogenic signalling, *scrib-* clones generated during the 3rd instar larval stage of development had significantly diminished in size following pupation. Interestingly, *scrib-/-* clone size was significantly increased when Jun-N-terminal Kinase (JNK) was blocked, suggesting that the surrounding WT tissue was able to outcompete the mutant clones by initiating JNK-mediated apoptosis, resulting in a proliferative advantage over the *scrib-/-* tissue (Brumby & Richardson, 2003). Similarly, Chen et al 2012 showed that when clones of cells mutant for *scrib* were generated, they were capable of forming large neoplastic tumours in animals where all cells within the imaginal discs were mutant. However when *scrib-* cells were surrounded by wild-type cells, the mutant cells were eliminated by apoptosis through a JNK-dependent mechanism (C. L. Chen et al., 2012). In these examples, overgrowth is seen when the whole tissue is mutant, but cell competition prevents these when clones of mutant tissue are generated in an otherwise WT tissue. This is mainly seen in the imaginal discs which is a highly proliferative epithelium. As originally characterised in *Drosophila*, cell competition describes the apoptotic elimination of cells surrounded by a population of more rapidly proliferating cells. Local cell-cell interactions between two populations of cells that differ in the expression of a particular protein lead to

the elimination of 'loser' cells through apoptosis, whilst surviving 'winner' cells proliferate to replace the outcompeted cells (C. L. Chen et al., 2012; Morata & Ripoll, 1975; Simpson & Morata, 1981). Therefore cell competition between mutant and wild-type cells acts as a tumour suppressing mechanism.

In previous *Drosophila* studies, the wing and eye imaginal discs have served as the prime model for studying cancer epithelial cells; however using this system it is difficult to image the behaviour of individual cells, thus precluding a more detailed analysis of tumour development. To overcome these limitations Baum and Georgiou (Georgiou et al., 2008) (Georgiou & Baum, 2010) developed an *in vivo* system which could be used to generate and visualise both clones of mutant tissue and individual well-spaced epithelial cells within the dorsal thorax (notum) of the fly, thus allowing mutant cells to be followed in the hours/days following tumour induction. This system has already enhanced our understanding of junction integrity and the cooperation of apical polarity proteins with Rho-GTPases in controlling epithelial cell morphology and dynamic protrusion formation. We now aim to use this system to help improve our understanding of the early events in tumourigenesis and tumour progression.

As mentioned previously **(Section 1.5)** the aim of my PhD was to carry out a large-scale genetic screen using the *Drosophila* notum as a model for tumour progression, and from this to identify and characterise genes that significantly perturb tumour progression. To do this RNAi would be used to knock-down genes specifically within neoplastic tumours (mutant clones) on the back of the fly to see how the gene of

interest affects tumour progression. The aim of this chapter was to initially determine the genetic background with which to carry out the enhancer/suppressor screen. As members of the Scribble complex have been shown to act as neoplastic tumour suppressors, these would be investigated as potential genetic backgrounds for the screen. The second aim of this chapter was to carry out cooperative experiments with Notch in the genetic background to validate our novel *in vivo* assay. Finally, to carry out a pilot screen on a selection of well-characterised cancer associated genes with the following main objectives:

- Optimise conditions for the full RNAi-screen

- Verify that RNAi-mediated knockdown was working as expected

- Verify that two independent RNAi lines for the same gene give the same phenotype

- Verify that phenotypes seen in the pilot screen correspond with expected phenotypes for well characterised cancer-associated genes.

- To generate a database to qualitatively analyse a range of tumour phenotypes for each gene studied relative to the mutant background alone.

3.2 Results

3.2.1 Scribble complex mutants generate tumours of varying phenotypes.

Using the MARCM system clones of GFP-marked cells homozygous mutant for *scrib*, *dlg* and *lgl* were generated in the notum of the fly. The behaviour of transformed cells was then observed in high temporal and spatial resolution in living pupae (20-24APF). The GFP fusion protein Moe-GFP (consisting of the actin binding domain of Moesin fused to GFP) was utilised to target GFP to the actin cytoskeleton, allowing for a detailed analysis of epithelial cell morphology.

Wild-type columnar epithelial cells form an organised monolayer on the back of the fly **(Fig. 3.2 A)** and cells are characteristically polygonal in shape at the apex. In comparison, *lgl*, *dlg* and *scrib* mutant clones showed significant alterations to clone size and cell morphology **(Fig. 3.2 B-D)**. Compared to WT, *lgl* homozygous mutant clones were slightly larger (ns), however clones mutant for *dlg* or *scrib* were significantly small or non-existent (p<0.0001) **(Fig. 3.2 E)**. 80% of animals expressing a mutant allele for *dlg* presented no clones in the notum at 20hrs APF. Although *dlg* and *scrib* mutant clones were consistently small in size it is important to note that when pupae were screened at 0APF large clones could be seen using a Leica GFP microscope. Visually, all three mutant phenotypes lost their typical polygonal shape (indicating effects on apicobasal polarity) resulting in epithelial multi-layering and hyper-proliferation **(Fig. 5.2 B'-D')**. This morphological alteration was notably more severe in *dlg* and *scrib* mutants. A variety of other cancer-associated phenotypes

were observed and subsequently quantified. Cells undergoing division could be identified through their characteristic cortical cell rounding, and quantified as another indicator of hyper-proliferation. For evidence of these cells dividing into two daughter cells **see Fig. 3.3** in which UAS-His-RFP was used to label chromosomes. In this transgene, RFP is linked to the coding region of histone 2A, labelling chromosomes with RFP, so they can be visualised during time-lapse movies of cells undergoing cell division. As seen, *dlg* and *scrib* had significantly fewer (p<0.0001) visible cells undergoing cell division compared to *lgl-/-* and WT **(Fig. 3.2 F)**. *lgl-/-* tumours also presented an invasive phenotype, whereby individual GFP-positive cells could be seen beneath the epithelium **(Fig. 3.2 G)**. Another phenotype associated with loss of apicobasal polarity is the length of the cell, from the cell apex to the basal protrusions. WT cells are typically 10-13μm long, whereas *dlg-/-* and *lgl-/-* showed an increase in cell length, most significantly was *lgl-/-* (p=0.0002), whereby cells were on average 16μm long **(Fig. 3.2 H)**.

Pnr-Gal4 drives expression in a central stripe of the dorsal thorax epithelium, and was used together with MARCM to positively label WT, *lgl-/-* and *dlg-/-* mutant clones with GFP. However due to *scribble* and *Pnr* being on the same chromosome, an Actin-Gal4 driver was used to label *scrib* mutant clones. Interestingly, in *scrib* mutant cells, Moe-GFP appeared to be mis-localised and was observed throughout the cell, including the nucleus **(Fig. 3.2 C')**.

From this data, it could be inferred that *lgl-/-* was the ideal genetic background for a enhancer/suppressor screen as large mutant clones could be generated and mutant

tissue was slightly multi-layered and invasive. Therefore both positive and negative effects on this phenotype could be easily detected.
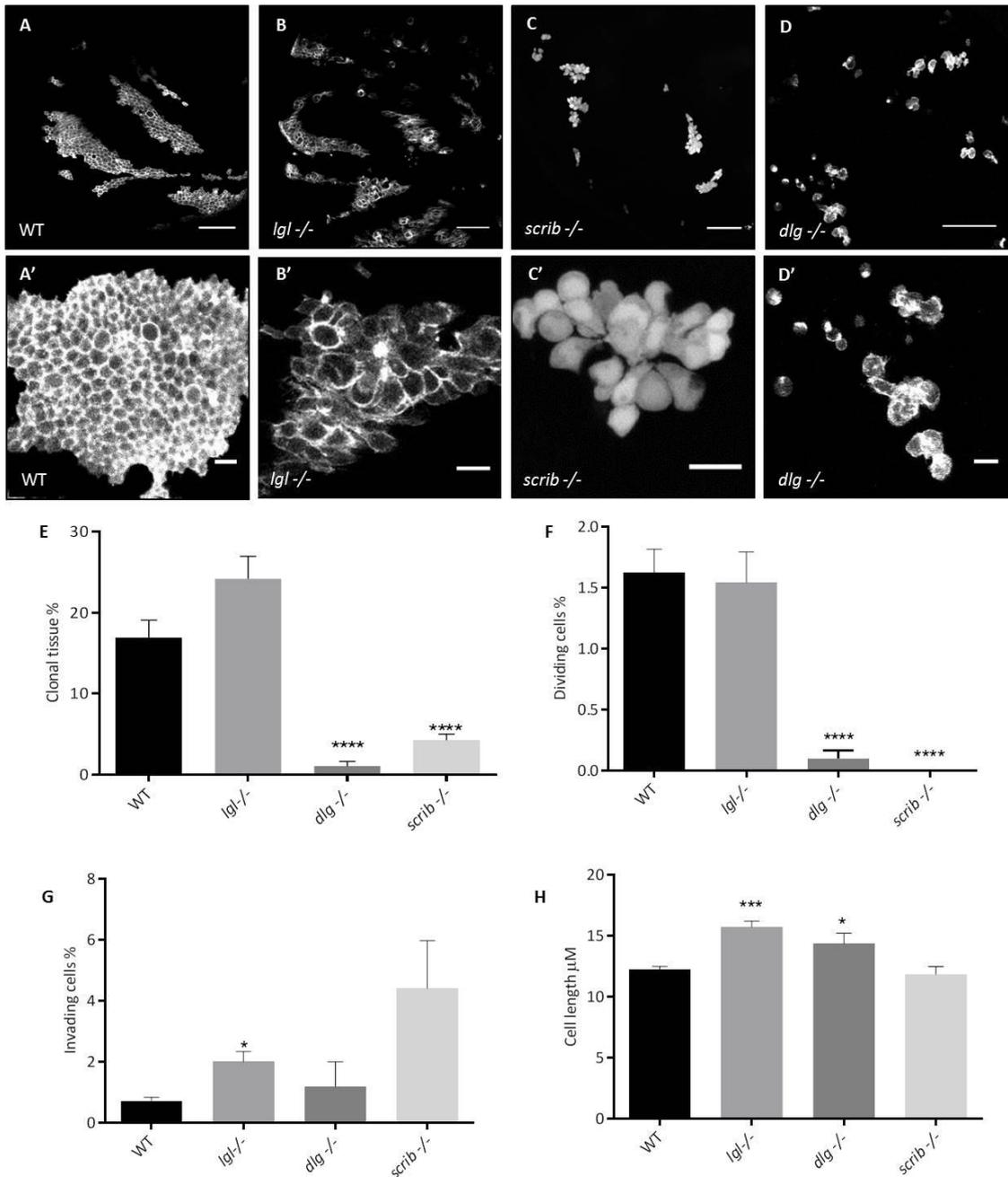
Figure 3.2 Homozygous Scribble complex mutants

Confocal images of the *Drosophila* notum **(A-D)** GFP positive clones of wild-type tissue in the notum **(A,A')**, and clones homozygous mutant for the tumour suppressor gene *lgl* **(B,B')**, *scrib* **(C,C')** and *dlg* **(D,D')**. Scale bars = 50μm (A-D), or 10μm (A'-D'). **(E-H)** Graphs represent the mean percentage of notum covered in GFP-labelled clonal tissue **(E)**, the percentage of cells undergoing cell division **(F)**, the percentage of invading cells detected beneath the epithelium **(G)** and epithelial cell length, in μm **(H)**. Graphs indicate the mean with error bars representing standard error (SEM). Each genotype was compared to WT using a students' t-test and considered significant when p<0.05. WT (n=10 animals), *lgl* (n=32), *dlg* (n=14), *scrib* (n=13).
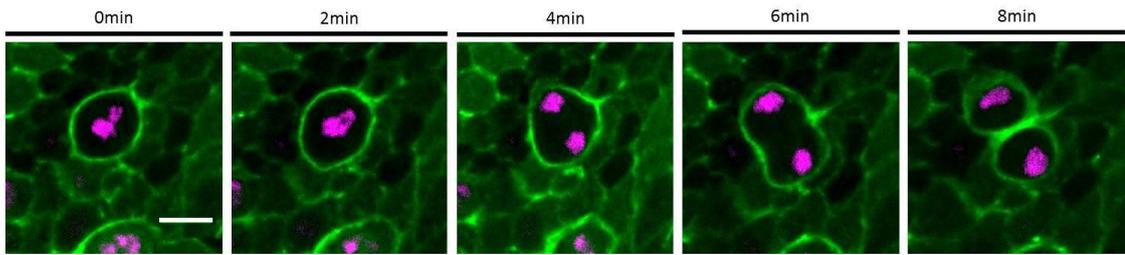
Figure 3.3 Time-lapse of mitotic wild-type cell

Clones of GFP-marked cells expressing Ubi-His-RFP (magenta) in wild-type background. Tissue expresses the transgene His-RFP in which RFP is linked to the coding region of histone 2A. Snap shots taken from an 8min time lapse. Scale bar = 50μm.

3.2.2 Oncogenic Notch cooperates with mutant *lgl* in the *Drosophila* notum

As mentioned **(Section 3.1)**, it has been shown in other *Drosophila* systems that the scribble polarity tumour suppressor gene cooperates with oncogenes to enhance the malignant phenotype (Brumby & Richardson, 2003; Pagliarini & Xu, 2003).

To test that *lgl* has the same cooperative effects in this system, an activated allele of the Notch oncogene (N[intra]) was overexpressed specifically within the *lgl* homozygous mutant tissue. Through diverse effects on differentiation, survival and proliferation, Notch signalling is a key player in normal development. Truncated forms of Notch can act as potent oncogenes and are seen in a large number of cervical and colon carcinomas.

Overexpressing Notch in an otherwise WT animal generated clones relatively similar to WT, with the epithelium remaining highly structured and organised **(Fig. 3.4 A)**. As seen previously **(Fig. 3.2 B')** *lgl* homozygous mutant tissue was multi-layered and invasive, however these phenotypes were dramatically enhanced when oncogenic Notch was simultaneously expressed specifically within the *lgl* mutant tissue **(Fig. 3.4 B-F)**. The cooperation between the *lgl* tumour suppressor mutation and oncogenic Notch resulted in significantly more invasive tumours when compared with lgl-/- (p=0.0233), WT (0.0009) and N[intra] (p=0.0354) genotypes. Additionally, cells were seen growing on-top of one another in multi-layers **(Fig. 3.4 B)**. Also, significantly more cells could be seen undergoing cell division when compared to *lgl*-/- (p=0.0461), WT (p=0.0045) and N[intra] (p=0.0009) **(Fig. 3.4 E)**.

These results illustrate that *lgl-/-* cooperates with the Notch oncogene to enhance the neoplastic cancer phenotype in the *Drosophila* notum.
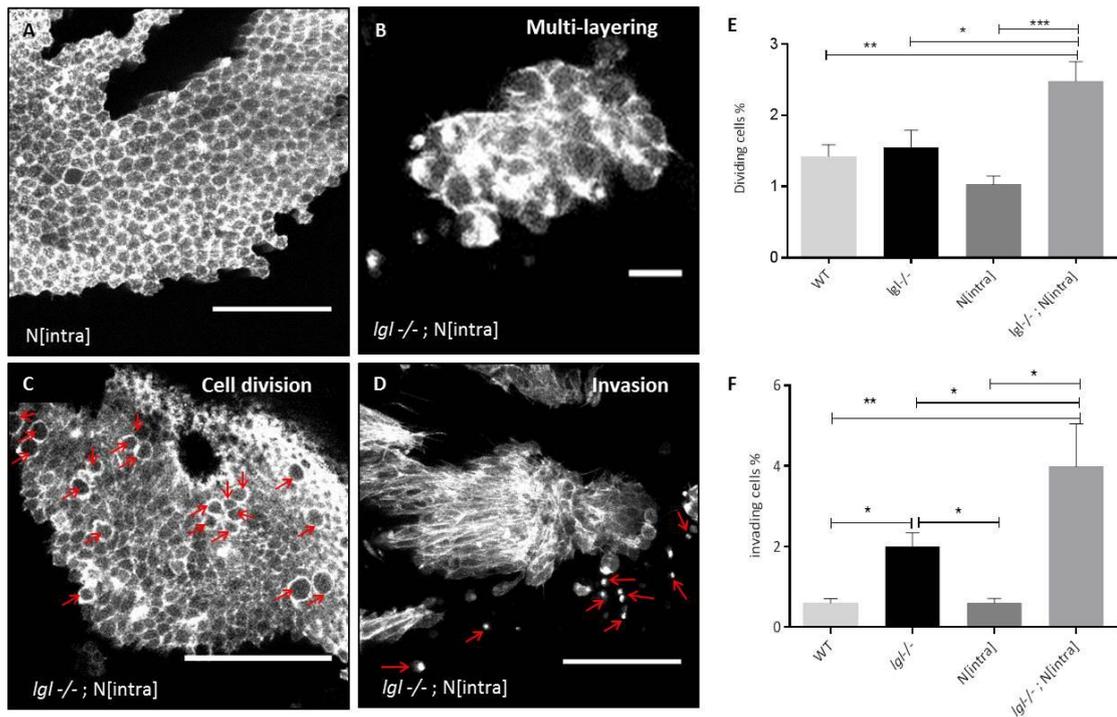
Figure 3.4 Cooperation between oncogenic Notch and mutant *lgl* in the notum

(A-D) Confocal images of mutant clones in the notum. GFP-marked clones expressing oncogenic Notch (N[intra]) **(A)** or simultaneously expressing N[intra] in homozygous *lgl* mutant clones **(B-D).** Scale bars = 50μm (A, C, D) or 10μm (B). Red arrows indicate cells undergoing cell division **(C)** and invading cells beneath the epithelium **(D)**. **(E-F)** The mean and standard error for each genotype was plotted whereby the number of dividing/invading cells is illustrated as a percentage of the total number of mutant cells within the notum for each animal. WT (n=10 animals), *lgl-/-* (n=32), N[intra] (n=7), *lgl-/-*; N[intra] (n=11).

3.2.3 Pilot screen

During an initial pilot screen a wealth of candidate genes previously implicated in cancer were studied to ensure the system was working efficiently. These candidates included various oncogenes, tumour suppressor genes, MMPs, and regulators of cell morphogenesis, with a range of biological functions **(Fig. 3.5 A)**. In addition, genes which are not normally expressed in the notum were knocked down; acting as negative controls. UAS-RNAi constructs were used to knockdown candidate genes specifically within the *lgl* homozygous mutant tissue. Where possible two different RNAi constructs were used for each gene. In total, the pilot consisted of 67 RNAi lines targeting 46 well-known genes **(Appendix B)**.

During the pilot screen a database of cancer-associated phenotypes was set up whereby each RNAi was qualitatively analysed. For each animal, a range of phenotypes was given a score relative to the *lgl* mutant phenotype alone. The phenotypes scored included: clonal tissue coverage, invading cells, dividing cells, apex size, multi-layering, basal protrusions, and cell length. Qualitative analysis for each RNAi was carried out blind by at least two scientists in the Georgiou lab due to the subjective nature of the analysis. Thresholds were then applied to each phenotype to identify 'hits' which appeared to dramatically enhance or suppress the tumour phenotype. **Fig. 3.5 B-G** outlines the biological processes of the genes which were identified as 'hits' for various categories. Genes identified for notably affecting the size of clonal tissue were shown to have a known role in regulating proliferation and growth, similar to those genes with strong multi-layering phenotypes. Highly invasive genes picked up in the screen were mostly involved in polarity, cell migration

and adhesion, whereas hits for apex defects and elongated/defected basal protrusions often came from knocking down known actin regulators.
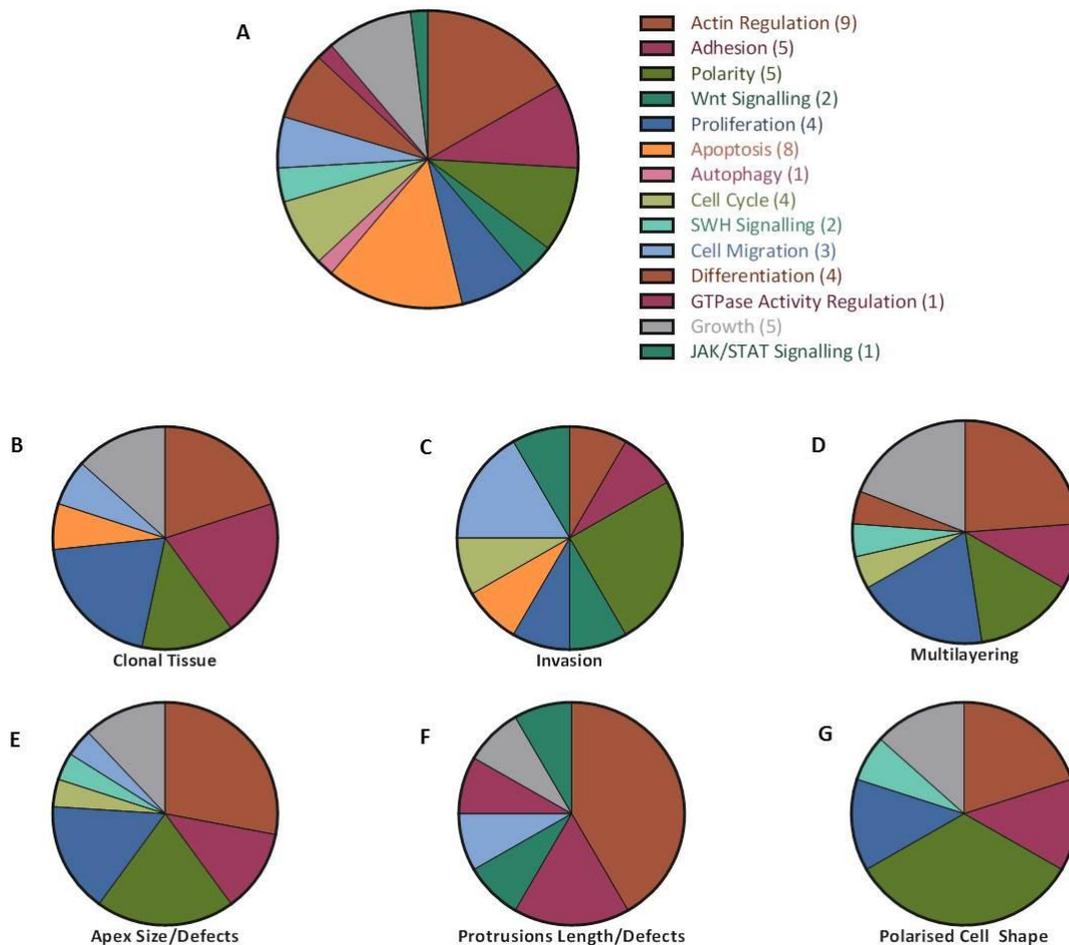
Figure 3.5 Biological functions and phenotypes of pilot screen genes

Break down of the main biological functions of all the genes studied in the pilot screen **(A).** Breakdown of the biological functions of the genes identified as 'hits' for various phenotypes, including percentage of mutant tissue within the notum **(B)**. Hits for invasion **(C)**, multilayering **(D)**, apex size/defects **(E)**, protrusion length/defects **(F)** and loss of polarised cell shape **(G)**. n=46 (a few genes listed under more than 1 main biological function, see **Appendix B** for full list of pilot genes and their biological functions).

Throughout the pilot screen a range of phenotypes were observed, including hyper-proliferation, multi-layering, invasion, apoptosis and effects on subcellular structures (junctions, microvilli, basal protrusions) **(Fig. 3.6).** For example, knocking down APC in the *lgl* mutant background produced highly invasive multi-layered clones with basal bundles and long thick protrusions **(Fig. 3.6 A-A")**, whilst knocking down TSC1 resulted in phenotypes indicative of hyper-proliferation such as clonal tissue overgrowth and increased cell division **(Fig. 3.6 B-B').** Misshapen, blebbing dividing cells were also observed with Moesin RNAi **(Fig. 3.6 C).** Knocking down p53 in the *lgl* mutant background drastically affected apex integrity, and clones were multi-layered and highly invasive **(Fig. 3.6 D-D")**. Whilst loss of E-cadherin drastically reduced the size of clonal tissue in the notum and altered cell morphology. As well as affecting the length of basal protrusions, often these actin rich membrane extensions had other morphological changes such as branching **(Fig. 3.6 E-E").** Negative controls failed to give drastic phenotypes **(see Appendix G (CD)** for scoring system of all genes).

A range of the observed phenotypes where then quantified for a variety of different genotypes **(Fig. 3.7)**. When compared to the *lgl* mutant phenotype alone, it could be seen that loss of APC (p=0.0004), E-cadherin (p=0.0124) and pTEN (p=0.0475) in the mutant background significantly enhanced the invasive phenotype of the tumour **(Fig. 3.7 A).** TSC1 RNAi however had a negative effect on the invasive properties of the tumour (p=0.0450). Knocking down E-cad, APC and P53 all produced clones of mutant tissue significantly smaller than that seen in *lgl* mutants alone (p=0.0161, 0.0242, 0.0343 respectively) **(Fig. 3.7 B)**. Cell length was also significantly increased with APC, TSC1, and pTEN RNAi (p=0.0006, 0.0068, 0.0022 respectively) **(Fig. 3.7 D)**.

It is also important to note that the quantitative analysis of various phenotypes reflected the qualitative analysis previously carried out. For example, when scoring for invasion animals where given a score of whole numbers from -2 to 2, where an invasive phenotype comparable to *lgl* alone was given a score of 0; increased invasion was given a positive score and decreased invasion a negative score. Across all the animals where APC was knocked down the average score for invasion was 0.65 which exceeded the >0.5 threshold set to identify hits for invasion. Similarly quantification of invasion in these animals showed a significant difference compared to *lgl* alone. Therefore it could be confidently assumed that qualitative analysis of genes would be sufficient to pick up 'hits' from the screen which could be further quantified and characterised.
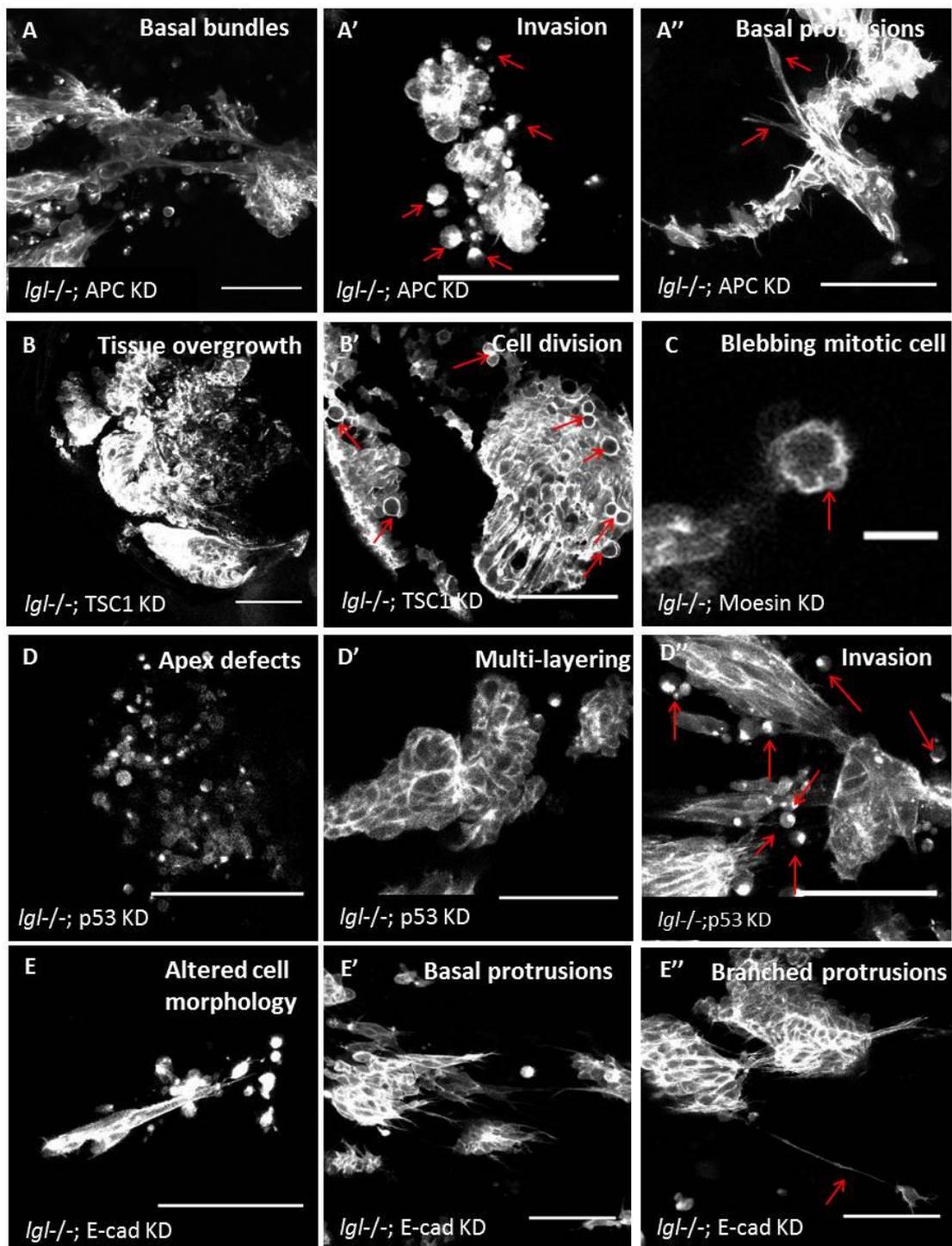
Figure 3.6 Cancer-associated phenotypes observed in the pilot screen

lgl homozygous mutant tissue expressing UAS-RNAi hairpin constructs for APC **(A-A'')**, TSC1 **(B-B')**, Moesin **(C)**, P53 **(D-D'')**, and E-cadherin **(E-E'')**. Scale bars = 50µm in all panels except **(C)** where the scale bar is 10µm. Various cancer-associated phenotypes were observed: basal bundles **(A)**, polarised invading cells **(A',D'',E)**, long thick filipodia **(A'')**, clonal tissue overgrowth **(B)**, large numbers of cells undergoing cell division **(B')**, blebbing of mitotic cells **(C)**, apex defects **(D)**, multi-layering **(D')**, and long branched basal protrusions **(E'-E'')**.
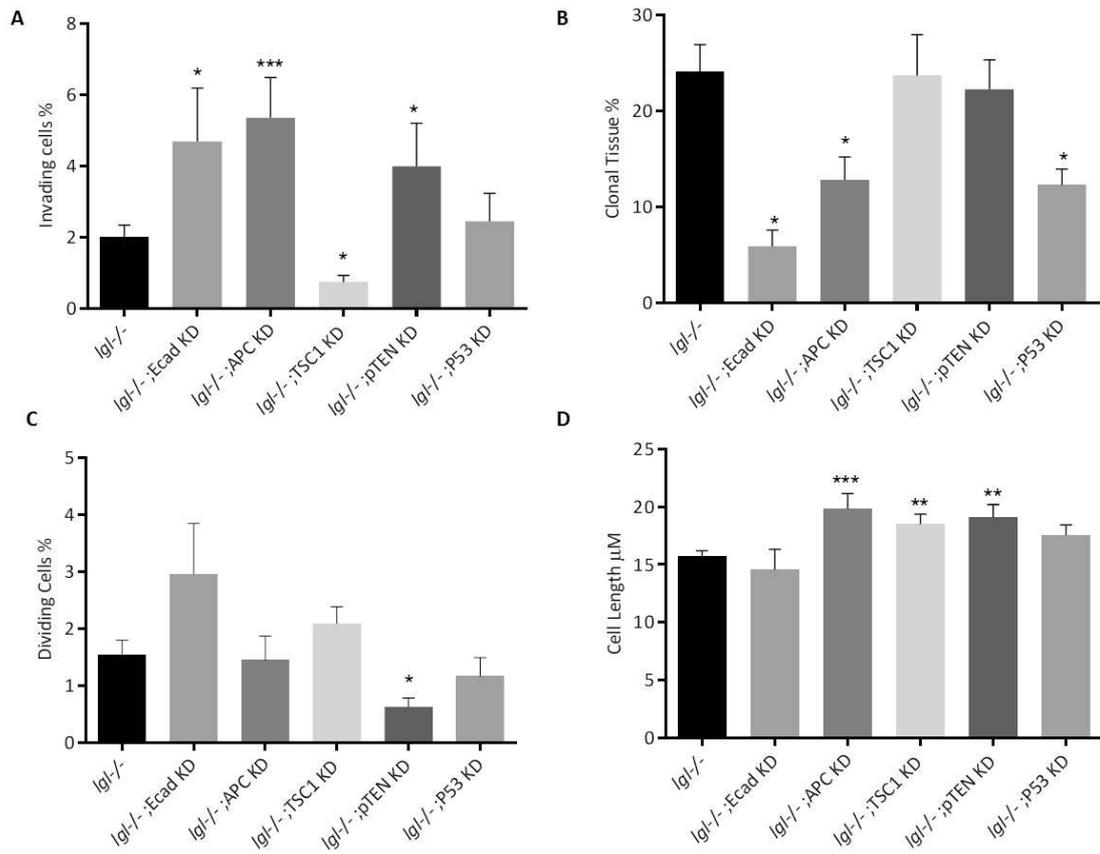
Figure 3.7 Quantification of phenotypes when expressing UAS-RNAi targeting various cancer associated genes.

Quantification of some cancer phenotypes for a small population of genes studied in the pilot screen compared to *lgl* homozygous mutant background. Looking at the percentage of mutant cells that had invaded and detected beneath the epithelium **(A)**, the percentage of the notum that was mutant **(B)**, the number of mutant cells undergoing cell division **(C)** and the length of mutant cells, in μm **(D)**. Bars represent the mean (n=10 animals) and error bars represent SEM.

Where possible two independent RNAi lines, ideally from different sources, were used to target each gene to ensure observed phenotypes were in fact a direct result of RNAi knockdown rather than through off target effects. In **Fig. 3.8 A-C'** two independent RNAi lines (35731-GD and 5671R-1) were used to target pTEN mRNA for degradation. When either of these RNAi constructs was expressed in the *lgl* mutant background similar phenotypes were observed, from multi-layering **(Fig. 3.8 A-A')**, long basal protrusions **(Fig. 3.8 B-B')** and invading cells detected beneath the epithelium **(Fig. 3.8 C-C')**. Similarly, when comparing RNAi lines for other genes such as APC, TSC1 and P53 there was no significant difference between the phenotypes observed **(Fig. 3.8 D-G)**. In only two instances was there a low significance ($p < 0.05$) between RNAi lines: the percentage of invading cells between RNAi lines targeting APC, and dividing cells between RNAi lines for TSC1.

On the whole, this showed that RNAi lines for the same gene generated similar phenotypes, and so it could be inferred with confidence that in this system the phenotypes observed were due to the knockdown of particular genes and not off-target effects.

	Multi-layering	Long Basal Protrusions	Invasion
RNAi line: 35731	A — *lgl*-/- ; pTEN KD	B — *lgl*-/- ; pTEN KD	C — *lgl*-/- ; pTEN KD
RNAi line: 5671R-1	A' — *lgl*-/- ; pTEN KD	B' — *lgl*-/- ; pTEN KD	C' — *lgl*-/- ; pTEN KD
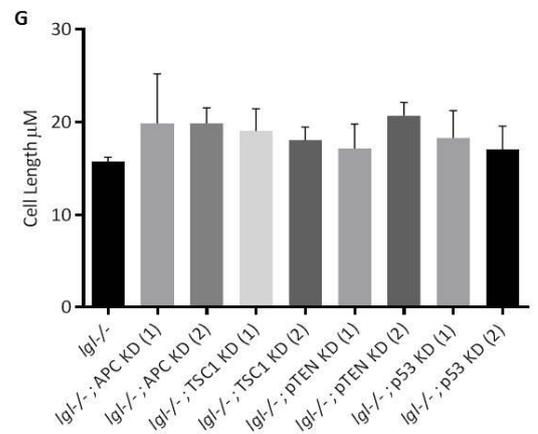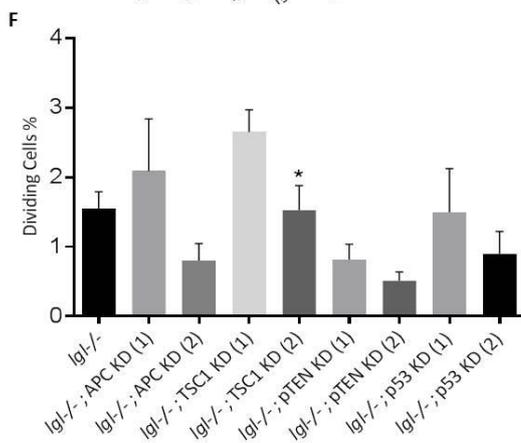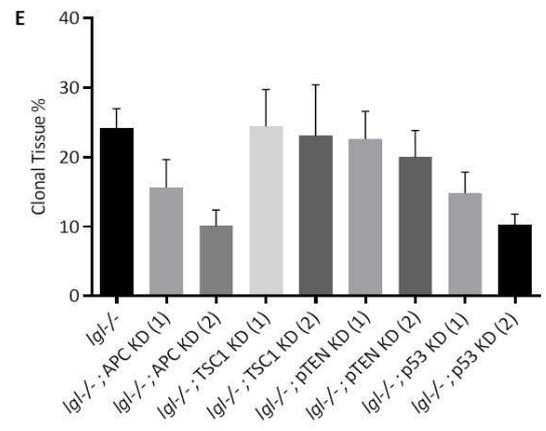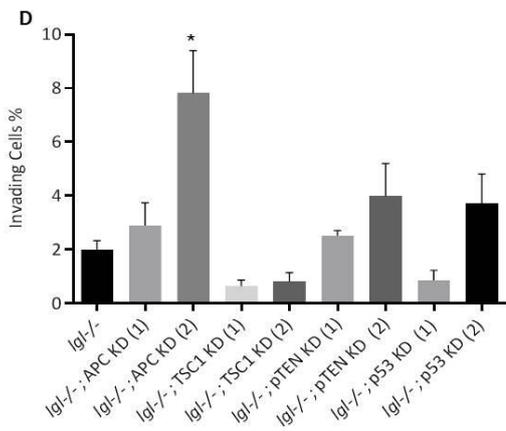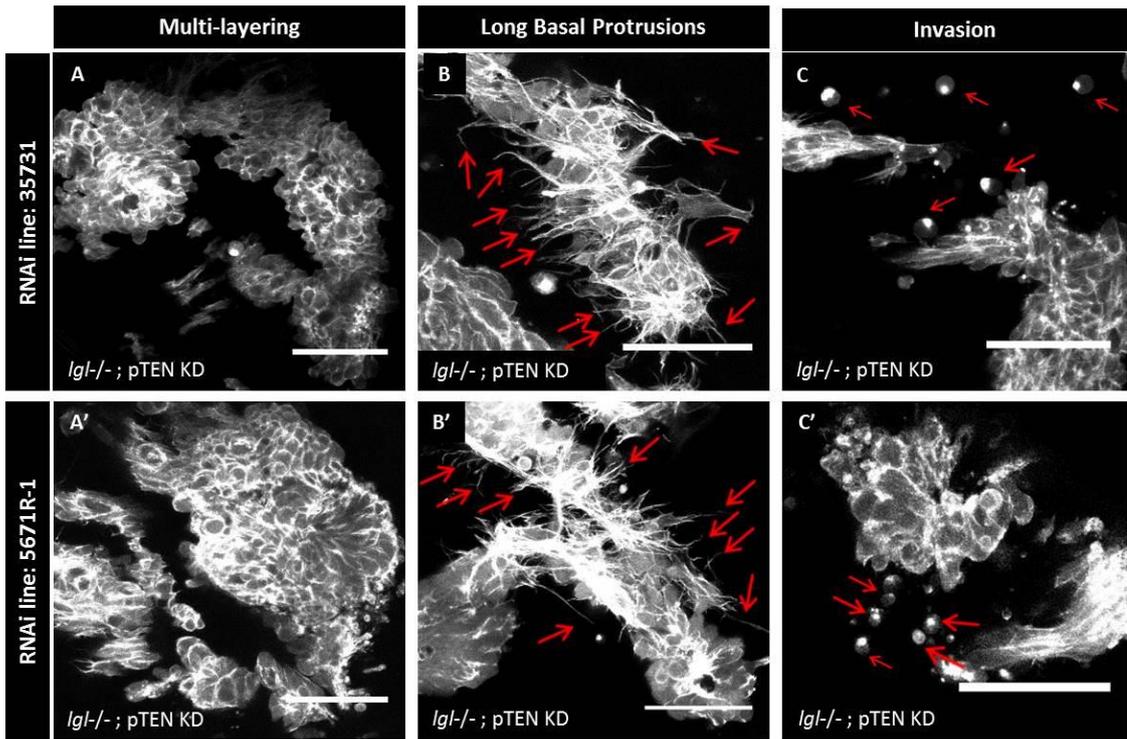
Figure 3.8 Comparing independent RNAi lines targeting the same gene

lgl homozygous mutant tissue simultaneously expressing two independent RNAi lines targeting pTEN (35731-GD, 5671R-1). For both, tissue was multi-layered **(A,A')**, with long basal protrusions **(B,B')** and invasive **(C,C').** Scale bars = 50µm. When independent RNAi lines for the same gene were compared to each other for various phenotypes **(D-G)**, in most cases there was no significant different between the RNAi lines (n=10 animals for each RNAi line).

During the pilot screen 10 animals were imaged per RNAi line, analysis of these animals indicated that in most cases animals with the same genotype gave similar phenotypes. To help visualise this we collaborated with several mathematicians who generated a statistical code and carried out some preliminary analysis. Here I show the results using hierarchical clustering, which uses a statistical code to cluster genes with similar phenotypes. To do this, Gowers statistical analysis was applied which analyses the distance between mean values for each gene. Using R studio, a statistical code was designed, whereby a hierarchical clustering algorithm groups animals according to Gowers analysis. **Figure 3.9** shows the dendrograms produced using this analysis. All animals analysed during the pilot screen are illustrated in **Fig. 3.9 A**, and by zooming into separate branches on the hierarchy tree **(Fig. 3.9 B)** clustering of these animals can be analysed. In this example, animals with KD for P53, Rbf1, Mnt, TSC1, Ex-IR, PAK3, Scrib, pTEN and cytochrome c can be seen clustering together. Analysis of these dendograms gives confidence that in most cases animals give similar phenotypes.

As mentioned, qualitative analysis of each animal was carried out blind by at least two scientists in the Georgiou lab due to the subjective nature of the analysis. Scoring by individual scientists was often checked against one another to ensure there were no major discrepancies. Additionally, mathematical analysis could be applied to determine if there were any particular phenotypes where scientists scored particularly differently. To do this, paired t-tests were used to compare the means of each phenotype then false discovery rate (FDR) analysis was used to calculate correction of the p-value. This analysis therefore looks at all the errors in the

hypothesis across multiple comparisons. From this various phenotypes scored were identified which appeared to result in regular scoring differences between scientists. These included clonal tissue, dividing cells, basal protrusions and clonal shape. Although this test indicates there is error between scientists scoring for these categories, it is important to note that in most instances the scores given still reflect positive or negative regulation of the phenotype. Some discrepancies are obviously expected due to the subjective nature of this analysis.

A

B

Rbf1 (1)
P53 (1)
P53 (2)
Mnt (1)
Cdc42 (1)
Mnt (1)
TSC1 (4)
Ex-IR (1)
PAK3 (1)
WASp (1)
PAK3 (2)
Ex-IR (2)
tkv (5)
TSC1 (3)
TSC1 (2)
SLIK (1)
tkv (4)
Rbf1 (2)
Ex-IR (3)
Ral (1)
TSC1 (1)
APC (1)
Mad (1)
tkv (3)
tkv (2)
tkv (1)
Scrib (IR) (1)
Scrib (IR) (2)
Scrib (IR) (3)
pTEN (1)
pTEN (2)
Mad (2)
Cytochrome c (3)
Cytochrome c (2)
Cytochrome c (1)
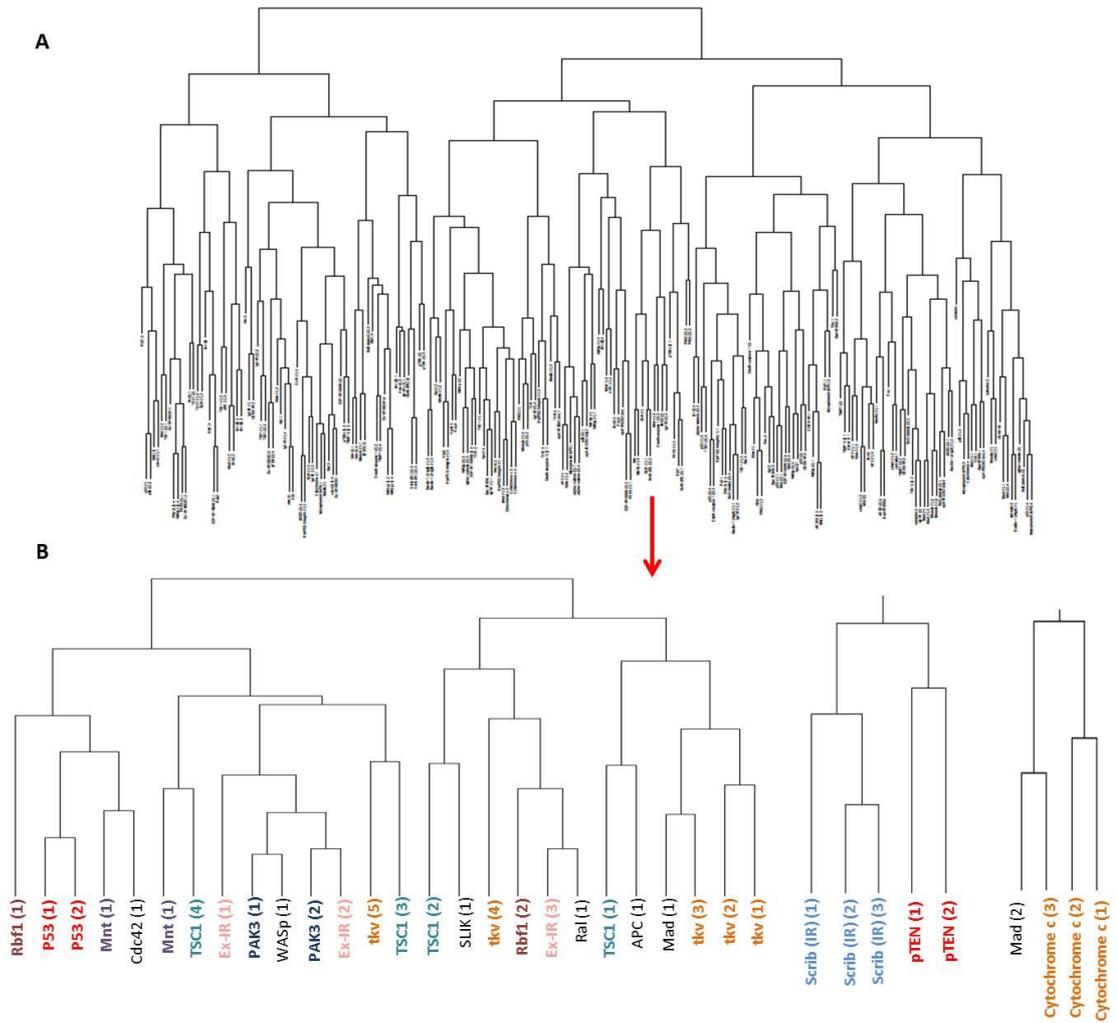
Figure 3.9 Hierarchial dendograms of pilot genes screened

(A) Dendrogram of all pilot genes screened using Gowers analysis. **(B)** Magnification of branches. Each end branch represents a single animal imaged during the screen. Animals expressing the same RNAi can be seen grouping together.

110

3.3 Discussion

Cancer is defined by an uncontrolled proliferation of abnormal cells that are able to invade other tissues, most commonly deriving from an epithelium. In order for these cancer cells to form secondary tumours at distant sites, they must lose cell-cell contacts within the epithelial sheet and breach the basement membrane for local invasion. Using the MARCM system, neoplastic tumours were generated in the notum of the fly, which were homozygous mutant for *lgl*, *dlg* or the *scrib* tumour suppressor genes. Clones of mutant tissue surrounded by wild-type tissue were generated to prevent lethality and to more accurately mimic tumour development in humans, thus allowing the role of the microenvironment to also be investigated.

Alterations to cell morphology and clonal size were observed in mutant tissue for all three tumour suppressor genes when compared to clones of labelled wild-type tissue. Unlike the *lgl-/-* mutant clones generated in the notum, both *dlg-/-* and *scrib-/-* mutant clones were comparatively smaller in size, which was likely to be a result of cell competition between the wild-type and homozygous mutant tissue. Why cell competition was seen in *dlg* and *scrib* mutant clones and not in *lgl-/-* clones in this system is still not fully understood. However, unlike Dlg and Scrib, Lgl has been shown to regulate the Salvador/Warts/Hippo pathway associated with cell survival and proliferation. This pathway comprises a kinase cascade leading to the phosphorylation of the transcriptional co-activator Yorkie (Yki), subsequently excluding it from the nucleus. Dephosphorylated Yki can translocate to the nucleus where binding to the scalloped (sd) transcription factor results in the transcription of genes associated with cell survival and proliferation (Grzeschik et al., 2010).

111

This may explain why *lgl* mutant cells were not outcompeted by the surrounding wild-type cell population in this system. Additionally, unlike previous work which looks at the imaginal discs, the notum is not a highly proliferative epithelium which could explain why larger clones are seen with *lgl-/-*. Therefore, in contrast to previous work, the notum can provide a 'one hit' system, inducing many of the hallmarks of cancer through a single mutation to a tumour suppressor, and so can be used as a model system to characterise the critical events underlying tumour cell invasion.

lgl-/- mutant clones exhibited cancer phenotypes including invasion, multi-layering and loss of polarised cell shape. As these phenotypes were not severe, and the mutant tissue wasn't being eliminated through cell competition, *lgl-/-* provided the ideal genetic background with which to carry out an enhancer suppressor screen as both increases and decreases in the cancer phenotype could be identified.

Although all mutant cells were labelled with Moe-GFP which targets GFP to the actin cytoskeleton, unexpected nuclear staining was seen in *scrib-/-* clones. Unlike *dlg* and *lgl* mutant clones (which used Pnr-Gal4), we saw a mislocalisation of Moe-GFP throughout the cell in *scrib-/-* clones (which used Actin-Gal4). Actin-Gal4 drives the expression of UAS-transgenes throughout the animal and in our experiment could result in Moe-GFP being driven too strongly, which could explain the mislocalisation of Moe-GFP in *scrib-/-* clones. This mislocalisation would make it difficult to study the actin cytoskeleton during the screening process. Additionally, global expression of Moe-GFP would mean that other cell types might be labelled in the notum (other

than epithelial cells) so it would be difficult to accurately identify invading cells that had come from the mutant tissue. In order to overcome this problem, other members of the Georgiou lab recombined mutant scrib with the Pnr-Gal4 driver which are both located on the 3rd chromosome. As expected, driving Moe-GFP with Pnr-Gal4 in *scrib -/-* clones resulted in the correct localisation of GFP to the actin cytoskeleton. *scrib-/-* mutant clones were still invariably small and so this was still not an ideal genetic background to use in an enhancer/ suppressor screen, however this recombined stock could be useful in any pertinent secondary screens.

It has been shown in other systems that the *Scrib* tumour suppressor gene requires cooperation with activated alleles of oncogenic *Ras* or *Notch* in order to produce dramatic overgrowth and become invasive (Brumby & Richardson, 2003; Pagliarini & Xu, 2003). Additionally *lgl* mutant clones were shown to require cooperation with oncogenic cMyc in the imaginal wing disc (Froldi et al., 2010). Similarly, in our system this cooperative effect was mimicked when the phenotype of *lgl-/-* tumours was enhanced when combined with an oncogenic form of Notch. However, unlike these previous *Drosophila* studies in which *scrib-/-*, *dlg-/-* and *lgl-/-* mutant clones remained hyperplastic without oncogenic cooperation, *lgl-/-* was sufficient to generate invasive tumours without over-expressing Notch in our system.

During the pilot screen of over 40 well characterised cancer-related genes, a range of phenotypes was observed, including, hyper-proliferation, multi-layering, invasion, apoptosis and effects on subcellular structures. By observing these phenotypes a

database comprising various categories was set up in order to qualitatively score tumour behaviour relative to *lgl* alone during the main candidate screen.

One of the aims of the pilot was to determine whether expected phenotypes for well characterised genes were replicated in this system. This was true in most cases, for example, a highly invasive phenotype was seen when expressing UAS-RNAi targeting E-cadherin, APC or TSC1. All three of these genes have been linked to progressive, invasive tumours through the loss of cell polarity, cooperation with potent oncogenes and regulation of the mTOR pathway respectively (Canel, Serrels, Frame, & Brunton, 2013; Y. Chen, Wei, Liu, & Guan, 2014). E-cadherin is an integral member of AJs and thus is pivitol for the maintenance of cell polarity; the loss of which effects epithelial integrity and promotes invasion (Canel et al., 2013). APC is most commonly known for its role in suppressing Wnt signalling which plays a key role in controlling cell proliferation and differentiation. APC inactivation studies have also linked APC to the maintenance of cell adhesion (Bienz & Hamada, 2004) and knockout mouse models result in defects in cell migration as well as proliferation and differentiation (Sansom et al., 2004). TSC1 knockout experiments in mice have resulted in significantly increased rates of proliferation compared to the control, along with invasive and migratory activity (Y. Chen et al., 2014). This was replicated in the pilot screen where large tissue over-growth, and invasive tumours were generated. Similarly, other phenotypes observed in Moesin, pTEN and p53 were also validated based on previous literature. For example, blebbing mitotic cells were seen when knocking down Moesin which is known for its role in the regulation of cell rounding during mitosis, where loss of Moesin results in disorganisation of the actin cortex resulting in

a phenotype comparable to apoptotic blebbing (Kunda, Pelling, Liu, & Baum, 2008). Misshapen dividing cells were also observed during the pilot screen when knocking down the kinase Slik which facilitates the phosphorylation of Moesin during the regulation of cell shape during mitosis (Carreno et al., 2008). An invasive phenotype was observed in pTEN KD animals which is also seen clinically, whereby pTEN mutations are highly prevelent in metastatic human tumours (Gonzalez-Angulo et al., 2011). In previous *Drosophila* studies, clones of pTEN mutant cells have an increased proliferation rate caused by antagonistic effects on Dp110 (*Drosophila* homologue of PI3-K), linking pTENs growth regulatory activity to the TOR pathway (H. Huang et al., 1999). pTEN has also been shown to modulate cell polarity and cytoskeletal arrangements in both *Drosophila* and mouse models (Tamura et al., 1998); additionally overexpression studies in vitro result in suppression of tumour cell invasion and metastasis (Tamura, Gu, Takino, & Yamada, 1999). It is unsurprising that severe phenotypes were seen for p53 since p53 activity is affected in more than 50% of all human cancers, and plays a critical role in signalling pathways that affect proliferation, EMT and apoptosis (Muller & Vousden, 2014).

As well as generating expected phenotypes based on previous literature when knocking down well known cancer-related genes, two independent RNAi lines for the same gene typically produced similar phenotypes. This confidently validated that RNAi was working efficiently in this system. Occasionally there was some variation between RNAi lines for the same gene, whereby phenotypes were more severe in one RNAi line compared to the other. Because, the two independent RNAi lines target different regions of the mRNA for the target gene, it is possible that one RNAi

transgene is having more of an effect than the other. One RNAi construct could be knocking down expression better than the other, or possibly one RNAi transgene may not work at all. For the full candidate screen, two RNAi lines for the same gene would still be used where possible to ensure phenotypes were not being missed. However it should be stressed that the vast majority of RNAi transgenes targeting the same gene gave the same phenotypes **(Fig. 3.8).** Additionally, our original plan was to image 10 animals per RNAi line for the entire candidate screen. However, since results from the pilot screen showed a low level of phenotypic variation in transformed cells between animals, we decided to reduce the number of animals to be imaged per RNAi line to 5, thereby maximising speed and efficiency.

In conclusion, from this chapter it was inferred that *lgl-/-* was the best genetic background with which to carry out an enhancer/suppressor screen, and that RNAi was working efficiently in this system. It was also shown that the co-operative nature between tumour suppressor genes and oncogenes could be replicated and studied. A range of phenotypes were seen, which lead to the generation of a database for the qualitative analysis of candidate genes relative to *lgl* alone. Additionally, I have shown that hits for specific phenotypes identified from this qualitative analysis were shown to be statistically significant when quantified. All of these results suggested that the large-scale candidate screen would likely pick up genes that are important in tumour progression.

CHAPTER 4

Large-scale RNAi screen

Chapter 4 Large-scale RNAi screen

4.1 Introduction

Following on from the pilot screen **(Chapter 3)** a large-scale candidate screen was carried out. This screen is now almost complete, with all animals (a total of 478 genes/696 RNAi lines) now imaged. Since the qualitative analysis of these genes is still in progress, I will be presenting a preliminary analysis of 418 genes (610 RNAi lines). Therefore, the data presented here represent a preliminary analysis of those genes that have been qualitatively analysed to date.

Drosophila has become the favourable model organism with which to carry out large-scale genetic screens, as they are cheap and easy to look after. Additionally, they have a relatively short life-cycle and are genetically easy to manipulate. In this way *Drosophila* screens are commonly used to identify novel genes which play an important role in a specific biological process. As mentioned in **Chapter 1**, notably the most famous genetic screen in *Drosophila* was carried out by Eric Wieschaus and Christiane Nusslein-Volhard in 1980. In this paper they did a mutagenesis screen in the embryo to identify essential patterning genes. Since this pioneering research, many different genetic screens have been carried out in *Drosophila.* In recent years RNAi mediated screens have identified new genes and regulatory networks involved in a variety of processes (Nusslein-Volhard & Wieschaus, 1980). For example in 2015 Zeng et al carried out a genome-wide screen which identified more than 400 genes involved in stem cell renewal in the adult *Drosophila* intestine (Zeng et al., 2015).

To select genes for the screen various online databases including the catalogue of somatic mutations in cancer (COSMIC) and publications from the Getz lab (Lawrence et al., 2014) were used to generate a compilation of genes commonly mutated or down-regulated in human cancers. As our model system is the *Drosophila* notum, only genes mis-regulated in epithelial cancers were selected. Genes were then listed in order of mutation rate, and cross-referenced to check for *Drosophila* orthologues and available transgenic RNAi lines. RNAi lines had to be on the X or 3rd chromosome as the *lgl[4]* mutation is on the 2nd. A range of genes were then selected from those well-characterised to those completely uncharacterised. For our screen, genes were ordered in batches to keep fly maintenance to a minimum, with a total of 7 batches being ordered from the Vienna *Drosophila* stock centre (VDRC), the National Institute of Genetics, Japan (NIG) and Bloomington, Indiana stock centres, over 3 years. This screen would not have been possible without the development of these near genome-wide RNAi stock centres. Overall, our screening process is relatively cheap and timely (we developed the methodology to generate labelled tumours on the back of the fly in the living animal and to image these tumours in high temporal and spatial resolution). The screening process required several members of the Georgiou Lab, all of whom were involved in every aspect of the process. For each transgenic RNAi line a minimum of 5 animals were imaged, with approximately 60 animals being imaged each week during high-throughput screening periods enabling us to analyse hundreds of genes relatively quickly.

Recently, the relative ease with which human cancer samples can be sequenced and deep-sequenced has resulted in a plethora of publications and online databases that demonstrate genetic and/or epigenetic aberrations in different human cancers. However, for many of these implicated genes, little is known about their function other than the fact that they have been shown to be commonly mutated/down-regulated in a variety of human cancers. This screen will hopefully address this current lack of knowledge.

4.1.1 Aims

The main aim of the candidate screen was to identify novel genes that promote or inhibit the tumour progression in the *Drosophila* notum. Once identified, potential hits would be characterised in more detail to improve our understanding of their role in tumour progression, and may provide potential targets for future therapeutic intervention. See **Chapters 5 and 6** for examples of genes that gave interesting phenotypes that were deemed suitable for further characterisation.

4.2 Results

4.2.1 Qualitative scoring of phenotypes

During the pilot screen, a database was generated with which to record and qualitatively score different phenotypes for every animal imaged. This was developed further following analysis of the pilot screen. Using the database, for a single animal all observed phenotypes would be given a numerical score. For each phenotype this score represents the extent of similarity/difference between the animal being analysed compared to *lgl-/-* animals. **Table 4.1** outlines the main phenotypes analysed and the scoring system used. The phenotypes which best represented that seen in *lgl-/-* animals were given a score of 0, with positive and negative effects being scored accordingly. Confocal images demonstrating this scoring system are shown in **Figure 4.1.** As scoring was a qualitative measure, done by eye, a minimum of two researchers from the Georgiou Lab scored and analysed every animal imaged.

Table 4.1 Scoring system for various phenotypes

Phenotype	Description	Scoring system
Clonal Tissue	<5% (**Fig. 4.2.1 A**)	-2
	5-15% (**Fig. 4.2.1 A'**)	-1
	15-35% (**Fig. 4.2.1 A''**)	0
	35-50% (**Fig. 4.2.1 A'''**)	1
	>50% (**Fig. 4.2.1 A''''**)	2
Invasion	No invading cells (**Fig. 4.2.1 B**)	-2
	Few invading cells (**Fig. 4.2.1 B'**)	-1
	Moderate level of invasion (**Fig. 4.2.1 B''**)	0
	Increased invasion (**Fig. 4.2.1 B'''**)	1
	Highly invasive (**Fig. 4.2.1 B''''**)	2
Basal protrusions	Absent filopodia (**Fig. 4.2.1 C**)	-3
	Short filopodia (**Fig. 4.2.1 C'**)	-2

	Normal filopodia (**Fig. 4.2.1 C''**)	-1
	Long filopodia (**Fig. 4.2.1 C'''**)	0
	Very long filopodia (**Fig. 4.2.1 C''''**)	1
Apex size	Very small (**Fig. 4.2.1 D**)	-2
	Small (**Fig. 4.2.1 D'**)	-1
	Normal (**Fig. 4.2.1 D''**)	0
	Large (**Fig. 4.2.1 D'''**)	1
Multi-layering Multi-layering was scored based on the amount of clonal tissue that had become multi-layered	0%	-1
	0-15%	0
	15-50%	1
	>50%	2
Cell length	<7μm	-2
	8-12μm	-1
	13-19 μm	0
	>20μm	1

Figure 4.1 Scoring system applied for various phenotypes

Confocal images of phenotypes observed during the screen and their relative scores compared to the *lgl-/-* phenotype. **(A-A''')** Clonal tissue and **(B-B''')** invasion (invading cells highlighted by red arrows) scale bars 100 μm. **(C-C''')** Basal protrusions (red arrows) and **(D-D''')** apex size. Scale bars 20 μm.

4.2.2 Correlation between qualitative scores and quantitative analysis

Quantitative analysis was performed on a selection of genes chosen at random in order to compare the quantitative measurements to the qualitative scores given. The aim of this analysis was to ensure that the scoring system employed to analyse genes was truly representative and could be relied upon when selecting genes for characterisation.

Figure 4.2 compares the scores given for various phenotypes to the quantitative measurements made for the same animals. When mutant GFP positive tissue was measured as a percentage of the whole notum **(Fig. 4.2 A)** there was a positive correlation between increases in the score given and percent of clonal tissue. Similarly this positive correlation was observed when quantification was carried out for other measurable phenotypes including invasion **(Fig. 4.2 B)**, length of basal protrusions **(Fig. 4.2 C)**, apical area **(Fig. 4.2 D)** and cell length **(Fig. 4.2 E)**. While overall there appeared to be a good correlation between the scores and actual measurements taken, as expected from such subjective analysis there was some variation between the quantitative results taken for animals with a given score. Despite this, the positive trend provides confidence that strong phenotypes will be easily identified during the screen.

From these results the scoring system appeared to be a good method for the relatively quick analysis of genes from the screen. By using the database generated in **Chapter 3** queries could then be designed to filter genes based on the scores given for an individual phenotype or groups of phenotypes with the aim of identifying

potential hits which could be characterised further (see **Appendix G (CD)** for all

genes and scores given for a range of phenotypes).
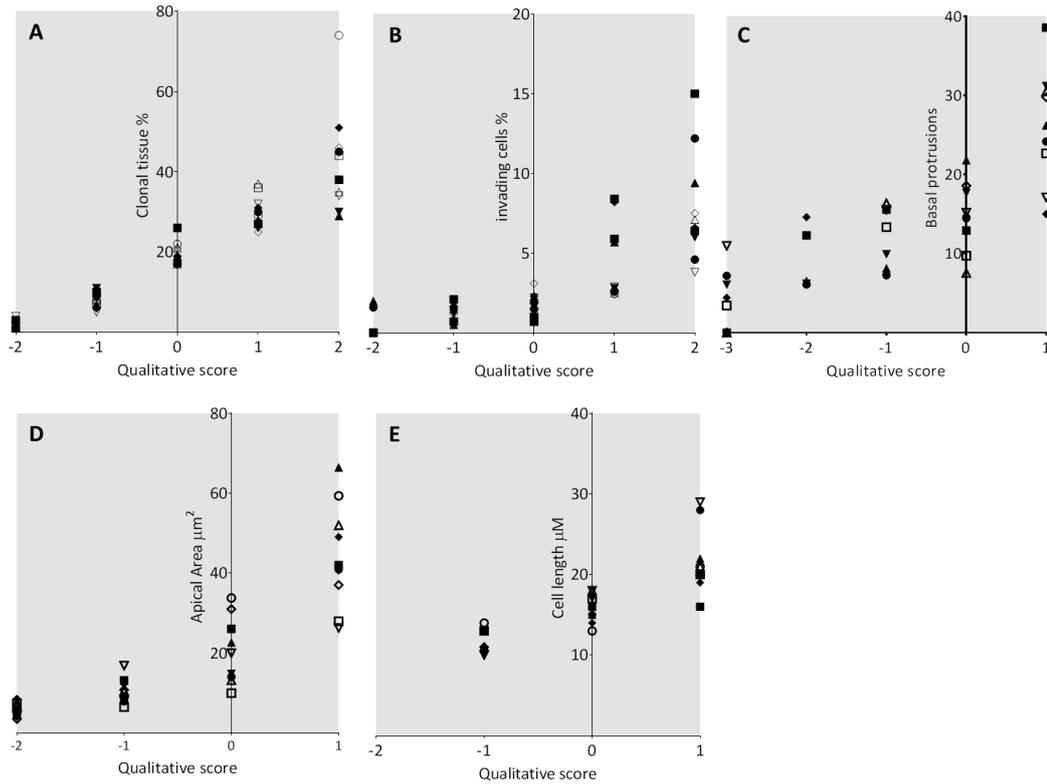
Figure 4.2 Correlation between qualitative scores and quantitative results.

Graphs comparing the scores given to individual animals against the quantitative measure for a range of phenotypes including: **(A)** clonal tissue, **(B)** invasion, **(C)** basal protrusions, **(D)** apical area and **(E)** cell length. Different shapes represent different animals. n=10.

4.2.3 Gene ontology of genes analysed during the screen

So far 696 transgenic RNAi lines have been imaged, targeting 478 genes within the *Drosophila* genome. The human orthologues of these genes have all been shown to be mis-regulated in human cancers. As the qualitative analysis hasn't been finished for all animals to date, here I will be presenting analysis on 418 genes (610 RNAi lines) **(see Appendix C for full list of genes)**. Of these, 45 genes had no known function, and remain uncharacterised in *Drosophila*. The remaining 373 genes have a known role in a variety of processes, based on the gene ontology (GO) terms assigned to them within the Gene Ontology Consortium database (Ashburner et al., 2000). **Figure 4.3** shows the range of GO terms associated with the genes that have so far been studied in this screen. These vary from epigenetic processes to cell adhesion, tissue/organ morphogenesis, immune response and neuron/axon development.

In total, we qualitatively analysed more than 20 different phenotypes and using our database it is very easy to identify genes which dramatically affect a particular phenotype. **Figure 4.4** shows the key GO terms for strong hits for 8 major categories. Of all the genes with no known function, 7 were identified as strong hits for invasion, 4 with apex defects, 4 highly multi-layered, 18 had elongated cells and 2 with very long basal protrusions.

From the screen, 80 genes were identified as hits for invasion (average score >0.5). Approximately half of these genes were shown to have known roles in cell adhesion, intracellular signalling pathways, actin/cytoskeletal regulation, cell cycle and cell/tissue/organ developmental processes; as you would expect from invasive

tumours **(Fig. 4.4 A)**. Top hits for invasion included regulators of metabolic processes (Trax, PP2C1), neurogenesis (Su(var)2-HP2, CG11180), cell migration and axon guidance (Unc5), spermatogenesis (CG7379), adhesion (Dscam3), circadian cycle (iH), epithelium morphogenesis (Ser) and intracellular signalling (RhoGap19D, Sdr). CG4393 was also identified as a top hit for invasion which has no known function.

Half of the genes identified as hits for invasion were also identified as hits for multi-layering. Multi-layering was the most common phenotype observed, with approximately 200 genes being scored an average of 1 or greater. Genes with the most severe multi-layered phenotypes (>1.8) included regulators of adhesion (ft, 18w, shot, mys), polarity (jbug), neurogenesis (cap), epigenetic processes (CTCF) and cell death (iz). Additionally, the uncharacterised gene CG10600 was identified within the top 20 hits for multi-layering.

Along with multi-layering, genes that scored highly for invasion also showed apical constriction. Additionally, apical constriction and multi-layering were often seen together without increased invasion. A large proportion of hits for both multi-layering **(Fig. 4. 4 B)** and apex defects **(Fig. 4. 4 C)** also had known roles in cell adhesion as well as epigenetic processes. Just over 40 genes were given an average score of <-1 for apex size, indicating that target gene KD is causing apical constriction. These genes included regulators of neurogenesis (Su(var)2-HP2), axon guidance and adhesion (Cont, troll, sema1b, Nrg, lea), polarity (Jbug) and cell cycle/cell death (bel). Genes identified with unknown function were CG15395 and CG42672.

Other phenotypes analysed included clonal tissue **(Fig. 4.4 D)**, basal protrusions **(Fig. 4.4 E)** and cell length **(Fig. 4.4 F)**. Both decreases and increases in clonal tissue were seen during the screen. More commonly gene KD caused a reduction in clonal tissue with more than 170 genes resulting in an average score of -1. On the other hand only 10 genes resulted in increased clonal tissue (>0.5). Top hits for reduced clonal tissue (<-1.5) included regulators of cell cycle (mrtf, Smr), development/adhesion (how, Cont) and morphogenesis/polarity (Cora). While increased clonal tissue included regulators of synaptic growth and neurogenesis (NGI2, MKK) and epigenetic processes (fd102c, nub). While no genes with unknown function were shown to increase clonal tissue within the notum, CG5645 was identified as a hit for reducing clonal tissue.

Only 27 genes from the screen resulted in a phenotype exhibiting long basal protrusions. Of these, unsurprisingly many were involved in regulating the actin cytoskeleton and cell-adhesion (kug, Pvr, Rac2, p120ctn). Additionally, top hits included those with known roles in axon guidance and migration (side, unc5); epigenetic processes (lilli) and cell cycle (myb, Trf4-2), as well as CG11593 and CG10669 with unknown function.

More than 120 genes were identified (score >0.5) which caused cells to elongate along the apicobasal axis. 4 genes with no known function (CG16868, CG10600, CG33288, and CG5550) were identified within the top 10 hits for increased cell length, while other genes within the top 10 included regulators of cell cycle (myb), cell death (fkh) and renal protein absorption (cubn).
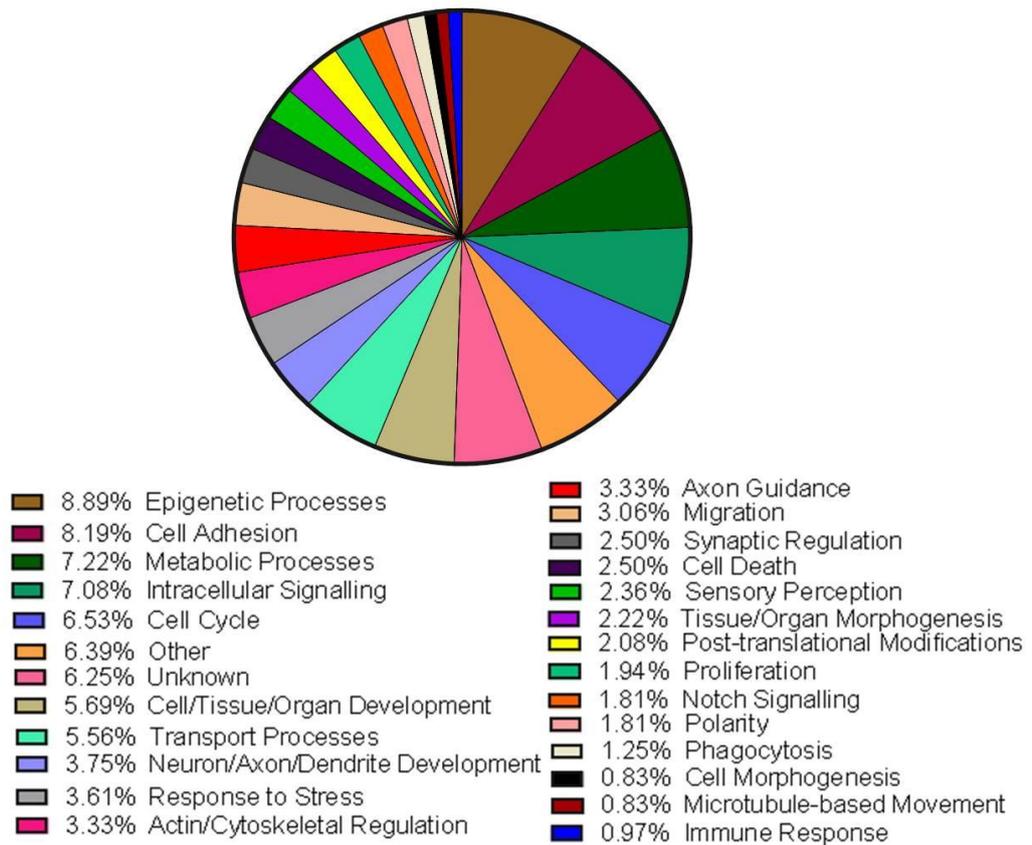
Figure 4.3 GO terms for all genes analysed during the screen

Breakdown of the main biological functions of the 418 genes studied during the screen, including the percentage of genes involved in each process.
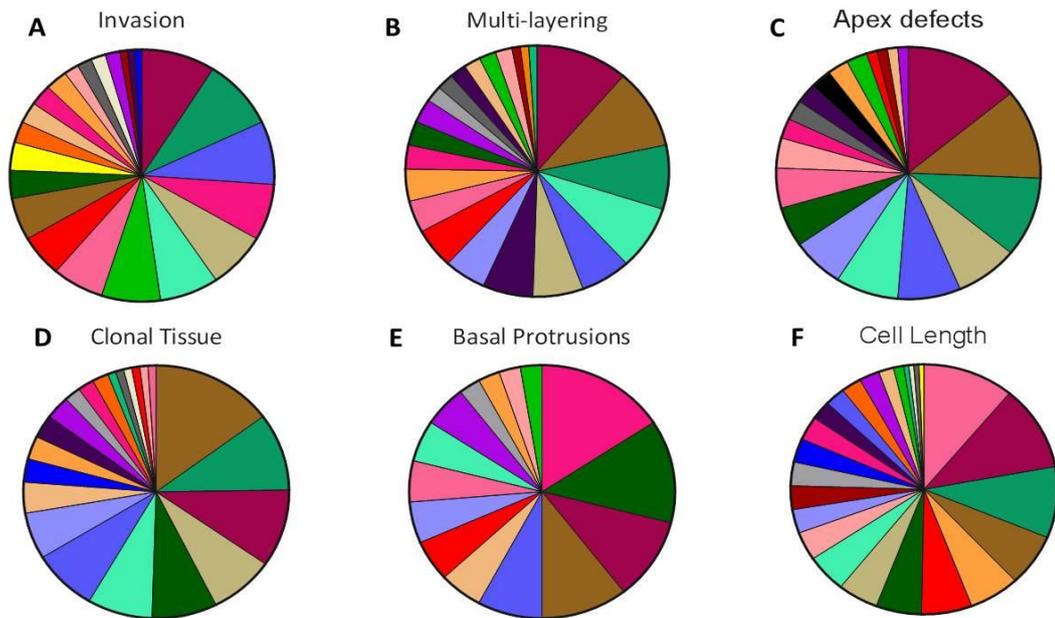
	GO term	Invasion	Multi-layering	Apex Defect	Clonal Tissue	Basal Filopodia	Cell Length
	Actin Regulation	7.2	3.1	2.6	2.0	15.8	3.1
	Epigenetic Processes	5.4	10.3	11.5	14.9	10.5	6.7
	Cell Adhesion	9.01	11.3	14.1	9.9	10.5	11.0
	Metabolic Processes	3.6	3.1	5.1	7.9	13.2	5.5
	Intracellular Signalling	9.01	8.2	10.3	9.9		9.2
	Transport Processes	7.2	8.2	7.7	7.9	5.3	4.9
	Axon Guidance	5.4	5.2	1.3	1.0	5.3	6.1
	Polarity	1.8	2.1	3.8	1.0	2.6	3.7
	Cell Cycle	8.1	6.2	7.7	7.9	7.9	2.5
	Other	2.7	4.1	2.6	2.9	2.6	6.1
	Neuron/Axon/Dendrite Development	2.7	5.2	6.4	4.9	5.3	3.1
	Migration	2.7	2.1	1.3		5.3	1.9
	Synaptic Regulation	1.8	2.1	2.6	1.0		0.6
	Cell Death	0.9	6.2	2.6	2.9		2.5
	Response to Stress		2.1		2.0	2.6	3.1
	Sensory Perception	7.2	2.1	2.6		2.6	1.2
	Tissue/Organ Morphogenesis	1.8	3.1	1.3	2.9	5.3	2.5
	PTMs	3.6					0.6
	Notch Signalling	2.7	1.0		2.0		2.5
	Phagocytosis	1.8			1.0		0.6
	Cell Morphogenesis			2.6			
	Microtubule movement	0.9	1.0	1.3			3.1
	Immune Response	0.9	2.1		2.9		3.1
	Development	7.2	6.2	7.7	7.9		4.9
	Proliferation		1.0		1.0		0.6
	Unknown	6.3	4.1	5.1	1.0	5.3	11.0

131

Figure 4.4 GO terms for genes which were identified as hits for a variety of phenotypes

Breakdown of the biological functions of the genes identified as 'hits' for various cancer-associated phenotypes, **(A)** hits for invasion, **(B)** multilayering, **(C)** apex size/defects, **(D)** clonal tissue, **(E)** protrusion length/defects and **(F)** cell length. (See **Appendix G (CD)** for the full list of genes and scores given). Numbers represent the percentage of genes within each category identified as a hit.

4.2.5 Identification of protein complexes involved in tumour progression

To further visualise and analyse hits identified during the screen the online tool COMPLEAT was used to carry out a complex enrichment analysis on all the genes screened. From this, a variety of protein complexes were identified which have a potential role in tumour progression **(Fig. 4.5)**. The graphs in **Figure 4.5 A-A''** compare the average scores given to each gene for various phenotypes. In each graph the black nodes represent genes which were hits for both phenotypes plotted. By zooming in to each node COMPLEAT identifies protein complexes using both literature and various protein-protein interaction prediction tools. **Figure 4.5 B-B''''** illustrates a few of these protein complexes which were identified during the screen. Protein complexes identified included: Smcb-Scmcf-PW29; ITGB11-RAP1A-PKD1; ITGB1-NRP1; and G protein complexes.

KD of genes which are members of the SMEA4A-PlexinD1 (Sema-1a, Sema-2a and plexB), ITGB11-RAP1A-PKD1 (mys, pdk2 and R) and ITGB1-NRP (mys, Ddr, Hml and CG34380) protein complexes resulted in multi-layered, invasive tumours with apical constriction. Apex defects and multi-layering were seen together when members of the Smcb-Scmcf-PW29 (SMC1 and Cap) and G protein complexes (GBeta13F) were knocked-down in the *Drosophila* notum.

Other protein complexes identified from the enrichment analysis included the ITGA5-ITGB1-FN1-TGM2 complex (mys, Dscam, Tg, Ptp4E) and NRP2-VEGFR3 complex (Ddr,

CG34380, Hml, Pvr) which were enriched in all datasets for invasion, multi-layering and apex size.

Comparing invasion against cell length **(Fig. 4.6)** produced a linear graph, whereby increases in cell length were accompanied with increased invasion. Protein complexes enriched in a positive direction along the X and Y axis included ITGB11-RAP1A-PKD1 and ITGB1-NRP1which were identified previously **(Fig. 4.5).** Conversely, decreased cell length resulted in a rescue of the invasion phenotype seen in *lgl-/-* animals. Protein complexes that had reduced cell length and decreased invasion included P300-CBP-p270-SWI/SNF complex (Osa, brm and Snr) and ETS2-SMARCA4-INI1 complex (brm and Snr).
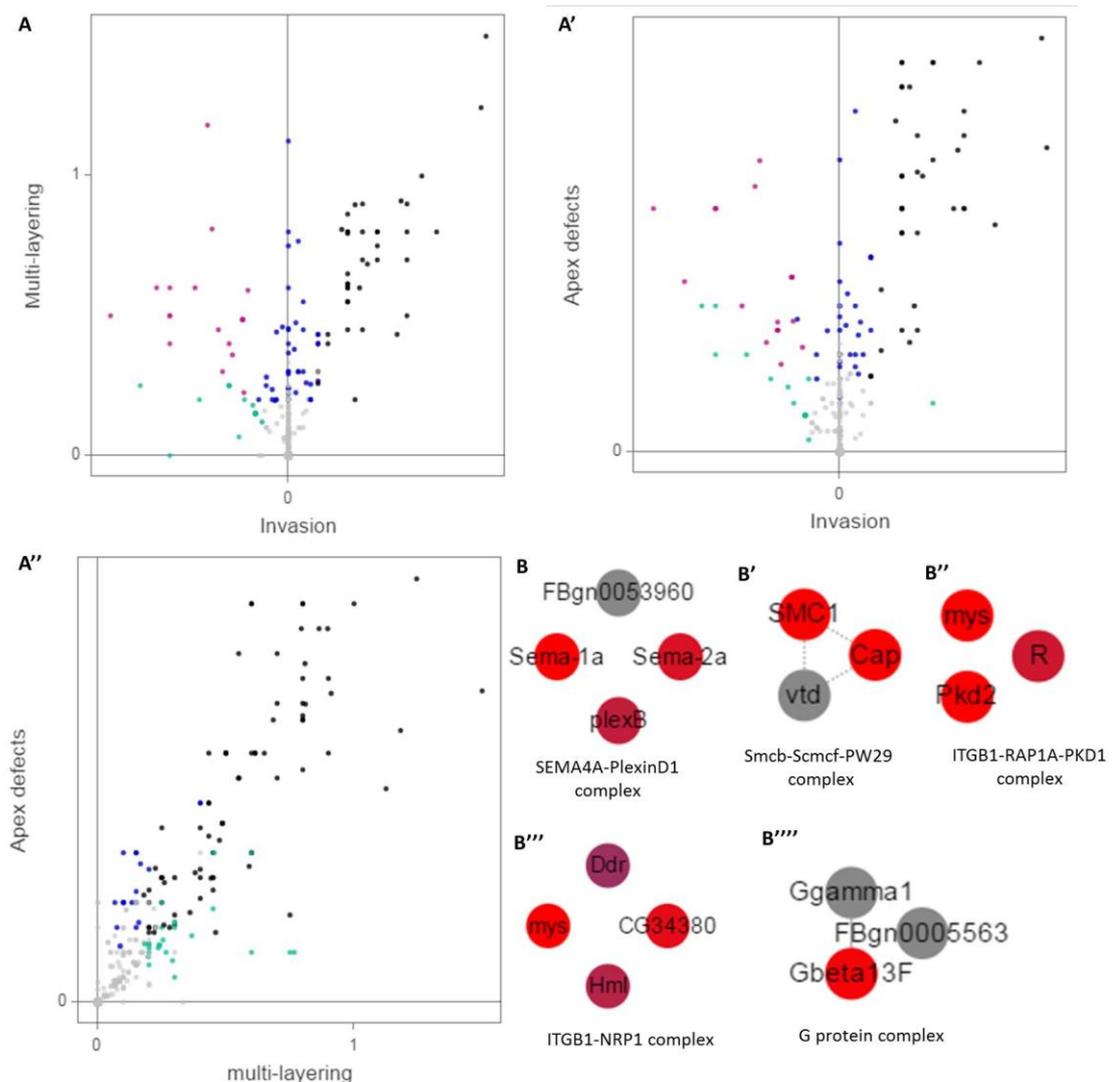
Figure 4.5 Protein enrichment analysis of identified genes

COMPLEAT online database results comparing all genes from the screen. **(A-A'')** Graphs compare different phenotypes, with black nodes representing genes/gene clusters which were enriched in both phenotypes along the X and Y axis; blue nodes represent genes with positive scores only for the phenotype on the Y axis, cyan for those on the X, pink for those enriched in both X and Y but in opposite directions and grey representing the genes which were not enriched in either phenotype. **(B-B'''')** Protein complexes identified from the enrichment analysis. Grey nodes represent members of the protein complex which were not included in the screen, while the other nodes represent the genes which were studied and identified as hits. A gene given maximum score is coloured red and colour is seen as a gradient from red to blue where genes were given a minimum score for that phenotype.
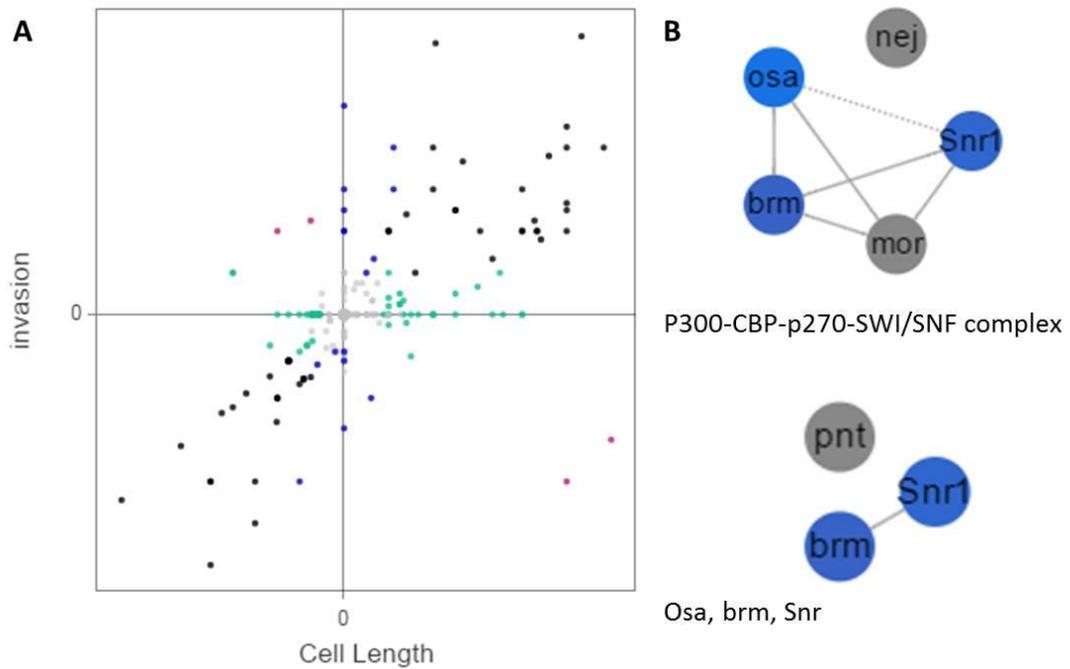
Figure 4.6 Protein enrichment analysis comparing invasive phenotypes and cell length

COMPLEAT online database result comparing all genes from the screen. **(A)** Graph comparing invasion against cell length, with black nodes representing genes/gene clusters which were enriched in both phenotypes along the X and Y axis; blue nodes represent genes with positive scores only for the phenotype on the Y axis, cyan for those on the X, and grey representing the genes which were not enriched in either phenotype. **(B)** Protein complexes identified from the enrichment analysis. Grey nodes represent members of the protein complex which were not included in the screen, while the blue nodes represent genes which gave negative scores for both cell length and invasion.

4.2.6 Hierarchical clustering of genes

Another way in which genes from the screen could be analysed was through the use of hierarchical clustering, which tries to cluster genes which have a similar phenotype. To do this, Gowers statistical analysis was applied (described in **Chapter 3)**. A hierarchical clustering algorithm then groups genes according to this Gowers analysis. **Figure 4.7** shows the dendrograms produced using this analysis. All genes analysed during the screen are illustrated in **Fig. 4.7 A**, and by zooming into separate branches on the hierarchy tree **(Fig. 4.2.7 B, C)** clustering of these genes can be analysed. 3 separate branches of the dendrogram are shown in **Figure 4.7 C-C''**. In **Fig. 4.7 C** the genes Rab40, CG11180 and wrapper have all been shown to regulate processes involved in neuron/axon/dendrite development, while CG7708 has no known function. Similarly in **Fig. 4.7 C'** trol and Gbeta13F have known roles in polarity, while CG10623 is unknown. Finally in **Fig. 4.7 C''**, all the genes except CG10669 (function unknown) regulate the actin cytoskeleton and adhesion. It is hoped that by using analyses such as this, that the function of unknown genes can be hypothesised based on the function of genes they cluster with.

A

B

C

| Rab40 (1) | CG7708 (2) | CG11180 (2) | Wrapper (2) | | trol (2) | CG10623 (1) | Gbeta13F (1) | | RunxA | Exn (2) | Asp (2) | CG10669 (2) |

Neuron/axon/dendrites
development

Polarity

Actin regulation/adhesion

138
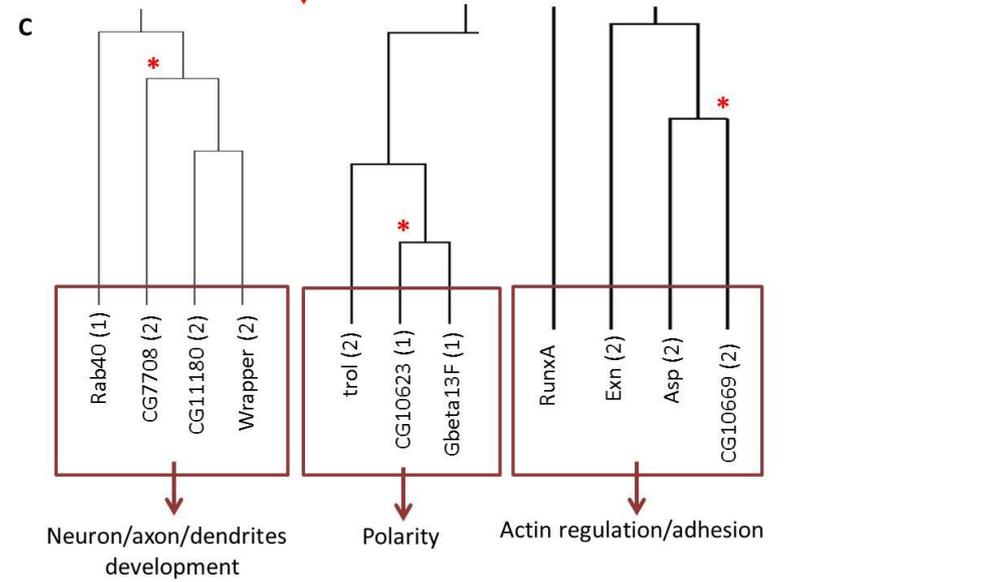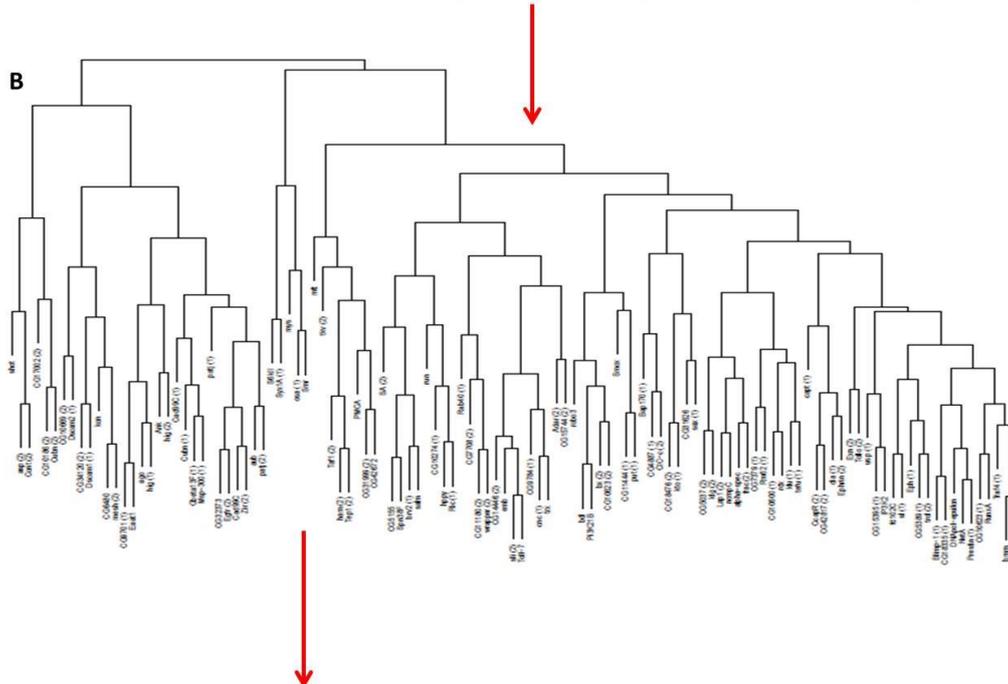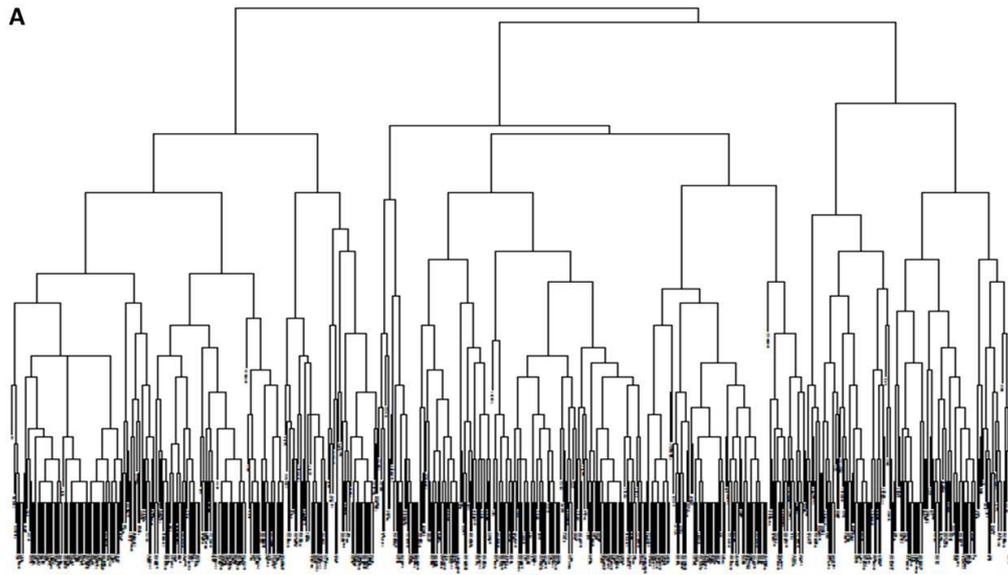
Figure 4.7 Hierarchial dendograms of genes screened

(A) Dendrogram of all genes screened using Gowers analysis. (B-C) magnification of branches produced, with a group of genes shown to cluster together in **C**. red asterisks indicates genes with no known function.

4.3 Discussion

Drosophila has become an increasingly popular model organism to study tumour progression and due to its easy genetic manipulation and short life-cycle is often used in large-scale screens. Baum and Georgiou developed an *in vivo* system using the *Drosophila* notum in which both clones of mutant tissue and individual epithelial cells could be studied in real time (Georgiou et al., 2008) (Georgiou & Baum, 2010). This system provides a 'one hit' system in which most of the hallmarks of cancer can be induced by mutating a single tumour suppressor gene. By generating clones on the back of the fly which were homozygous mutant for *lgl,* a screen was established such that transgenic RNAi was used to knockdown additional genes in the mutant background.

Our screen has been ongoing for a number of years, with many people contributing to its progress. Additionally, we collaborated with several mathematicians for the statistical clustering analysis. The initial aim of our genetic screen was to screen through 500 genes and from this identify novel genes which affect tumour progression in this system. Although I have only presented a preliminary analysis on 418 genes (610 RNAi lines) in this chapter, an additional 60 genes had been imaged but still need to be qualitatively analysed or re-imaged. Therefore we have almost (but not quite) completed the screen and as such are only just beginning to explore the mathematical tools available to analyse this data. Additionally, with the vast collection of highly complex confocal images obtained during the screen, it will be a challenge to extract as much information as possible from these images. From my preliminary analysis more than 200 genes were identified which were shown to

drastically promote tumour progression in this system. It is unsurprising that so many genes were identified as hits for multi-layering, as the human orthologues of these genes have already been implicated in cancer. Of these, 80 genes were able to promote invasion. While this system in the notum has already improved our understanding on how polarity proteins and Rho-GTPases regulate junction integrity and dynamic protrusions, we have now used it to identify genes which seem to play a role in tumour progression and invasion.

Using COMPLEAT protein enrichment analysis various protein complexes were identified which promoted tumour progression during the screen. Some of these complexes have already been implicated in cancer, including the NRP2-VEGFR3 complex. This complex has recently been linked to progression of breast cancer and metastasis (Luo et al., 2016) and during our screen was identified from genes which promoted invasion, multi-layering and apical constriction. Additionally, the complex SME4A-plexinD, a hit for invasion, multi-layering and apical constriction during the screen and known to regulate axon guidance during neuronal development, has been implicated in metastatic breast cancers (Luchino et al., 2013). This strengthens confidence in the results from the screen and the scoring method used.

During the screen a relationship was observed between invasion and cell length phenotypes whereby gene KD, which caused cell elongation along the apicobasal axis, also promoted invasion. Elongation along the apicobasal axis is a common feature in developing tissues and organs, such as *Drosophila* ventral furrow formation. This cell elongation is generated by changes in actomyosin contractility

regulated by Rho-GTPases (Kolsch, Seher, Fernandez-Ballester, Serrano, & Leptin, 2007). Additionally, these changes in actomyosin contractility are seen during EMT as cells adopt a mesenchymal-like morphology and front-rear polarity for cell migration (Etienne-Manneville & Hall, 2002). Therefore it is unsurprising that changes in cell shape were observed when gene KD promoted invasion. Invasive phenotypes were also regularly observed alongside apical constriction and multi-layering. Apical constriction is an essential process during certain developmental processes such as gastrulation. This process is also largely regulated by Rho family members, along with the developmental transcription factors twist and snail (Kolsch et al., 2007). Moreover in this process snail has been shown to induce myosin contraction by down-regulating transcription of E-cadherin (Oda, Tsukita, & Takeichi, 1998). In cancer, EMT is associated with the upregulation of both snail and twist. In this way, apical constriction and cell elongation, allow cells to be extruded from the epithelial sheet and invade/metastasise.

Using hierarchical clustering and Gowers statistical analysis genes could be grouped based on the distance between them (calculated using mean phenotypic scores). From this predictions can be made for genes with unknown function. For example, CG7708 was found to be clustered with genes involved in neuron/axon/dendrite development, based on the phenotypes observed during the screen. Therefore it could be hypothesised that this gene could be involved in similar processes. In this way COMPLEAT and statistical analysis can be used to make simple predictions.

The database and scoring system implemented provided a relatively simple approach to analysing such a large data-set, and by setting up different queries, data could be easily extracted to find genes which promote or inhibit specific phenotypes. Moreover, it is hoped that when publishing the screen data, this database will be made available online for the public and science community. It is hoped that by the end of the year we will reach our initial aim of 500 genes. Additionally it is hoped to use other statistical methods to better present the data. For example, to look at the amount of variance in the data principal component analysis (PCA) could be used, which allows high dimensional data to be viewed in 2D. Additionally, linear discriminant analysis could also be used for classification of the data such that further predictions could be made based on phenotypes observed in unknown genes. Similar to PCA, this analysis is used to reduce the dimension of the data by seeking axes which would give the best separation of classes. For classification the data can then be separated into two groups, where one group acts as a training set and the other as a predicting set. In this way we will be able to see how accurate predictions of gene function will be based on our data set. Clustering could also be used to identify novel genes in a known signalling pathway, or could implicate a novel pathway that had not previously been implicated in cancer before.

From our screen a wealth of candidate genes were identified which dramatically affect a variety of phenotypes and from these, genes would be selected to characterise. Selection of genes depends on several factors, including the scientific interests of the researcher, the availability of tools for characterisation and current knowledge from literature. Moreover, genes could be selected which were identified

as hits for a single phenotype or present with multiple phenotypes. Using *Drosophila* there are many ways in which these genes can be characterised. For example, genes could be knocked-down in the notum without the *lgl[4]* mutation. If available, mutants for these genes could be analysed. Immunohistochemistry techniques could be used to stain for a variety of junction, adhesion, cytoskeletal and polarity proteins. Individual mutant epithelial cells could be labelled and followed in real time for a more detailed analysis of cell morphology. Additionally, the human orthologues of these genes could be studied in vitro and assayed for a number of processes such as invasion, metastasis, proliferation and adhesion. I have carried out a number of these experiments, which are detailed in the following results chapters.

CHAPTER 5

Characterisation of two novel *Drosophila* genes: CG7379 and CG10600

Chapter 5 Characterisation of two novel *Drosophila* genes: CG7379 and CG10600

5.1 Introduction

From the RNAi-mediated screen **(Chapter 4)**, I identified two novel *Drosophila* genes: CG7379 and CG10600 as giving strong and interesting phenotypes. This introduction will give an overview of what is currently known about these two genes and their human orthologues.

5.1.1 CG7379

CG7379 is a *Drosophila* gene spanning 1787nt on the right arm of chromosome 3 and encodes a single polypeptide consisting of 433 amino acids (Attrill et al., 2016). Very little is known about CG7379 as it remains uncharacterised in *Drosophila*, however it has been proposed to play a role in chromatin modification and zinc binding due to its highly conserved zinc finger and PHD-type domains (He, Helbing, Wagner, Sensen, & Riabowol, 2005). The human orthologue of CG7379 is ING1 and other ING family members (Inhibitor of growth proteins N-terminal histone-binding). Although there is only 28% identity and 41% similarity between these protein orthologues, it is important to note that there is a very strong identity of 78% within the conserved PHD smart domains **(Fig. 5.1)** (Flockhart et al., 2012).
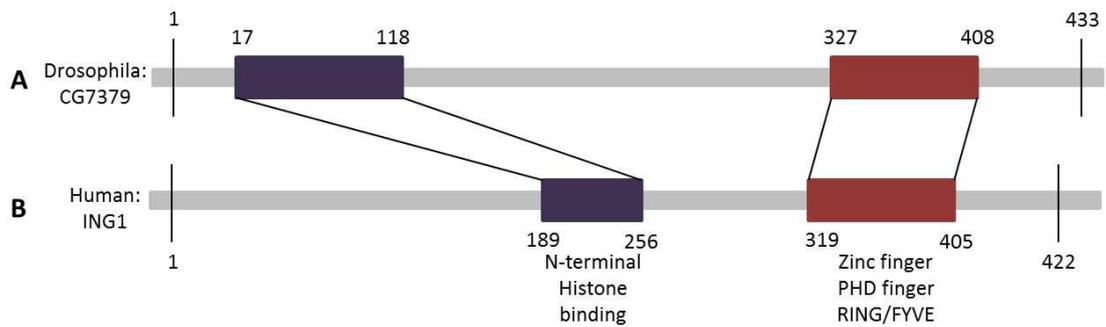
Figure 5.1 Conserved domains between CG7379 and ING

Conserved N-terminal and PHD/zinc finger protein domains between *Drosophila* CG7379 **(A)** and its human orthologue ING1 **(B).**

In mammals there are five isoforms of the ING family proteins, all of which contain C-terminal PHD-type zinc finger domains enabling these proteins to specifically bind post-translational modifications (PTMs) (typically lysine-4-trimethylated H3) on histone residues. Covalent PTMs to histones results in the alteration of chromatin structure and therefore can have a dynamic impact on gene expression (Garkavtsev, Demetrick, & Riabowol, 1997; Loewith, Meijer, Lees-Miller, Riabowol, & Young, 2000; Soliman & Riabowol, 2007). There are many ways that histone residues can be modified, from acetylation, methylation, ubiquitylation to phosphorylation; for example when lysine residues 4 and 36 on H3 undergo methylation the euchromatin structure opens up allowing transcriptional activation. However when lysines 9 and 27 on H3 are methylated a 'closed' heterochromatin structure is formed which is associated with gene silencing. Often, these PTMs result in the recruitment of 'reader' and 'effector' proteins which subsequently effect downstream biological functions (Litt, Simpson, Recillas-Targa, Prioleau, & Felsenfeld, 2001; Santos-Rosa et

al., 2002). Therefore histone PTMs and the reader and effector proteins acting downstream play an important role in gene expression and regulation.

Allelic loss and reduced expression of ING family members has been reported in a variety of mammalian cancers including glioma (Garkavtsev et al., 2004) and breast cancer (S. Kim, Chin, Gray, & Bishop, 2004). As well as their obvious role in modifying gene expression, ING proteins have been linked to the regulation of cell proliferation through p53 dependent and independent pathways. DNA damage and stress signals trigger a cascade of PTMs which regulate p53 activity for DNA repair, apoptosis or cell cycle arrest/senescence (Garkavtsev et al., 1998; Tsang et al., 2003). ING has been shown to enhance p300 acetylase activity resulting in the acetylation of p53 and subsequent transcription of regulators of cell cycle arrest and apoptosis, including p21 and bax. Additionally, ING expression has been shown to increase prior to cell death (Shiseki et al., 2003). Moreover, p53 mutants compromise apoptosis induced via ING activity, while c-myc has been shown to cooperate with ING to regulate cell death. Therefore, in cancer it has been proposed that loss/misregulation of ING provides cancer cells protection from apoptosis (Helbing, Veillette, Riabowol, Johnston, & Garkavtsev, 1997).

Recently, several studies with ING3 and ING4 have shown that overexpression of these proteins can inhibit migration of hepatoma cells in culture, whilst downregulation promotes cell motility. With their known role in regulating apoptosis it is no surprise that downregulation of these proteins enables cancer cells to survive, however these recent studies have also highlighted INGs potential role in regulating

cell matrix proteins for increased motility (M. Lu et al., 2012). For example, ING1 was shown to interact with liprin-alpha-1 which directs LAR to localise to and regulate cell focal adhesions. Additionally, overexpression of ING4 was shown to decrease activity of various Matrix-metalloproteinases (MMPs) which have a known role in cell migration as well as apoptosis (Shen et al., 2007). In September 2015, further evidence came to light linking ING5 to cell migration and EMT *in vitro*. In MDA-MB-231 and MCF-7 breast cancer cell lines, overexpression of ING5 was shown to increase epithelial markers and decrease mesenchymal markers both at an mRNA and protein level (Zhao, Ju, Wang, Ma, & Zhao, 2015). Although our understanding of how ING family proteins affects these regulators of EMT still remains unclear, these recent studies make ING a very interesting gene to study with regards to tumour progression.

5.1.2. CG10600

CG10600 is an uncharacterised *Drosophila* gene located on the 2nd chromosome and encodes a 1377 amino acid protein with unknown function. Its human orthologue is round spermatid basic protein 1-like (RSBN1L) which is so called due to its expression in the nucleus of haploid round spermatids (Takahashi et al., 2004). Protein alignment of RSBN1L and CG10600 show 31% identity with 47% similarity (Hu et al., 2011). Both CG10600 and RSBN1L have no known or predicted protein domains. It has been suggested that RSBN1L may regulate transcription of haploid germ cells, however, like CG10600, the function and biological processes of RSBN1L remain largely unknown (Takahashi et al., 2004). Interestingly, RSBN1L is seen to be downregulated

in a variety of cancers including gliomas, breast and colorectal cancer (Attrill et al., 2016; Flockhart et al., 2012).

5.2 Results

5.2.1 Knocking-down CG7379 in *lgl-/-* mutant background promotes invasion

During the candidate screen **(Chapter 4)** CG7379 was identified as a strong hit for invasion and multi-layering. Two independent RNAi lines both from VDRC (27988, 27989) were used to knock-down CG7379 in *lgl* homozygous mutant tissue (*lgl -/-*) in the notum, with both RNAis producing very similar phenotypes. The most striking phenotype observed when knocking down CG7379 (CG7379 KD) was the excessive number of epithelial cells seen beneath the epithelium **(Fig. 5.2 A, B).** When quantified as a percentage of the total number of mutant cells in the notum, there was a significant (p=0.0066) increase in the amount of invasion when knocking down CG7379 during the screen compared to *lgl-/-* alone. The mutant tissue was also consistently multi-layered, at a greater extent to that seen in *lgl-/-* mutants alone, with cells losing their typical polarised cell shape and growing on top of one another in a disorderly fashion. **Table 5.1** outlines the average scores for CG7379 for each of the main phenotypes analysed during the screen. From this it can be seen that CG7379 was a hit only for invasion and multi-layering. See **Section 4.1** for a summary of the scoring system used when qualitatively analysing each cancer-associated phenotype.

Table 5.1 Qualitative scores for various phenotypes in *lgl-/-*; CG7379KD animals

Phenotype:	Clonal Tissue	Invasion	Multi-layering	Basal Protrusions	Division	Cell Length	Apex size
Score:	-0.18	1.36	0.73	-1.5	0	0	0

To see whether the phenotype observed was a direct result from knocking down CG7379 or whether it was a cooperative effect seen with the *lgl* mutation, UAS-RNAi targeting CG7379 was also expressed in a completely WT background. From this, the tumours produced appeared relatively normal, forming an organised mono-layered epithelium with no multi-layering. Some invading cells were detected beneath the epithelium, which was significant (p=0.0004) compared to WT **(Fig. 5.2 D)**. CG7379 KD was therefore sufficient to promote invasion, with the phenotype being exacerbated when combined with the *lgl* mutation.

In conclusion, this data shows that CG7379 cooperates with Lgl in a tumour suppressive manner, and when downregulated in cancer cells promotes tumour progression and invasion.
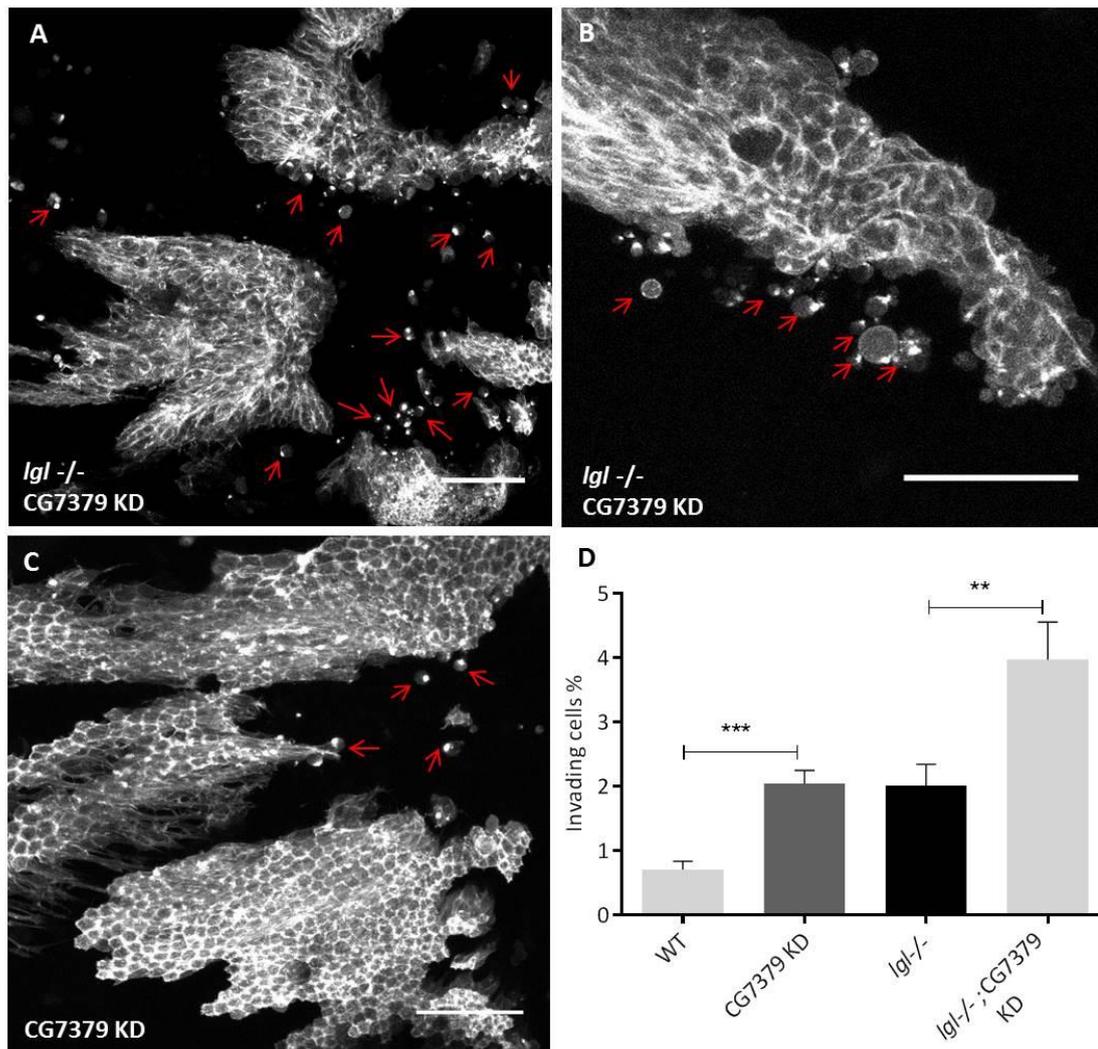
Figure 5.2 Cooperation between _lgl_ mutant and CG7379 knockdown promotes invasion

(A-C) UAS-RNAi was used to knockdown CG7379 specifically within _lgl_ homozygous mutant tissue in the _Drosophila_ notum during a large scale genetic screen. **(A,B)** Mutant clones were multi-layered and highly invasive**. (C)** RNAi targeting CG7379 expressed in a WT background without the _lgl_ mutation. Red arrows indicate invading cells. Note the lack of multilayering in C. Scale bars represent 50μm. **(D)** Quantification of the percentage of invading cells in the mutant tissue with mean and SEM. (_lgl-/-_ n=32, CG7379 n=5, _lgl-/-_;CG7379 n=10 from two different RNAis: 27988,27989 VDRC).

5.2.2 Blocking apoptosis in *lgl* -/- tissue does not mimic the phenotype seen when knocking down CG7379

The human orthologue of CG7379, ING, has been implicated in metastatic cancers and shown to act as a tumour suppressor, protecting tissues from cancerous cells by regulating the apoptotic cascade (Helbing et al., 1997). Therefore there are two potential hypotheses for the increased invasion seen when knocking down CG7379 in the *Drosophila* notum: 1) *lgl* mutant cells regularly invade but die through apoptosis. If KD of CG7379 inhibits apoptosis this could therefore lead to more surviving invading cells. 2) KD of CG7379 actually promotes invasion of epithelial cells.

To see whether the invasive phenotype seen previously **(Fig. 5.3)** was a result of cancer cells evading apoptosis, P35 was overexpressed specifically within the *lgl-/-* mutant tissue using the MARCM system. Overexpression of P35 has been shown to block apoptosis and is widely used in *Drosophila* cell death studies (Hay et al, 1994). However, when P35 was expressed specifically in the *lgl-/-* mutant tissue, the invasive phenotype observed with CG7379 KD was not replicated. Compared to *lgl-/-* alone, expressing P35 had no significant effect on invasion, however when compared to *lgl-/-* with CG7379 KD (compare panel A in **Fig. 5.2** to panel A **Fig. 5.3**) there was a significant decrease in invasion (p=0.0036) **(Fig. 5.3 B).**

As blocking apoptosis in the *lgl* -/- mutant background wasn't sufficient to promote invasion, it could be inferred that knocking down CG7379 in the mutant background promotes invasion through other independent mechanisms.

154
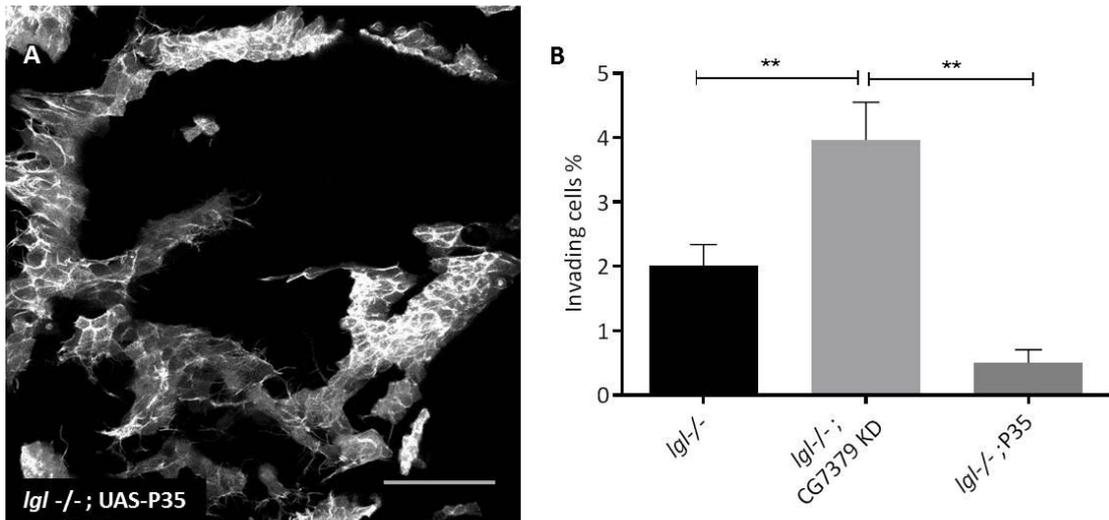
Figure 5.3 Over-expressing P35 in *lgl* -/- mutant background has no effect on invasion

(A) The *Drosophila* baculovirus P35 was expressed under UAS control specifically in *lgl* -/- homozygous mutant tissue on the back of the fly. Scale bar represents 50μm. **(B)** Quantification of the percentage of invading cells in mutant tissue in *lgl-/-*, *lgl-/-*; CG7379 KD and *lgl-/-*; UAS-P35 animals. Mean and SEM plotted. (*lgl* -/- n=32, CG7379 RNAi, *lgl-/-* n=10, UAS-P35, *lgl-/-* n=5).

5.2.3 CG7379 plays an important role in junctional integrity and cell adhesion

As shown, CG7379 KD appeared to promote invasion independently of its possible role in regulating apoptosis **(Section 5.2.2)**. CG7379 must therefore be able to regulate other factors linked to invasion and EMT, such as adhesion and polarity. To study the function of CG7379 in epithelial adhesion, *Drosophila* nota were fixed and stained with E-cadherin antibodies to give a detailed analysis of adherens junction (AJ) architecture and junction integrity. For this experiment, clones were generated in the *Drosophila* notum, which were either homozygous mutant for *lgl* (*lgl-/-*), expressing RNAi targeting CG7379 (CG7379 KD) or simultaneously homozygous mutant for *lgl* and expressing an RNAi construct targeting CG7379 (*lgl-/-*; CG7379 KD).

As a control, *lgl* -/- clones were fixed and stained for E-cadherin; which compared to the surrounding WT tissue was mostly unaffected **(Fig. 5.4 A).** Strikingly, CG7379 KD clones showed a severely altered distribution of E-cadherin levels **(Fig. 5.4 B).** Ectopic accumulation or a significant reduction of E-cadherin was frequently observed along disjointed AJs within the mutant clone. Similarly, in *lgl-/-*; CG7379 KD clones **(Fig. 5.4 C)** E-cadherin appeared fragmented along cell junctions and poorly distributed. Quantification of E-cadherin intensity levels at epithelial junctions **(Fig. 5.4 D)** confirmed that knocking down CG7379, with or without an accompanying *lgl* mutation, significantly ($p<0.0001$) reduced overall E-cadherin levels of at the junction.

Along with reduced E-cadherin levels, several junction abnormalities were also observed **(Fig. 5.4 E).** These included junctional breaks and atypical ectopic junction extensions, which were never observed in WT tissue or *lgl-/-* mutant clones. Ectopic junction extensions were seen a significant number of times in both CG7379 KD clones (p=0.0371) and *lgl-/-*; CG7379 KD clones (p=0.0484) **(Fig. 5.4 F).**

Additionally, in CG7379 KD clones (where E-cadherin levels were sufficient to visualise junctions) there was a significant expansion of apex area in the XY plane (p<0.00001) **(Fig. 5.4 G, H).**

Overall knocking down CG7379, with or without an accompanying *lgl* mutation, had significant effects on E-cadherin levels and junctional architecture, suggesting that CG7379 has a role in regulating cell adhesion and junction integrity.
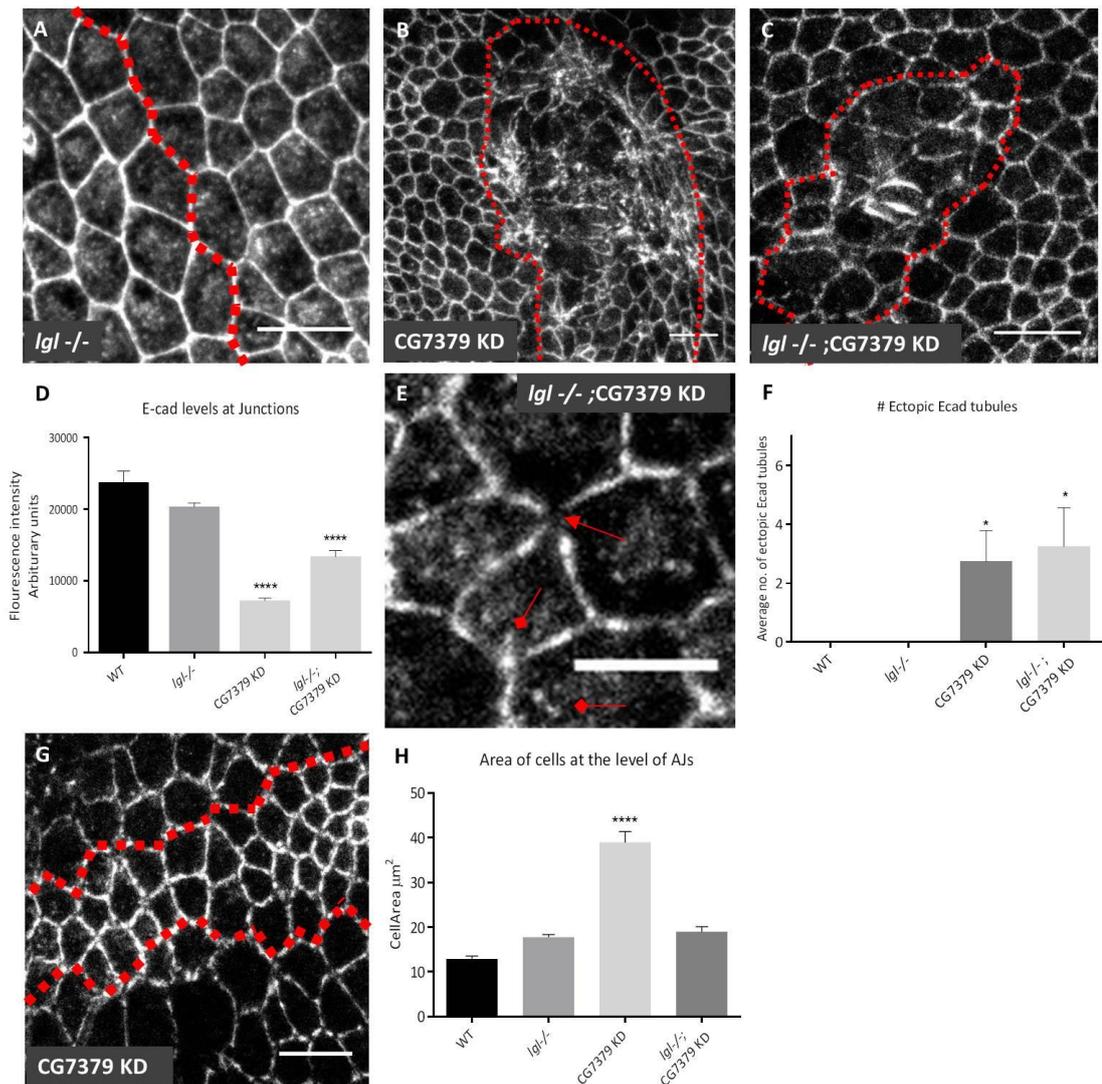
Figure 5.4 E-cadherin localisation and junction integrity is affected in Drosophila mutants

Positively marked clones from *Drosophila* nota were fixed and stained for E-cadherin. **(A)** E-cadherin staining in WT (left side of dotted line) and *lgl* homozygous mutant tissue (right side of dotted line). **(B)** Disruption to junction integrity in a CG7379 KD clone (highlighted by red dashed line). **(C)** An *lgl-/-*; CG7379 KD clone (highlighted by red dashed line) showing disruption to AJ integrity and mislocalisation of E-Cadherin. **(D)** Quantification of E-cadherin fluorescence intensity at cell junctions (n ≥40 from ≥4 animals for each genotype). **(E)** E-cadherin staining showed disruption to adherens junction architecture with junctional breaks (red triangle arrowhead) and ectopic junctional extensions (red square arrowheads) in *lgl-/-*; CG7379 KD cells. **(F)** Quantification of ectopic E-cadherin extensions (average number per animal) (n= 4 animals for each genotype). **(G)** Larger apical area was observed in CG7379 KD cells (mutant cells are in top and bottom sections). This effect is quantified in **(H)** (n≥40

from ≥4animals for each genotype). Graphs represent mean with error bars showing SEM. Scale bars in panels A, B, C and F = 50µm; scale bar in E = 5µm.

As shown, CG7379 KD clones severely altered E-cadherin localisation, with significant alterations to AJ architecture. To further study the effect of CG7379 on junction integrity, *Drosophila* nota were fixed and stained for Armadillo (β-catenin) - another core component of the AJs.

Similar to E-cadherin, Armadillo localisation was unaffected in *lgl-/-* mutant tissue when compared to WT **(Fig. 5.5 A).** However in CG7379 KD clones **(Fig. 5.5 B)** Armadillo levels were significantly reduced (p<0.0001) **(Fig. 5.5 E)** and although junction architecture remained intact, a significant number (p<0.0001) of ectopic puncta were observed within the mutant tissue **(Fig. 5.5 F).** Similarly, in *lgl-/-*; CG7379 KD clones **(Fig. 5.5 C)** Armadillo levels were also significantly reduced (p<0.0001) with a significant number of puncta (p<0.0001) observed **(Fig. 5.5 E-F)**. Additionally, Armadillo was often seen abnormally distributed **(Fig. 5.5 D),** no longer localising to the junctions.

These results show that CG7379 appears to maintain junction integrity and cell adhesion by regulating the junctional proteins E-cadherin and Armadillo (β-catenin).
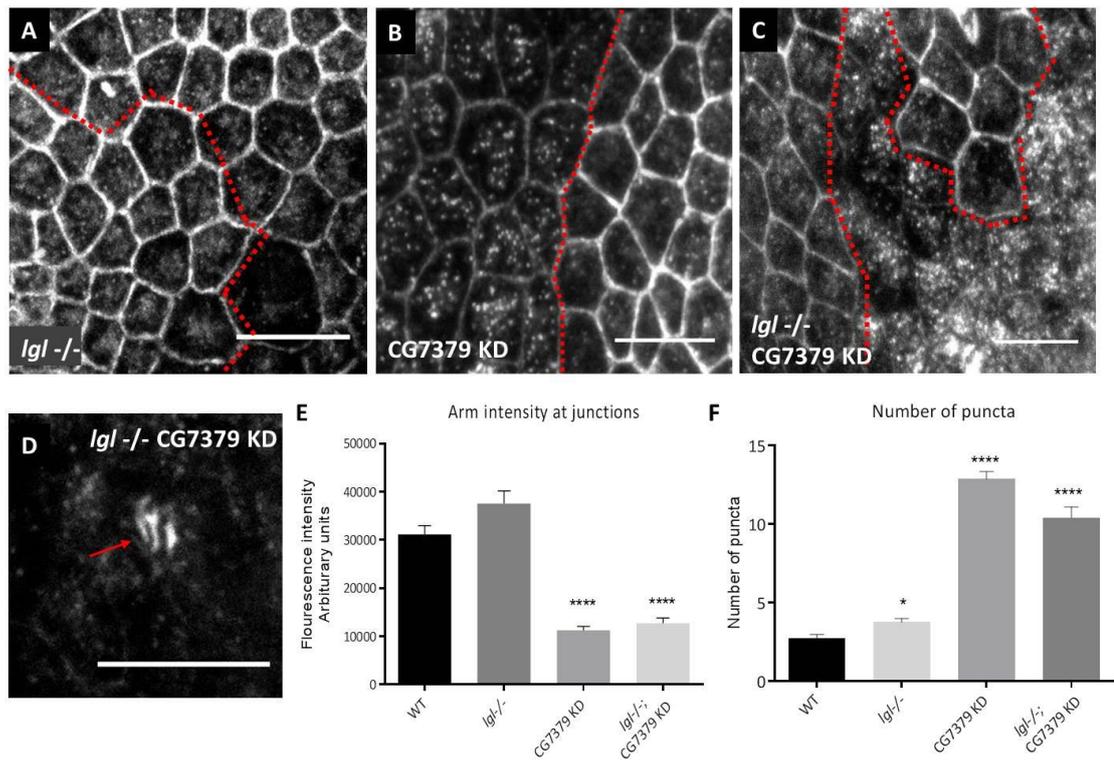
Figure 5.5 Armadillo localisation is affected in *Drosophila* mutants

Drosophila nota were fixed and stained for Armadillo (β-catenin). **(A)** Armadillo levels are similar in WT (left side of dotted line) and *lgl-/-* mutant tissue (right side of dotted line). **(B-C)** Armadillo levels are reduced and puncta are observed in CG7379 KD (B; left side of dotted line) and *lgl-/-*; CG7379 KD clones (C; right side of dotted line). Scale bars = 50µm. **(D)** Abnormal distribution of Armadillo below the junction in an *lgl-/-*; CG7379 KD clone; scale bar = 10µm. **(E-F)** Quantification of Armadillo fluorescence intensity at the AJ **(E)** and the number of ectopic puncta observed **(F)** (n>20 from ≥4 animals for each genotype) Bars represent mean with error bars plotting SEM.

5.2.4 CG7379 is important for septate junction integrity but does not affect apicobasal polarity

As shown, CG7379 is important in maintaining epithelial junction integrity and cell adhesion through its effects on E-cadherin and Armadillo, which localise to the AJs. To see whether CG7379 also has an integral role in apicobasal polarity, tissue was fixed and stained for various polarity proteins essential to regulating and maintaining epithelial polarity. FasIII localises to the septate junctions in *Drosophila* (the *Drosophila* equivalent of tight junctions in mammals) which is basally located to the AJs in flies (see **Fig. 1.5** and **Fig. 1.8**). FasIII is a homophillic adhesion molecule required for cell-cell signalling and polarity (Wells et al., 2013).

In CG7379 KD clones a significant (p=0.0294) reduction in FasIII localisation at the septate junctions was observed when compared to WT tissue **(Fig. 5.6 A, C)**. Moreover, this phenotype was significantly (p<0.0001) enhanced in *lgl-/-*; CG7379 KD clones **(Fig. 5.6 B, C)**.

This data shows that knocking down CG7379 is sufficient to affect FasIII levels at the septate junction, and that FasIII localisation is further disrupted when loss of CG7379 is combined with lgl loss.
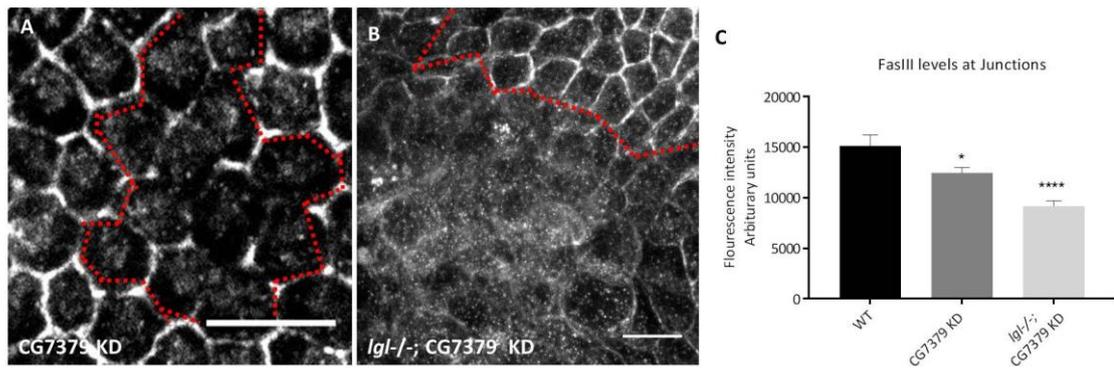
Figure 5.6 FasIII localisation is affected in CG7379 KD clones

Drosophila nota were fixed and stained with an antibody against FasIII. **(A)** FasIII localisation is disrupted in CG7379 mutant clones (mutant tissue highlighted by red dotted line) and **(B)** severely disrupted in *lgl-/-; CG7379* KD clones (mutant tissue below red dotted line). Scale bars represent 10µm. **(C)** Quantification of the fluorescence intensity of FasIII at septate junctions. Mean and SEM plotted (n≥4 with ≥50 junctions measured for each genotype).

5.2.5 CG7379 KD has no effect on the localisation of aPKC or Baz

aPKC and Bazooka (Par-3) are members of the apical Par complex and, through a process of mutual exclusion with other polarity complexes, their localisation is restricted apically to maintain apicobasal polarity (see **section 1.4**). To further investigate whether CG7379 is important for the regulation of apicobasal polarity in epithelial cells, nota were fixed and stained for either aPKC or Baz.

In *lgl -/-* mutant tissue both aPKC and Baz localisation was unaffected when compared to WT tissue **(Fig. 5.7 A)**. Similarly, CG7379 KD clones also had no effect on aPKC or Baz. Moreover, aPKC and Baz localisation was still unaffected in *lgl-/-*; CG7379 KD clones **(Fig. 5.7 C, D)**. These results suggest that CG7379 does not affect apicobasal polarity through effects on aPKC or Baz.
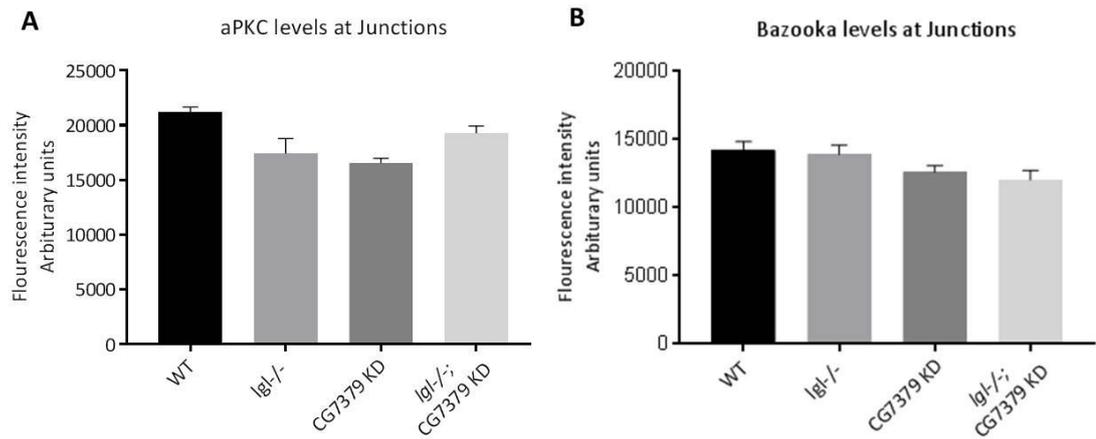
Figure 5.7 aPKC and Baz levels are unaffected in CG7379 KD animals

Quantification of the mean fluorescence intensity of **(A)** aPKC and **(B)** Baz in WT, *lgl-/*, CG7379 KD and *lgl-/-*; CG7379 KD animals. Mean and SEM plotted (n=4 with >40 junctions measured for each genotype).

5.2.6 Knocking-down ING1 in human cell lines promotes invasion

As previously shown, knocking down CG7379 in the *Drosophila* epithelium promotes tumour cell invasion. To see whether knocking down the mammalian orthologue ING1 in human cells has the same effect, and whether this cancer-promoting role is limited to epithelial cells, an *in vitro* invasion assay was performed.

RT-PCR was performed on cells harvested from breast adenocarcinoma (MDA-MB-231) and adult glioblastoma (U87) using ING1 and GAPDH specific primers to check the relative expression levels of ING1 in these cells **(Appendix D).** As moderate levels **(see Appendix D** for CT values compared to control) of ING1 were expressed in both cell lines, MDA-MB-231 and U87 cells were then transfected with siRNA targeting human ING1. Percentage knockdown was calculated using the $\Delta\Delta C_T$ method (see **section 2.3.4**) normalising to scrambled (non-silencing) and GAPDH controls. Using this formula the level of ING1 knockdown was 91% and 86% in U87 and MDA-MB-231 cells respectively **(Fig. 5.8).** Transfected cells were seeded in invasion chambers containing an ECM-like membrane and after 48hours cells that had invaded through the ECM- matrix were detached, lysed and dyed with a fluorescent label. Fluorescence intensity correlates directly to the extent of invasion. As seen in **Fig. 5.9 C**, knocking down ING1 in both U87 and MDA-MB-231 cell lines significantly (p=0.0228 and p=0.0042 respectively) increased the amount of invasive cells that had breached the membrane relative to untransfected cells. Relative to U87 and MDA-MB-231 untransfected controls, ING1 KD increased invasion by 88% and 325% respectively.

These results confirm that the down-regulation of ING1 (the human orthologue of CG7379) alone is sufficient to significantly promote invasion in mammalian epithelial cancers as seen in the *Drosophila* model. Additionally, knocking down ING1 was sufficient to promote invasion in glioblastomas.
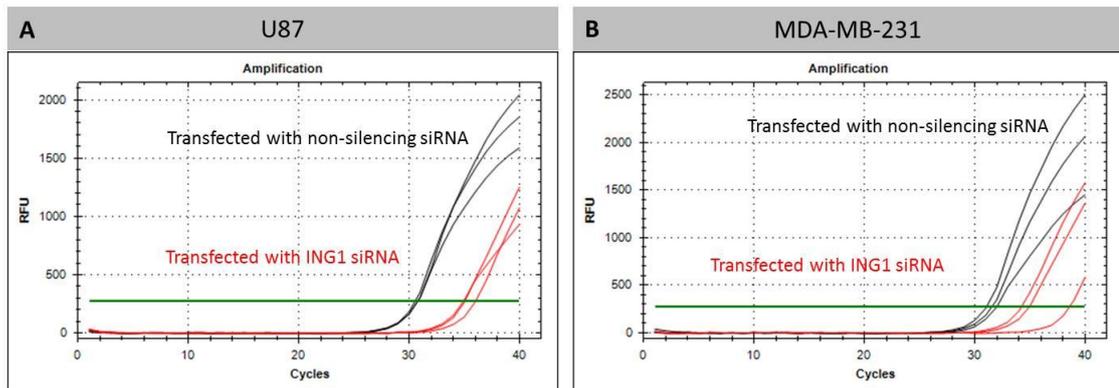
Figure 5.8 RT-PCR data for ING1 expression

(A) U87 and **(B)** MDA-MB-231 cells were transfected with siRNA targeting ING1 or scrambled (non-silencing) siRNA. After 48hrs, cells were harvested for RNA purification, and RT-PCR was carried out using ING1 specific primers (shown) and GAPDH (house-keeping) primers (for KD calculations). 3 replicates were carried out for each.
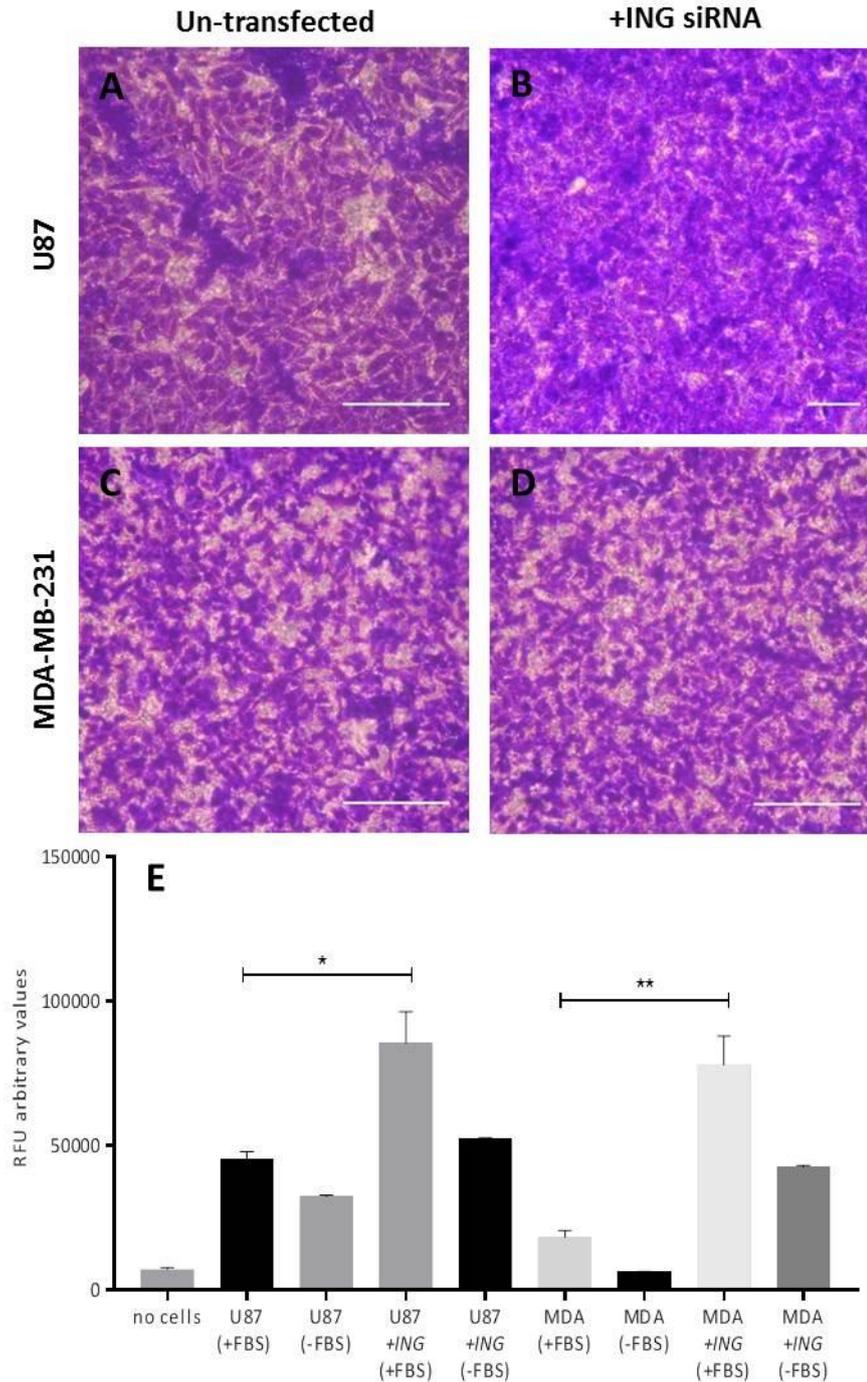
Figure 5.9 Knocking-down ING promotes invasion in mammalian cell lines

Invasion assay performed on the glioblastoma cell line U87 and the breast adenocarcinoma cell line MDA-MB-231. These cell lines were transfected with siRNA targeting the human ING1 gene, and data compared relative to un-transfected control cells. Invasion assays were performed with and without the chemoattractant FBS. **(A-D)** Both qualitative **(A-D)** and quantitative invasion assays **(E)** were conducted separately. In the qualitative assay, invasion was determined via an in situ stain. In

the quantitative assay invasion was determined by lysing cells and labelling with a dye to establish relative fluorescence. (n=3) Mean and SEM plotted. Significance was calculated when comparing treated cells to untreated cells from the same cell line (+FBS).

5.2.7 Loss of CG10600 cooperates with the *lgl[4]* mutation to promote tumour progression in the *Drosophila* notum.

Following the candidate screen **(Chapter 4)** a second novel gene, CG10600, was identified as a hit for a range of different cancer-associated phenotypes. Two distinct RNAi lines from two independent stock centres 10600R2 (NIG) and 31276 (VDRC) were used to knock-down CG10600 specifically within *lgl* -/- mutant clones. Both RNAi lines presented with similar phenotypes and in total 10 animals were imaged. Correct development of the *Drosophila* notum was often impaired, where the lateral dorsal epithelium failed to close properly during embryogenesis **(Fig. 5.10 A)**. Clones of mutant tissue were typically multi-layered **(Fig. 5.10 B)** and presented a significantly (p=0.0473) invasive phenotype **(Fig. 5.10 C, G)**. Additionally basal filopodia were significantly longer (p<0.0001) in *lgl*-/- ; CG10600 KD animals compared to *lgl* -/- alone **(Fig. 5.10 D, H)** and cells were significantly (p=0.0071) elongated along the apicobasal axis **(Fig. 5.10 I)**. Refer to **Figure 4.1** for examples of phenotypes and their relative scores, or **Appendix G (CD)** to see the average phenotypic scores given for every gene.

To see whether the observed phenotypes were due to a cooperative effect between CG10600 and Lgl, CG10600 was knocked down in a WT background. In animals of this genotype **(Fig. 5.10 E)** clones retained their typical mono-layered structure with no changes in cell length or invasion. Interestingly CG10600 KD was sufficient to produce basal filopodia which were significantly longer (P<0.0001) than WT **(Fig. 5.10 F, H)**.

These results show that knocking down CG10600 cooperates with the *lgl* mutation to promote the tumour phenotype, but is unable to elicit such an effect on its own.
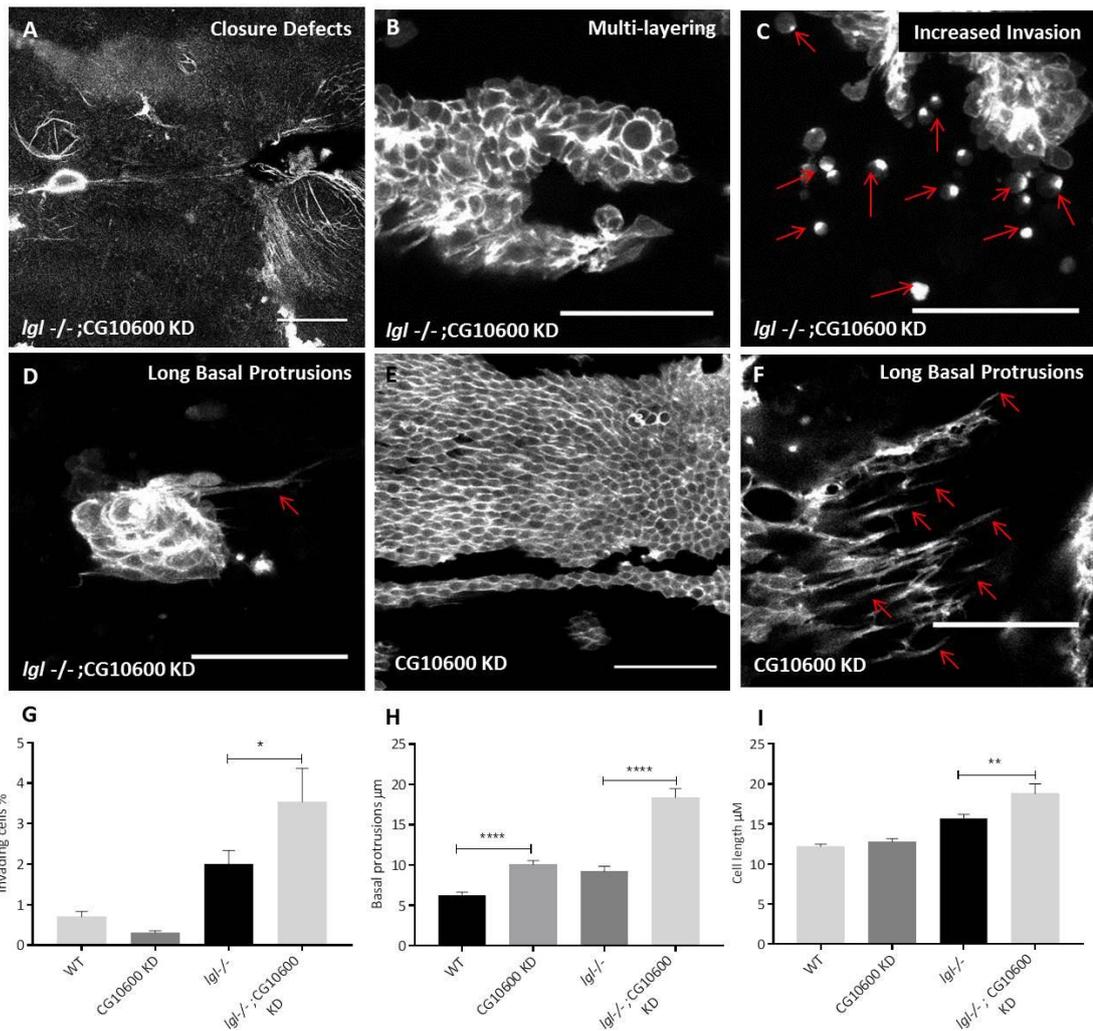
Figure 5.10 Cooperation between *lgl* mutant and CG10600 knockdown promotes tumour progression

Live confocal imaging of animals expressing an RNAi-hairpin targeting CG10600 specifically within *lgl* homozygous mutant clones in the *Drosophila* notum. Animals presented a variety of strong phenotypes including **(A)** dorsal closure defects, **(B)** multi-layering, **(C)** invasion and **(D)** long basal filopodia. **(E)** CG10600 KD without the *lgl* mutation had no multi-layering and appeared relatively WT, however **(F)** long basal filopodia were seen. Scale bars = 50μm. Quantification of **(G)** invasion, **(H)** epithelial cell length and **(I)** length of basal protrusions. For invasion and cell length n=10 animals. For basal protrusions >50 cells were measured from ≥4 animals. Error bars represent SEM.

5.2.8 CG10600 cooperates with Lgl to regulate E-cadherin expression

As CG10600 was shown to cooperate with Lgl to suppress tumour progression, nota were fixed and stained for E-cadherin to study the effects of CG10600 on AJ stability and cell adhesion.

As expected **(Fig. 5.10),** knocking down CG10600 in a WT background had no effect on E-cadherin expression levels or localisation **(Fig. 5.11 A, D).** However when CG10600 was knocked down specifically within *lgl*-/- mutant clones, the cooperative effect resulted in a significant (p=0.0008) decrease in mean E-cadherin intensity at the AJs compared to the surrounding WT tissue **(Fig. 5.11 B, E)**. E-cadherin appeared very fragmented along the junctions with junctional breaks, and even missing junctions **(Fig. 5.11 C)**.

These results confirm that CG10600 requires cooperation with Lgl to act as a tumour suppressor in this system. Moreover, the cooperation between CG10600 and Lgl maintains E-cadherin localisation, junction integrity and cell adhesion.
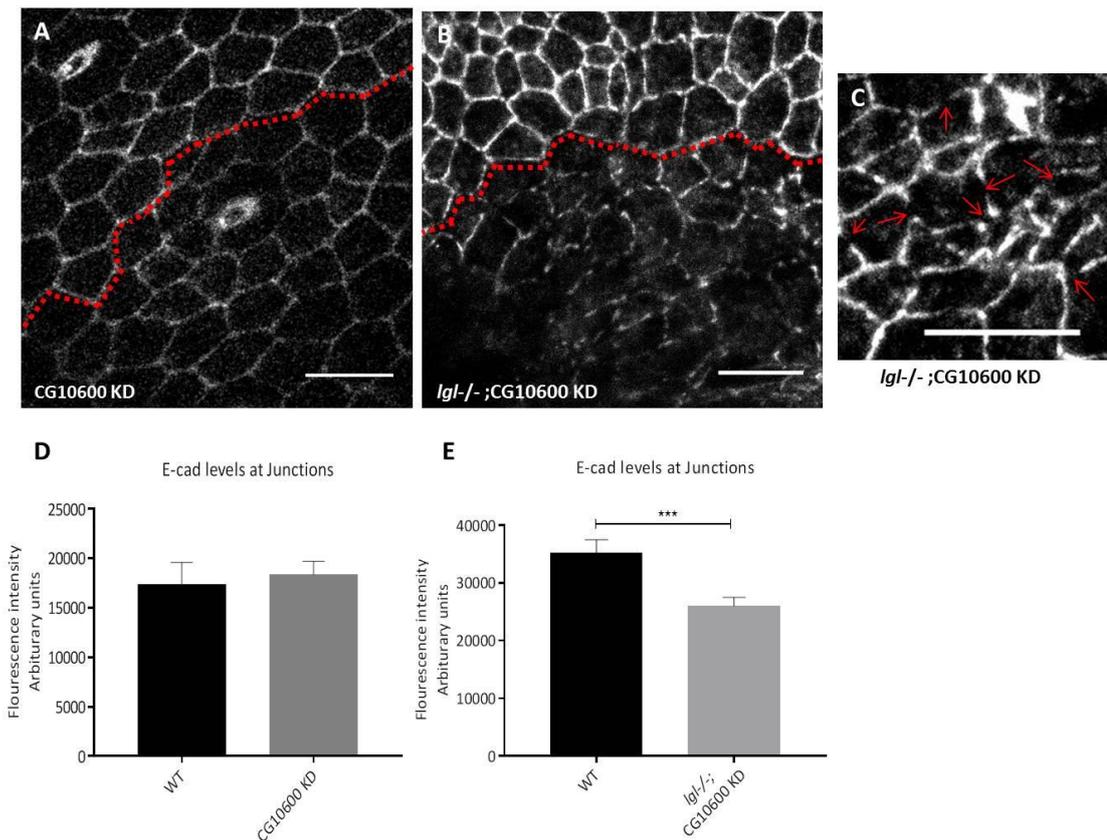
Figure 5.11 Cooperation between CG10600 and Lgl regulates E-cadherin localisation and adherens junction integrity

Drosophila nota were fixed and stained for E-cadherin. **(A-C)** Positively marked clones expressing an RNAi construct targeting CG10600 in an otherwise WT animal (**A**; below red dashed line) and specifically within *lgl* homozygous mutant clones (**B**; below red dashed line)**.** Junction breaks indicated by red arrows in **C**. Scale bars = 10μm. **(D-E)** Quantification of mean fluorescence intensity of E-cadherin at cell junctions (n>20 cells from ≥4 animals for each genotype) Bars represent mean with error bars plotting SEM.

5.2.9 Knocking-down RSBN1L in a mammalian breast adenocarcinoma cell line promotes invasion

As previously shown, knocking down CG10600 in the *Drosophila* epithelium, in combination with an underlying *lgl* mutation, promotes tumour cell invasion. To see whether knocking down human RSBN1L (the human orthologue of CG10600) has the same effect, the same *in vitro* invasion assay from **Section 5.2.7** was performed.

RT-PCR was performed on cells harvested from breast adenocarcinoma (MDA-MB-231) and adult glioblastoma (U87) using RSBN1L and GAPDH specific primers to check the relative expression levels of RSBN1L in these cells **(Appendix D).** As moderate levels (**see Appendix D** for CT values compared to control) of RSBN1L were expressed in both cell lines, MDA-MB-231 and U87 cells were transfected with siRNA targeting human RSBN1L. Percentage knockdown was calculated using the $\Delta\Delta C_T$ method (**see section 2.3.4**) normalising to scrambled (non-silencing) and GAPDH controls. Using this formula the level of RSBN1L knockdown was 67% and 86% in U87 and MDA-MB-231 cells respectively **(Fig. 5.12). Fig. 5.13** shows that knocking down RSBN1L in adult glioblastoma derived U87 cells had no effect on the invasive properties of the cells. However, in the breast adenocarcinoma derived MDA-MB-231 cell line, knocking down RSBN1L was sufficient to significantly (p=0.0003) promote invasion.

These results confirm that the down-regulation of RSBN1L (CG10600), promotes invasion in mammalian epithelial cancers as seen in the *Drosophila* model. However, RSBN1L KD was not sufficient to promote invasion in the glioblastoma cell line.

However, it can't be excluded that RSBN1L has a role in glioblastoma migration which precedes invasion.
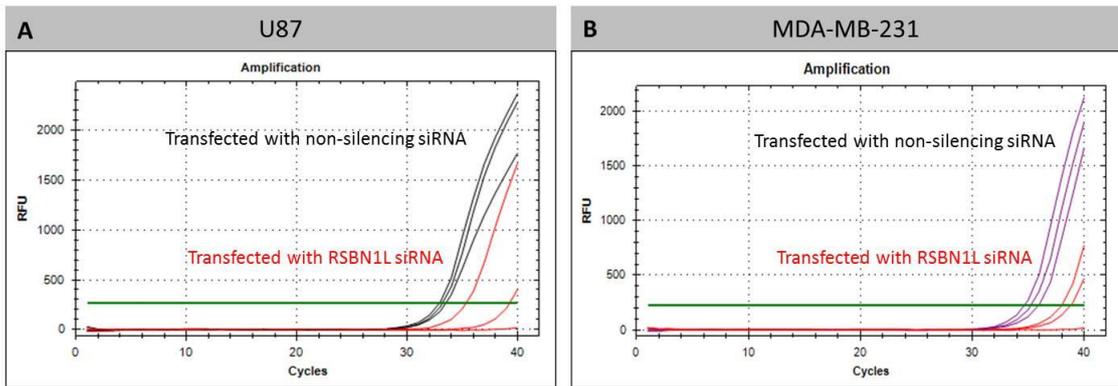
Figure 5.12 RT-PCR data for RSBN1L expression

U87 **(A)** and MDA-MB-231 **(B)** cells were transfected with siRNA targeting RSBN1L or scrambled (non-silencing) siRNA. After 48hrs, cells were harvested for RNA purification, and RT-PCR was carried out using RSBN1L specific primers (shown) and GAPDH (house-keeping) primers (for KD calculations). 3 replicates were carried out for each.
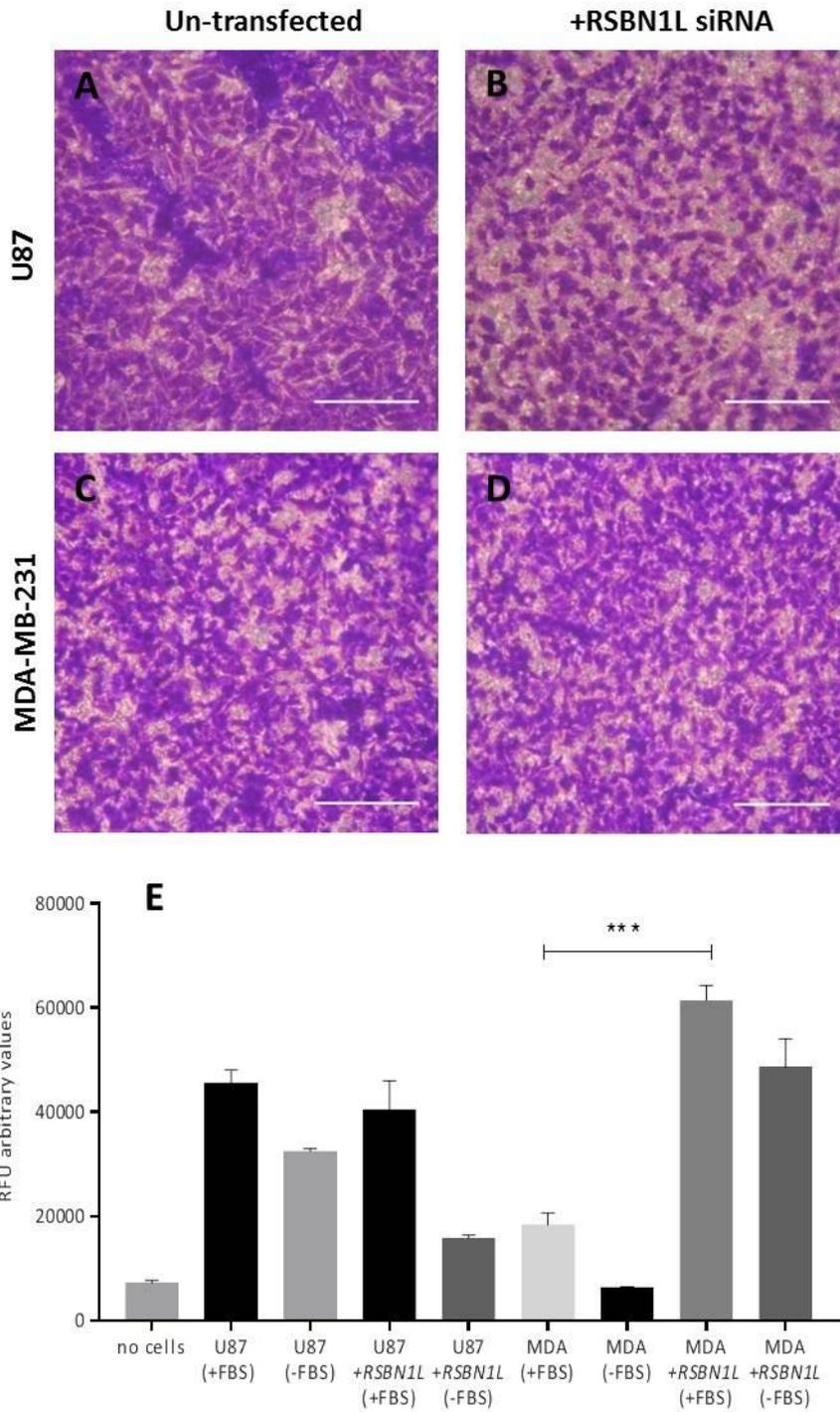
Figure 5.13 RSBN1L KD promotes invasion in MDA-MB-231 cells

These cell lines were transfected with siRNA targeting the human RSBN1L gene, and data compared relative to un-transfected control cells. Invasion assays were performed with and without the chemoattractant FBS. Both qualitative **(A-D)** and quantitative invasion assays **(E)** were conducted separately. In the qualitative assay, invasion was determined via an in situ stain. In the quantitative assay invasion was determined by lysing cells and labelling with a dye to establish relative fluorescence.

(n=3) Mean and SEM plotted. Significance was calculated when comparing treated cells to untreated cells from the same cell line (+FBS).

5.2.10 Using CRISPR/Cas technology to generate CG10600 mutants.

As CG10600 has not been characterised in *Drosophila*, experiments were limited by the lack of tools available. Therefore CRISPR/Cas technology was employed in order to generate a mutant for this gene to further study its role in cancer progression. CRISPR/Cas technology is a relatively novel approach to genome engineering and requires two main components: a guide RNA (gRNA) and an RNA-guided endonuclease (cas9). The gRNA is composed of CRISPR RNA (crRNA), a 20nt sequence complementary to the target DNA, and transactivating RNA (tracrRNA). The gRNA directs cas9 to the genomic target region creating double stranded DNA breaks. The error prone non-homologous end joining repair system then generates mutations in the DNA when repairing the cleaved site (Port, Chen, Lee, & Bullock, 2014).

Figure 5.14 illustrates the CRISPR/Cas methodology employed in order to generate a CG10600 mutant allele. CRISPR/Cas requires a guide RNA (gRNA) core which includes a 20 nucleotide target sequence. Additionally the complementary strand (protospacer) must be directly followed by a short protospacer adjacent motif (PAM) which is specifically recognised by the cas9 nuclease. **Fig. 5.14 A** illustrates the target region found in the N- terminal coding sequence of CG10600 (see **Appendix E** for the full sequence of CG10600 and location of target sequence). Annealed oligos were then cloned into the pCFD3 vector backbone **(Fig. 5.14 B)**. The vector contains a gRNA scaffold sequence- which along with PAM sequence is essential for cas9 recognition, and the vermillion gene which produces a red eye phenotype when the plasmid has been successfully inserted in the *Drosophila* genome. DNA containing the plasmid and ligated oligos were then sent for micro-injection (Department of

Genetics, Cambridge), where it was injected into 200 embryos with the genotype: y1 P(nos-cas9, w+)M(3xP3-RFPattP)ZH-2A w* **(Fig. 5.14 C)**. In this fly stock cas9 is downstream of the nanos (nos) promoter and is therefore expressed specifically in the germline (and the attp landing is marked with RFP under the 3xP3 artificial promoter). Of the surviving larvae, 41 flies emerged which possessed the vermilion red eye phenotype and of these sequencing revealed 2 which contained single nucleotide mutations (see **Section 5.2.11**).

Figure 5.14 CRISPR/Cas methodology

(A) The 20nt target sequence in CG10600 genomic DNA on *Drosophila* chromosome 2, and PAM sequence. Oligos are designed with GTCG and CAAA sequence overhangs which recognise and bind to sticky ends created when pCFD3 is digested with BbsI. **(B)** Oligos were annealed together and cloned into BbsI digested pCFD3. **(C)** Purified plasmids containing the insert were then injected into the embryos of cas9

expressing flies and *Drosophila* stocks containing potential induced CRISPR mutations over the CyO-GFP balancer chromosome were generated.

5.2.11 Identification of CG10600 mutant

CRISPR-induced mutations produce a double peak in the chromatogram following direct sequencing. These double peaks are observed when the DNA contains heterozygous alleles, indicating a single nucleotide change. Following sequencing, two stocks were observed which contained heterozygous double peaks. **(Fig. 5.15 B, D)**. The first mutant generated (CG10600*1) contained a large number of base changes throughout the genomic DNA resulting in numerous amino acid changes in the translated polypeptide **(Appendix F).** After a few generations the CRISPR* stock was checked to see if the CyO balancer was maintained within the stock. As the stock contained flies with both straight and curly wings, this showed that the mutations are not homozygous lethal. The second mutant (CG10600*2) contained only one single base pair change, which when translated produces the same amino acid (Aspartate) as found in the WT amino acid sequence, therefore it is a silent mutation.

Further sequencing of CG10600*1 is still required to determine the amino acid sequence for the full protein, which will be carried out by current lab members. Once fully sequenced, programmes such as discovery studio will be used to predict how protein structure is altered compared to WT and how this might affect function.
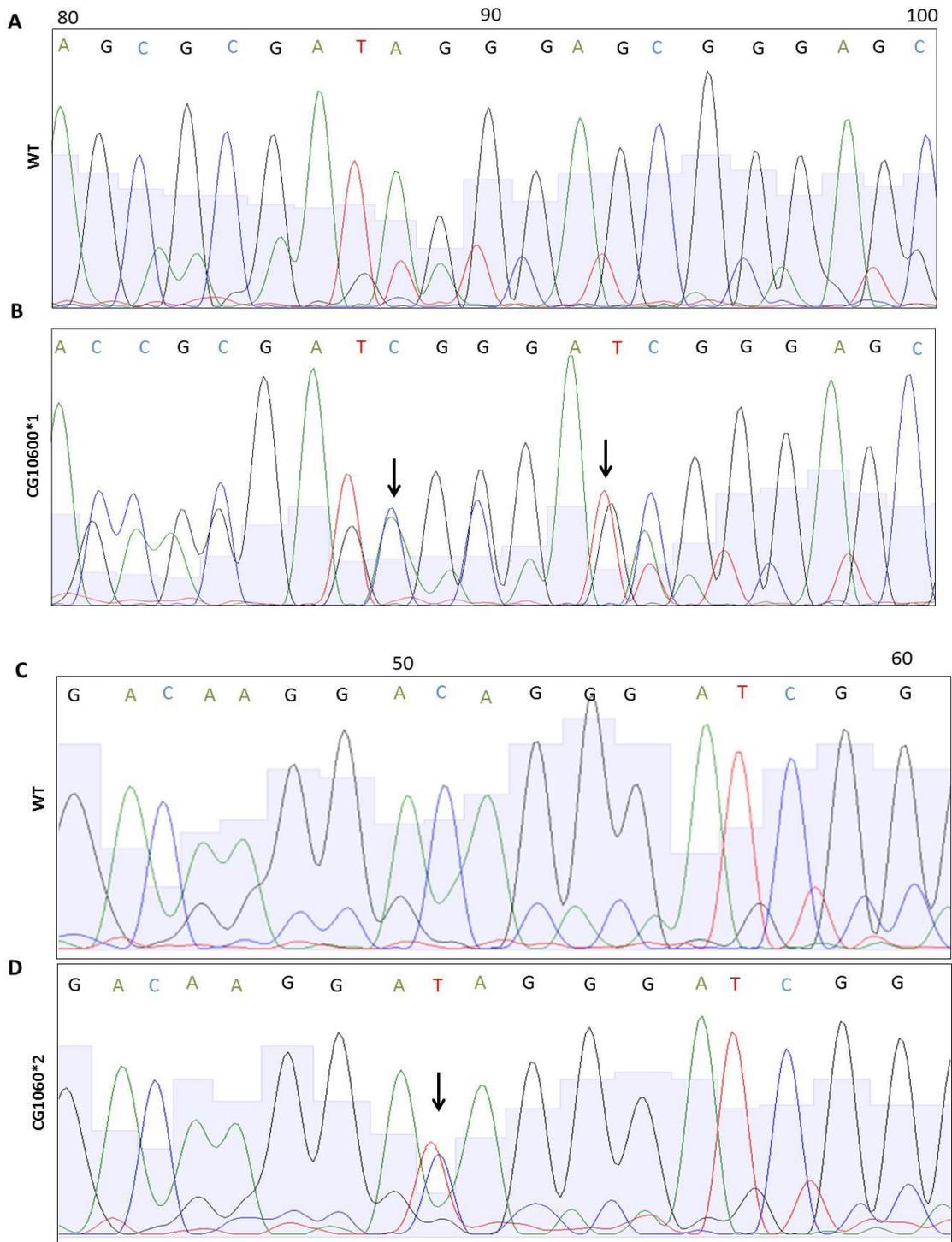
Figure 5.15 Identification of CG10600 mutants using Sanger sequencing technology.

(A, C) Chromatograms of WT, **(B)** CG10600*1 and **(D)** CG10600*2. Arrows indicate double peaks in the mutant chromatograms.

5.3 Discussion

Following the genetic screen two genes were identified as hits due to their dramatic effect on tumour progression. Both genes, CG7379 and CG10600, had never been characterised in *Drosophila*.

5.3.1 CG7379 promotes invasion through its effects on junction integrity and cell polarity

CG7379 was identified as a hit from the genetic screen due to its dramatic effect on the organisation and invasiveness of epithelial cells in the *Drosophila* notum. Knocking down CG7379 in both WT and *lgl* homozygous mutant clones in the living epithelium resulted in a significant percentage of the clonal tissue becoming invasive, with the epithelium becoming multi-layered and disorganised. The phenotype was enhanced when RNAi targeting CG7379 was combined with the *lgl[4]* mutation. This data showed that CG7379 acts a tumour suppressor, which, when down-regulated can promote tumour cell invasion and cooperates with Lgl in this system.

The human orthologue of CG7379 is ING1 which is a relatively well characterised member of ING family proteins and has been widely linked to cancer. ING has been linked to cancer through its links to P53 and its consequent role in the regulation of the cell cycle and apoptosis. Cell attrition in populations of rapidly proliferating cells acts as a natural defence against cancer. Therefore, one of the hallmarks of cancer, as first described in 2000 by Hanahan and Weinberg, is that cancer cells must be able

to evade this programmed cell death. Therefore, one possible reason for observing high numbers of invading cells in *lgl-/-*; CG7379 KD animals could be due to an inhibition of apoptosis in the absence of CG7379 function, i.e. there are normally many invading cells in *lgl-/-* mutant clones, however they usually apoptose; when combined with a loss of CG7379 function these cells do not undergo apoptosis and therefore we see many surviving invading cells. To test whether the invasive phenotype seen when knocking down CG7379 was due to reduced levels of apoptosis, I overexpressed P35 specifically in *lgl* homozygous mutant clones, to block apoptosis. However, preventing cell death in the *lgl* mutant background was not able to mimic the phenotype seen when knocking down CG7379 under the same conditions. In fact my data showed that blocking apoptosis in the *lgl* mutant background reduced the level of invasion; one possible explanation for this could be that as cells apoptose they delaminate, and some of the cells seen beneath the epithelial sheet in *lgl* mutant animals are actually dying cells. Therefore blocking apoptosis reduces the number of cells detected beneath the epithelium. From these results it is likely that evasion of cell death is not the cause of the increased number of cells observed beneath the epithelial sheet in *lgl-/-;* CG7379 KD animals. Hence, I believe CG7379 suppresses tumour cell invasion in the epithelium through other mechanisms outside of promoting apoptosis. In order for tumour cells to invade the basement membrane and surrounding local tissue many poorly understood physiological changes must occur. The Cadherin-Catenin complex which makes up the AJ forms adhesive connections between adjacent cells in the epithelium and is important for junction integrity and planar cell polarity. E-cadherin, which couples adjacent cells and is a core component of AJs, has long been identified as a

suppressor of invasion and is downregulated in many cancers. Meanwhile, Armadillo (β-catenin) links the cytoplasmic tail of E-cadherin to α-catenin, creating a network which directly interacts with the actin cytoskeleton (Gumbiner, 1996).

Knocking down the novel *Drosophila* gene CG7379, either with or without an accompanying *lgl* mutation, significantly altered E-cadherin levels and localisation. AJ breaks and increased levels of E-cadherin in ectopic junctional tubules have been observed previously in mutants that affect the turnover of junctional material (Georgiou et al., 2008). In order to maintain stable AJs, E-cadherin and other junctional proteins are actively recycled through clathrin-mediated endocytosis. This turnover of cadherin is essential to the dynamic nature of the AJs, allowing the junction to be regularly assembled and disassembled during cell division, death, development and invasion. Genes that affect cadherin junctional recycling have been shown to include regulators of the actin cytoskeleton, p120 and cdc42 (Davis, Ireton, & Reynolds, 2003; Georgiou et al., 2008). Therefore this data suggests that CG7379 may be involved in regulating this recycling process. CG7379 has been proposed to play a role in chromatin remodelling; therefore CG7379 KD could result in a decreased expression of genes that regulate this recycling process. Additionally, the human orthologue of CG7379, ING1, has recently been shown to regulate senescence and the Rb pathway by inducing expression of Intersectin-2, which is involved in endocytosis (Rajarajacholan, Thalappilly, & Riabowol, 2013). This supports the hypothesis that CG7379 might be affecting junction integrity and E-cadherin localisation by regulating endocytosis.

Further experiments would therefore be used to test whether the ectopic E-cadherin structures in CG7379 KD animals are due to a reduction in the endocytosis-mediated recycling of junctional material. For this, fluorescently labelled dextran could be used in an assay for endocytosis, as it is preferentially internalised by early endosomes. By staining WT, *lgl-/-*, CG7379KD and *lgl-/-*; CG7379 KD nota with E-cadherin and fluorescently-labelled dextran, we would be able to see whether E-cadherin and dextran co-localise in ectopic structures and if endocytosis is down-regulated in CG7379 mutants.

Knocking down CG7379 significantly reduced Arm, and similar to E-cadherin, ectopic puncta were highly abundant. Future experiments could test whether E-cadherin and Armadillo ectopic puncta and tubules remain at the cell surface due to reduced junctional turnover. To do this, nota could be fixed and stained for antibodies against extracellular E-cad/Arm in PBS before permeabilisation with PBT. Then, following permeabilisation, tissue would be stained for intracellular proteins and imaged. Additionally, CG7379 KD significantly reduced levels of the cell adhesion glycoprotein FasIII. FasIII is an important component of the septate junction (SJ) - which mediates transport between adjacent cells, and maintains junction integrity. As both Arm and FasIII were seen to be mislocalised, loss of CG7379 is likely to affect cell adhesion as a whole, and is not limited to E-cadherin or the AJ. To strengthen this claim, mutant nota could be stained for other proteins known to localise to the SJ including dlg, coracle and neurexin. Moreover, phalloidin could be used to stain actin filaments to see how CG7379 KD affects the actin cytoskeleton.

Apicobasal polarity is fundamental to normal cell physiology and defines distinct apical, lateral and basal domains for the asymmetric organisation of cell components (Fristrom, 1988; Januschke & Gonzalez, 2008). Various polarity proteins are important in the formation and maintenance of this epithelial polarity and are commonly mis-regulated in cancer (see **Section 1.3**). The data showed that both aPKC and Baz expression/localisation was not visibly affected when knocking down CG7379. These results indicate that CG7379 appears to be involved in the regulation of junction integrity at both the AJ and SJ, but does not seem to be directly involved in regulating apicobasal polarity. To look further at the dynamics of the AJ and SJ in CG7379 KD animals' fluorescence recovery after photobleaching (FRAP) experiments could be used to study the recovery of junctional components at sites which have been photobleached. FRAP has become a popular tool in *Drosophila* for studying junction dynamics for example, De Beco et al (2009) used this approach to show how endocytosis regulates levels of E-cadherin at the AJs, while others have used FRAP to study the dynamics of core SJ components (de Beco, Gueudry, Amblard, & Coscoy, 2009; Oshima & Fehon, 2011). To verify that apicobasal is unaffected, mutant nota could be stained for other polarity proteins such as crumbs.

The human orthologue ING1 was also shown to promote invasion when knocked down in mammalian cancer cell lines *in vitro*. As well as promoting invasion in the epithelial breast adenocarcinoma cell line MDA-MB-231, knocking-down ING1 also promoted invasion in an adult glioblastoma (U87). These data suggest that ING1 anti-invasive functions in different types of cancer, and indicates that the research carried out in the *Drosophila* model can be translated to human cancers in this instance. To

see whether ING1 KD promotes invasion by impairing junction integrity *in vitro* as seen *in vivo*, transfected cells could similarly be fixed and stained for E-cadherin (AJ), ZO-1 (TJ) and other junctional proteins. For this, both transformed and untransformed cell lines (MDCK) would be used. MDCK cells forms polarised cysts in vitro and are commonly used to study epithelial cell polarity and junction integrity (A. Z. Wang, Ojakian, & Nelson, 1990). Additionally, experiments could be repeated with other ING family members: ING2-5, which have also been implicated in tumour progression.

All together these results show that CG7379 has a strong tumour suppressive function by regulating junction stability and the adhesive properties of epithelial cells. Moreover, downregulation of CG7379 in cancer is sufficient to significantly promote invasion through its effects on junctional integrity and cell-cell adhesion. Future experiments will involve investigating the role of CG7379 in maintaining junction integrity as a defence against invasion and whether these junction defects are mimicked with ING1 KD in mammalian cells. Additionally, apoliner, TUNEL assay and anti-caspase-3 staining could be used to verify that the invasion seen in CG7379 KD is not a result of decreased apoptosis.

5.3.2 CG10600 cooperates with Lgl to enhance the tumour phenotype

During the genetic screen CG10600 was a strong hit for a range of cancer-associated phenotypes. These included multi-layering, invasion, reduced apical area, dorsal closure defects and long basal filopodia. These phenotypes however were not prominent when CG10600 was knocked down alone without the *lgl4* mutation. This

suggests that CG10600 cooperates with Lgl in the *Drosophila* notum to prevent tumour progression.

This cooperative affect was also shown to alter E-cadherin localisation, and resulted in breaks in the AJs. This shows that CG10600 and Lgl cooperate in a tumour suppressive mechanism by regulating E-cadherin and AJ stability. To further study the effects of CG10600 KD on AJ stability, other AJ markers could be stained for, including Arm and α-catenin. Additionally, it would be interesting to see whether unlike CG7379, CG10600 directly affects apicobasal polarity by staining for antibodies against aPKC and Baz. As CG10600 KD affected E-cadherin expression, both a live-cadherin marker (ubi promoter drives expression of E-cadherin tagged GFP) and FRAP experiments could also be used to look at junction dynamics in real time. The effects seen with CG10600 were less severe than those seen with CG7379; therefore it is likely that loss of junction integrity is not the main cause of the invasive phenotype.

The human orthologue RSBN1L, which remains uncharacterised in mammalian systems, was shown to promote invasion when knocked-down in a breast adenocarcinoma cell line, but had no effect on invasion in adult glioblastomas. This suggests that the tumour suppressing function of RSBN1L (CG10600) may be restricted to epithelial cancers. With breast adenocarcinomas arising from epithelial cells and glioblastomas found in the brain deriving from astrocytes, these differing cell types might suggest why RSBN1L appeared to promote invasion in a tissue specific manner. The switch between hyper-proliferative and invasive phenotypes in glioblastomas has been shown to arise mostly from signals from the tumour

microenvironment. Additionally, increased anaerobic glycolysis, hypoxic conditions and increases in angiogenic factors are largely associated with increases in glioblastoma invasiveness (Xie, Mittal, & Berens, 2014). Recently, evidence has bought to light similarities between pathways associated with EMT in epithelial derived cancers and those associated with glioblastoma invasion (Zarkoob, Taube, Singh, Mani, & Kohandel, 2013). These include c-myc, EGF, TGFβ, NFκB and WNT signalling (Bruna et al., 2007; Dhruv et al., 2013; P. Lu et al., 2016; Q. Meng et al., 2014; Zheng, Yang, Aldape, He, & Lu, 2013). As previously discussed ING1 is capable of promoting tumour invasion in both U87 and MDA-MB-231 cells and so it could be hypothesised that this gene acts through signalling pathways and processes involved in both tumour types. RSBN1L KD on the other hand was only sufficient to promote invasion in MDA-MB-231 cells which could suggest that this gene acts via different molecular mechanisms and signalling pathways associated with the adaptive EMT process and not glioblastoma invasion.

As CG10600 has never been characterised in *Drosophila*, experiments were limited by the lack of tools available. Therefore to try and improve our understanding on how this gene works CRISPR/Cas technology was used to generate a mutant, which could then be studied further. Further sequencing will be required to determine the full amino acid sequence of the mutant generated and predict protein structure using tools such as discovery studio and I-Tasser. From this, the mutant could be recombined with an FRT site to allow a clonal analysis of mutant tissue in a WT background using MARCM. The mutant allele could then be put in trans with a deficiency to see if the phenotype is enhanced and test if the mutant is a null.

In conclusion, two novel genes: CG7379 and CG10600, previously uncharacterised in *Drosophila* were found to promote tumour progression in the *Drosophila* notum. Early characterisation experiments show that a loss CG7379 function promotes invasion, not by a loss of apoptosis, but through a loss of AJ and SJ integrity. CG10600 KD was also shown to affect AJ integrity to a limited extent and new mutant fly stocks were generated to further our understanding of CG10600 function. Both genes were also shown to promote invasion when knocked down in human cancers in vitro, strongly implicating these genes as possible candidates to study further as potential future therapeutic targets.

CHAPTER 6

Characterisation of Dscam

Chapter 6 Characterisation of Dscam

6.1 Introduction

Drosophila Dscam (Down syndrome cell adhesion molecule) is a member of the immunoglobulin superfamily, which is integral to various neurological processes in the central nervous system including synaptic adhesion, self-avoidance and axon guidance. Dscam is a transmembrane protein consisting of 10 immunoglobulin (Ig) domains, 6 fibronectin III extracellular domains, and a divergent cytoplasmic tail **(Fig. 6.1)**. There are 4 Dscam homologues in *Drosophila* (Dscam1-4); however it has been shown that more than 38,000 isoforms can be generated through alternative splicing of Dscam1 alternative exons. Dscam1 genomic DNA spans approximately 60,000bps containing 115 exons, of these 21 are constituently expressed while the remaining 95 exons are divided into 4 cassettes, which are alternatively spliced. Exon clusters 4, 6, 9 and 17 contain 12, 48, 33 and 2 exons respectively which can be alternatively spliced into the mRNA exclusively **(Fig. 6.1).**
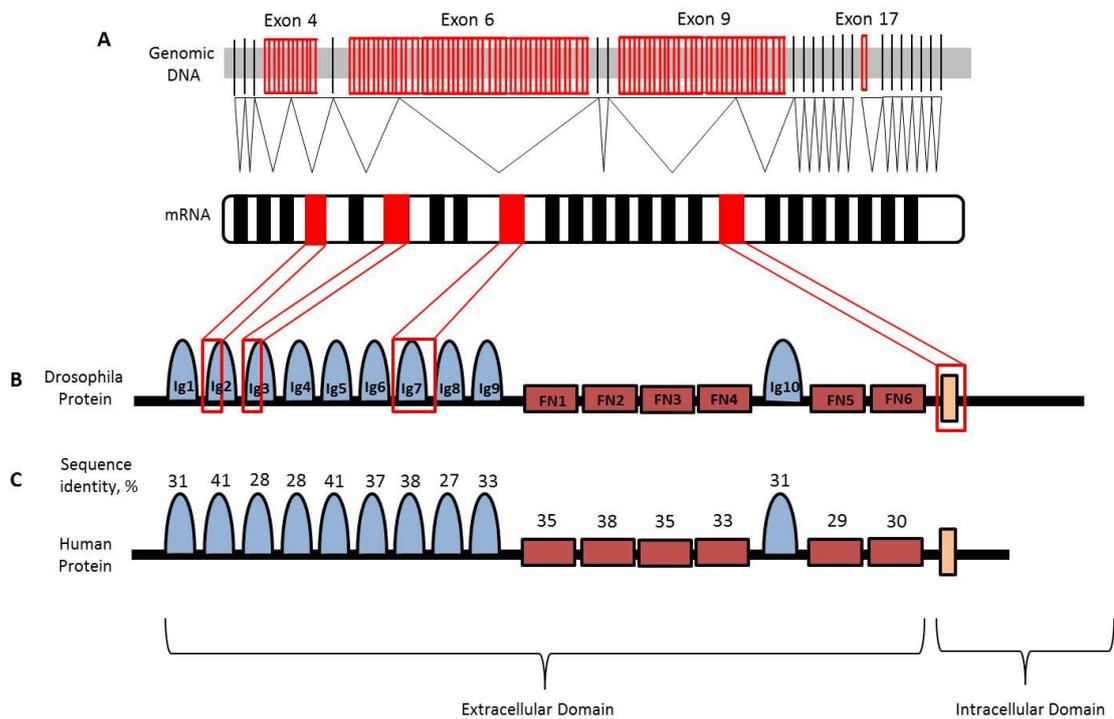
Figure 6.1 Alternative splicing in *Drosophila* Dscam1 and sequence homology with human Dscam

Drosophila Dscam1 contains 4 exon cassettes which are alternatively spliced into mRNA resulting in >38,000 different possible isoforms **(A)**. The translated protein contains the same immunoglobulin and fibronectin domains however Ig2, Ig3, Ig7 and the transmembrane domain have differing amino acid sequences due to alternative splicing **(B)**. Human Dscam shows high sequence identity in the extracellular domain, however little sequence identity in the intracellular domain **(C).**

Isoform diversity is thought to be vital to the function and specificity of Dscam1 in the nervous system (Schmucker et al., 2000). Neurites belonging to the same neuron express a subset of the same Dscam1 alternatively spliced isoforms. During self-avoidance identical isoforms selectively recognise each other and exhibit homophillic interactions causing them to repel one another. Meanwhile neighbouring cells can contact one another as they express differing Dscam1 isoforms (B. E. Chen et al., 2006; Hattori et al., 2007; Hughes et al., 2007; Matthews et al., 2007; Soba et al., 2007). Similarly, Dscam2 mediates tiling whereby neurites belonging to the same cell

type promotes homophillic repulsion (Millard, Flanagan, Pappu, Wu, & Zipursky, 2007).

There are two closely related mammalian orthologues of *Drosophila* Dscam: DSCAM and DSCAM1L. These genes function in a similar manner; however unlike their *Drosophila* orthologues do not undergo alternative splicing. It has been suggested that unlike the highly conserved diversity generated in *Drosophila* via alternative splicing, the human genome has developed other mechanisms to generate such diversity. For example, combinations of genes and large protein families in mammals give rise to a diverse number of neuronal cell surface proteins which are similar to Dscam, capable of recognising each other and functioning in a similar way (Wojtowicz, Flanagan, Millard, Zipursky, & Clemens, 2004).

Compared to *Drosophila*, human DSCAM shows 29% identity with 46% similarity, and high homology within the extracellular domain. The intracellular domains however show no sequence similarities **(Fig. 6.1)** (Schmucker et al., 2000). Both *Drosophila* and mammalian Dscam regulate the actin cytoskeleton through Pak/JNK signalling. In humans, Dscam directly interacts with Pak1 whilst *Drosophila* Dscam activates Pak1 through the SH3/SH2 adaptor protein Dock. Pak1 is a Ser/Thr kinase linking small GTP-binding proteins (cdc42 and Rac) to JNK signalling. Rac1 has been shown to enhance the interactions between DSCAM and Pak1. Once activated, Pak1 phosphorylates MAP2K1 which in turn activates downstream MAPKs and JNK, mediating a variety of physiological processes including cytoskeletal reorganisation and migration **(Fig. 6.2)** (Qu et al., 2013; Schmucker et al., 2000).
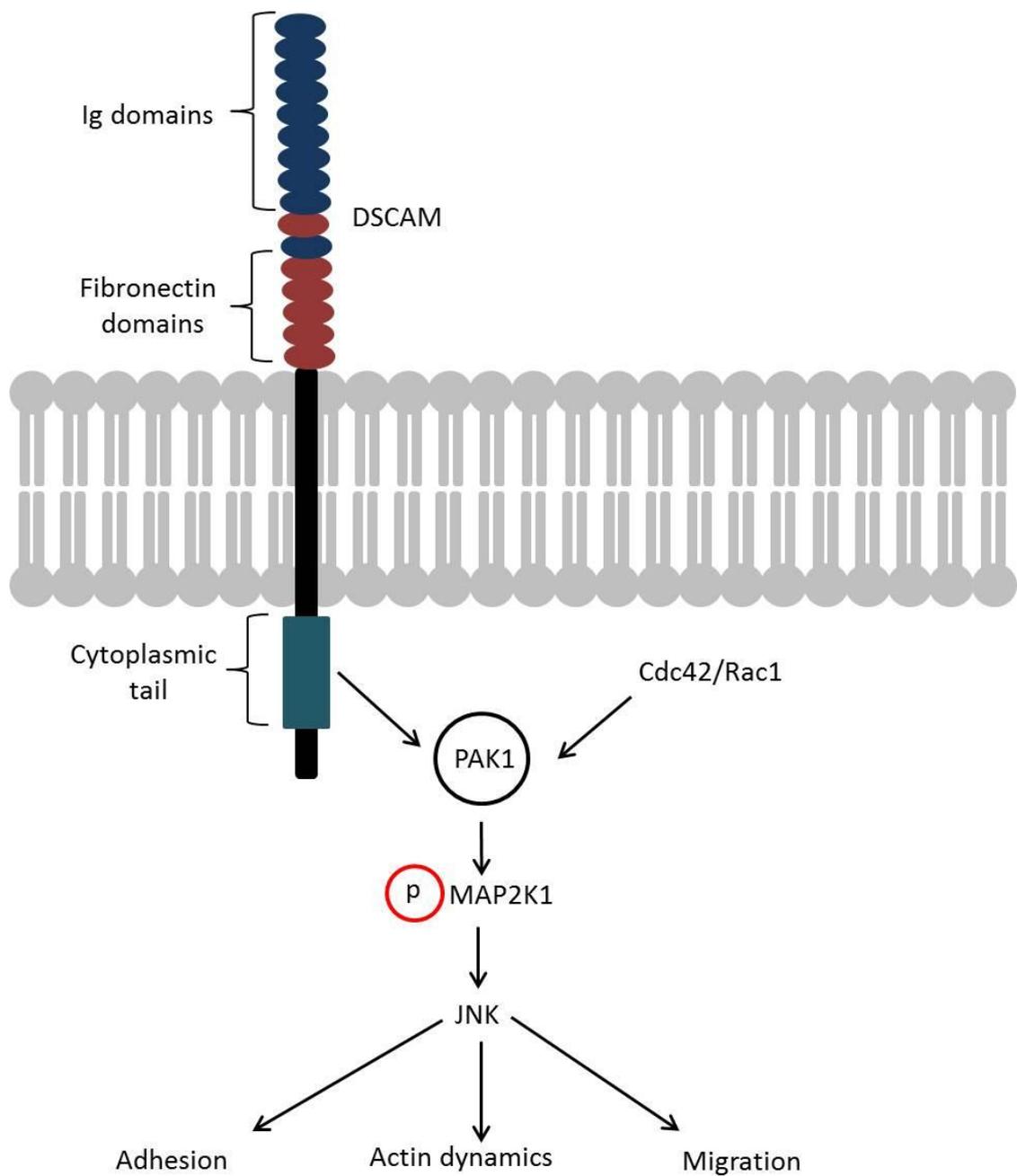
Figure 6.2 Dscam activates PAK1 and JNK signalling

The cytoplasmic tail of transmembrane DSCAM interacts with PAK1 (enhanced by cdc42/Rac1) which leads to phosphorylation and activation of MAPK and JNK signalling pathways.

Additionally, in the nervous system Dscam has been shown to act as a Netrin-1 receptor to regulate neuronal guidance (Ly et al., 2008). Netrin-1 is a chemoattractant which is secreted by plate floor cells in the nervous system which

guides the neuronal growth cone for axon extension (Kennedy et al, 1994). Inactivation of Netrin-1 signalling and loss of similar receptors including DCC (deleted in colorectal cancer) have been implicated in the progression of malignant tumours (Ly et al., 2008).

Clinically, DSCAM is implicated in many human diseases including Downs syndrome and congenital heart disease. Expression of DSCAM is regularly increased in various neurological disorders; however it is has also been shown to be deregulated in mammalian cancer including breast cancer (Miano et al., 2016). While Dscam has been studied extensively in the nervous system; no-one has really looked at how it may function with regards to tumour progression or its biological role outside of the nervous system.

6.2 Results

6.2.1 Knocking-down Dscam1 in *lgl-/-* mutant background promotes tumour progression

During the candidate screen **(Chapter 4)** Dscam was identified as a strong hit due to its drastic effect on tumour progression. Dscam remains extremely novel with regards to its role within the epithelium and cancer development, however due to it being well characterised in the nervous system there are many genetic tools available, making it an ideal candidate to study further.

A variety of strong cancer-related phenotypes were observed when RNAi targeting Dscam1 was expressed alongside the *lgl[4]* mutation on the back of the fly. Many animals presented with developmental problems whereby the lateral dorsal epithelium failed to close properly during embryogenesis **(Fig. 6.3 A)**, whilst the size of clonal tissue and percentage coverage within the notum was significantly reduced (p=0.0135) compared to *lgl-/-* mutant animals. Moreover, the apex of mutant cells were significantly smaller (p<0.0001) and malformed compared to cells in *lgl-/-* mutant clones alone **(Fig. 6.3 B, H)**. As expected from neoplastic tumours, clones were highly disorganised with cells growing on-top of one another in a multi-layered fashion **(Fig. 6.3 C)** to a greater extent than that seen in *lgl-/-* animals. See **Appendix G (CD)** for all genes and scores given for each phenotype. Compared to WT there was also a significant (p=0.0022) increase in invasion in *lgl-/-*; Dscam1 KD animals however, there was no significant difference when compared to *lgl-/-* **(Fig. 6.3 D, J).** Additionally, basal filopodia were significantly longer (p<0.0001) than those seen in

lgl-/- mutants **(Fig. 6.3 E, I).** See **Section 4.2.1** for examples of phenotypes and their scores compared to *lgl-/-*.

Additionally, using confocal microscopy, movies could be made to see how tumours behaved over time. **Figure 6.4** shows an invasive cell which was detected beneath the epithelium. Due to being labelled with UAS-Moe-GFP the cell would have originated from the *lgl-/-*; Dscam1 KD mutant clone. As shown, the invasive mutant cell had long intermediate lamellipodia and basal filopodia, which, when followed over time were very dynamic. Additionally, the invading cell moves laterally in the XY plane. The images in **Fig. 6.4** taken every minute show these actin rich protrusions extending and retracting over time. While invading cells with protrusions were only seen a few times in *lgl-/-*; Dscam1 KD animals, they were never seen in *lgl-/-* alone. For comparison, **Fig. 6.5** shows a delaminated *lgl-/-* mutant cell which moves laterally in the XY plane but does not have any lamellipodia or filopodia.

Dscam1 was then knocked down in a WT background to see whether the phenotypes observed were due to Dscam1 alone or cooperation between Dscam1 and Lgl. Knocking down Dscam1 alone, without the *lgl[4]* mutation however was not sufficient to produce neoplastic tumours on the back of the fly **(Fig. 6.3 F)**, with cells appearing characteristically WT. This data suggests that in this system, Dscam1 cooperates with Lgl to promote tumour progression.

All together these results show that Dscam1 and Lgl cooperate in the *Drosophila* notum in a tumour suppressive manner, and when both are down-regulated produce

203

highly disorganised tumours with constricted apices and long dynamic basal filopodia.

A **Closure Defect**
lgl-/- ;Dscam1 KD

B **Small Apices**
lgl-/- ;Dscam1 KD

C **Multi-layering**
lgl-/-;Dscam1 KD

D **Invasion**
lgl-/- ;Dscam1 KD

E **Long Basal Protrusions**
lgl-/- ;Dscam1 KD

F
Dscam1 KD

G

H

I

J
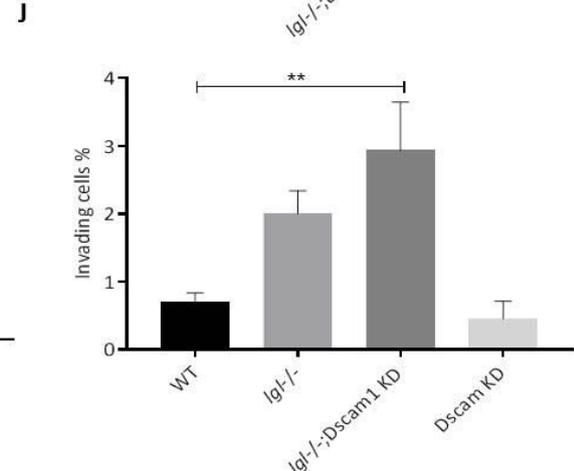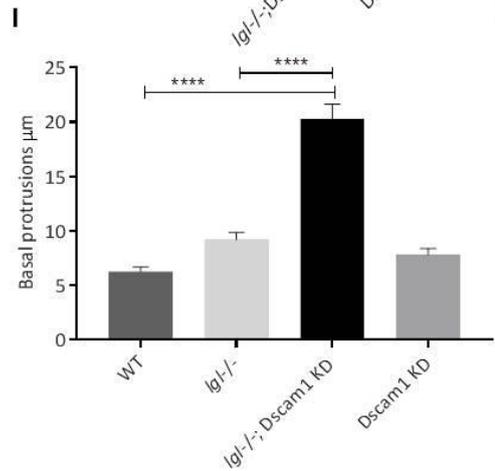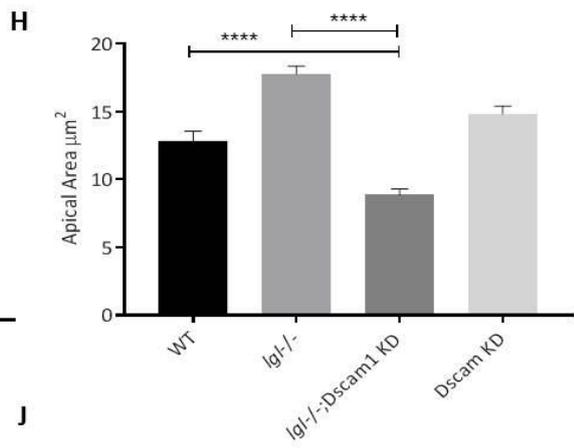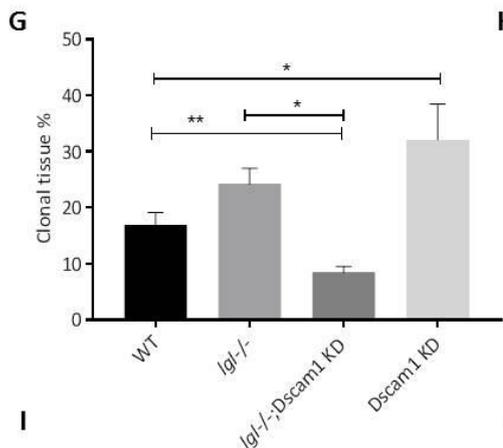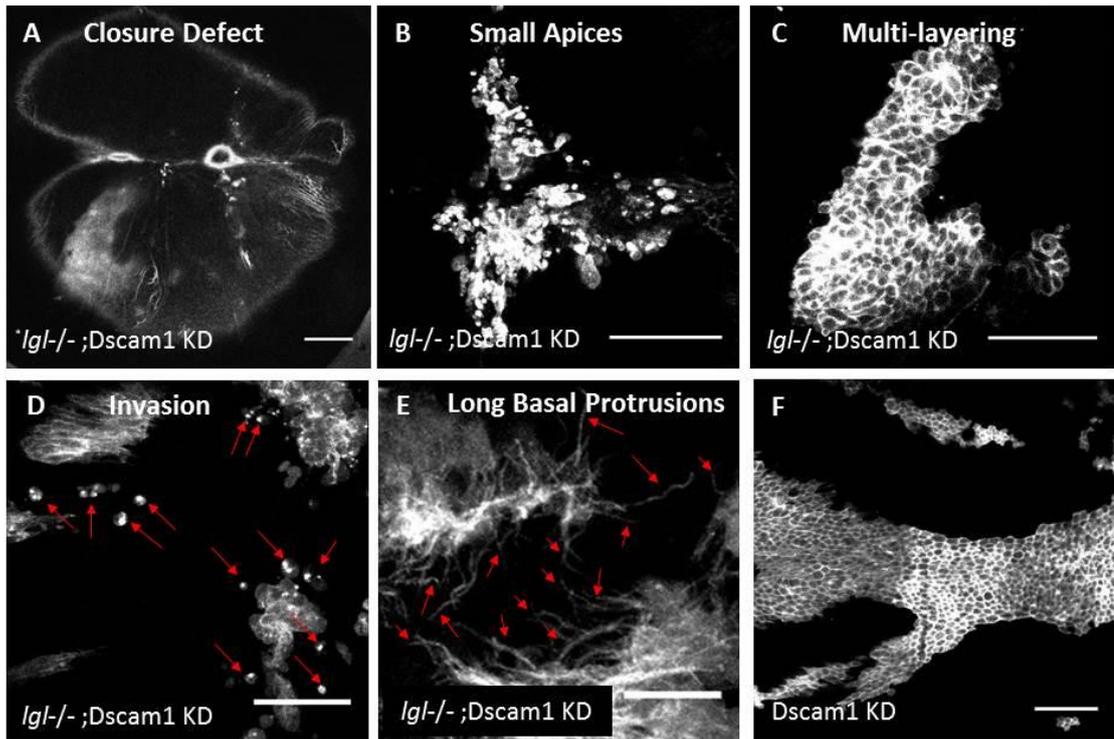
Figure 6.3 Cooperation between *lgl* mutant and *Dscam1* knockdown promotes tumour progression

(A-E) Live confocal images of clones within the notum expressing RNAi targeting Dscam1 specifically within *lgl-/-* mutant tissue. A range of cancer-associated phenotypes were observed including, **(A)** dorsal closure defects, **(B)** reduced apex size, **(C)** multi-layering, **(D)** invasion (red arrows) and **(E)** long basal filopodia (red arrows). **(F)** Confocal image of animal expressing RNAi targeting *Dscam1* without the *lgl[4]* mutation. **(G-J)** Quantification of **(G)** the percentage of clonal tissue within the notum, **(H)** apex area, **(I)** length of basal filopodia and **(J)** invasion. Mean and SEM plotted. Number of animals quantified for clonal tissue: WT: n=10, *lgl-/-*: n=32, *lgl-/- ;Dscam1*-RNAi: n=7, *Dscam1*-RNAi: n=5. For apex size and basal filopodia, n≥5 animals with >40 cell apices/protrusions measured for each genotype. Scale bars = 50μm.

t=1 t=2 t=3 t=4

t=5 t=6 t=7 t=8

t=9 t=10 t=11 t=12

t=13 t=14 t=15 t=16

t=17 t=18 t=19 t=20

t=21 t=22 t=23 t=24

t=25 t=26 t=27 t=28
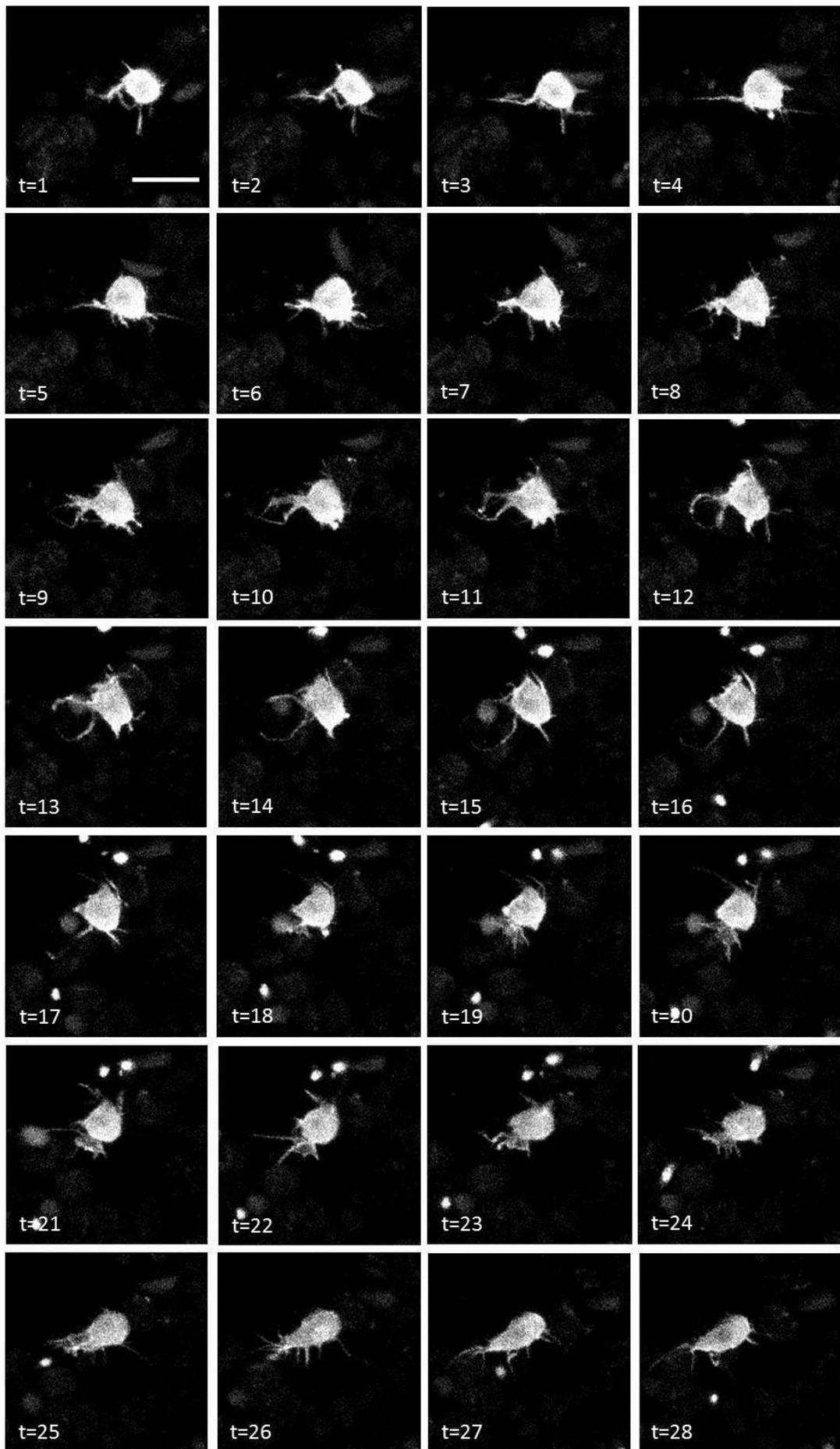
Figure 6.4 Delaminated *lgl-/-*; Dscam1-RNAi mutant cell has dynamic protrusions

Confocal snap-shots from a real-time movie following a mutant cell that had delaminated from the epithelium. Cell is homozygous mutant for *lgl* with Dscam1 simultaneously knocked down. Snapshots taken every minute (t); scale bar = 20μm.
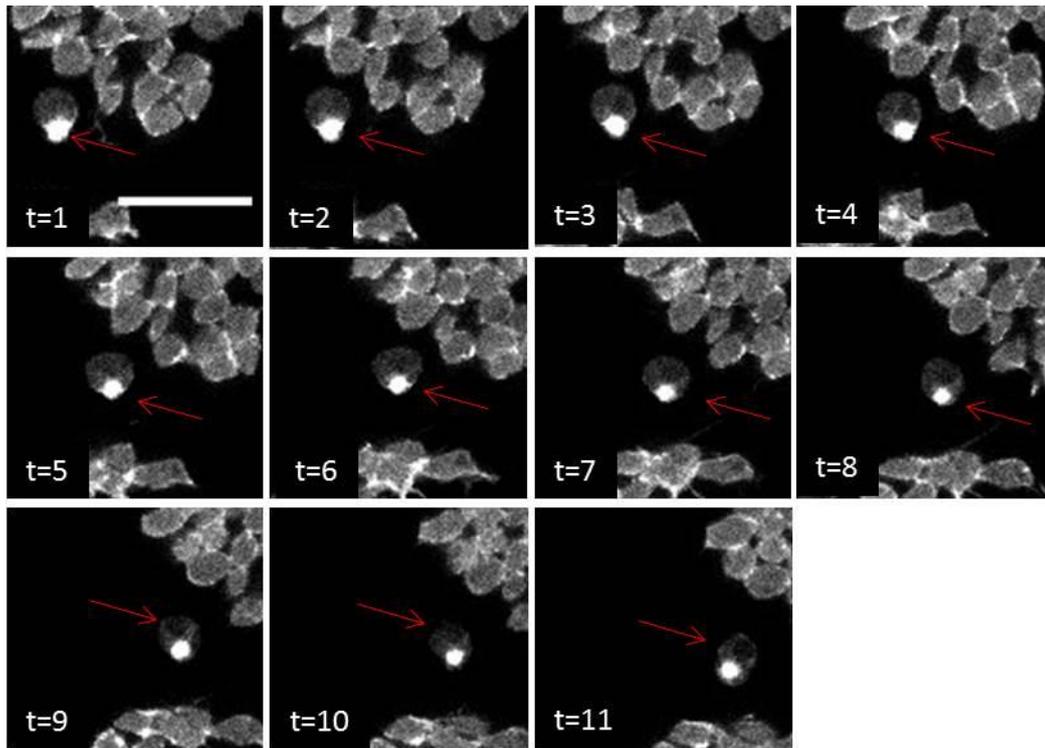
Figure 6.5 Delaminated *lgl*-/- mutant cell

Confocal snap-shots from a real-time movie following a mutant cell that had delaminated from the epithelium. Cell is *lgl* homozygous mutant. Snapshots taken every minute (t); scale bar = 20µm.

6.2.2 *Drosophila* Dscam homologues produce similar phenotypes

In *Drosophila* there are 4 different Dscam homologues (Dscam1-4). To see whether different homologues function in a similar manner to Dscam1 in this system, Dscam2 and 4 were separately knocked down in *lgl* homozygous mutant tissue within the notum.

Like Dscam1, both Dscam2 and Dscam4 enhanced the tumour phenotype when knocked-down in the *lgl-/-* mutant background, generating clones with similar phenotypes. For example, knocking down these Dscam homologues resulted in multi-layering of the mutant epithelium **(Fig. 6.6 A)** and reduced clonal tissue coverage within the notum (when compared to *lgl-/-* alone, p=0.0396) **(Fig. 6.6 B, H).** Compared to Dscam1 KD, expressing RNAi targeting Dscam2 and Dscam4 had significantly larger apices (Dscam2>Dscam1 p=0.0025, Dscam4>Dscam1 p=0.0027), however both still caused significant (p<0.0001) apical constriction compared to *lgl-/-* alone. **(Fig. 6.6 C, G)**. Both Dscam2 and Dscam4, like Dscam1, seem to play a role in regulating cellular membrane extensions, as knockdown of either homologue generated significantly long basal filopodia compared to *lgl-/-* (*lgl-/-*; Dscam2 KD>*lgl-/-* p<0.0001, *lgl-/-*;Dscam4 KD>*lgl-/-* p=0.0002) **(Fig. 6.6 D, H)**. However, these basal protrusions were significantly shorter than those generated when knocking down Dscam1 (*lgl-/-*; Dscam2 KD>*lgl-/-*; Dscam1 KD p=0.0039, *lgl-/-*; Dscam4 KD>*lgl-/-*; Dscam1 KD p=0.0097). Additionally, one animal with the genotype *lgl-/-*; Dscam2 had filopodia with sheet like protrusions extending from it **(Fig. 6.6 E)**.

These data indicate that all *Drosophila* homologues of Dscam cooperate with *lgl* in this system and when knocked down promote tumour progression, notably causing apical constriction, multi-layering and elongated basal protrusions compared to *lgl-/-* alone. However, compared to Dscam1 KD, the effects seen knocking down Dscam2 and Dscam4 were not as severe. Additionally, like Dscam1, Dscam2 and Dscam4 KD without the *lgl* mutation was not sufficient to promote tumour progression.
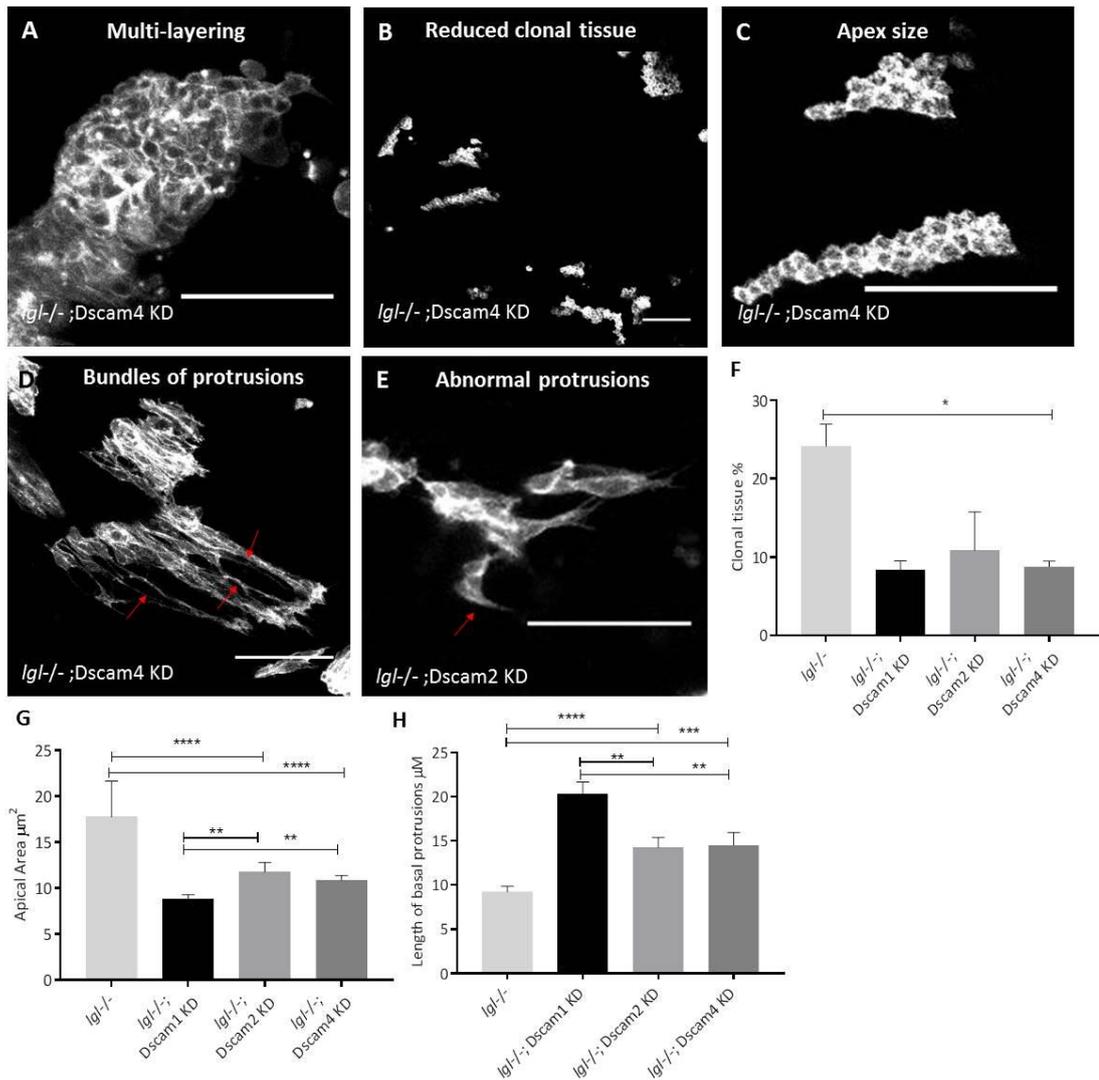
Figure 6.6 Dscam homologues produce similar phenotypes

Live confocal imaging of *lgl* homozygous mutant clones labelled with GFP in the notum, simultaneously expressing RNAi targeting either **(A-D)** Dscam4 or **(E)** Dscam2. A range of phenotypes were observed including, **(A)** multi-layering, **(B)** reduced clone size, **(C)** small apices, **(D)** bundles of basal protrusions and **(E)** abnormal protrusions. Scale bars = 50μm. **(F-H)** Quantification of **(F)** clonal tissue coverage, **(G)** size of epithelial apex, and **(H)** length of basal protrusions (n≥5 animals, with >40 cell apices/protrusions measured for each genotype).

6.2.3 Subcellular localisation of Dscam

Dscam has previously only been implicated and studied in the nervous system, where it has been shown to localise to the cell membrane, soma, and dendrites (Soba et al., 2007). To determine the subcellular localisation of Dscam1 in epithelial cells WT tissue was fixed and stained with an antibody labelling Dscam (N-terminal). To do this Neuralised-Gal4, which drives expression of UAS-constructs (including GFP) in sensory organ precursor cells and their progeny, was used to specifically label WT individual well-spaced cells on the back of the fly and the whole notum was fixed and stained for both E-cadherin and Dscam.

In WT tissue Dscam appeared to localise cytoplasmically at the level of the AJs, however did not colocalise with E-cadherin at the junctions **(Fig. 6.7 A-A'')**. Moving basally through the cell, Dscam localises to the cortex **(Fig. 6.7 B-B'')** and then finally to the basal filopodia **(Fig. 6.7 C-C'')**. This differential localisation of Dscam throughout the columnar epithelial cell may explain why a variety of different phenotypes were seen when knocking-down Dscam in the notum- from apical defects to elongated basal filopodia.
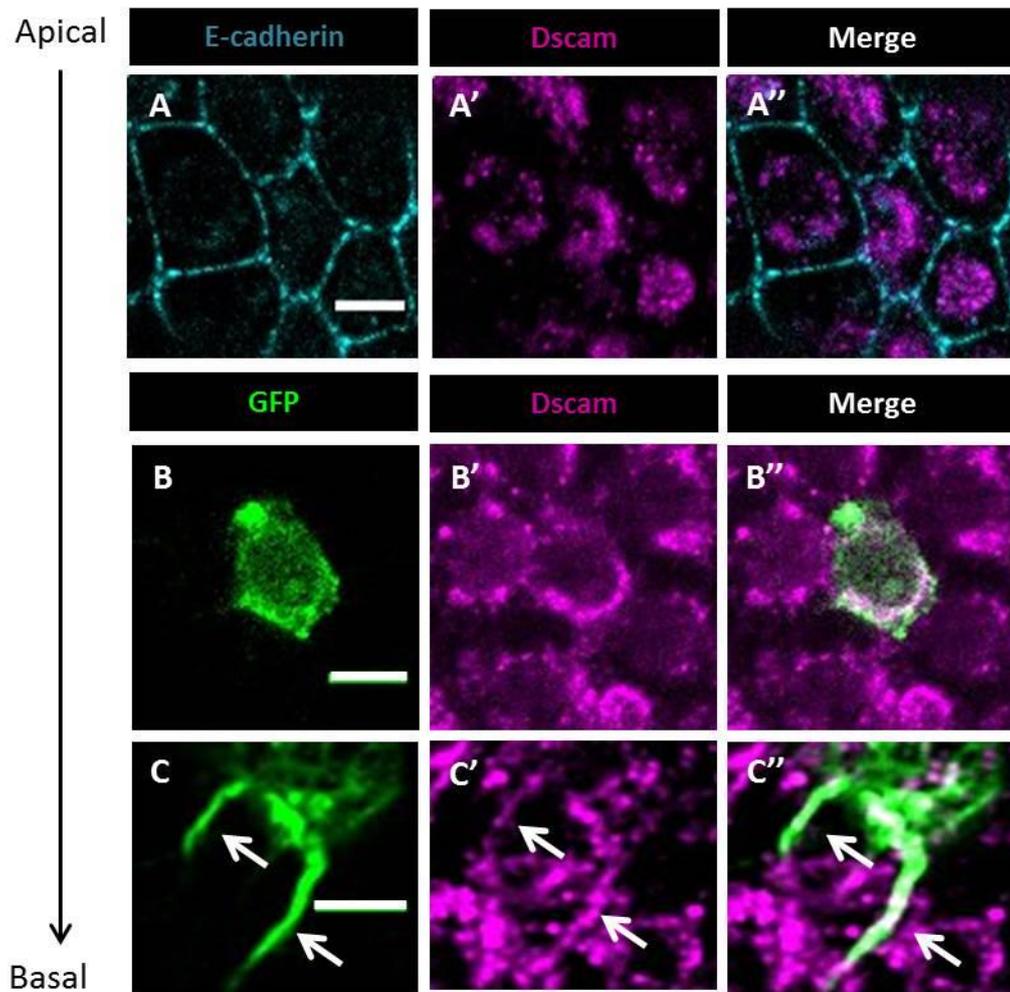
Figure 6.7 Subcellular localisation of Dscam in epithelial cells

WT *Drosophila* nota (Neu-Gal4, UAS-Moe-GFP) were fixed and stained for GFP, E-cadherin and N-terminal Dscam. **(A-A'')** Confocal snapshots were taken at the level of the AJs, **(B-B'')** at an intermediate level and **(C-C'')** at the level of the basal protrusions. Arrows indicate basal filopodia where Dscam is highly expressed. Scale bars represent 5µm.

6.2.4 Dscam is highly expressed in bristle precursor cells in the notum

As there are limitations using RNAi, for further analysis of the Dscam phenotype, two Dscam mutants- *Dscam^{21}* and *Dscam^{23}*- were analysed, both of which have previously been described as strong loss-of-function mutant alleles when studied in olfactory receptor neurons (Hummel et al., 2003; Schmucker et al., 2000). Initially negatively marked clones expressing *Dscam^{21}* and *Dscam^{23}* mutant alleles were generated in the notum and stained for E-cadherin and Dscam.

Surprisingly, Dscam staining within the mutant clone appeared largely unaltered **(Fig. 6.8 A-A''')**. However this could be due to the antibody recognising the mutant protein - if the protein is truncated. While the mutant alleles have been described as loss-of-function, exactly what these mutants are is not described in the published literature. It has been shown that the antibody binds to the constant region of Dscam (binding to variable exons 4, 6 and 9 were ruled out using western blotting). Additionally, previous published work with these mutants and the same antibody showed no staining or detection on a western blot (Hummel et al., 2003; Schmucker et al., 2000). Unexpectedly, in WT tissue Dscam expression was notably upregulated in the machrochaetes **(Fig. 6.8 B-B'')**, whilst in Dscam mutant machrochaetes this expression was drastically reduced **(Fig. 6.8 C-C'')**.

Sensory organ precursor (SOP) cells are evenly spaced within the *Drosophila* dorsal thorax and from 12APF they divide and give rise to the microchaetes (small bristles) on the back of the fly. This process requires various signalling pathways including EGFR and Notch signalling, which when disrupted often lead to irregular bristle

growth and spacing. However when both Dscam mutants were analysed they appeared to have no effect on bristle development.

As it is still unclear why Dscam was seen to be upregulated in the bristle precursor cells and then lost in the Dscam mutants, a much more detailed analysis would be required, including immunostaining using a different Dscam antibody to ensure results were not due to non-specific labelling.
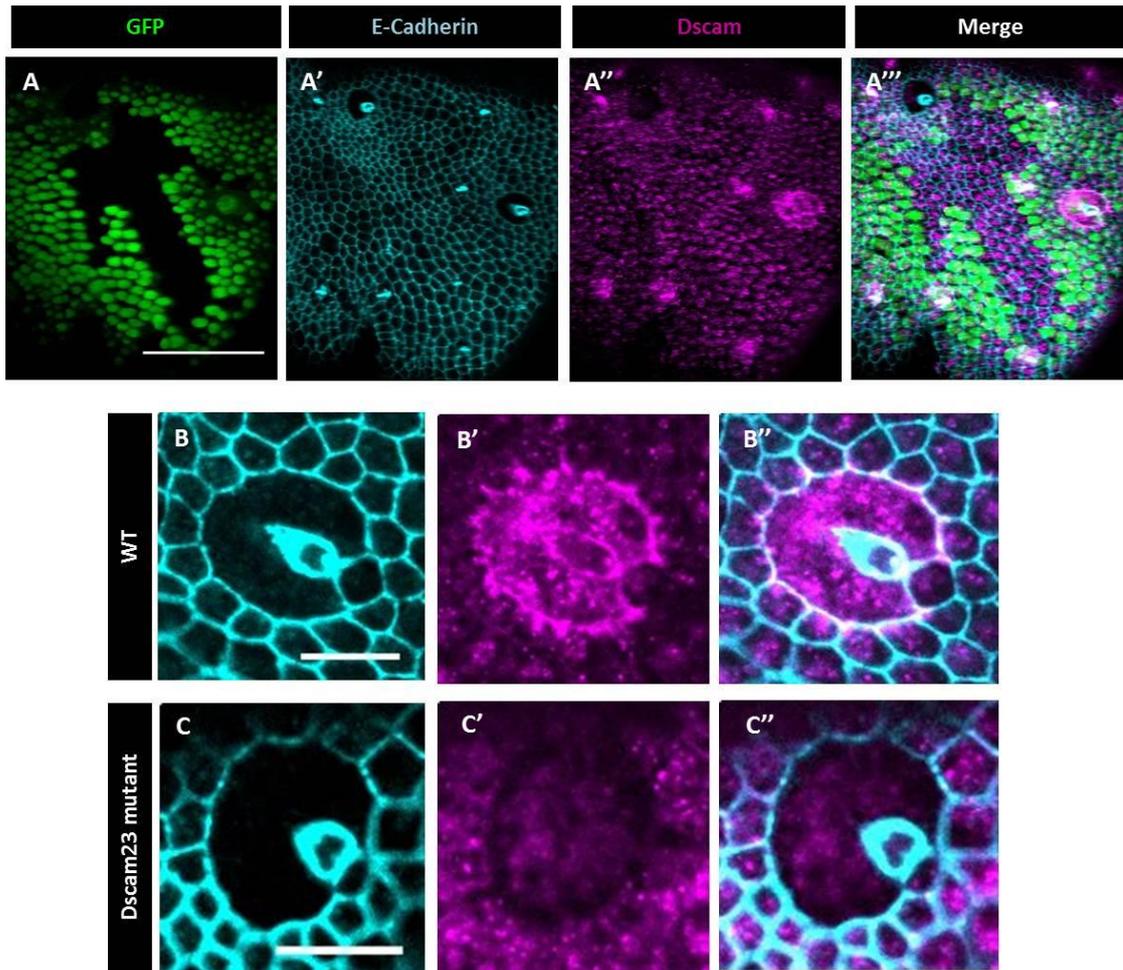
Figure 6.8 Dscam localisation appears normal in *Dscam^23* mutant epithelial cells, while expression is reduced in mutant macrochaetes

(A-A''') Negatively marked clones were generated in the notum homozygous mutant for *Dscam^23*, with surrounding WT tissue labelled with nuclear GFP (nls-GFP). Scale bar = 50μm. **(B-C)** Nota were fixed and stained at 15h APF for GFP, E-cadherin and Dscam. Close up of WT **(B-B'')** and Dscam[23] mutant **(C-C'')** macrochaetes. Scale bars = 10μm.

6.2.5 Knocking-down Lgl in $Dscam^{21}$ mutant background promotes tumour progression.

In **Section 6.2.1** Dscam was shown to cooperate with Lgl in this system when studying the genotype *lgl-/-*;Dscam1-RNAi. For comparison, $Dscam^{21}$ homozygous mutant clones were generated with or without the simultaneous knockdown of Lgl using RNAi.

In $Dscam^{21}$ mutant clones, the phenotype produced was similar to that seen when Dscam1 was knocked down using RNAi, with clonal tissue forming an organised monolayer **(Fig. 6.9 A).** Additionally, $Dscam^{21}$ had no significant effect on the percentage of clonal tissue within the notum, however when combined with Lgl-RNAi, clonal tissue was significantly reduced (p=0.0215) **(Fig. 6.9 G).** However, this was not to the same extent as that seen with *lgl-/-*;Dscam1 KD, whereby clonal tissue was significantly (p<0.0001) smaller.

As seen previously, knocking down Dscam1 in *lgl-/-* mutant tissue resulted in apical constriction. This phenotype was not mimicked when Dscam1 was knocked down in WT tissue alone **(Fig. 6.3)**. Strangely, $Dscam^{21}$ mutant apices were significantly larger than WT (p<0.0001) **(Fig. 6.9 B, H).** When combined with Lgl KD, $Dscam^{21}$ mutant apices were significantly reduced in size (p<0.0001), however the apex size in $Dscam^{21}$; Lgl KD was still not significantly different from WT **(Fig. 6.9 E, H)**.

Also, $Dscam^{21}$ mutant cells had long actin rich basal filopodia **(Fig. 6.9 C, I)** which were significantly (p=0.0018) longer than WT cells. Moreover, when combined with

218

Lgl KD the length of basal protrusions was significantly increased (p=0.0009) **(Fig. 6.9 F, I)**. However, once again this phenotype was not as severe as that seen with *lgl-/-*; Dscam1 KD, which presented with significantly (p<0.0001) longer basal protrusions **(Fig. 6.9 I)**. Additionally, closure defects were often seen in *Dscam^{21}*; Lgl KD animals **(Fig. 6.9 D)**.

In conclusion, knocking down Lgl in the *Dscam^{21}* mutant background enhances the tumour phenotype compared to *Dscam^{21}* alone, however is rarely significant compared to WT. Moreover, the phenotype for *lgl-/-*; Dscam1 KD is more severe than that seen with *Dscam^{21}*; Lgl KD.

A Clonal Tissue   **B** Normal Apices   **C** Long Basal Protrusions

Dscam21

D Closure Defect   **E** Small Apices   **F** Long Basal Protrusions

*Dscam21;*Lgl KD

G

Clonal tissue %

WT   Dscam1 KD   *Dscam21*   Dscam21; Lgl KD   *lgl-/-;*Dscam1 KD

*

****

H

Apical Area µm^2

WT   Dscam KD   *Dscam21*   Dscam21; Lgl KD   *lgl-/-;*Dscam1 KD

**** **** ****

I

Length of basal protrusions µM

WT   Dscam1 KD   *Dscam21*   Dscam21; Lgl KD   *lgl-/-;* Dscam1 KD

** *** ****

220
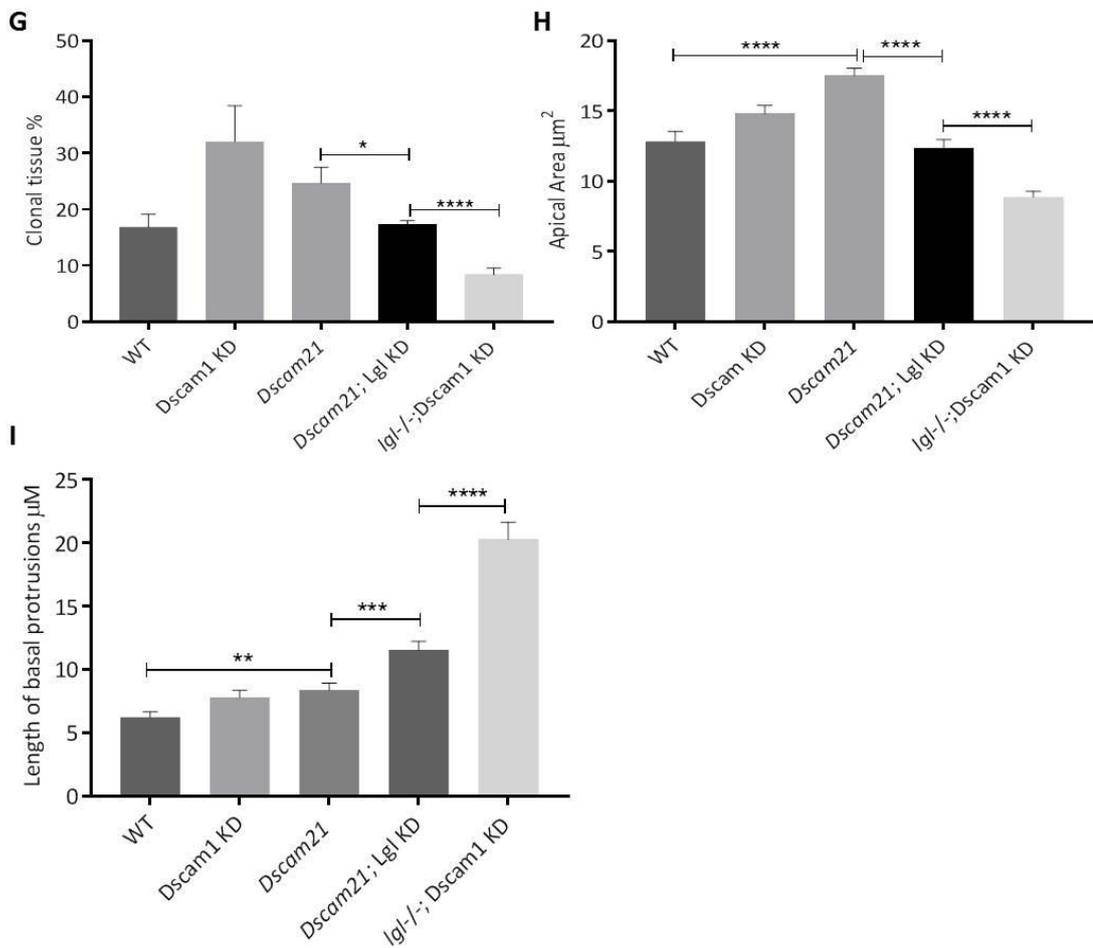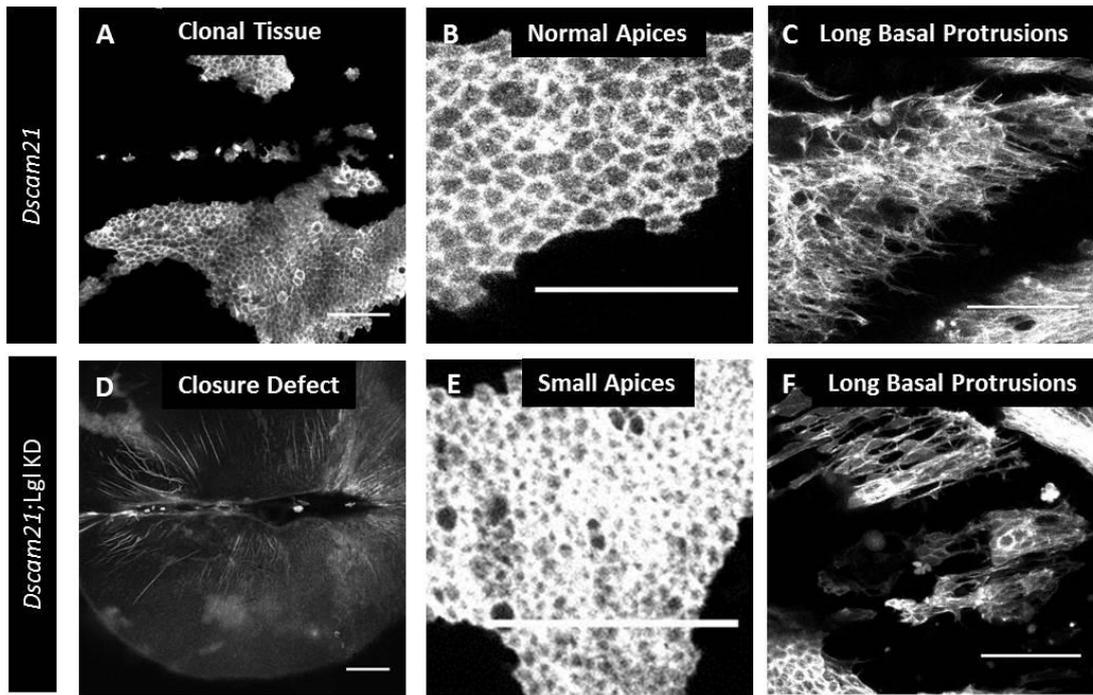
Figure 6.9 Cooperative effect seen between *Dscam*[21] and Lgl-RNAi

Dscam[21] homozygous mutant clones were generated without **(A-C)** or with **(D-F)** the simultaneous knockdown of *lgl* using RNAi. Scale bars = 50μm. Quantification of **(G)** clonal tissue coverage, **(H)** apex size, and **(I)** length of basal protrusions. Error bars represent SEM. (n=5 animals and ≥40 apices and filopodia measured for each genotype).

6.2.6 Dscam affects cell adhesion and junction integrity

The main phenotypes observed when combining *lgl* and *dscam* mutants were multi-layering, apical constriction and elongated filopodia. These phenotypes are all associated with localised changes in cell shape. Therefore the AJ protein E-cadherin was studied to see whether Dscam plays a role in regulating and maintaining cell adhesion and junction integrity, which are typically disrupted during tumour progression and lead to a loss of polarised cell shape. For this experiment, *Dscam*[21]; WT and *Dscam*[21]; Lgl KD clones were generated in the notum and fixed and stained for E-cadherin. Additionally a live-cadherin reporter (ubi promoter drives expression of E-cadherin tagged GFP) was used to study E-cadherin in real time in *Dscam*[21] mutants.

Cells within *Dscam*[21] mutant clones showed a significant (p<0.0001) reduction in E-cadherin expression at the AJs. Moreover, E-cadherin at junctions at the edge of a clone (i.e. a junction shared between mutant and wild-type cells) was also significantly reduced (P<0.0001) compared to WT **(Fig. 6.10 A-B).** Bright ectopic puncta were observed in *Dscam*[21] mutant tissue, however, strangely these were also seen in WT tissue with no significant difference in the average number of puncta per cell **(Fig. 6.10 B')**. Similarly, ectopic E-cadherin junction extensions and junctional breaks were observed, however, again this was not significant compared to WT **(Fig. 6.10 B''-B''')**. When using live-cadherin to label AJs, *Dscam*[21] mutant junctions showed a significant (p=0.0002) decrease in E-cadherin levels compared to WT. However, unlike that seen with the E-cadherin antibody, border junctions had normal E-cadherin levels **(Fig. 6.10 C-D)**. Additionally, atypical ectopic E-cadherin rich

junction extensions were seen in mutant tissue which was significant compared to WT (p=0.0167) **(Fig. 6.10 D')**. Using the live-cadherin, no junctional breaks or ectopic puncta were seen.

When *Dscam*[21] was combined with Lgl KD these phenotypes were significantly enhanced. Compared to the surrounding WT tissue *Dscam*[21]; Lgl KD mutant junctions had a significant reduction (p<0.0001) in E-cadherin levels **(Fig. 6.10 E, F)**. E-cadherin in junctions bordering WT and mutant tissue was also significantly reduced (P<0.0001). Indicating that loss of E-cadherin from mutant side of the clone boundary was sufficient to decrease E-cadherin on the WT side. Additionally, a significant number of junctional breaks (p=0.0452) and ectopic E-cadherin junction extensions (p=0.0139) were seen in mutant tissue compared to WT **(Fig. 6.10 E'-E'', F-F'').**

The data shows that a loss of Dscam function leads to a mild disruption to E-cadherin localisation and junction integrity in epithelial cells. *Dscam*[21] alone was sufficient to reduce E-cadherin in the mutant tissue, however this affect was enhanced when combined with Lgl KD, again emphasising the cooperative effect between Lgl and Dscam in *Drosophila* epithelial cells.
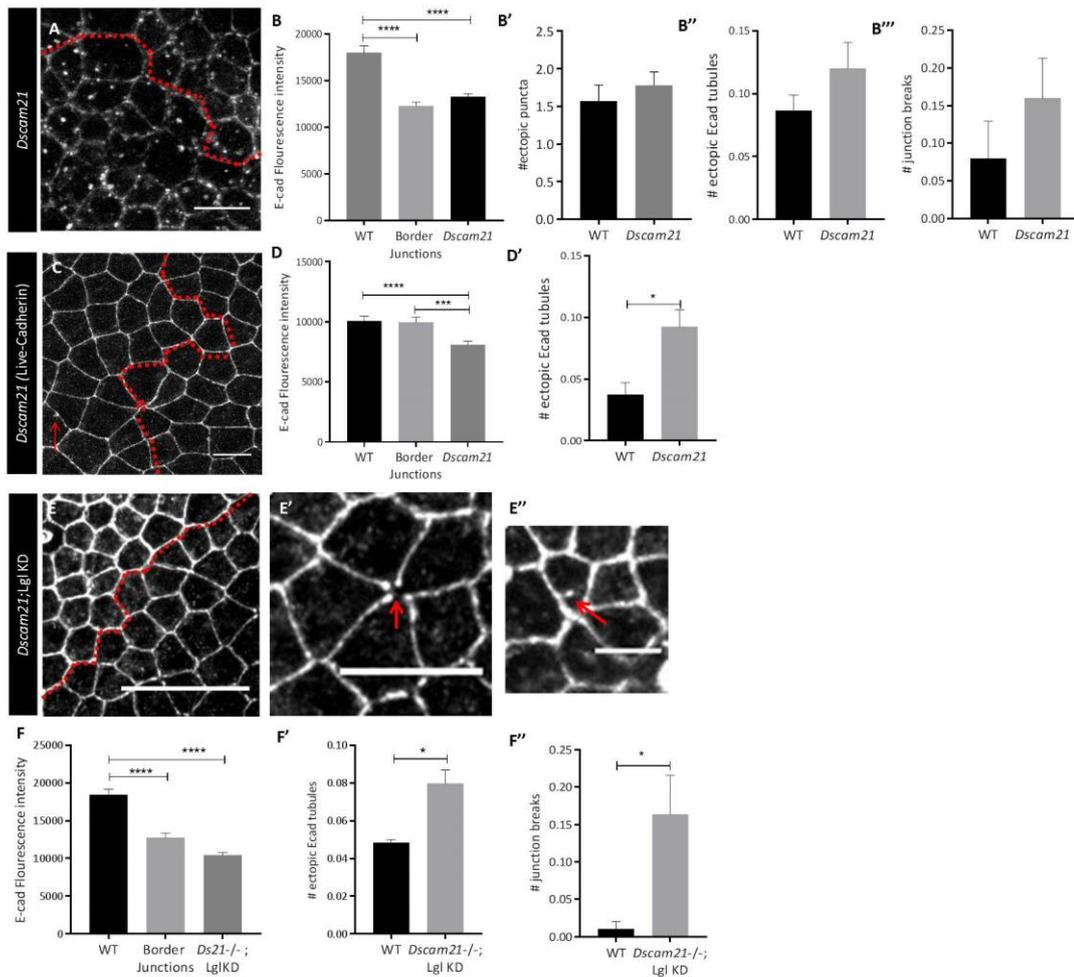
Figure 6.10 E-cadherin localisation is mildly disrupted in Dscam mutants

Visualising E-cadherin in Dscam mutant clones in the *Drosophila* nota. Tissue was either fixed and stained with antibodies against E-cadherin **(A, E-E')** or live animals were imaged expressing a live-cadherin construct in the notum **(B)**. **(A-C)** *Dscam^{21}* mutant clones (A, below red dashed line; B, to the left of the red dashed line) show mildly reduced E-Cadherin levels in the mutant cells. Arrow indicates ectopic E-cadherin tubules. **(E)** *Dscam^{21}*; Lgl KD clone (mutant cells are to the right of the red dashed line). **(E'-E'')** Higher magnification image of a *Dscam^{21}*; *lgl* KD clone. Arrow indicates a junctional break **(E')** and ectopic E-cadherin extensions **(E'')**. Scale bars = 10μm in panels A, C, E'; 20μm in panel E and 5 μm in panel E''. **(B,D,F)** Quantification of fluorescence intensity of E-cadherin in WT, border and mutant AJs. **(B')** Quantification of the average number of ectopic puncta per cell. **(B'',D',F')** Quantification of the average number of ectopic E-cadherin tubules. **(B''',D', F'')**. Quantification of average number of junction breaks. (n=4 for live-cadherin; n=3 for fixed staining; with ≥50 junctions measured for fluorescence intensity and ≥50 cells analysed for junction breaks, puncta and ectopic tubules).

6.2.7 Cooperation between *Dscam* and Lgl affect apicobasal polarity

The polarity proteins Bazooka (Baz) and aPKC localise to and regulate assembly of AJs in *Drosophila* epithelial cells and are vital in regulating and maintaining apicobasal polarity. To further analyse whether Dscam is important for apicobasal polarity, RNAi was used to knock down Lgl in *Dscam^{21}* mutant clones, and nota were fixed and stained with antibodies against Baz and aPKC.

In *Dscam^{21}*; *lgl* KD mutant animals Baz staining was significantly (p<0.0001) reduced at AJs compared to WT **(Fig. 6.11 A', C)**. Additionally, within the mutant tissue junctional breaks and atypical ectopic Baz rich tubules were often seen. Moreover, when looking at E-cadherin staining from the same animal, these junction breaks were often seen in the same place **(Fig. 6.11 A-A'')**.

Similarly, aPKC staining was also significantly (p<0.0001) reduced in *Dscam^{21}*; Lgl KD mutant tissue compared to WT **(Fig. 6.11 B, D)**.

These results strengthen the previous data with E-cadherin, suggesting that the combined loss of Lgl and Dscam function leads to a mild disruption to junction integrity and apicobasal polarity.
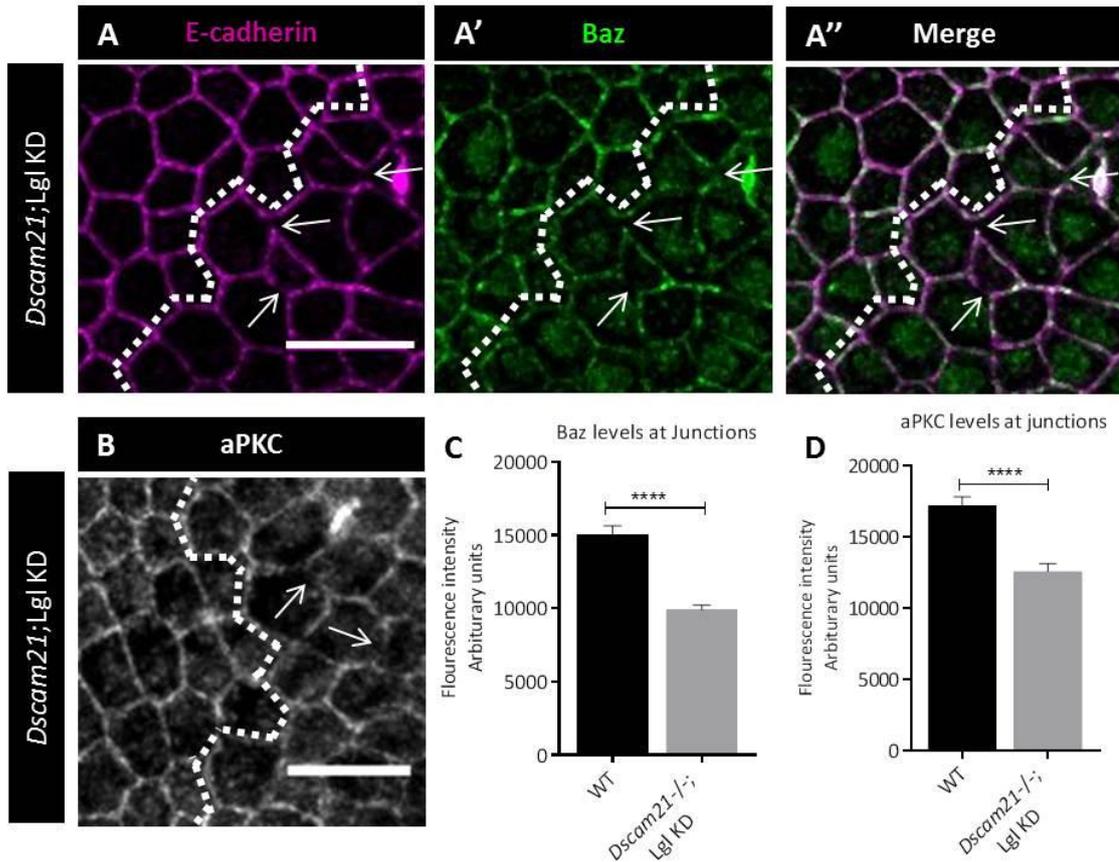
Figure 6.11 Reduced Baz and aPKC levels in *Dscam²¹*; Lgl KD mutant clones

(A-B) Positively marked clones from *Drosophila* nota were fixed and stained for **(A)** E-cadherin, **(A')** Baz and **(B)** aPKC in WT (left side of dotted line) and *Dscam²¹*; Lgl KD mutant tissue (right side of dotted line). **(A-B)** Disruption to junction integrity in mutant tissue highlighted by arrows. Scale bars represent 10μm. **(C-D)** Quantification of Baz and aPKC fluorescence intensity. (n ≥40 from 4 animals for aPKC and 2 animals for Baz)**.** Mean and SEM plotted.

6.3 Discussion

Due to its drastic effect on tumour progression the complex molecule Dscam was identified as a hit from a genetic screen carried out in **Chapter 4**. Dscam has been widely researched with regards to its well-characterised roles in the nervous system including self-avoidance and axon guidance (Schmucker et al., 2000). However, with regards to cancer development and progression Dscam remains extremely novel and uncharacterised. This chapter suggests that Dscam may also act in a tumour suppressive manner in the *Drosophila* epithelium.

In the *Drosophila* notum, Dscam1 was shown to cooperate with Lgl in a tumour suppressive manner, affecting several cancer-associated processes. Additionally, Dscam2 and Dscam4 produced similar phenotypes. For example, down regulating Dscam1, Dscam2 or Dscam4 in *lgl* homozygous mutant tissue resulted in aberrant tissue multi-layering, developmental defects, elongated dynamic basal filopodia and apical constriction. These phenotypes were however more severe in *lgl-/-; Dscam1* KD clones. This suggests that a combined loss of Dscam1 and lgl function could lead to a disruption of cell-cell adhesion and/or the apical actomyosin cytoskeleton, which regulates cell shape and contractility. To further analyse the role of Dscam in cell-cell adhesion, *Dscam* mutant and Dscam KD tissue could be fixed and stained for adhesion molecules such as α- and β-catenin along with integrins such as myospheroid (β-integrin). Moreover, to look in more detail at the effects Dscam has on the actin cytoskeleton, a Life-actin fused protein could be expressed, allowing the actin cytoskeleton to be studied in the living epithelium of *Dscam* mutants.

During developmental processes such as gastrulation Rho family members and actomyosin contractility has been shown to regulate apical constriction (Barrett, Leptin, & Settleman, 1997; Dawes-Hoang et al., 2005). Therefore it could be hypothesised that the apical constriction seen in Dscam KD clones could be due to Dscam regulating actin filaments and signalling through the Rho-kinase pathway. In support of this claim, in 2000, Schmucker et al, identified Dscam as a regulator of axon guidance in the nervous system through its interactions with Dock and Pak which have been shown previously to regulate the actin cytoskeleton through interactions with Rho family GTPases; RhoA, Rac and cdc42 (Hing, Xiao, Harden, Lim, & Zipursky, 1999; Nobes & Hall, 1995).

Similarly, the significant elongation of basal filopodia seen suggests that Dscam cooperates with Lgl to directly or indirectly regulate the actin cytoskeleton. Filopodia are extensions of the plasma membrane encompassing bundles of actin filaments which serve many functions in epithelial cells, including adhesion and probing environmental signals (Vignjevic, Peloquin, & Borisy, 2006). These actin-rich protrusions are regulated directly by small Rho-GTPases and polarity protein complexes as well as indirectly via signalling pathways including Wnt/β-catenin which in turn regulate the filopodial binding protein fascin. In cancer, filopodia are linked to invasive tumours whereby changes in the actin cytoskeleton, typically driven by localised changes in Rho family GTPases, produces actin-rich protrusions which drive front-rear polarity and increase cell motility (Etienne-Manneville & Hall, 2002; Nobes & Hall, 1995; Vignjevic et al., 2007). Dscam could therefore regulate basal protrusions by directly affecting the actin cytoskeleton or indirectly by regulating Rho-GTPases.

Additionally, in the nervous system Dscam is important in the regulation of self-avoidance and in recent loss-of-function studies in the mouse retina resulted in elongated dendritic processes, crossing over and generation of dendrite bundles (Li et al., 2015).

Within the last 5 years, two genetic techniques; Flybow (Hadjieconomou et al., 2011) and LOLLIbow (Boulina, Samarajeewa, Baker, Kim, & Chiba, 2013); have been developed which enable multi-colour labelling of sub-populations of cells in *Drosophila*. In this way both Flybow and LOLLIbow could be used to visualise protrusions in *Dscam* mutants and Dscam KD animals in real time. This would enable us to look more closely at protrusion morphology and dynamics in Dscam mutants and how protrusions from neighbouring cells interact with one another.

Immunofluorescent analysis of both *Dscam*[21] and *Dscam*[21]; Lgl KD mutant tissue showed decreased levels of E-cadherin, Baz and aPKC at AJs. In *Dscam*[21] mutant animals a decrease in E-cadherin levels was seen along the mutant/WT boundary, however this was not reproduced when using the live-cadherin. As the live-cadherin construct drives GFP-tagged E-cadherin expression, this could result in excess E-cadherin levels in the cell, which may account for the difference seen at the mutant/WT boundary. The *Dscam*[21]; Lgl KD genotype was also sufficient to reduce E-cadherin levels along the mutant/WT clone boundary. Additionally, ectopic E-cadherin rich puncta and junction extensions, along with junction breaks were also observed. Ectopic E-cadherin puncta were only observed in animals with *Dscam*[21] mutant clones and labelled with an antibody against E-cadherin. Moreover these

puncta were also observed in WT tissue, which suggests that the observed puncta are simply an artefact of the antibody staining. As it is unlikely that the effects on junction integrity are a major contributor to the overall cancer phenotype, future experiments would involve looking at other possible explanations. For example, in the nervous system Dscam has been shown to regulate JNK signalling through its interactions with Pak and Dock, therefore it may be enhancing the cancer phenotype in our system through JNK signalling (Qu et al., 2013; Schmucker et al., 2000). Additionally FRET-probes could be used to analyse changes in Rho GTPase activity in Dscam mutant animals.

Altogether, these data suggest that in the epithelium a loss of Dscam1 function could affect cell adhesion, junction integrity and apicobasal polarity, which could contribute to the phenotypes observed when Dscam1 and lgl function is lost (e.g. multi-layering and apical constriction) as these phenotypes are associated with the deregulation of these processes. Additionally, elongated dynamic filopodia are associated with a front-rear polarity which requires a loss of apicobasal polarity within columnar epithelial cells. All of this together suggests that Dscam may be important in the regulation of cell polarity and the localisation of polarity determinants within the cell. However, it should be stressed that the effects on junction integrity and cell polarity are relatively mild and are unlikely to be the underlying cause of the extreme phenotypes observed in *lgl-/-;* Dscam1 KD clones.

Although a much more detailed analysis is required, this chapter outlines some preliminary experiments which show that loss of Dscam promotes tumour

progression in the epithelium. These results suggest that Dscam function is not limited to the nervous system, however has a novel cooperative role in preventing tumour progression in the epithelium.

CHAPTER 7

Discussion

Chapter 7 Discussion

Tumour invasion and metastasis is a complex process that is poorly understood yet accounts for more than 90% of cancer related deaths. By improving our knowledge of the complex processes and genetic determinants involved in tumour progression, targeted cancer therapy could be advanced and improved. *Drosophila* has become an increasingly popular model organism with which to characterise the cell biological events that underlie tumour progression. These events are usually associated with aberrant changes in cell polarity and adhesion through the concomitant misexpression of various genes. During my PhD I carried out a pilot screen to optimise conditions for a large-scale genetic screen in the *Drosophila* notum. From the full genetic screen I then identified novel genes which affect tumour progression.

7.1 Optimisation and screening

The aim of the pilot screen was to (i) optimise conditions for the full screen by identifying a suitable genetic background for the enhancer/suppressor screen. (ii) Carry out cooperative experiments with Notch. (iii) Verify RNAi KD was working as expected and independent RNAi lines produce similar phenotypes and (iv) to generate a database and scoring system for qualitative analysis of animals.

My preliminary studies showed that by inducing mutant clones which were homozygous mutant for the *lgl* tumour suppressor gene, neoplastic tumours could be generated on the back of the fly. These tumours covered approximately 15-35% of the notum and were mildly multi-layered and invasive. In contrast, clones of cells

homozygous mutant for *dlg* and *scrib* were invariably small. A possible explanation for this is that *dlg* and *scrib* mutant cells were being eliminated through JNK-mediated apoptosis as a result of cell competition (as described in **Section 3.3)**. This theory has been demonstrated by other researchers who show that mutations in these tumour suppressor genes require cooperation with oncogenic Ras or Notch in order to preferentially grow and not be out competed by surrounding wild-type cells (Brumby & Richardson, 2003; C. L. Chen et al., 2012; Pagliarini & Xu, 2003). Unlike these other systems, my data showed that clones homozygous mutant for *lgl* were sufficient to grow and survive. One possible explanation for this is that unlike Dlg and Scrib*,* Lgl interacts with the Salvador/warts/hippo pathway such that loss of Lgl results in the mislocalisation of hippo and subsequent activation of the transcription factor Yki. Yki in turn transcriptionally activates genes involved in cell growth and survival (Grzeschik et al., 2010). In other *Drosophila* tissues such as the imaginal discs, *lgl* mutant clones are small and outcompeted by WT cells like those seen with *dlg* and *scrib* mutant clones in our system. To my knowledge, this is the first demonstration of a tissue in the fly where large *lgl* mutant clones can be generated, implying that these *lgl* clones have not been outcompeted by wild type tissue. One possible explanation for this is that in other studied tissues, such as the imaginal discs, mutant clones find themselves in a highly proliferative environment; however in the notum the tissue is not so proliferative. From these results, it was determined that *lgl* mutant clones would make an ideal genetic background with which to carry out an enhancer/suppressor screen. In addition, I showed that when combined with an activated allele of Notch, the tumour phenotype was enhanced, mimicking the cooperative effect seen in other systems (Brumby & Richardson, 2003).

By carrying out a pilot screen on more than 40 well-characterised cancer-related genes we demonstrated that it was possible to replicate expected phenotypes with known genes in our system and RNAi lines for the same gene were giving similar phenotypes. Analysis of all the genes screened during the pilot led the way to the generation of a database for large data-sets. Using this database all genes, animals and phenotypes could be qualitatively analysed. Moreover, this way of recording data allowed me to quickly identify genes which appeared to drastically enhance or suppress the tumour phenotype. By the end of the pilot screen, our data provided evidence that our system was working efficiently, and could be used for the high-throughput screening of a host of candidate genes.

The main aim of the full candidate screen was to identify novel genes that promote or inhibit the tumour phenotype. These genes were selected using various online tools and databases, identifying genes that were previously found to be commonly mutated or down regulated in a range of human epithelial cancers (Lawrence et al., 2014). Identified genes were then checked for their suitability for the screen, i.e. that they had *Drosophila* orthologues and available RNAi constructs on the correct chromosomes. For some of these genes the fly gene and/or the human orthologue were well characterised, while others were completely unknown. These uncharacterised genes had only been implicated in human cancer from genetic/epigenetic analyses of human tumours. This range of well-characterised and unknown genes was selected to help make sense of gene clustering based on phenotype. In 3 years, over 418 genes (610 RNAi lines) were screened and

qualitatively analysed using the database generated from the pilot screen. Using our method of subjective analysis, genes that generated strong phenotypes when knocked-down could be easily identified. For example, approximately 200 genes increased tumour multi-layering and 80 genes increased tumour invasiveness. From this screen I selected 3 genes with interesting phenotypes to characterise further, these were: CG7379, CG10600 and Dscam.

7.2 Characterisation of identified hits

We are now in the process of completing this large-scale screen and have identified numerous novel conserved genes that affect tumour progression in this *Drosophila* system. To take this further, (i) interesting hits will be characterised in the fly, with the aim of understanding the molecular basis of their effect on tumour progression, and (ii) the role of homologous genes will be tested in mammalian cell culture models of metastasis and, potentially, in other in vivo cancer models. In this way, we may identify promising targets for the development of anti-cancer therapies. The three genes I chose to characterise, in addition to providing promising results, also illustrate many of the techniques that will be used to characterise these genes in the future.

7.2.1 CG7379 has anti-invasive properties and regulates junction integrity and cell-adhesion

I selected the *Drosophila* gene CG7379 to characterise further as KD in the *lgl* mutant background significantly enhanced the invasive phenotype. While CG7379 remained

uncharacterised in *Drosophila,* its human orthologue ING1 is a well-characterised gene implicated in human cancer; however its known role is in the regulation of apoptosis (Garkavtsev et al., 1998; Helbing et al., 1997; Tsang et al., 2003). By overexpressing p35 my data shows that the invasive phenotype observed in *CG7379* mutant clones is not due to reduced apoptosis – rather my results suggest that CG7379 function is important to maintain cell-cell junction integrity. To verify this, further experiments would involve testing the level of apoptosis is WT, *lgl -/-* alone, *CG7379-/-* alone and *lgl-/-*; CG7379 KD using TUNEL staining, or live constructs such as apoliner. Analysis of E-cadherin and Arm (β-catenin) in CG7379 KD clones with and without an accompanying *lgl* mutation showed that junction integrity was compromised, with reduced E-cadherin and Arm at AJs, junctional breaks and ectopic E-cadherin rich junction extensions. To maintain cell-cell adhesion, E-cadherin and Arm form a cadherin-catenin complex that is at the core of the adherens junction. These essential components of the AJ are often misregulated in human cancers to allow tumour cells to break cell contacts and invade. Therefore by affecting AJ protein localisation and regulation in cells, this may explain why CG7379 KD promotes tumour invasion. Researchers have shown that deregulation of endocytosis mediated junctional turnover of E-cadherin results in E-cadherin ectopic structures, similar to those seen in my results (Georgiou et al., 2008). Therefore I propose that CG7379 maintains junction integrity and cell adhesion by regulating the turnover of junction material required during important processes such as cell division. In support of my hypothesis, the human orthologue ING1 has been shown to regulate senescence by mediating expression of intersectin-2 which is a member of endocytic machinery (Rajarajacholan et al., 2013). My data suggests that by regulating

expression of genes involved in endocytosis, CG7379 might be acting in a tumour suppressive manner to maintain junction integrity. My work also demonstrated that CG7379 KD impaired integrity of the septate junctions as well as the adherens junctions, as FasIII localisation was also disrupted. On the other hand my data showed that CG7379 KD did not affect the localisation/expression of key players in apicobasal polarity. Future experiments would involve looking at cell-cell junctions in mammalian cells in vitro, including non-transformed epithelial cell lines such as MDCK cells. For this ING1 would be knocked down in MDCK cells to test whether the same effect is observed on junctions in mammalian cells. If these junction defects are mimicked in mammalian cells, this would reveal a new aspect of ING1 function that has not been demonstrated before and is relevant to its role in cancer.

By carrying out in vitro experiments on ING1, I showed that the human orthologue also has anti-invasive properties in different mammalian tissues and KD was sufficient to promote invasion in both glioblastoma and adenocarcinoma cells. This suggests that this gene doesn't act in a tissue specific manner and is likely to be involved in pathways common to both tissue types.

In summary my data characterising CG7379 strongly suggests that this gene has important anti-invasive properties by regulating junction integrity and cell adhesion. Previous research on ING1 has suggested that this gene may function in a tumour suppressive manner by regulating apoptotic machinery, p53 and the cell cycle. However, I propose that this gene might also be involved in endocytosis mediated recycling of E-cadherin and junctional material in order to regulate dynamic junction

stability and adhesion. In order for research to progress I suggest that experiments should be carried out which test whether CG7379 is directly involved in the expression/regulation of endocytosis-mediated E-cadherin junctional turnover. Additionally, I would investigate further the effects CG7379 KD has on junction dynamics using FRAP, and study junction integrity and adhesion in vitro. Once it is fully understood how CG7379/ING1 KD promotes invasion, this gene could potentially become a therapeutic target.

7.2.2 CG10600 regulates cell adhesion and junction integrity

The second gene that I chose to characterise further was CG10600 (human orthologue: RSBN1L). This gene was completely novel and remained uncharacterised in both *Drosophila* and humans making it an extremely interesting and novel gene to characterise. I chose this gene as it affected a wide range of phenotypes from developmental defects, apical constriction, multi-layering, invasion and basal filopodia. Similar to CG7379, my data showed that E-cadherin levels were reduced with CG10600 KD, however this was only seen when combined with the *lgl[4]* mutation. Therefore I suggest that while CG7379 KD is sufficient to impair junction integrity and adhesion, CG10600 requires cooperation with additional tumour suppressors. In comparison to the E-cadherin phenotype observed with CG7379 KD, CG10600 KD did not result in ectopic E-cadherin structures or puncta, suggesting that the process of endocytosis is not being impaired in these mutants.

While I was able to demonstrate that CG10600 has a potential tumour suppressive function by regulating junction integrity and adhesion through E-cadherin,

unfortunately I did not investigate the effect of CG10600 KD on apicobasal polarity. My in vitro invasion assay demonstrated that RSBN1L (CG10600) has anti-invasive functions in a tissue specific manner, able to promote invasion in the adenocarcinoma cell line. While RSBN1L did not promote invasion in U87 cells, as metastasis precedes invasion in glioblastomas, this data does not preclude RSBN1L from being involved in metastasis in these tumours. To answer this question an assay for cell migration should be carried out, such as classic Boydon chamber or scratch assays. With cell adhesion being impaired in *Drosophila*, it would worth looking at adhesion, junction integrity, cell polarity and cytoskeletal regulators in both glioblastomas and adenocarcinomas in vitro.

Due to the lack of genetic tools available for studying this gene, I used CRISPR/cas technology to generate a mutant. While I was able to identify many single nucleotide changes in the mutant, I was not able to fully sequence the entire mutant to the end/new stop codon. Therefore for research to progress further, I suggest that current lab members finish sequencing the mutant and determine how these nucleotide changes might affect protein structure and function.

While my data on CG10600 is limited, it is clear that this gene is important for cell adhesion and may have clinical relevance due to its anti-invasive function in adenocarcinomas. Still, there are many unanswered questions, such as what is the extent of tissue specificity? Does this gene regulate apicobasal polarity? How does this gene regulate cell adhesion? One of the advantages of investigating a completely novel gene is its novelty; however the major disadvantage is the lack of tools

240

available for its study. Therefore, uncovering the molecular mechanisms by which this gene affects tumour progression will be a more long-term project.

7.2.3 Dscam regulates the actin cytoskeleton, cell adhesion and apicobasal polarity.

The final gene that I chose to characterise further was *Dscam*. I showed that *Dscam1, Dscam2* and *Dscam4* all cooperate with *lgl* in our system to promote tumour progression. However, KD of *Dscam1* produced more severe phenotypes compared to the other two homologues. KD of *Dscam* in the *lgl* mutant background produced similar phenotypes to that seen with *CG10600* KD including developmental defects, apical constriction and multi-layering. The most striking phenotype observed however, was long basal filopodia.   When studied in the nervous system, Dscam has been shown to regulate axon guidance as well as the length and spacing of dendritic processes (Li et al., 2015; Schmucker et al., 2000). Taking this into consideration, it is unsurprising that in the epithelium we saw a phenotype associated with actin rich protrusions. I therefore propose that while in the nervous system Dscam is important in regulating the actin cytoskeleton in dendrites and axons, similarly in the epithelium Dscam is important in regulating the actin cytoskeleton in epithelial cells, thereby affecting protrusion morphology. Moreover, based on current literature I believe that it is highly likely to do this via interactions with Rho GTPases. In the nervous system the current model shows that Dscam interacts with Pak and Dock which in turn regulate Rac, Rho and cdc42 during axon guidance (Schmucker et al., 2000). Therefore I suggest that in epithelial cells, Dscam may function in a similar manner to mediate actin filaments associated with filopodia. Dynamic filopodia are associated

241

with EMT and front-rear polarity enabling epithelial cells to become motile and invade/metastasis. Therefore if Dscam is important in the regulation of these dynamic processes it may have clinical relevance in invasive/metastatic human cancers.

Analysis of E-cadherin in *lgl-/-*; Dscam1 KD animals showed that loss of both Dscam1 and Lgl together is sufficient to impair junction integrity. Moreover junction breaks were also observed with this genotype. Additionally, my data showed that in combination with the *lgl[4]* mutation *Dscam* KD affected apicobasal polarity as demonstrated by staining for aPKC and Baz.

Taking into consideration all of my data on Dscam and its known role in the nervous system, I suggest that Dscam regulates the actin cytoskeleton in the epithelium, and is important in regulating filopodial morphology. Additionally, I propose that Dscam plays a novel role in regulating cell-adhesion, junction integrity and apicobasal polarity in an epithelium. To verify this, Dscam mutant tissue could be fixed and stained for actin cytoskeletal proteins (e.g. phalloidin), or a live-actin construct could be expressed in the living epithelium. Additionally, genetic tools like Flybow or LOLLIbow could be used to more closely look at protrusion dynamics and morphology. Moreover, this would allow us to look at the how protrusions from different cells interact with each other.

7.3 Future work

Future work will not be limited to the characterisation of individual genes but will hope to determine whether genes that promote invasion in our in vivo *Drosophila* model also mediate cellular invasion in malignant human cancers. Moreover we hope to identify how tumour genetic background and the extent of tumour progression may affect gene transcription. Hits for invasion could also be used in a secondary screen using 3-D invasion assays. Human orthologues of these genes could be knocked down in human cancer cell lines to see if they promote invasion in the same way. Similarly, this work could be studied further using mouse models, looking at specific cancer types (dependant on which cell types were affected in vitro).

To improve our understanding of the effects candidate genes have on the transcriptome future work will involve extracting RNA from cells of the notum and carrying out RNA Seq analysis. RNA Seq enables complete transcriptome profiling to quantitate precise expression levels of isoforms and transcripts within a cell (Z. Wang, Gerstein, & Snyder, 2009). These experiments will involve generating GFP-positive tumours with different genetic backgrounds; then dissecting and fixing the tissue prior to dissociation. FACS could then be used to sort GFP positive and negative nuclei before performing whole transcriptome sequencing on total RNA from transformed cells together with control RNA derived from non-transformed cells. This will allow us to study genotypes that form neoplastic (invasive) tumours and compare those to hyperplastic (non-invasive) tumours. Once this protocol is established, the lab will possess a combination of tools (fly genetics, live cell imaging and

transcriptome sequencing) that would allow the genetic dissection of tumour progression in a living animal.

7.4 Conclusions

There is still much to learn about the genetic determinants involved in tumour progression, invasion and metastasis. By optimising a novel system in the *Drosophila* notum I have shown how novel genes involved in a variety of cancer-related and cell biological processes can be identified and characterised. I have further investigated 3 genes that I propose play a role in tumour progression. I have proposed that CG7379 acts in a tumour suppressive manner, affecting junction integrity and cell adhesion, potentially through endocytosis-mediated recycling of junction material. Moreover I have provided evidence that this gene's tumour suppressive function may not be limited to epithelial derived cancers. CG10600 on the other hand appeared to be more tissue specific and affect junction integrity and cell adhesion via alternate mechanisms to endocytosis. Finally, I suggested that Dscam has a novel role outside of the nervous system in regulating the actin cytoskeleton, adhesion and apicobasal polarity. My research emphasises the complexity of tumour progression and the multitude of genetic determinants involved, and it is hoped that future experiments can provide a molecular mechanism of how these genes function. In this way my research and continued work in the Georgiou lab may have clinical relevance.

CHAPTER 8

Bibliography

Chapter 8 Bibliography

Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., & Hoschuetzky, H. (1994). Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J Cell Sci, 107 (Pt 12)*, 3655-3663.

Aguirre-Ghiso, J. A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer, 7*(11), 834-846. doi:10.1038/nrc2256

Alberts, B., Johnson, A. and Lewis, J et al. (2002). *Molecular Biology of the Cell. 4th edition*: New York: Garland Science.

Anderson, J. M., & Van Itallie, C. M. (2009). Physiology and function of the tight junction. *Cold Spring Harb Perspect Biol, 1*(2), a002584. doi:10.1101/cshperspect.a002584

Arthur, W. T., & Burridge, K. (2001). RhoA inactivation by p190RhoGAP regulates cell spreading and migration by promoting membrane protrusion and polarity. *Mol Biol Cell, 12*(9), 2711-2720.

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

Ashkenazi, A., & Dixit, V. M. (1999). Apoptosis control by death and decoy receptors. *Curr Opin Cell Biol, 11*(2), 255-260.

Assemat, E., Bazellieres, E., Pallesi-Pocachard, E., Le Bivic, A., & Massey-Harroche, D. (2008). Polarity complex proteins. *Biochim Biophys Acta, 1778*(3), 614-630. doi:10.1016/j.bbamem.2007.08.029

Attrill, H., Falls, K., Goodman, J. L., Millburn, G. H., Antonazzo, G., Rey, A. J., . . . FlyBase, C. (2016). FlyBase: establishing a Gene Group resource for Drosophila melanogaster. *Nucleic Acids Res, 44*(D1), D786-792. doi:10.1093/nar/gkv1046

Barrett, K., Leptin, M., & Settleman, J. (1997). The Rho GTPase and a putative RhoGEF mediate a signaling pathway for the cell shape changes in Drosophila gastrulation. *Cell, 91*(7), 905-915.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., & Garcia De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol, 2*(2), 84-89. doi:10.1038/35000034

Beaucher, M., Goodliffe, J., Hersperger, E., Trunova, S., Frydman, H., & Shearn, A. (2007). Drosophila brain tumor metastases express both neuronal and glial cell type markers. *Dev Biol, 301*(1), 287-297. doi:10.1016/j.ydbio.2006.09.019

Bergers, G., & Benjamin, L. E. (2003). Tumorigenesis and the angiogenic switch. *Nat Rev Cancer, 3*(6), 401-410. doi:10.1038/nrc1093

Bernards, A., & Hariharan, I. K. (2001). Of flies and men--studying human disease in Drosophila. *Curr Opin Genet Dev, 11*(3), 274-278.

Betschinger, J., Mechtler, K., & Knoblich, J. A. (2003). The Par complex directs asymmetric cell division by phosphorylating the cytoskeletal protein Lgl. *Nature, 422*(6929), 326-330. doi:10.1038/nature01486

Bienz, M., & Hamada, F. (2004). Adenomatous polyposis coli proteins and cell adhesion. *Curr Opin Cell Biol, 16*(5), 528-535. doi:10.1016/j.ceb.2004.08.001

Bilder, D. (2003). PDZ domain polarity complexes. *Curr Biol, 13*(17), R661-662.

Bilder, D. (2004). Epithelial polarity and proliferation control: links from the Drosophila neoplastic tumor suppressors. *Genes Dev, 18*(16), 1909-1925. doi:10.1101/gad.1211604

Bilder, D., & Perrimon, N. (2000). Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. *Nature, 403*(6770), 676-680. doi:10.1038/35001108

Bishop, J. M., and Weinberg, R.A. (1996). *Molecular Oncology*: New York: Scientific American, Inc.

Boulina, M., Samarajeewa, H., Baker, J. D., Kim, M. D., & Chiba, A. (2013). Live imaging of multicolor-labeled cells in Drosophila. *Development, 140*(7), 1605-1613. doi:10.1242/dev.088930

Brabletz, T., Jung, A., Reu, S., Porzner, M., Hlubek, F., Kunz-Schughart, L. A., . . . Kirchner, T. (2001). Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc Natl Acad Sci U S A, 98*(18), 10356-10361. doi:10.1073/pnas.171610498

Brand, A. H., Manoukian, A. S., & Perrimon, N. (1994). Ectopic expression in Drosophila. *Methods Cell Biol, 44*, 635-654.

Brenner, D., & Mak, T. W. (2009). Mitochondrial cell death effectors. *Curr Opin Cell Biol, 21*(6), 871-877. doi:10.1016/j.ceb.2009.09.004

Brumby, A. M., & Richardson, H. E. (2003). scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila. *EMBO J, 22*(21), 5769-5779. doi:10.1093/emboj/cdg548

Bruna, A., Darken, R. S., Rojo, F., Ocana, A., Penuelas, S., Arias, A., . . . Seoane, J. (2007). High TGFbeta-Smad activity confers poor prognosis in glioma patients and promotes cell proliferation depending on the methylation of the PDGF-B gene. *Cancer Cell, 11*(2), 147-160. doi:10.1016/j.ccr.2006.11.023

Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., & Reddel, R. R. (1995). Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J, 14*(17), 4240-4248.

Cain, K., Bratton, S. B., & Cohen, G. M. (2002). The Apaf-1 apoptosome: a large caspase-activating complex. *Biochimie, 84*(2-3), 203-214.

Cancer Research UK. (2016). Retrieved from www.cancerresearchuk.org/health-proffessional/cancerstatistics

Canel, M., Serrels, A., Frame, M. C., & Brunton, V. G. (2013). E-cadherin-integrin crosstalk in cancer invasion and metastasis. *J Cell Sci, 126*(Pt 2), 393-401. doi:10.1242/jcs.100115

Cappuzzo, F., Ardizzoni, A., Soto-Parra, H., Gridelli, C., Maione, P., Tiseo, M., . . . Crino, L. (2003). Epidermal growth factor receptor targeted therapy by ZD 1839 (Iressa) in patients with brain metastases from non-small cell lung cancer (NSCLC). *Lung Cancer, 41*(2), 227-231.

Carmeliet, P., & Jain, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature, 407*(6801), 249-257. doi:10.1038/35025220

Carreno, S., Kouranti, I., Glusman, E. S., Fuller, M. T., Echard, A., & Payre, F. (2008). Moesin and its activating kinase Slik are required for cortical stability and microtubule organization in mitotic cells. *J Cell Biol, 180*(4), 739-746. doi:10.1083/jcb.200709161

Cavallaro, U., & Christofori, G. (2004). Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer, 4*(2), 118-132. doi:10.1038/nrc1276

Cesare, A. J., & Reddel, R. R. (2010). Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet, 11*(5), 319-330. doi:10.1038/nrg2763

Chen, B. E., Kondo, M., Garnier, A., Watson, F. L., Puettmann-Holgado, R., Lamar, D. R., & Schmucker, D. (2006). The molecular diversity of Dscam is functionally required for neuronal wiring specificity in Drosophila. *Cell, 125*(3), 607-620. doi:10.1016/j.cell.2006.03.034

Chen, C. L., Schroeder, M. C., Kango-Singh, M., Tao, C., & Halder, G. (2012). Tumor suppression by cell competition through regulation of the Hippo pathway. *Proc Natl Acad Sci U S A, 109*(2), 484-489. doi:10.1073/pnas.1113882109

Chen, Y., Wei, H., Liu, F., & Guan, J. L. (2014). Hyperactivation of mammalian target of rapamycin complex 1 (mTORC1) promotes breast cancer progression through enhancing glucose starvation-induced autophagy and Akt signaling. *J Biol Chem, 289*(2), 1164-1173. doi:10.1074/jbc.M113.526335

Chipuk, J. E., Moldoveanu, T., Llambi, F., Parsons, M. J., & Green, D. R. (2010). The BCL-2 family reunion. *Mol Cell, 37*(3), 299-310. doi:10.1016/j.molcel.2010.01.025

Christofori, G. (2006). New signals from the invasive front. *Nature, 441*(7092), 444-450. doi:10.1038/nature04872

Christofori, G., & Semb, H. (1999). The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. *Trends Biochem Sci, 24*(2), 73-76.

Ciccia, A., & Elledge, S. J. (2010). The DNA damage response: making it safe to play with knives. *Mol Cell, 40*(2), 179-204. doi:10.1016/j.molcel.2010.09.019

Cohen, M., Georgiou, M., Stevenson, N. L., Miodownik, M., & Baum, B. (2010). Dynamic filopodia transmit intermittent Delta-Notch signaling to drive pattern refinement during lateral inhibition. *Dev Cell, 19*(1), 78-89. doi:10.1016/j.devcel.2010.06.006

Coleman, M. L., Marshall, C. J., & Olson, M. F. (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. *Nat Rev Mol Cell Biol, 5*(5), 355-366. doi:10.1038/nrm1365

Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G., Greider, C. W., Harley, C. B., & Bacchetti, S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J, 11*(5), 1921-1929.

Dahman, C. (2008). *Drosophila Methods and Protocols*: New York: Springer-Velgag. Chapter 1.

Daksis, J. I., Lu, R. Y., Facchini, L. M., Marhin, W. W., & Penn, L. J. (1994). Myc induces cyclin D1 expression in the absence of de novo protein synthesis and links mitogen-stimulated signal transduction to the cell cycle. *Oncogene, 9*(12), 3635-3645.

Dameron, K. M., Volpert, O. V., Tainsky, M. A., & Bouck, N. (1994). Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1. *Science, 265*(5178), 1582-1584.

Davis, M. A., Ireton, R. C., & Reynolds, A. B. (2003). A core function for p120-catenin in cadherin turnover. *J Cell Biol, 163*(3), 525-534. doi:10.1083/jcb.200307111

Dawes-Hoang, R. E., Parmar, K. M., Christiansen, A. E., Phelps, C. B., Brand, A. H., & Wieschaus, E. F. (2005). folded gastrulation, cell shape change and the control of myosin localization. *Development, 132*(18), 4165-4178. doi:10.1242/dev.01938

de Beco, S., Gueudry, C., Amblard, F., & Coscoy, S. (2009). Endocytosis is required for E-cadherin redistribution at mature adherens junctions. *Proc Natl Acad Sci U S A, 106*(17), 7010-7015. doi:10.1073/pnas.0811253106

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev, 19*(18), 2100-2110. doi:10.1101/gad.1346005

DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G., & Thompson, C. B. (2008). The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab, 7*(1), 11-20. doi:10.1016/j.cmet.2007.10.002

DeNardo, D. G., Andreu, P., & Coussens, L. M. (2010). Interactions between lymphocytes and myeloid cells regulate pro- versus anti-tumor immunity. *Cancer Metastasis Rev, 29*(2), 309-316. doi:10.1007/s10555-010-9223-6

Deng, C. X. (2006). BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Res, 34*(5), 1416-1426. doi:10.1093/nar/gkl010

Dhruv, H. D., McDonough Winslow, W. S., Armstrong, B., Tuncali, S., Eschbacher, J., Kislin, K., . . . Berens, M. E. (2013). Reciprocal activation of transcription factors underlies the dichotomy between proliferation and invasion of glioma cells. *PLoS One, 8*(8), e72134. doi:10.1371/journal.pone.0072134

Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., . . . Dickson, B. J. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature, 448*(7150), 151-156. doi:10.1038/nature05954

Drees, F., Pokutta, S., Yamada, S., Nelson, W. J., & Weis, W. I. (2005). Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell, 123*(5), 903-915. doi:10.1016/j.cell.2005.09.021

Duffy, M. J. (1996). The biochemistry of metastasis. *Adv Clin Chem, 32*, 135-166.

Dvorak, H. F. (1986). Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. *N Engl J Med, 315*(26), 1650-1659. doi:10.1056/NEJM198612253152606

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes Dev, 12*(15), 2245-2262.

Ebnet, K. (2008). Organization of multiprotein complexes at cell-cell junctions. *Histochem Cell Biol, 130*(1), 1-20. doi:10.1007/s00418-008-0418-7

Etienne-Manneville, S., & Hall, A. (2002). Rho GTPases in cell biology. *Nature, 420*(6916), 629-635. doi:10.1038/nature01148

Fidler, I. J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer, 3*(6), 453-458. doi:10.1038/nrc1098

Flockhart, I. T., Booker, M., Hu, Y., McElvany, B., Gilly, Q., Mathey-Prevot, B., . . . Mohr, S. E. (2012). FlyRNAi.org--the database of the Drosophila RNAi screening center: 2012 update. *Nucleic Acids Res, 40*(Database issue), D715-719. doi:10.1093/nar/gkr953

Foley, K. P., & Eisenman, R. N. (1999). Two MAD tails: what the recent knockouts of Mad1 and Mxi1 tell us about the MYC/MAX/MAD network. *Biochim Biophys Acta, 1423*(3), M37-47.

Fraser, P., & Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature, 447*(7143), 413-417. doi:10.1038/nature05916

Fridman, J. S., & Lowe, S. W. (2003). Control of apoptosis by p53. *Oncogene, 22*(56), 9030-9040. doi:10.1038/sj.onc.1207116

Fristrom, D. (1988). The cellular basis of epithelial morphogenesis. A review. *Tissue Cell, 20*(5), 645-690.

Froldi, F., Ziosi, M., Garoia, F., Pession, A., Grzeschik, N. A., Bellosta, P., . . . Grifoni, D. (2010). The lethal giant larvae tumour suppressor mutation requires dMyc oncoprotein to promote clonal malignancy. *BMC Biol, 8*, 33. doi:10.1186/1741-7007-8-33

Fynan, T. M., & Reiss, M. (1993). Resistance to inhibition of cell growth by transforming growth factor-beta and its role in oncogenesis. *Crit Rev Oncog, 4*(5), 493-540.

Garkavtsev, I., Demetrick, D., & Riabowol, K. (1997). Cellular localization and chromosome mapping of a novel candidate tumor suppressor gene (ING1). *Cytogenet Cell Genet, 76*(3-4), 176-178.

Garkavtsev, I., Grigorian, I. A., Ossovskaya, V. S., Chernov, M. V., Chumakov, P. M., & Gudkov, A. V. (1998). The candidate tumour suppressor p33ING1 cooperates with p53 in cell growth control. *Nature, 391*(6664), 295-298. doi:10.1038/34675

Garkavtsev, I., Kozin, S. V., Chernova, O., Xu, L., Winkler, F., Brown, E., . . . Jain, R. K. (2004). The candidate tumour suppressor protein ING4 regulates brain tumour growth and angiogenesis. *Nature, 428*(6980), 328-332. doi:10.1038/nature02329

Gateff, E. (1978). Malignant neoplasms of genetic origin in Drosophila melanogaster. *Science, 200*(4349), 1448-1459.

Gay, L. J., & Felding-Habermann, B. (2011). Contribution of platelets to tumour metastasis. *Nat Rev Cancer, 11*(2), 123-134. doi:10.1038/nrc3004

Georgiou, M., & Baum, B. (2010). Polarity proteins and Rho GTPases cooperate to spatially organise epithelial actin-based protrusions. *J Cell Sci, 123*(Pt 7), 1089-1098. doi:10.1242/jcs.060772

Georgiou, M., Marinari, E., Burden, J., & Baum, B. (2008). Cdc42, Par6, and aPKC regulate Arp2/3-mediated endocytosis to control local adherens junction stability. *Curr Biol, 18*(21), 1631-1638. doi:10.1016/j.cub.2008.09.029

Gonzalez-Angulo, A. M., Ferrer-Lozano, J., Stemke-Hale, K., Sahin, A., Liu, S., Barrera, J. A., . . . Meric-Bernstam, F. (2011). PI3K pathway mutations and PTEN levels in primary and metastatic breast cancer. *Mol Cancer Ther, 10*(6), 1093-1101. doi:10.1158/1535-7163.MCT-10-1089

Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., & de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell, 97*(4), 503-514.

Grifoni, D., Garoia, F., Schimanski, C. C., Schmitz, G., Laurenti, E., Galle, P. R., . . . Strand, D. (2004). The human protein Hugl-1 substitutes for Drosophila lethal giant larvae tumour suppressor function in vivo. *Oncogene, 23*(53), 8688-8694. doi:10.1038/sj.onc.1208023

Grinberg, A. V., Hu, C. D., & Kerppola, T. K. (2004). Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells. *Mol Cell Biol, 24*(10), 4294-4308.

Grivennikov, S. I., Greten, F. R., & Karin, M. (2010). Immunity, inflammation, and cancer. *Cell, 140*(6), 883-899. doi:10.1016/j.cell.2010.01.025

Grzeschik, N. A., Amin, N., Secombe, J., Brumby, A. M., & Richardson, H. E. (2007). Abnormalities in cell proliferation and apico-basal cell polarity are separable in Drosophila lgl mutant clones in the developing eye. *Dev Biol, 311*(1), 106-123. doi:10.1016/j.ydbio.2007.08.025

Grzeschik, N. A., Parsons, L. M., Allott, M. L., Harvey, K. F., & Richardson, H. E. (2010). Lgl, aPKC, and Crumbs regulate the Salvador/Warts/Hippo pathway through two distinct mechanisms. *Curr Biol, 20*(7), 573-581. doi:10.1016/j.cub.2010.01.055

Gumbiner, B. M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell, 84*(3), 345-357.

Gumbiner, B. M. (2005). Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol, 6*(8), 622-634. doi:10.1038/nrm1699

Guo, W., & Giancotti, F. G. (2004). Integrin signalling during tumour progression. *Nat Rev Mol Cell Biol, 5*(10), 816-826. doi:10.1038/nrm1490

Gupta, G. P., & Massague, J. (2006). Cancer metastasis: building a framework. *Cell, 127*(4), 679-695. doi:10.1016/j.cell.2006.11.001

Hadjieconomou, D., Rotkopf, S., Alexandre, C., Bell, D. M., Dickson, B. J., & Salecker, I. (2011). Flybow: genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. *Nat Methods, 8*(3), 260-266. doi:10.1038/nmeth.1567

Hanahan, D., & Folkman, J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell, 86*(3), 353-364.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell, 100*(1), 57-70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013

Harbour, J. W., & Dean, D. C. (2000). The Rb/E2F pathway: expanding roles and emerging paradigms. *Genes Dev, 14*(19), 2393-2409.

Hartsock, A., & Nelson, W. J. (2008). Adherens and tight junctions: structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta, 1778*(3), 660-669. doi:10.1016/j.bbamem.2007.07.012

Hattori, D., Demir, E., Kim, H. W., Viragh, E., Zipursky, S. L., & Dickson, B. J. (2007). Dscam diversity is essential for neuronal wiring and self-recognition. *Nature, 449*(7159), 223-227. doi:10.1038/nature06099

He, G. H., Helbing, C. C., Wagner, M. J., Sensen, C. W., & Riabowol, K. (2005). Phylogenetic analysis of the ING family of PHD finger proteins. *Mol Biol Evol, 22*(1), 104-116. doi:10.1093/molbev/msh256

Helbing, C. C., Veillette, C., Riabowol, K., Johnston, R. N., & Garkavtsev, I. (1997). A novel candidate tumor suppressor, ING1, is involved in the regulation of apoptosis. *Cancer Res, 57*(7), 1255-1258.

Heldin, C. H., Miyazono, K., & ten Dijke, P. (1997). TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature, 390*(6659), 465-471. doi:10.1038/37284

Herranz, N., Pasini, D., Diaz, V. M., Franci, C., Gutierrez, A., Dave, N., . . . Peiro, S. (2008). Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Mol Cell Biol, 28*(15), 4772-4781. doi:10.1128/MCB.00323-08

Hess, K. R., Varadhachary, G. R., Taylor, S. H., Wei, W., Raber, M. N., Lenzi, R., & Abbruzzese, J. L. (2006). Metastatic patterns in adenocarcinoma. *Cancer, 106*(7), 1624-1633. doi:10.1002/cncr.21778

Hing, H., Xiao, J., Harden, N., Lim, L., & Zipursky, S. L. (1999). Pak functions downstream of Dock to regulate photoreceptor axon guidance in Drosophila. *Cell, 97*(7), 853-863.

Hoesel, B., & Schmid, J. A. (2013). The complexity of NF-kappaB signaling in inflammation and cancer. *Mol Cancer, 12*, 86. doi:10.1186/1476-4598-12-86

Hsu, P. P., & Sabatini, D. M. (2008). Cancer cell metabolism: Warburg and beyond. *Cell, 134*(5), 703-707. doi:10.1016/j.cell.2008.08.021

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., & Mohr, S. E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics, 12*, 357. doi:10.1186/1471-2105-12-357

Huang, H., Potter, C. J., Tao, W., Li, D. M., Brogiolo, W., Hafen, E., . . . Xu, T. (1999). PTEN affects cell size, cell proliferation and apoptosis during Drosophila eye development. *Development, 126*(23), 5365-5372.

Huang, R. Y., Guilford, P., & Thiery, J. P. (2012). Early events in cell adhesion and polarity during epithelial-mesenchymal transition. *J Cell Sci, 125*(Pt 19), 4417-4422. doi:10.1242/jcs.099697

Huber, A. H., Stewart, D. B., Laurents, D. V., Nelson, W. J., & Weis, W. I. (2001). The cadherin cytoplasmic domain is unstructured in the absence of beta-catenin. A possible mechanism for regulating cadherin turnover. *J Biol Chem, 276*(15), 12301-12309. doi:10.1074/jbc.M010377200

Hughes, M. E., Bortnick, R., Tsubouchi, A., Baumer, P., Kondo, M., Uemura, T., & Schmucker, D. (2007). Homophilic Dscam interactions control complex dendrite morphogenesis. *Neuron, 54*(3), 417-427. doi:10.1016/j.neuron.2007.04.013

Hummel, T., Vasconcelos, M. L., Clemens, J. C., Fishilevich, Y., Vosshall, L. B., & Zipursky, S. L. (2003). Axonal targeting of olfactory receptor neurons in Drosophila is controlled by Dscam. *Neuron, 37*(2), 221-231.

Januschke, J., & Gonzalez, C. (2008). Drosophila asymmetric division, polarity and cancer. *Oncogene, 27*(55), 6994-7002. doi:10.1038/onc.2008.349

Joberty, G., Petersen, C., Gao, L., & Macara, I. G. (2000). The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. *Nat Cell Biol, 2*(8), 531-539. doi:10.1038/35019573

Johnson, D. I. (1999). Cdc42: An essential Rho-type GTPase controlling eukaryotic cell polarity. *Microbiol Mol Biol Rev, 63*(1), 54-105.

Johnson, J. P. (1991). Cell adhesion molecules of the immunoglobulin supergene family and their role in malignant transformation and progression to metastatic disease. *Cancer Metastasis Rev, 10*(1), 11-22.

Jones, R. G., & Thompson, C. B. (2009). Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes Dev, 23*(5), 537-548. doi:10.1101/gad.1756509

Joyce, J. A., & Pollard, J. W. (2009). Microenvironmental regulation of metastasis. *Nat Rev Cancer, 9*(4), 239-252. doi:10.1038/nrc2618

Junttila, M. R., & de Sauvage, F. J. (2013). Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature, 501*(7467), 346-354. doi:10.1038/nature12626

Kallay, L. M., McNickle, A., Brennwald, P. J., Hubbard, A. L., & Braiterman, L. T. (2006). Scribble associates with two polarity proteins, Lgl2 and Vangl2, via distinct molecular domains. *J Cell Biochem, 99*(2), 647-664. doi:10.1002/jcb.20992

Kim, M. Y., Oskarsson, T., Acharyya, S., Nguyen, D. X., Zhang, X. H., Norton, L., & Massague, J. (2009). Tumor self-seeding by circulating cancer cells. *Cell, 139*(7), 1315-1326. doi:10.1016/j.cell.2009.11.025

Kim, S., Chin, K., Gray, J. W., & Bishop, J. M. (2004). A screen for genes that suppress loss of contact inhibition: identification of ING4 as a candidate tumor suppressor gene in human cancer. *Proc Natl Acad Sci U S A, 101*(46), 16251-16256. doi:10.1073/pnas.0407158101

Kinzler, K. W., & Vogelstein, B. (1997). Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature, 386*(6627), 761, 763. doi:10.1038/386761a0

Kishida, S., Yamamoto, H., Ikeda, S., Kishida, M., Sakamoto, I., Koyama, S., & Kikuchi, A. (1998). Axin, a negative regulator of the wnt signaling pathway, directly interacts with adenomatous polyposis coli and regulates the stabilization of beta-catenin. *J Biol Chem, 273*(18), 10823-10826.

Kleiner, D. E., & Stetler-Stevenson, W. G. (1999). Matrix metalloproteinases and metastasis. *Cancer Chemother Pharmacol, 43 Suppl*, S42-51.

Klymkowsky, M. W., & Savagner, P. (2009). Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. *Am J Pathol, 174*(5), 1588-1593. doi:10.2353/ajpath.2009.080545

Kolsch, V., Seher, T., Fernandez-Ballester, G. J., Serrano, L., & Leptin, M. (2007). Control of Drosophila gastrulation by apical localization of adherens junctions and RhoGEF2. *Science, 315*(5810), 384-386. doi:10.1126/science.1134833

Kunda, P., Pelling, A. E., Liu, T., & Baum, B. (2008). Moesin controls cortical rigidity, cell rounding, and spindle morphogenesis during mitosis. *Curr Biol, 18*(2), 91-101. doi:10.1016/j.cub.2007.12.051

Kuphal, S., Wallner, S., Schimanski, C. C., Bataille, F., Hofer, P., Strand, S., . . . Bosserhoff, A. K. (2006). Expression of Hugl-1 is strongly reduced in malignant melanoma. *Oncogene, 25*(1), 103-110. doi:10.1038/sj.onc.1209008

Lamouille, S., Xu, J., & Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol, 15*(3), 178-196. doi:10.1038/nrm3758

Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., . . . Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature, 505*(7484), 495-501. doi:10.1038/nature12912

Le, T. L., Yap, A. S., & Stow, J. L. (1999). Recycling of E-cadherin: a potential mechanism for regulating cadherin dynamics. *J Cell Biol, 146*(1), 219-232.

Lee, S. S., Weiss, R. S., & Javier, R. T. (1997). Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein. *Proc Natl Acad Sci U S A, 94*(13), 6670-6675.

Lee, T., & Luo, L. (2001). Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. *Trends Neurosci, 24*(5), 251-254.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell, 88*(3), 323-331.

Li, S., Sukeena, J. M., Simmons, A. B., Hansen, E. J., Nuhn, R. E., Samuels, I. S., & Fuerst, P. G. (2015). DSCAM promotes refinement in the mouse retina through cell death and restriction of exploring dendrites. *J Neurosci, 35*(14), 5640-5654. doi:10.1523/JNEUROSCI.2202-14.2015

Lim, S., & Kaldis, P. (2013). Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development, 140*(15), 3079-3093. doi:10.1242/dev.091744

Lin, D., Edwards, A. S., Fawcett, J. P., Mbamalu, G., Scott, J. D., & Pawson, T. (2000). A mammalian PAR-3-PAR-6 complex implicated in Cdc42/Rac1 and aPKC signalling and cell polarity. *Nat Cell Biol, 2*(8), 540-547. doi:10.1038/35019582

Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M. N., & Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J, 20*(9), 2224-2235. doi:10.1093/emboj/20.9.2224

Lodish, H., Berk, A. and Zipursky, S.L. (2000). *Molecular Cell Biology 4th edition* (4th edition ed.): New York: W. H. Freeman.

Loewith, R., Meijer, M., Lees-Miller, S. P., Riabowol, K., & Young, D. (2000). Three yeast proteins related to the human candidate tumor suppressor p33(ING1) are associated with histone acetyltransferase activities. *Mol Cell Biol, 20*(11), 3807-3816.

Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., . . . Schlessinger, J. (1992). The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell, 70*(3), 431-442.

Lu, M., Chen, F., Wang, Q., Wang, K., Pan, Q., & Zhang, X. (2012). Downregulation of inhibitor of growth 3 is correlated with tumorigenesis and progression of hepatocellular carcinoma. *Oncol Lett, 4*(1), 47-52. doi:10.3892/ol.2012.685

Lu, P., Wang, Y., Liu, X., Wang, H., Zhang, X., Wang, K., . . . Hu, R. (2016). Malignant gliomas induce and exploit astrocytic mesenchymal-like transition by activating canonical Wnt/beta-catenin signaling. *Med Oncol, 33*(7), 66. doi:10.1007/s12032-016-0778-0

Luchino, J., Hocine, M., Amoureux, M. C., Gibert, B., Bernet, A., Royet, A., . . . Mann, F. (2013). Semaphorin 3E suppresses tumor cell death triggered by the plexin D1 dependence receptor in metastatic breast cancers. *Cancer Cell, 24*(5), 673-685. doi:10.1016/j.ccr.2013.09.010

Luo, M., Hou, L., Li, J., Shao, S., Huang, S., Meng, D., . . . Zhao, X. (2016). VEGF/NRP-1axis promotes progression of breast cancer via enhancement of epithelial-mesenchymal transition and activation of NF-kappaB and beta-catenin. *Cancer Lett, 373*(1), 1-11. doi:10.1016/j.canlet.2016.01.010

Ly, A., Nikolaev, A., Suresh, G., Zheng, Y., Tessier-Lavigne, M., & Stein, E. (2008). DSCAM is a netrin receptor that collaborates with DCC in mediating turning responses to netrin-1. *Cell, 133*(7), 1241-1254. doi:10.1016/j.cell.2008.05.030

Malanchi, I., Santamaria-Martinez, A., Susanto, E., Peng, H., Lehr, H. A., Delaloye, J. F., & Huelsken, J. (2012). Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature, 481*(7379), 85-89. doi:10.1038/nature10694

Malumbres, M., & Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer, 9*(3), 153-166. doi:10.1038/nrc2602

Margolis, B., & Skolnik, E. Y. (1994). Activation of Ras by receptor tyrosine kinases. *J Am Soc Nephrol, 5*(6), 1288-1299.

Matsumine, A., Ogai, A., Senda, T., Okumura, N., Satoh, K., Baeg, G. H., . . . Akiyama, T. (1996). Binding of APC to the human homolog of the Drosophila discs large tumor suppressor protein. *Science, 272*(5264), 1020-1023.

Matthews, B. J., Kim, M. E., Flanagan, J. J., Hattori, D., Clemens, J. C., Zipursky, S. L., & Grueber, W. B. (2007). Dendrite self-avoidance is controlled by Dscam. *Cell, 129*(3), 593-604. doi:10.1016/j.cell.2007.04.013

Mayo, L. D., Turchi, J. J., & Berberich, S. J. (1997). Mdm-2 phosphorylation by DNA-dependent protein kinase prevents interaction with p53. *Cancer Res, 57*(22), 5013-5016.

Mechler, B. M., McGinnis, W., & Gehring, W. J. (1985). Molecular cloning of lethal(2)giant larvae, a recessive oncogene of Drosophila melanogaster. *EMBO J, 4*(6), 1551-1557.

Meek, D. W. (1998). Multisite phosphorylation and the integration of stress signals at p53. *Cell Signal, 10*(3), 159-166.

Meng, Q., Zhi, T., Chao, Y., Nie, E., Xu, X., Shi, Q., . . . Yu, R. (2014). Bex2 controls proliferation of human glioblastoma cells through NF-kappaB signaling pathway. *J Mol Neurosci, 53*(2), 262-270. doi:10.1007/s12031-013-0215-1

Meng, W., & Takeichi, M. (2009). Adherens junction: molecular architecture and regulation. *Cold Spring Harb Perspect Biol, 1*(6), a002899. doi:10.1101/cshperspect.a002899

Miano, V., Ferrero, G., Reineri, S., Caizzi, L., Annaratone, L., Ricci, L., . . . De Bortoli, M. (2016). Luminal long non-coding RNAs regulated by estrogen receptor alpha in a ligand-independent manner show functional roles in breast cancer. *Oncotarget, 7*(3), 3201-3216. doi:10.18632/oncotarget.6420

Millard, S. S., Flanagan, J. J., Pappu, K. S., Wu, W., & Zipursky, S. L. (2007). Dscam2 mediates axonal tiling in the Drosophila visual system. *Nature, 447*(7145), 720-724. doi:10.1038/nature05855

Miller, K., Wang, M., Gralow, J., Dickler, M., Cobleigh, M., Perez, E. A., . . . Davidson, N. E. (2007). Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med, 357*(26), 2666-2676. doi:10.1056/NEJMoa072113

Miyashita, T., Krajewski, S., Krajewska, M., Wang, H. G., Lin, H. K., Liebermann, D. A., . . . Reed, J. C. (1994). Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression in vitro and in vivo. *Oncogene, 9*(6), 1799-1805.

Morais-de-Sa, E., Mirouse, V., & St Johnston, D. (2010). aPKC phosphorylation of Bazooka defines the apical/lateral border in Drosophila epithelial cells. *Cell, 141*(3), 509-523. doi:10.1016/j.cell.2010.02.040

Morata, G., & Ripoll, P. (1975). Minutes: mutants of drosophila autonomously affecting cell division rate. *Dev Biol, 42*(2), 211-221.

Morgan, D. O. (2006). *The cell cycle: principles of control.*: London: New Science Press.

Moses, H. L., Yang, E. Y., & Pietenpol, J. A. (1990). TGF-beta stimulation and inhibition of cell proliferation: new mechanistic insights. *Cell, 63*(2), 245-247.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., . . . Wu, J. R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A, 85*(18), 6622-6626.

Muller, P. A., & Vousden, K. H. (2014). Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell, 25*(3), 304-317. doi:10.1016/j.ccr.2014.01.021

Nair, S. K., & Burley, S. K. (2003). X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell, 112*(2), 193-205.

Nance, J., Munro, E. M., & Priess, J. R. (2003). C. elegans PAR-3 and PAR-6 are required for apicobasal asymmetries associated with cell adhesion and gastrulation. *Development, 130*(22), 5339-5350. doi:10.1242/dev.00735

Narod, S., Lynch, H., Conway, T., Watson, P., Feunteun, J., & Lenoir, G. (1993). Increasing incidence of breast cancer in family with BRCA1 mutation. *Lancet, 341*(8852), 1101-1102.

Negrini, S., Gorgoulis, V. G., & Halazonetis, T. D. (2010). Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol, 11*(3), 220-228. doi:10.1038/nrm2858

Nelson, W. J. (2009). Remodeling epithelial cell organization: transitions between front-rear and apical-basal polarity. *Cold Spring Harb Perspect Biol, 1*(1), a000513. doi:10.1101/cshperspect.a000513

Nobes, C. D., & Hall, A. (1995). Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell, 81*(1), 53-62.

Normanno, N., De Luca, A., Bianco, C., Strizzi, L., Mancino, M., Maiello, M. R., . . . Salomon, D. S. (2006). Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene, 366*(1), 2-16. doi:10.1016/j.gene.2005.10.018

Nusslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature, 287*(5785), 795-801.

Oda, H., Tsukita, S., & Takeichi, M. (1998). Dynamic behavior of the cadherin-based cell-cell adhesion system during Drosophila gastrulation. *Dev Biol, 203*(2), 435-450. doi:10.1006/dbio.1998.9047

Ohtani, K., DeGregori, J., & Nevins, J. R. (1995). Regulation of the cyclin E gene by transcription factor E2F1. *Proc Natl Acad Sci U S A, 92*(26), 12146-12150.

Oshima, K., & Fehon, R. G. (2011). Analysis of protein dynamics within the septate junction reveals a highly stable core protein complex that does not include the basolateral polarity protein Discs large. *J Cell Sci, 124*(Pt 16), 2861-2871. doi:10.1242/jcs.087700

Pagliarini, R. A., & Xu, T. (2003). A genetic screen in Drosophila for metastatic behavior. *Science, 302*(5648), 1227-1231. doi:10.1126/science.1088474

Pardali, K., Kowanetz, M., Heldin, C. H., & Moustakas, A. (2005). Smad pathway-specific transcriptional regulation of the cell cycle inhibitor p21(WAF1/Cip1). *J Cell Physiol, 204*(1), 260-272. doi:10.1002/jcp.20304

Patel, S. D., Ciatto, C., Chen, C. P., Bahna, F., Rajebhosale, M., Arkus, N., . . . Shapiro, L. (2006). Type II cadherin ectodomain structures: implications for classical cadherin specificity. *Cell, 124*(6), 1255-1268. doi:10.1016/j.cell.2005.12.046

Pokutta, S., Herrenknecht, K., Kemler, R., & Engel, J. (1994). Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur J Biochem, 223*(3), 1019-1026.

Pokutta, S., & Weis, W. I. (2000). Structure of the dimerization and beta-catenin-binding region of alpha-catenin. *Mol Cell, 5*(3), 533-543.

Pollard, T. D., & Beltzner, C. C. (2002). Structure and function of the Arp2/3 complex. *Curr Opin Struct Biol, 12*(6), 768-774.

Pollard, T. D., & Borisy, G. G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell, 112*(4), 453-465.

Polyak, K., & Weinberg, R. A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer, 9*(4), 265-273. doi:10.1038/nrc2620

Port, F., Chen, H. M., Lee, T., & Bullock, S. L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc Natl Acad Sci U S A, 111*(29), E2967-2976. doi:10.1073/pnas.1405500111

Potter, C. J., Turenchalk, G. S., & Xu, T. (2000). Drosophila in cancer research. An expanding role. *Trends Genet, 16*(1), 33-39.

Powell, S. N., & Kachnic, L. A. (2003). Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation. *Oncogene, 22*(37), 5784-5791. doi:10.1038/sj.onc.1206678

Qian, B. Z., & Pollard, J. W. (2010). Macrophage diversity enhances tumor progression and metastasis. *Cell, 141*(1), 39-51. doi:10.1016/j.cell.2010.03.014

Qu, C., Li, W., Shao, Q., Dwyer, T., Huang, H., Yang, T., & Liu, G. (2013). c-Jun N-terminal kinase 1 (JNK1) is required for coordination of netrin signaling in axon guidance. *J Biol Chem, 288*(3), 1883-1895. doi:10.1074/jbc.M112.417881

Rajarajacholan, U. K., Thalappilly, S., & Riabowol, K. (2013). The ING1a tumor suppressor regulates endocytosis to induce cellular senescence via the Rb-E2F pathway. *PLoS Biol, 11*(3), e1001502. doi:10.1371/journal.pbio.1001502

Rak, J., Mitsuhashi, Y., Sheehan, C., Tamir, A., Viloria-Petit, A., Filmus, J., . . . Kerbel, R. S. (2000). Oncogenes and tumor angiogenesis: differential modes of vascular endothelial growth factor up-regulation in ras-transformed epithelial cells and fibroblasts. *Cancer Res, 60*(2), 490-498.

Ridley, A. J., Schwartz, M. A., Burridge, K., Firtel, R. A., Ginsberg, M. H., Borisy, G., . . . Horwitz, A. R. (2003). Cell migration: integrating signals from front to back. *Science, 302*(5651), 1704-1709. doi:10.1126/science.1092053

Rodriguez-Boulan, E., & Macara, I. G. (2014). Organization and execution of the epithelial polarity programme. *Nat Rev Mol Cell Biol, 15*(4), 225-242. doi:10.1038/nrm3775

Royer, C., & Lu, X. (2011). Epithelial cell polarity: a major gatekeeper against cancer? *Cell Death Differ, 18*(9), 1470-1477. doi:10.1038/cdd.2011.60

Rubin, G. M., & Lewis, E. B. (2000). A brief history of Drosophila's contributions to genome research. *Science, 287*(5461), 2216-2218.

Sabeh, F., Shimizu-Hirota, R., & Weiss, S. J. (2009). Protease-dependent versus -independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. *J Cell Biol, 185*(1), 11-19. doi:10.1083/jcb.200807195

Sahai, E., & Marshall, C. J. (2003). Differing modes of tumour cell invasion have distinct requirements for Rho/ROCK signalling and extracellular proteolysis. *Nat Cell Biol, 5*(8), 711-719. doi:10.1038/ncb1019

Sansom, O. J., Reed, K. R., Hayes, A. J., Ireland, H., Brinkmann, H., Newton, I. P., . . . Winton, D. J. (2004). Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev, 18*(12), 1385-1390. doi:10.1101/gad.287404

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., . . . Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature, 419*(6905), 407-411. doi:10.1038/nature01080

Savage, D. G., & Antman, K. H. (2002). Imatinib mesylate--a new oral targeted therapy. *N Engl J Med, 346*(9), 683-693. doi:10.1056/NEJMra013339

Schimanski, C. C., Schmitz, G., Kashyap, A., Bosserhoff, A. K., Bataille, F., Schafer, S. C., . . . Strand, D. (2005). Reduced expression of Hugl-1, the human homologue of Drosophila tumour suppressor gene lgl, contributes to progression of colorectal cancer. *Oncogene, 24*(19), 3100-3109. doi:10.1038/sj.onc.1208520

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., . . . Zipursky, S. L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell, 101*(6), 671-684.

Schneiderman, H. A., & Gateff, E. (1967). Control systems in insect development. *Science, 158*(3800), 534-535. doi:10.1126/science.158.3800.534-c

Semenza, G. L. (2010). Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene, 29*(5), 625-634. doi:10.1038/onc.2009.441

Shay, J. W., & Wright, W. E. (2000). Hayflick, his limit, and cellular ageing. *Nat Rev Mol Cell Biol, 1*(1), 72-76. doi:10.1038/35036093

Shen, J. C., Unoki, M., Ythier, D., Duperray, A., Varticovski, L., Kumamoto, K., . . . Harris, C. C. (2007). Inhibitor of growth 4 suppresses cell spreading and cell migration by interacting with

a novel binding partner, liprin alpha1. *Cancer Res, 67*(6), 2552-2558. doi:10.1158/0008-5472.CAN-06-3870

Sherr, C. J., & Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev, 13*(12), 1501-1512.

Shiseki, M., Nagashima, M., Pedeux, R. M., Kitahama-Shiseki, M., Miura, K., Okamura, S., . . . Harris, C. C. (2003). p29ING4 and p28ING5 bind to p53 and p300, and enhance p53 activity. *Cancer Res, 63*(10), 2373-2378.

Siliciano, J. D., Canman, C. E., Taya, Y., Sakaguchi, K., Appella, E., & Kastan, M. B. (1997). DNA damage induces phosphorylation of the amino terminus of p53. *Genes Dev, 11*(24), 3471-3481.

Simpson, P. (2007). The stars and stripes of animal bodies: evolution of regulatory elements mediating pigment and bristle patterns in Drosophila. *Trends Genet, 23*(7), 350-358. doi:10.1016/j.tig.2007.04.006

Simpson, P., & Morata, G. (1981). Differential mitotic rates and patterns of growth in compartments in the Drosophila wing. *Dev Biol, 85*(2), 299-308.

Singh, A., & Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene, 29*(34), 4741-4751. doi:10.1038/onc.2010.215

Soba, P., Zhu, S., Emoto, K., Younger, S., Yang, S. J., Yu, H. H., . . . Jan, Y. N. (2007). Drosophila sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron, 54*(3), 403-416. doi:10.1016/j.neuron.2007.03.029

Soliman, M. A., & Riabowol, K. (2007). After a decade of study-ING, a PHD for a versatile family of proteins. *Trends Biochem Sci, 32*(11), 509-519. doi:10.1016/j.tibs.2007.08.006

Sporn, M. B. (1996). The war on cancer. *Lancet, 347*(9012), 1377-1381.

Suzuki, A., & Ohno, S. (2006). The PAR-aPKC system: lessons in polarity. *J Cell Sci, 119*(Pt 6), 979-987. doi:10.1242/jcs.02898

Takahashi, T., Tanaka, H., Iguchi, N., Kitamura, K., Chen, Y., Maekawa, M., . . . Nishimune, Y. (2004). Rosbin: a novel homeobox-like protein gene expressed exclusively in round spermatids. *Biol Reprod, 70*(5), 1485-1492. doi:10.1095/biolreprod.103.026096

Tamm, I., Wang, Y., Sausville, E., Scudiero, D. A., Vigna, N., Oltersdorf, T., & Reed, J. C. (1998). IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas (CD95), Bax, caspases, and anticancer drugs. *Cancer Res, 58*(23), 5315-5320.

Tamura, M., Gu, J., Matsumoto, K., Aota, S., Parsons, R., & Yamada, K. M. (1998). Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. *Science, 280*(5369), 1614-1617.

Tamura, M., Gu, J., Takino, T., & Yamada, K. M. (1999). Tumor suppressor PTEN inhibition of cell invasion, migration, and growth: differential involvement of focal adhesion kinase and p130Cas. *Cancer Res, 59*(2), 442-449.

Tanentzapf, G., & Tepass, U. (2003). Interactions between the crumbs, lethal giant larvae and bazooka pathways in epithelial polarization. *Nat Cell Biol, 5*(1), 46-52. doi:10.1038/ncb896

Teng, M. W., Swann, J. B., Koebel, C. M., Schreiber, R. D., & Smyth, M. J. (2008). Immune-mediated dormancy: an equilibrium with cancer. *J Leukoc Biol, 84*(4), 988-993. doi:10.1189/jlb.1107774

Tepass, U., Theres, C., & Knust, E. (1990). crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila epithelial cells and required for organization of epithelia. *Cell, 61*(5), 787-799.

Thiery, J. P., Acloque, H., Huang, R. Y., & Nieto, M. A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell, 139*(5), 871-890. doi:10.1016/j.cell.2009.11.007

Thornberry, N. A., & Lazebnik, Y. (1998). Caspases: enemies within. *Science, 281*(5381), 1312-1316.

Tsai, J. H., & Yang, J. (2013). Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev, 27*(20), 2192-2206. doi:10.1101/gad.225334.113

Tsang, F. C., Po, L. S., Leung, K. M., Lau, A., Siu, W. Y., & Poon, R. Y. (2003). ING1b decreases cell proliferation through p53-dependent and -independent mechanisms. *FEBS Lett, 553*(3), 277-285.

Valastyan, S., & Weinberg, R. A. (2011). Tumor metastasis: molecular insights and evolving paradigms. *Cell, 147*(2), 275-292. doi:10.1016/j.cell.2011.09.024

Valiente, M., Andres-Pons, A., Gomar, B., Torres, J., Gil, A., Tapparel, C., . . . Pulido, R. (2005). Binding of PTEN to specific PDZ domains contributes to PTEN protein stability and phosphorylation by microtubule-associated serine/threonine kinases. *J Biol Chem, 280*(32), 28936-28943. doi:10.1074/jbc.M504761200

Vaux, D. L., Cory, S., & Adams, J. M. (1988). Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature, 335*(6189), 440-442. doi:10.1038/335440a0

Vignjevic, D., Peloquin, J., & Borisy, G. G. (2006). In vitro assembly of filopodia-like bundles. *Methods Enzymol, 406*, 727-739. doi:10.1016/S0076-6879(06)06057-5

Vignjevic, D., Schoumacher, M., Gavert, N., Janssen, K. P., Jih, G., Lae, M., . . . Robine, S. (2007). Fascin, a novel target of beta-catenin-TCF signaling, is expressed at the invasive front of human colon cancer. *Cancer Res, 67*(14), 6844-6853. doi:10.1158/0008-5472.CAN-07-0929

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nat Med, 10*(8), 789-799. doi:10.1038/nm1087

Volpert, O. V., Dameron, K. M., & Bouck, N. (1997). Sequential development of an angiogenic phenotype by human fibroblasts progressing to tumorigenicity. *Oncogene, 14*(12), 1495-1502. doi:10.1038/sj.onc.1200977

Wang, A. Z., Ojakian, G. K., & Nelson, W. J. (1990). Steps in the morphogenesis of a polarized epithelium. I. Uncoupling the roles of cell-cell and cell-substratum contact in establishing

plasma membrane polarity in multicellular epithelial (MDCK) cysts. *J Cell Sci, 95 (Pt 1)*, 137-151.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet, 10*(1), 57-63. doi:10.1038/nrg2484

Weinberg, R. A. (1995). The retinoblastoma protein and cell cycle control. *Cell, 81*(3), 323-330.

Weinberg, R. A. (2014). *The Biology of Cancer*: Taylor & Francis Inc , Garland Publishing Inc.

Wells, R. E., Barry, J. D., Warrington, S. J., Cuhlmann, S., Evans, P., Huber, W., . . . Zeidler, M. P. (2013). Control of tissue morphology by Fasciclin III-mediated intercellular adhesion. *Development, 140*(18), 3858-3868. doi:10.1242/dev.096214

Werb, Z. (1997). ECM and cell surface proteolysis: regulating cellular ecology. *Cell, 91*(4), 439-442.

Wojtowicz, W. M., Flanagan, J. J., Millard, S. S., Zipursky, S. L., & Clemens, J. C. (2004). Alternative splicing of Drosophila Dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding. *Cell, 118*(5), 619-633. doi:10.1016/j.cell.2004.08.021

Woods, D. F., & Bryant, P. J. (1991). The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell, 66*(3), 451-464.

Woods, D. F., Hough, C., Peel, D., Callaini, G., & Bryant, P. J. (1996). Dlg protein is required for junction structure, cell polarity, and proliferation control in Drosophila epithelia. *J Cell Biol, 134*(6), 1469-1482.

Wyllie, A. H., Kerr, J. F., & Currie, A. R. (1980). Cell death: the significance of apoptosis. *Int Rev Cytol, 68*, 251-306.

Xie, Q., Mittal, S., & Berens, M. E. (2014). Targeting adaptive glioblastoma: an overview of proliferation and invasion. *Neuro Oncol, 16*(12), 1575-1584. doi:10.1093/neuonc/nou147

Xu, T., & Rubin, G. M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. *Development, 117*(4), 1223-1237.

Yamanaka, T., Horikoshi, Y., Suzuki, A., Sugiyama, Y., Kitamura, K., Maniwa, R., . . . Ohno, S. (2001). PAR-6 regulates aPKC activity in a novel way and mediates cell-cell contact-induced formation of the epithelial junctional complex. *Genes Cells, 6*(8), 721-731.

Yang, F., Sun, L., Li, Q., Han, X., Lei, L., Zhang, H., & Shang, Y. (2012). SET8 promotes epithelial-mesenchymal transition and confers TWIST dual transcriptional activities. *EMBO J, 31*(1), 110-123. doi:10.1038/emboj.2011.364

Yap, A. S., Niessen, C. M., & Gumbiner, B. M. (1998). The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interaction with p120ctn. *J Cell Biol, 141*(3), 779-789.

Yilmaz, M., & Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev, 28*(1-2), 15-33. doi:10.1007/s10555-008-9169-0

Zarkoob, H., Taube, J. H., Singh, S. K., Mani, S. A., & Kohandel, M. (2013). Investigating the link between molecular subtypes of glioblastoma, epithelial-mesenchymal transition, and CD133 cell surface protein. *PLoS One, 8*(5), e64169. doi:10.1371/journal.pone.0064169

Zeitlinger, J., & Bohmann, D. (1999). Thorax closure in Drosophila: involvement of Fos and the JNK pathway. *Development, 126*(17), 3947-3956.

Zeng, X., Han, L., Singh, S. R., Liu, H., Neumuller, R. A., Yan, D., . . . Hou, S. X. (2015). Genome-wide RNAi screen identifies networks involved in intestinal stem cell regulation in Drosophila. *Cell Rep, 10*(7), 1226-1238. doi:10.1016/j.celrep.2015.01.051

Zetter, B. R. (1998). Angiogenesis and tumor metastasis. *Annu Rev Med, 49*, 407-424. doi:10.1146/annurev.med.49.1.407

Zhang, W., & Liu, H. T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res, 12*(1), 9-18. doi:10.1038/sj.cr.7290105

Zhao, Q. Y., Ju, F., Wang, Z. H., Ma, X. Z., & Zhao, H. (2015). ING5 inhibits epithelial-mesenchymal transition in breast cancer by suppressing PI3K/Akt pathway. *Int J Clin Exp Med, 8*(9), 15498-15505.

Zheng, Y., Yang, W., Aldape, K., He, J., & Lu, Z. (2013). Epidermal growth factor (EGF)-enhanced vascular cell adhesion molecule-1 (VCAM-1) expression promotes macrophage and glioblastoma cell interaction and tumor cell invasion. *J Biol Chem, 288*(44), 31488-31495. doi:10.1074/jbc.M113.499020

APPENDIX

Appendix A: Plasmid map for pCFD3

pCFD3-dU6:3gRNA was a gift from Simon Bullock (Addgene plasmid # 49410) (Port et al., 2014).
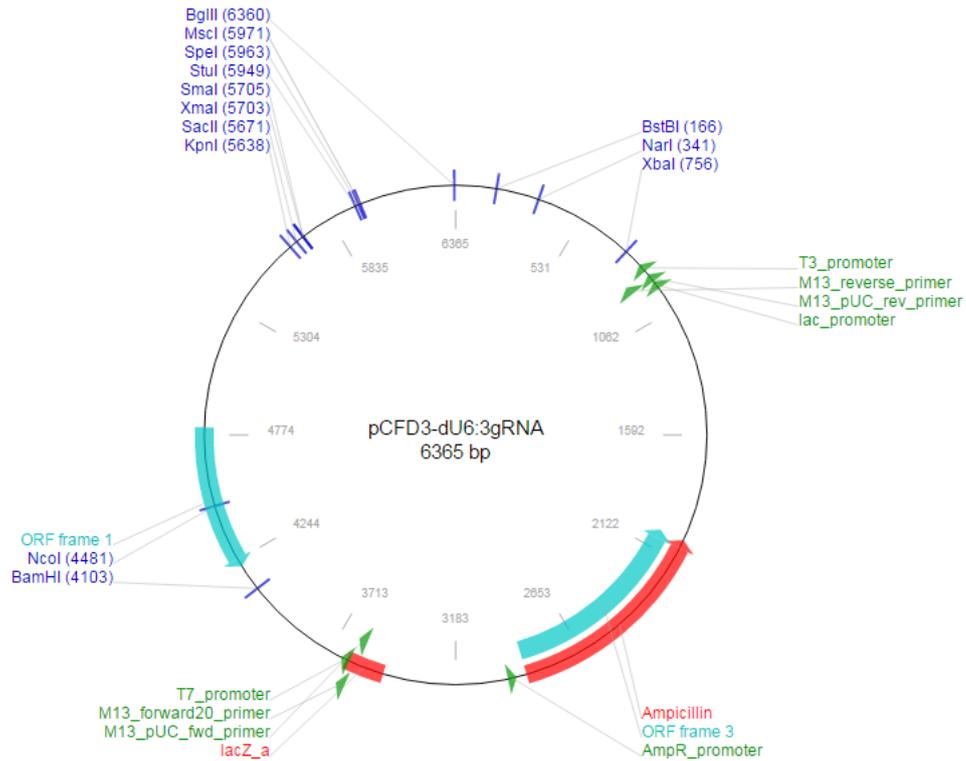
Appendix B: Pilot genes

Gene	RNAi	Source	GO Terms
Abi	36142-GD (III)	VDRC	actin cytoskeleton regulation, adhesion; lamellipodium assembly
APC	51469-GD (III)	VDRC	Polarity. Adhesion
	1451R-1 (III)	NIG	
Apc2	22290-GD (X)	VDRC	Cell adhesion; cell cycle. polarity;
Archipelago	15010R-3 (III)	NIG	Cell cycle. Regulation of hypoxic reaction
Arm	11579R-1 (III)	NIG	Polarity and adhesion
Baz-IR058	5055R-1 (III)	NIG	Polarity and adhesion
Bcl2	47515-GD (III)	VDRC	apoptosis
	12397R-1 (III)	NIG	
Beclin-1 (Atg6)	22122-GD (III)	VDRC	Autophagy.
cdc42	12530R-2 (III)	NIG	GTPase/Polarity/actin.cell migration
CDK5-IR	8203R-3 (III)	NIG	adult behavior;centrosome localization;
Cytochrome c	17128-GD (III)	VDRC	Apoptosis
dE2f2	45743-GD (X)	VDRC	Cell cycle.
	1071R-1 (III)	NIG	
Death caspase	5370R-2 (III)	NIG	organelle organization; autophagy
Dlg	41136-GD (III)	VDRC	Polarity.
	1725R-1 (III)	NIG	
E2F-IR15886	15886-GD (III)	VDRC	Cell cycle. Growth; proliferation; apoptosis.
	15887-GD (III)	VDRC	
fas	42237-GD (III)	VDRC	Morphogenesis of an epithelium; polarity of embryonic epithelium; dorsal closure.
if (aPS2) Integrin	44885-GD	VDRC	Cell migration. regulation of glucose metabolic process;
lgl-IR	51247-GD (III)	VDRC	Polarity. Cell cycle
	51249-GD (III)	VDRC	
Mad	12399R-1 (III)	NIG	Cell cycle. Dorsal closure
	12399R-2 (III)	NIG	
Merlin	7161-GD (III)	VDRC	Proliferation
mew (aPS1) Integrin	1771R-1 (III)	NIG	cell migration irrespective of tissue type
Mnt	10971-GD (III)	VDRC	regulation of glucose metabolic process; growth; cell cycle;
	2856R-2 (X)	NIG	
Moesin	10701R-1 (III)	NIG	Differentiation. cell polarity;
	10701R-3 (X)	NIG	
Myospheroid (bPS)	1560R-1 (III)	NIG	Protrusion-matrix adhesion; proliferation. cell adhesion;

NF1	35877-GD (III)	VDRC	response to stress; growth; metabolic process;
	8318R-1 (III)	NIG	
p53	10692-GD (III)	VDRC	Response to DNA damage. Apoptosis
	14873R-1 (III)	NIG	
PAK3	39843-GD (III)	VDRC	Actin regulation.
	14895R-1 (III)	NIG	
pTEN	35731-GD (III)	VDRC	cell cycle; extracellular matrix organization;
	5671R-1 (III)	NIG	
Puc	3019-GD (X)	VDRC	Actin regulation; apoptosis. cell cycle
Ral	2849R-1 (III)	NIG	Actin regulation. JNK cascade; Notch signalling pathway; Ras protein signal transduction; dorsal closure;
Rbf1	10696-GD (III)	VDRC	cell proliferation; response to DNA damage; cell cycle, Apoptosis
salvador	13831R-2 (III)	NIG	Growth. regulation of cell death;
Scrib	5462R-2 (III)	NIG	Polarity. cell-cell junction organization; polarity;
	45556-GD (III)	VDRC	
SLIK	4527R-2 (III)	NIG	establishment of mitotic spindle orientation; actin cytoskeleton organization;
	4527R-3 (III)	NIG	
TIMP1	6281R-2 (III)	NIG	Wing disc morphogenesis; delamination.
tkv	14026R-1 (III)	NIG	Cell fate determination; cell-cell adhesion; organ growth.
	14026R-3 (III)	NIG	
TOR	5092R-1 (III)	NIG	Growth. cell fate determination;
TSC1	22252-GD (III)	VDRC	growth; epithelium development; autophagy;
	6147R-1 (III)	NIG	
upd-IR3282	5993R-2 (III)	NIG	Jak/STAT.
WASp	13757-GD (III)	VDRC	Actin regulation.
	1520R-2 (III)	nIG	
α-catenin	20123-GD (III)	VDRC	cell adhesion; AJ organization; epithelium development; establishment or maintenance of cell polarity
Kirre	6695-GD (III)	VDRC	Negative control – muscle precursor cells, embryo
	6696-GD (III)	VDRC	
Otu	47431-GD (III)	VDRC	Negative control – expressed in female germline only
	12743R-3	NIG	

Appendix C: Screen genes

Gene	RNAi	Source	Main GO Terms
18w	8896R3 (III)	NIG	immune response; cell adhesion
	30498 (III)	Bloomington	
Ac78C	55312 (III)	Bloomington	Response to stimulus in sensory perception
Actn	4376R1 (III)	NIG	actin regulation
	7760 (III)	VDRC	
Adar	7763 (III)	VDRC	locomotor behaviour
	28311 (III)	Bloomington	
Ago	15010R-3 (III)	NIG	cell cycle; response to hypoxia;
Akt1	2902 (III)	VDRC	cell death
	4006R-3 (III)	NIG	
Alk	11446 (X)	VDRC	neurogenesis
alpha-spec	31209 (III)	Bloomington	cell adhesion; neurogenesis
Ank	25945 (III)	VDRC	signal transduction; cytoskeletal regulation
Ank2	26121 (III)	VDRC	structural constituent of cytoskeleton; cytoskeletal protein binding
	5115 (III)	VDRC	
Apoltp	51937 (III)	Bloomington	regulation of lipid transport
Arm	11579R-1 (III)	NIG	cell adhesion; cell morphogenesis;
Asp	2910 (III)	VDRC	actin regulation; establishment of mitotic spindle
	28741 (III)	Bloomington	
AstC-R2	50000 (III)	VDRC	G-protein coupled receptor signalling pathway.
	25940 (III)	Bloomington	
Asx	31192 (III)	Bloomington	mitotic cell cycle, histone modifications
ATP7	31083 (III)	Bloomington	ion transport and homeostasis
ATPsyn-beta	11154R-1 (III)	NIG	regulation of glucose metabolic process
	37812 (X)	VDRC	
Atx-1	21870 (III)	VDRC	Photoreceptor cell maintenance.
Aub	30124 (III)	VDRC	regulation of nuclear-transcribed mRNA catabolic process
babo	3825 (III)	VDRC	metabolic process and signalling
	853 (III)	VDRC	
Bap170	3274R-2 (III)	NIG	regulation of transcription
	34582 (III)	VDRC	
Bbg	15974 (III)	VDRC	cell migration; innate immune response
Bdl	31974 (III)	Bloomington	cell adhesion
Bel	6299 (III)	VDRC	mitotic sister chromatid segregation; autophagy
bent	46252 (III)	VDRC	filament assembly, sarcomere organization
Blimp-1	5249R-2 (III)	NIG	regulation of transcription
	34978 (III)	VDRC	
bmm	37880 (III)	VDRC	lipid homeostasis
Bnk	1480R-3 (III)	VDRC	actin filament organization
Brm	37720 (III)	VDRC	transport; metabolic process
	37721 (III)	VDRC	

Bro	13679 (X)	VDRC	regulation of transcription
brv2	31670 (III)	Bloomington	calcium ion transport,
Btl	32134R-1 (III)	NIG	epithelium migration, metabolic process
	27108 (X)	VDRC	
Bun	28322 (III)	Bloomington	Apoptosis
Bx	29454 (III)	Bloomington	morphogenesis of an epithelium
Ca-alpa1T	26251 (III)	Bloomington	Calcium ion transport.
ca-alpha1D	51491 (III)	VDRC	Calcium ion transport.
	52644 (III)	VDRC	
Cac	48093 (X)	VDRC	calcium ion transport
Calx	28306 (III)	Bloomington	Response to stress, transport
calypso	47743 (III)	VDRC	Histone modifications for ubiquitination
	28904 (X)	VDRC	
Cap	39207 (III)	VDRC	Chromatid cohesion, DNA repair, neurogenesis
Capt	5061R-4 (III)	VDRC	regulation of cell morphogenesis
	21995 (III)	VDRC	
Capu	32922 (III)	Bloomington	actin filament organization
CcapR	14767 (III)	VDRC	G-protein coupled receptor signalling pathway
	31490 (III)	Bloomington	
cdep	2008R3 (III)	NIG	regulation of Rho and Race signal transduction
	31168 (III)	Bloomington	
CG10186	30767 (III)	VDRC	Sensory perception of pain.
	10186R-1 (III)	NIG	
CG10211	12352 (III)	VDRC	response to oxidative stress
CG10274	10274R-2 (X)	VDRC	Unknown
	52407 (III)	VDRC	
CG10348	10348R1 (III)	NIG	Sensory perception of pain.
CG10585	31272 (III)	VDRC	Isoprenoid biosynthetic process.
	31273 (III)	VDRC	
CG10600	10600R-2 (III)	NIG	Unknown
	31276 (III)	VDRC	
CG10621	31291 (III)	VDRC	Unknown
CG10623	10623R-3 (III)	NIG	Unknown
	10623R-4 (X)	NIG	
CG10669	10669R-1 (III)	NIG	Unknown
	10669R-4 (III)	NIG	
CG10702	3691 (III)	VDRC	transmembrane receptor protein tyrosine kinase signalling pathway
	27052 (III)	VDRC	
CG10916	31379 (III)	VDRC	intracellular signal transduction, protein ubiquitination
CG10931	52107 (X)	VDRC	Histone modifications
CG11180	28629 (III)	Bloomington	Neurogenesis.
CG11247	52116 (III)	VDRC	Neurogenesis.
CG11318	11318R1 (III)	NIG	G-protein coupled receptor signalling pathway.
	51792 (III)	Bloomington	
CG11444	11444R-1 (III)	VDRC	Unknown
	31501 (III)	VDRC	

Gene	Line	Source	Function
CG11593	11593R2 (III)	NIG	Unknown
CG11836	43308 (III)	Bloomington	Proteolysis.
CG12268	12268R1 (III)	NIG	Unknown
	1166 (III)	VDRC	
CG12484	12484R1 (III)	NIG	lateral inhibition
	12484R2 (III)	NIG	
CG12950	12950R1 (III)	NIG	Unknown
	10011 (III)	VDRC	
CG1311	12242 (III)	VDRC	Unknown
CG13185	18782 (III)	VDRC	Ribosomal large subunit assembly; cellular response to starvation.
CG13579	13579R1 (III)	NIG	G-protein coupled receptor signalling pathway.
	9366 (III)	VDRC	
CG13597	13597R-4 (X)	VDRC	mRNA splicing
	32225 (III)	VDRC	
CG14446	14446R2 (III)	NIG	Unknown
	32487 (III)	Bloomington	
CG14838	14838R3 (III)	NIG	Unknown
CG15117	47530 (III)	VDRC	metabolic process;
CG15269	51506 (III)	Bloomington	Unknown
CG15395	15395R-3 (III)	VDRC	Unknown
	39929 (III)	VDRC	
CG15556	1791 (III)	VDRC	G-protein coupled receptor signalling pathway
CG15744	28516 (III)	Bloomington	G-protein coupled receptor signalling pathway
CG16721	25201 (III)	VDRC	CAMP-mediated signalling.
CG16868	16868R1 (III)	NIG	Unknown
	29617 (III)	Bloomington	
CG17002	17002R-2 (III)	VDRC	Unknown
CG1724	30315 (III)	VDRC	Protein transport; protein targeting to mitochondrion.
	30317 (III)	VDRC	
CG17739	36542 (III)	VDRC	Unknown
CG18335	18335R-1 (III)	VDRC	Unknown
	52471 (III)	VDRC	
CG18476	6527 (III)	VDRC	Unknown
	17822R-1 (III)	VDRC	
CG2135	16625 (III)	VDRC	Carbohydrate metabolic process.
CG25C	28369 (III)	VDRC	Morphogenesis, dorsal closure
CG30268	42705 (III)	VDRC	Oxidation-reduction process.
CG30280	34743 (III)	VDRC	G-protein coupled receptor signalling pathway.
	34744 (III)	VDRC	
CG30417	25619 (III)	VDRC	Unknown
	48182 (X)	VDRC	
CG31626	1975R-2 (III)	VDRC	Unknown

	11082 (X)	VDRC	
CG31731	31731R-2 (III)	VDRC	Transport.
CG31999	31999R1 (III)	NIG	Cell adhesion.
	31587 (III)	Bloomington	
CG32066	44825 (III)	VDRC	Unknown
CG32238	29424 (III)	VDRC	Protein polyglutamylation.
CG32333	26118 (III)	VDRC	Sensory perception of pain.
CG32373	32373R-1 (III)	VDRC	synaptic target recognition
CG33288	51189 (X)	VDRC	Unknown
	51190 (III)	VDRC	
CG3339	41917 (III)	VDRC	microtubule-based movement
	41918 (III)	VDRC	
CG34113	10989R2 (III)	NIG	Unknown
	22742 (III)	VDRC	
CG34120	11673 (III)	VDRC	Transport.
	1819R-1 (III)	VDRC	
CG34380	39446 (III)	VDRC	tyrosine kinase receptor signalling pathway
CG34422	40440 (III)	VDRC	phagocytosis
CG4096	44522 (III)	Bloomington	Epidermal growth factor receptor signalling pathway; proteolysis.
CG42255	32092R-3 (III)	VDRC	Unknown
CG42342	28648 (III)	Bloomington	Unknown
CG42672	31737 (III)	Bloomington	Unknown
CG42817	3849 (III)	VDRC	Unknown
	25882 (III)	Bloomington	
CG4360	4360R2 (III)	NIG	Unknown
	51813 (III)	Bloomington	
CG4393	4393R1 (III)	NIG	Unknown
	4393R4 (III)	NIG	
CG4438	50033 (III)	VDRC	Unknown
CG4587	4587R2 (III)	NIG	sensory perception of pain, voltage-gated calcium channel activity
	3839 (III)	VDRC	
CG4887	4887R-3 (III)	VDRC	sensory perception of pain, mRNA splicing,
	21970 (III)	VDRC	
CG4896	26652 (III)	VDRC	mRNA splicing
	48197 (III)	VDRC	
CG5037	42787 (III)	VDRC	Heme biosynthetic process.
	31169 (III)	Bloomington	
CG5155	31568 (III)	Bloomington	Sperm individualization.
CG5245	27381 (X)	VDRC	Unknown
CG5389	5389R-3 (X)	VDRC	ATP synthesis regulation of proton transport.
	22111 (III)	VDRC	
CG5550	31000 (III)	VDRC	Unknown
	31001 (III)	VDRC	

CG5645	9289 (X)	VDRC	Unknown
CG5984	52489 (X)	VDRC	Unknown
CG6153	22257 (III)	VDRC	Cellular response to DNA damage stimulus.
CG6272	34156 (III)	VDRC	neurogenesis; regulation of transcription,
	34158 (III)	VDRC	
CG6480	23447 (III)	VDRC	Histone modification, chromatin silencing.
CG7229	37671 (III)	VDRC	spermatogenesis
CG7379	27988 (III)	VDRC	Chromatin modifications
	27989 (III)	VDRC	
CG7394	9210 (III)	VDRC	neurogenesis; protein import into mitochondrial matrix
CG7420	46316 (III)	VDRC	negative regulation of protein secretion
	7420R-1 (III)	VDRC	
CG7607	7607R1 (III)	NIG	Unknown
CG7708	28613 (III)	Bloomington	Unknown
CG7911	23075 (III)	VDRC	chromatin remodelling;
CG8027	8027R-3 (III)	VDRC	Unknown
CG9095	23159 (III)	VDRC	Unknown
CG9279	45052 (III)	VDRC	microtubule-based movement;
CG9492	51725 (III)	Bloomington	microtubule-based movement
CG9572	44802 (III)	VDRC	Unknown
CG9701	9701R-2 (III)	VDRC	Carbohydrate metabolic process.
CG9784	9784R-1 (III)	VDRC	dephosphorylating;
	30098 (III)	VDRC	
Chd3	13636 (III)	VDRC	Chromatin/DNA binding;
cheer	3937R2 (III)	NIG	axon development; mitotic cell cycle
CIC-c	6466 (III)	VDRC	transmembrane transport.
	27034 (III)	Bloomington	
Cirl	8639R1 (III)	NIG	G-protein coupled receptor signalling pathway.
	34821 (III)	Bloomington	
Cnc	17894R-3 (X)	NIG	response to stress; metabolic process, nervous system development
	37674 (III)	VDRC	
Cnn	4832R1 (III)	NIG	nervous system development, chromosome segregation
Cont	40613 (III)	VDRC	septate junction assembly; axon ensheathment; cell adhesion;
	28923 (III)	Bloomington	
Cora	9788 (III)	VDRC	embryonic morphogenesis; cell polarity;
Crb	6383R2 (III)	NIG	Polarity; Notch signalling pathway.
	40869 (III)	Bloomington	
CSN8	50566 (X)	VDRC	Mitotic G2 DNA damage checkpoint; cell cycle.
CTCF	30713 (III)	VDRC	regulation of transcription,
Cubn	14613 (III)	VDRC	renal protein absorption;
	14614 (III)	VDRC	
cv-c	14847R2 (X)	NIG	actin cytoskeleton organization; Rho protein signal transduction;
	32329 (III)	VDRC	

CycD	29024 (III)	VDRC	JAK-STAT cascade; regulation of cell growth
Cype	33878 (III)	Bloomington	sensory perception of pain; mitochondrial electron transport,
Ddr	55906 (III)	Bloomington	transmembrane receptor protein tyrosine kinase signalling pathway;
Dgk	29459 (III)	Bloomington	G-protein coupled receptor signalling pathway
	36745 (III)	Bloomington	
Dhc36C	27451 (III)	VDRC	microtubule-based movement; glycerol-3-phosphate metabolic process
	5526R-1 (III)	NIG	
Dhc62B	15804R-1 (III)	NIG	Microtubule movement.
Dhc93AB	41947 (III)	VDRC	microtubule-based movement;
	3723R-2 (III)	NIG	
Dia	1768R-1 (III)	NIG	actin cytoskeleton reorganization; cell cycle;
Dis3	6413R-2 (III)	NIG	regulation of gene expression;
	35090 (III)	VDRC	
Dlt	41876 (III)	VDRC	lateral inhibition; morphogenesis of an epithelium; cell proliferation
DNApol-epsilon	6768R-2 (III)	NIG	cell cycle; response to DNA damage
Dob	14814 (III)	VDRC	Triglyceride catabolic process.
Dom	7787 (III)	VDRC	regulation of immune system process; metabolic process;
	41674 (III)	Bloomington	
Dp	7836 (III)	VDRC	regulation of transcription, lateral inhibition;
	15637R-2 (III)	NIG	
Ds	4313 (III)	VDRC	cell adhesion; polarity
Dscam1	25622 (III)	VDRC	axon guidance
	36233 (III)	VDRC	
Dscam2	1003 (III)	VDRC	cell adhesion
	41890 (X)	VDRC	
Dscam3	6685 (III)	VDRC	cell adhesion
	31190R-1 (III)	NIG	
Dscam4	25366 (III)	VDRC	cell adhesion
	42882 (III)	VDRC	
Dsor2	40026 (III)	VDRC	migration; mitotic G2 DNA damage checkpoint (cell cycle);
Dys	19230 (III)	VDRC	cell polarity;
E(z)	6502R-4 (III)	NIG	regulation of gene expression; cell cycle
	27646 (III)	VDRC	
Eaat1	3747R-2 (III)	NIG	L-glutamate transport;
Ebi	40862 (III)	VDRC	regulation of gene expression; neuron differentiation; response to oxidative stress;
	4063R-2 (III)	NIG	
Egfr	43267 (III)	VDRC	peripheral nervous system development; cell proliferation
	10079R-2 (III)	NIG	
Egg	12196R-1 (III)	NIG	regulation of DNA metabolic process;
	21172 (III)	VDRC	
Ehbp1	31143 (III)	Bloomington	Regulation of Notch signalling pathway.
eIF-2beta	9416 (III)	VDRC	translational initiation; response to stress;

Elongin-C	15302 (III)	VDRC	ubiquitin-dependent protein catabolic process; dendrite morphogenesis;
Emb	31353 (III)	Bloomington	Protein transport; regulation of response to hypoxia.
	34021 (III)	Bloomington	
Ent1	11907R2 (III)	NIG	Nucleoside transport.
Eph	1511R-1 (III)	NIG	regulation of glucose metabolic process; nervous system development;
	4771 (III)	VDRC	
Ephrin	1862R2 (III)	NIG	axon guidance;
	27039 (III)	Bloomington	
Exn	33373 (III)	Bloomington	cell-cell adhesion;
fd102C	11152R-2 (III)	NIG	regulation of transcription
Fkh	37062 (III)	VDRC	Apoptosis;
	49960 (III)	VDRC	
Fra	6557 (III)	VDRC	axon guidance; cell migration;
	29909 (III)	VDRC	
Frac	31577 (III)	Bloomington	axon target recognition
	31578 (III)	Bloomington	
Fred	46180 (III)	VDRC	cell-cell adhesion
	42621 (III)	Bloomington	
Ft	29566 (III)	Bloomington	adhesion;
	34970 (III)	Bloomington	
Fw	39575 (III)	VDRC	planar polarity; cell adhesion
Gbeta13F	10545R-2 (III)	NIG	polarity, G-protein coupled receptor signalling pathway; cell adhesion
	31257 (III)	VDRC	
Grk	17610R2 (III)	NIG	tissue migration;
	4331 (III)	VDRC	
Grn	33746 (III)	Bloomington	regulation of transcription; axon guidance;
	34578 (III)	Bloomington	
Gtp-bp	14877 (III)	VDRC	GTP binding
Gug	13687 (III)	VDRC	epithelium development;
Ham	15907R1 (III)	NIG	Regulation of transcription, sensory organ development; cell proliferation.
	32470 (III)	Bloomington	
Hbs	7449R1 (III)	NIG	cell-cell adhesion
Hig	13266 (X)	VDRC	cell adhesion; synaptic target recognition
	2040R-1 (III)	NIG	
Hiw	26998 (III)	VDRC	regulation of cellular component biogenesis;
	28031 (III)	Bloomington	
Hml	37005 (III)	VDRC	metabolic process;
	37006 (III)	VDRC	
How	10293R-1 (III)	NIG	cell development; cell migration; cell adhesion;
hppy	7097R-1 (III)	NIG	Apoptosis
Hsp83	7716 (III)	VDRC	metabolic process;
IA-2	11344R3 (III)	NIG	Protein dephosphorylating; regulation of secretion.
	7560 (III)	VDRC	
Idh	42915 (III)	VDRC	metabolic process.

	42916 (III)	VDRC	
Ih	35836 (III)	VDRC	circadian cycle; transport;
	29574 (III)	Bloomington	
IP3K2	12724R-2 (III)	NIG	autophagy cell death;
Ir8a	25813 (III)	Bloomington	transmission of nerve impulse; transport; calcium-mediated signaling.
Irp-1B (2)	30153 (III)	VDRC	regulation of translational initiation
Iz	27985 (III)	Bloomington	cell death; response to stress; regulation of transcription
Jbug	11605R2 (III)	NIG	planar polarity;
	39070 (III)	Bloomington	
Jim	11352R2 (III)	NIG	Regulation of chromatin silencing; dendrite morphogenesis.
	11352R3 (III)	NIG	
Jub	30806 (III)	bloomington	Hippo signalling; mitotic nuclear division.
Kirre	6695 (III)	VDRC	cell-cell adhesion
Kis	18326R-1 (III)	NIG	developmental growth;
kl-2	19181 (III)	VDRC	Cilium movement.
kl-3	53317 (III)	Bloomington	Cilium movement.
Klg	39515 (III)	VDRC	cell adhesion
	28746 (III)	Bloomington	
Klu	12296R-2 (III)	NIG	Regulation of Raps protein signal transduction; neurogenesis; proliferation.
	12296R-3 (III)	NIG	
Kon	10275R-1 (III)	NIG	epithelium development; filo podium assembly;
Krn	8056R4 (III)	NIG	epidermal growth factor receptor signalling pathway; cell proliferation; cell migration
Kst	33933 (III)	Bloomington	zonula adherents assembly; cell cycle
Kto	8491R-1 (III)	NIG	transcription; autophagy cell death; Wnt signalling pathway;
	23142 (III)	VDRC	
Kug	5098 (III)	VDRC	adhesion; polarity; actin cytoskeleton organization;
	7749R-1 (III)	NIG	
LanB1	23119 (III)	VDRC	tissue development;
	23121 (III)	VDRC	
Lap1	10255R1 (III)	NIG	Protein transport; cell polarity.
	27036 (III)	Bloomington	
Lar	36270 (X)	VDRC	adhesion; axon extension
Lea	5574R2 (III)	NIG	brain development; cell-cell adhesion;
	34589 (III)	Bloomington	
lectin-22C	15378R1 (III)	NIG	carbohydrate metabolic process
	15378R6 (III)	NIG	
Lid	42203 (III)	VDRC	JAK-STAT cascade; histone modifications
	42204 (III)	VDRC	
Lilli	13081 (III)	VDRC	lipid transport; transcription regulation
	8817R-4 (III)	NIG	
ln-R	991 (III)	VDRC	regulation of developmental growth; cell proliferation; response to stress;
Lpt	22171 (III)	VDRC	histone modification; transcription; phagocytosis;

lr25a	15627R2 (III)	NIG	Detection of chemical stimulus involved in sensory perception; transmission of nerve impulse; ion transport.
LRP1	8397 (III)	VDRC	calcium ion binding
	44579 (III)	Bloomington	
Mbc	51446 (III)	Bloomington	cytoskeleton organization;
Mbt	29379 (III)	Bloomington	cell-cell adhesion;
MED23	28361 (III)	VDRC	transcription
	28363 (III)	VDRC	
mei-38	51805 (III)	VDRC	mitotic spindle organization;
Mes-4	10837 (III)	VDRC	histone modifications
mesh	40940 (X)	VDRC	Cell-matrix adhesion; septate junction assembly.
	12074R-1 (III)	NIG	
mfas	37889 (III)	VDRC	axonogenesis;
Mgl	29324 (III)	Bloomington	regulation of endocytosis;
	33940 (III)	Bloomington	
mGluRA	11144R3 (III)	NIG	regulation of synaptic transmission, G-protein coupled receptor signalling pathway;
	1793 (III)	VDRC	
Mhc	17927R4 (III)	NIG	epithelial cell migration
	7164 (III)	VDRC	
Mkk4	26929 (III)	VDRC	neurogenesis; JNK cascade;
	9738R-1 (III)	NIG	
Mrtf	34503 (III)	VDRC	cell cycle; cellular response to hypoxia; regulation of transcription, cell migration; actin cytoskeleton organization;
Msk	38963 (III)	VDRC	cell development;
Msp-300	25906 (III)	VDRC	actin cytoskeleton organization;
	18252R-2 (III)	NIG	
Mtt	32376 (III)	Bloomington	G-protein coupled receptor signalling pathway
Mub	34870 (III)	Bloomington	mRNA splicing,
mus205	24472 (III)	VDRC	DNA replication; double-strand break repair via homologous recombination (DNA repair).
My	37710 (III)	VDRC	cell cycle; glucose metabolic process;
	9045R-1 (III)	VDRC	
Myd88	25399 (III)	VDRC	development; regulation metabolic process;
Mys	1560R-1 (III)	NIG	cell adhesion;
Na	3306 (III)	VDRC	circadian rhythm;
nAcrbeta-64	39421 (X)	VDRC	synaptic transmission, ion transport.
	31883 (III)	Bloomington	
Ndae1	4675R-1 (III)	NIG	transport
	3664 (III)	VDRC	
NetA	18657R-2 (III)	NIG	axon guidance; cell migration.
Net	17302 (III)	VDRC	cell cycle; lateral inhibition;
	19594 (X)	VDRC	
NGI2	28331 (III)	Bloomington	synaptic growth at neuromuscular junction

Nipped-A	2905R3 (III)	NIG	histone modifications; phagocytosis; regulation of transcription, Notch signaling pathway; cell cycle;
Nlg3	1062R1 (III)	NIG	histone modifications, phagocytosis; regulation of transcription, Notch signaling pathway; cell cycle;
Nmdar2	14794R1 (III)	NIG	sensory perception of touch; ion transport;
	3196 (III)	VDRC	
nompC	31512 (III)	Bloomington	sensory perception of touch; transport;
nopo	22013 (III)	VDRC	Cell death; mitotic cell cycle, protein ubiquitination.
Nos	6713R3 (III)	NIG	synapse assembly; DNA replication; nervous system development;
	50675 (III)	Bloomington	
Nrg	27201 (III)	VDRC	brain development; adhesion;
	27202 (III)	VDRC	
Nrm	979 (III)	VDRC	Synaptic target recognition; cell adhesion.
Nrx-1	36326 (III)	VDRC	synapse assembly; transport;
Nrx-IV	28715 (III)	Bloomington	cell junction assembly; adhesion;
Nub	6217 (III)	VDRC	dendrite morphogenesis; regulation of transcription
	43678 (X)	VDRC	
Nup62	44806 (III)	VDRC	phagocytosis.
	6251R-1 (III)	NIG	
Or33a	1461 (III)	VDRC	Sensory perception of smell.
	1462 (III)	VDRC	
Or59b	43808 (III)	VDRC	sensory perception of smell
	43809 (III)	VDRC	
Osa	7467R-1 (III)	NIG	Wnt signalling;
	7810 (III)	VDRC	
Otk	8967R4 (III)	NIG	cell adhesion
p120ctn	51729 (III)	Bloomington	Cell adhesion; regulation of actin cytoskeleton organization.
Pak	12553 (X)	VDRC	cell morphogenesis; adhesion; cell junction organization
Pan	3014 (III)	VDRC	morphogenesis of an epithelium; cell proliferation;
Para	6131 (III)	VDRC	mechanosensory behaviour;
	6132 (III)	VDRC	
Patj	31620 (III)	VDRC	homeostatic process;
	12021R-4 (III)	NIG	
Pdk2	31296 (III)	Bloomington	ion transport; spermatogenesis;
	51502 (III)	Bloomington	
pdm2	30708 (III)	VDRC	neurogenesis; regulation of transcription,
	52272 (III)	VDRC	
Pex13	39544 (III)	VDRC	unknown
	50697 (III)	Bloomington	
phl	31038 (III)	Bloomington	regulation of small GTPase mediated signal transduction; cell cycle;
Pi3K21B	33556 (III)	VDRC	cell proliferation;

|			
Pi3K92E	38985 (III)	VDRC	epithelium development; homeostatic process;
	4141R-1 (III)	NIG	
Piwi	6122R-1 (III)	NIG	cell fate determination;
Plc21C	4574R2 (III)	NIG	regulation of neuromuscular synaptic transmission;
	33719 (III)	Bloomington	
plexA	11081R2 (III)	NIG	axon guidance; sensory perception of pain;
plexB	6873 (III)	VDRC	axon guidance; regulation of Rho protein signal transduction
PMCA	31572 (III)	Bloomington	calcium ion transmembrane transport
Pp2A-29B	23886 (III)	VDRC	Transport; metabolic process; autophagy.
	49671 (III)	VDRC	
Pp2C1	40827 (III)	Bloomington	Protein dephosphorylating; oxidation-reduction process.
PpV	31690 (III)	VDRC	Protein dephosphorylating; mitotic cell cycle.
Prestin	5485R-1 (III)	NIG	transport
	5485R-3 (III)	NIG	
Ptp10D	1817R1 (III)	NIG	dephosphorylating; axon guidance; central nervous system development;
	1101 (III)	VDRC	
Ptp36E	34369 (III)	VDRC	protein dephosphorylating
	34368 (III)	VDRC	
Ptp4E	1012 (II)	VDRC	Protein dephosphorylating; axon guidance; CNS development.
	38369 (III)	Bloomington	
ptp52F	3116 (III)	VDRC	axon guidance;
Put	7904R-2 (III)	VDRC	cell cycle; dorsal closure; metabolic process
	37279 (III)	VDRC	
Pvr	8222R-2 (X)	VDRC	response to stress; actin cytoskeleton organization;
	13502 (III)	VDRC	
Pxn	15276 (III)	VDRC	Phagocytosis; response to oxidative stress; ECM organization.
	15277 (X)	VDRC	
Pyx	31297 (III)	Bloomington	ion transport;
	51836 (III)	Bloomington	
R	33437 (III)	VDRC	adhesion; cell-cell junction assembly; cell migration;
Rab40	1900R-1 (III)	VDRC	Rab protein signal transduction;
	29259 (III)	VDRC	
Rac1	49247 (X)	VDRC	cell proliferation; metabolic process; cell migration; cell adhesion;
Rac2	8556R-3 (III)	NIG	actin cytoskeleton organization; metabolic process; cell junction maintenance;
	50349 (III)	VDRC	
Ras85D	28129 (III)	VDRC	cell proliferation; mitotic cell cycle
Rdx	9924R-2 (X)	NIG	polarity; apoptosis
ref(2)P	10360R-1 (III)	NIG	Mitochondrion organization.
Rfabg	6879 (III)	VDRC	Wnt signaling pathway;
RhoGAP19D	43955 (X)	VDRC	Rho protein signal transduction;
	32361 (III)	Bloomington	
RhoGAPp190	28877 (III)	VDRC	Rho protein signal transduction;

RhoGEF3	31581 (III)	Bloomington	regulation of Rho protein signal transduction;
Ric	8418R-3 (III)	NIG	small GBase mediated signal transduction; response to oxidative stress;
	35930 (X)	VDRC	
Rim	7305R2 (III)	NIG	vesicle-mediated transport;
	39385 (III)	VDRC	
Rm62	10279R-2 (X)	NIG	mRNA splicing,
	46908 (III)	VDRC	
Robo	35768 (III)	Bloomington	cell-cell adhesion; epithelial cell migration,
robo3	29398 (III)	Bloomington	axon guidance
RpL5	38929 (III)	VDRC	ribosomal large subunit assembly; translation
Rst	951 (III)	VDRC	axon genesis; apoptosis; cell adhesion
	28672 (III)	Bloomington	
Run	28673 (III)	Bloomington	tissue development; regulation of transcription,
	34707 (III)	Bloomington	
RunxA	33353 (III)	Bloomington	dendrite morphogenesis; regulation of transcription,
S6k	10539R-2 (III)	NIG	Developmental growth; axon genesis; response to stress.
	18126 (X)	VDRC	
S6kII	17596R-1 (III)	NIG	circadian behaviour; metabolic process;
SA	3423R-2 (III)	NIG	neurogenesis; mitotic sister chromatid cohesion
	3423R-3 (III)	NIG	
Salm	33714 (III)	Bloomington	neuron differentiation; instar larval or pupal development; epithelial cell migration
Salr	4881R3 (III)	NIG	regulation of transcription, sensory organ development; neurogenesis;
Sax	1891R-1 (III)	NIG	response to stress; morphogenesis of an epithelium;
	42457 (III)	VDRC	
Scox	8885R2 (III)	NIG	regulation of transcription, sensory organ development; neurogenesis;
	7861 (III)	VDRC	
Sdk	33412 (III)	Bloomington	Eye development
Sdr	3837R1 (III)	NIG	Transmembrane receptor protein tyrosine kinase signalling pathway.
	44575 (III)	VDRC	
Sei	3606 (III)	VDRC	ion transport; sensory perception of sound;
sema-1a	18405R1 (III)	NIG	Organ morphogenesis; brain development.
Sema-1b	6446R3 (III)	NIG	axon guidance; embryonic morphogenesis
	28588 (III)	Bloomington	
Sema-2a	15810 (III)	VDRC	axon guidance; synaptic target inhibition;
Sema-2b	48057 (III)	VDRC	axon guidance; response to DNA damage stimulus;
	28932 (III)	Bloomington	
Sema-5c	5661R1 (III)	NIG	axon guidance; brain development
	1052 (X)	VDRC	
sens-2	21386 (III)	VDRC	nucleic acid binding;
Sep2	26413 (III)	VDRC	Cell cycle; regulation of glucose metabolic process.
Ser	27174 (X)	VDRC	morphogenesis of an epithelium;
Shn	34689 (III)	Bloomington	Cell morphogenesis; cell proliferation;

Shot	28336 (III)	Bloomington	cytoskeleton organization;
Sick	53315 (III)	Bloomington	axon genesis; actin filament organization;
Side	31062R2 (III)	NIG	Motor neuron axon guidance.
	1284 (III)	VDRC	
Sif	5406R3 (III)	NIG	actin cytoskeleton organization; regulation of Rho protein signal transduction;
Sirt4	40295 (X)	VDRC	histone modifications
SK	24597 (III)	VDRC	Regulation of membrane potential in photoreceptor cell.
	27238 (III)	Bloomington	
Sl	4200R-3 (III)	NIG	cell migration; signal transduction;
	7173 (III)	VDRC	
Slbo	27043 (III)	Bloomington	regulation of transcription, cell migration; regulation of glucose metabolic process; JAK-STAT cascade;
	53309 (III)	Bloomington	
Sli	31467 (III)	Bloomington	Adhesion; epithelium development; cell migration.
SMC1	6532 (III)	VDRC	regulation of gene expression; mitotic sister chromatid cohesion;
Smox	2262R-1 (III)	NIG	cell cycle; brain development; regulation of cellular metabolic process;
Smr	27068 (III)	Bloomington	transcription, mitotic cell cycle; glucose metabolic process; Notch signalling pathway;
	34087 (III)	Bloomington	
Snr1	12645 (III)	VDRC	cell proliferation; chromatin remodelling;
Sox15	45482 (III)	VDRC	regulation of transcription, apoptosis; cell proliferation;
spen	48846 (III)	VDRC	tyrosine kinase signalling pathway;
Spn38F	9334R-3 (III)	NIG	proteolysis;
Spp	7247 (III)	VDRC	cellular response to unfolded protein
Ss	10715 (III)	VDRC	metabolic process;
	33415 (III)	Bloomington	
Stj	25807 (III)	Bloomington	synaptic transmission; sensory perception of pain;
su(Hw)	34006 (III)	Bloomington	negative regulation of transcription,
Su(var)2-HP2	25972 (III)	Bloomington	neurogenesis; cell cycle;
	31346 (III)	Bloomington	
Syx1A	5448R-4 (III)	NIG	Metabolic processes
	33112 (III)	VDRC	
Taf1	41099 (III)	VDRC	cell cycle; histone modifications
	32421 (III)	Bloomington	
Tefu	6535R-1 (III)	NIG	mitotic G2 DNA damage checkpoint (cell cycle); regulation of immune response; regulation of transcription,
	22502 (III)	VDRC	
Tehao	17903 (III)	VDRC	Innate immune response; signal transduction.
	7121R-3 (III)	NIG	
Ten-a	11270R1 (III)	NIG	adhesion; synaptic growth at neuromuscular

	32482 (III)	VDRC	junction;
Tep1	30873 (III)	VDRC	Humoral response.
	32856 (III)	Bloomington	
Tep2	18589R1 (III)	NIG	humoral response; phagocytosis;
Tep4	13466 (III)	VDRC	humoral response
Tep5	31866 (III)	VDRC	Metabolic processes
Tg	7356R1 (III)	NIG	innate immune response;
	26100 (III)	VDRC	
Tim17B1	9455 (III)	VDRC	Protein import into mitochondrion.
Tim17B2	3838 (III)	VDRC	protein import into mitochondrion;
Tkv	14026R-1 (III)	NIG	embryonic morphogenesis;
	14026R-3 (III)	NIG	
Tol	2656 (III)	VDRC	axon guidance;
Toll-7	30488 (III)	Bloomington	Axon guidance; autophagy.
Tollo	9430 (III)	VDRC	innate immune response; regulation of glucose metabolic process;
	27098 (III)	VDRC	
Traf-like	34836 (III)	VDRC	signal transduction; defence response
Traf4	3048R-3 (III)	NIG	sister chromatid cohesion;
Trax	49383 (III)	VDRC	Metabolic processes
Trf4-1	41096 (III)	VDRC	apical constriction; adherents junction organization; dorsal closure;
Trf4-2	19801 (III)	VDRC	Sister chromatid cohesion.
	19799 (III)	VDRC	
Trim9	21405 (III)	VDRC	axon guidance
Trio	18214R1 (III)	NIG	actin cytoskeleton reorganization
	40137 (III)	VDRC	
Troll	22642 (X)	VDRC	Maintenance of epithelial cell apical/basal polarity; axon guidance.
	29440 (III)	Bloomington	
Trp1	8894 (III)	VDRC	posttranslational protein targeting to membrane,
	8895 (III)	VDRC	
Trr	10749 (X)	VDRC	hippo signalling; regulation of transcription,
Trx	31092 (III)	Bloomington	Regulation of response to DNA damage stimulus; cell migration; axon guidance.
Trx-2	36298 (III)	VDRC	Disulphide oxidoreductase activity.
TTLL3B	31456 (X)	VDRC	sperm individualization; protein polyglycylation
tweek	19306 (III)	VDRC	endocytosis
	26645 (III)	VDRC	
U2af38	3582R-2 (III)	NIG	Cellular homeostasis; cell fate determination
Uif	38365 (III)	Bloomington	Cellular homeostasis; Notch signalling pathway.
unc-13	2999R3 (III)	NIG	synaptic vesicle exocytosis; neurotransmitter secretion;
	33606 (III)	VDRC	
unc-5	8166R1 (III)	NIG	cell migration; axon guidance;
	8137 (III)	VDRC	
unc79	51471 (III)	Bloomington	locomotor behaviour
Unc-89	29412 (III)	VDRC	regulation of Rho protein signal transduction; sarcomere organization;
	29413 (III)	VDRC	

Usp	4380R-3 (III)	NIG	glucose metabolic process; response to starvation;
	16893 (III)	VDRC	
Utx	5640R-2 (III)	NIG	metabolic process; signal transduction;
Vap	44638 (III)	VDRC	Endocytosis; regulation of ATPase activity.
Vav	39059 (III)	Bloomington	cell morphogenesis; cell junction organization; adhesion;
Vhl	50727 (III)	Bloomington	cell migration; mitotic cell cycle; morphogenesis of an epithelium;
Vkg	16858R-3 (III)	NIG	Skeletal muscle tissue development.
	16986 (III)	NIG	
Vn	10491R2 (III)	NIG	brain development; notum development; apoptosis
	50358 (III)	VDRC	
Wb	3141 (III)	VDRC	Imaginal disc-derived wing morphogenesis.
	15288R-3 (III)	NIG	
Wit	10776R-2 (III)	NIG	signal transduction; synaptic growth
	42244 (III)	VDRC	
Wnt2	38077 (III)	VDRC	Axon extension; Wnt signalling pathway.
wrapper	10382R3 (III)	NIG	apoptotic process; axon development;
	29561 (III)	Bloomington	
Wry	8021 (III)	VDRC	Notch signalling pathway.
X11L-beta	1852R-3 (III)	NIG	protein localization
Xpd	41021 (X)	VDRC	regulation of chromosome segregation; cell division;
zfh2	13305 (III)	VDRC	wing disc development; CNS development; regulation of transcription,
Zir	28005 (III)	Bloomington	Phagocytosis; regulation of Rho protein signal transduction.
zormin	17563 (III)	VDRC	sensory perception of pain

Appendix D: Expression levels of ING1 and RSBN1L in MDA-MB231 and U87 cells
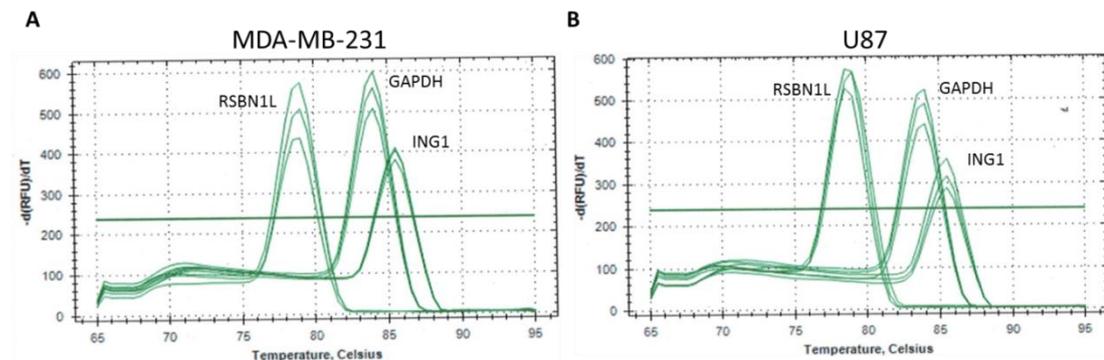
RT-PCR on cells harvested from **(A)** MDA-MB-231 and **(B)** U87 cell lines using RSBN1L, GAPDH and ING1 specific primers to check expression levels.

Table D1: CT values

	U87	MDA-MB-231
ING1	27.36	25.7
RSBN1L	29.59	28.57
GAPDH	18.22	19.26

Appendix E: CG10600 sequence

CRISPR target region highlighted.

```
CACATCTGAAGACTAGAGTTGAAAAACGACAGTGCGAGACGGCTTCTAGAACAAAAAAAAAAACTAT
TCGCGAATCGCTGAAAACGCCGCTGCGGAAATTCTGCCAGTCGAATGAAAATCGAGTGGAAGAGGG
CGCGGGTGAATTGCAAACGGAGCGTAAGATAAAAGACACGGCTAGCTGAGAGGCGTCACGTAGACG
GCCGACTGATCGATATTCAATACAGGCGTAGGGCAACACGCACACACACCCATACGCGCGTGTGTGT
GCCACTGCACACACAAACACATGGAGAACACTCGAGGTGGCGAAAATAATGTAGAATTGGCCAACG
GCGGCAAGGAATCCACGGGAAATGTACGAAAACCGGAGAAGAAGCCACTGGAAGGAGCGGTCGGG
GCTTCCAAGCAGCGGGAGACGCCTGAAAAGGGCGTAAACAAAGCGTCGGCAGTGGAAAACAACAAC
GCAAATGCCCGCGGAAATGCCAGTCCAACAACTACAAGCGCACAGCGAGAAGCAACAACAACAATTA
CGCCAGTGGAACCAACAGCAGCAGCAACAACTACAAACGGCAACACTTTACCCACACAACAGCAACA
ACAACTGCAGTCAGGCGATGGCAACGCAAATAACCCAACAACAAATGCACGTATAATGCTAGCTGCA
GCTGAAGAGGAACAGGCAGTAGCATCGTCGGTGGCAACAACTGCAGGGGCACCCACAGCCGCAACA
ACAACTGCAACAACAACCGCGTTGGCAGACACAACAAAAACAAGCACAAGTACGGACGAGAAGACA
GGTAAAGTAAAGCGAAATTCGCAGAACACGAATCCAATGTCAACTAATGTGATTTGCACACCCACCC
ACTCGCCAACGCCGCCCACAGAAGAGCGCAACCACAAAAGGAAGAAGAGAGACAGGGAGCGGGAT
GAAGACAAGAATCGGCGCAGCAACAACCACAACAGCAGCAGCAAGGATCGGGACAAGGATAGGGA
TCGGGACAAAGAGCGGGACAAGGAGCGCG==ATAGGGAGCGGGAGCGCGAC==AGGAATCGCGAGAGC
AGCTCCAAATCCTCCAAGGAGGCCGAACACAAGTCCAGCTCATCGCGCAGCAAGGAGAAGTATCATC
GCAGTGATAGGGACCGCGAACGGGAGAAGGAGCGCGATCGGGAGAAGGTGAGCCCAGGAACCGTT
TCGAAAGTGCAGAACGCATACAGAAACGTTGATTTATTGTAAACTCTGATCTGATTTCCAGAACATAC
ATTTCTTTAATAAGTAAATGTCAATTTCTAGTAAAACTAACAAACAAAGTATTGTCAGGGTCTTACCA
CCCATTTACTACGAATAAGCAAGACACGTGATCTTAAATACAGCGAGTCATAATTCTCAGAACCTTTTA
TTTTTGATGCGCTACATGTTTATAGTATACTTTCCAGTTTACATAGTGCTAAGTAAACCTCTACAAGATT
GCACAGCTGTCCATCAAAGTAGAAATAATGAAGCACTACAAATTTATATTTTCCAGATTCCTTCCTCGT
CGTCCAGCTCATCCTCGCACTCGAATCGAAGGAGAAGTTCCGACAGCAGCCGGAGCAGTCAGAACAG
CTCCTTCAAGTCGAAGGATCCAGGGGCAGATCCTGTTTTGCCTCAGGCGGTTGTTGCTGCTCTCGAAA
AGCAGATTAAGGTGGAGGAGGTCATCAAAAAAGAACCAGTAGAAACCAAACAAGAAGTGGATACAG
AACCTACACCAGTACTTAGTAACCTCATTAAAACTGAAGCGGAAATCAAAACGGAATCGACGGAGAA
GCCAAACAAACCTCTAGTAAATGGCGAGGCGATAGCGAAAACCCGAGATGAGGTTTCCAGACAGCT
AAATTTCGGAGACAAGGTCTCCAATTCCAAGTTAATGCCTTCGCCGGCACCGAAATCCAGTTCGGGAT
CGAACTCCACGGCCAGCAATCATTCCTCCAGCCAAAGTAGCAGCAGCCGCAAATCCTACTTATCGTCT
TCGTCCTCTTCATCCTCCAGCAACCATCACAAGAGTTCCTCTTCGTCATCGTCTCGCCACTCCTCATCTA
CATCTCGGGAGTGCTCCAAGTGCTACAAGCGCTCCAAAATTCGGCGAACCAGCGTGGGCGTCCAGTG
CTTACAGCATGCACCAGCAACTGGACCCTGGCAAACCGTTCAGAGCCTGCCGCCTAGGCCTAACCGA
AAACCACCGGCTGGCTTGGAGAACTTAAAGTACGGCTGCTACTTTCAGGTGGAGCAATATCCAAATG
GCGGGGCCTCCATTGTTCACCTGTATCAACATGAAATCGACGCATTGTCGCCAGATGAAATGGAGGA
GCTGGTGGATGAGTTCTTTGAGGTGTGCTTTGCCGAGGATGAGCAAGGGTATGCCCACCATGTGATG
GGTATTGTCCACGATGCCGCGAGGTATTTACCCGATCTACTCGAGCACATGGCCGAGAACTACTCAAC
GCTGACGGTAAAGGCGGGAGTTCTTGGACGAAACTCCGATATCGAGACCTGTACAATGTCGCAATAC
AACGAGCAGGTAAATCCATATATAAAGCAGTGCTAGCAAGAAAAATATGTAACAATTCTTGTTCGTTT
CTCACCTTTAGGTGGTTCGCAACTATTGCCAAGGAACCTTTCGATATGGACCACTGCATCAGATCAGT
TTGGTTGGCAAAGTGCATGAAGAGGTGGGCGGATACTTCCCAGATCTGTTGGGTCGTATTGAAATGA
ATCCATTTTTACGAAAGGTAAGCACACTCAAATGCTAAAAACACTTGACTAATTTAACTCTCTTATCTC
AGACCATGCCTTGGGCTTGTAATTCTATACTCCAAACGGATCCTCGTCAGTCGAACGATGGCCCCATT
CTCTGGATCCGTGCTGGGGAGCAGTTGGTCCCCACAGCAGAGTTGAATAGCAAGACCCCACTGAAAA
GGCAAAGGTAAGTAGCATTTGGAGCAATATTCTTAAAAACCATTAATTAACTTACACAATTCACATGC
AGGACTCGTATTAATGAATTGAGGAACCTTCAATATTTGCCGCGTTTGTCTGAGGCACGGGAGACGA
```

```
TGATCGAGGATCGGACCAAGGCGCATGCGGATCACGTGGGCCATGGCCACGAACGCATCACCACCG
CAGCAGTTGGTGAGATGCCTCTCATTGAAAAGCTGCTCCGAAATCGACTAATCGGCGTAATGTTTCTA
TTATCCGTCCTGCAGGCATACTGAAAGCTGTACACTGTGGTCAAACTTACTCGCAGAACCGCATCACC
AAAGACGTGGTCGCTTTTGCCGCCCAGGACTTTAACACAATGGTGGAAATGCTGCAGCTGGACCTGC
ACGAGCCGCCAATCTCACAGTGCGTACAGTGGATCGAGGACGCCAAACTTAACCAGCTGCGCCGCGA
GGGCGTCCGTTACGCTCGTATTCAGCTGTGCGACAATGACATATACTTCTTACCCCGCAACATAATCCA
CCAGTTTCGCACAGTGACAGCGGTAACGAGCGTAGGTAAGGGATCGAATGCAATTTTTCTCTTCATCG
TTCGATTAACCTACAAATTTTTACTCTGCAGCTTGGCACCTGCGACTAAGACAGTATTATCCCGGCCAG
GAAGTGATTAACGAGAAGAATAACCCTGTGCTAGCCGAAACACCTCACTACAAGGAAAAGCAAACGA
TCCTGCCAAACCCAATTAGCCACGACGAATGCGGCAAAAAGACACCATCAAAGCGTGCACACGATGG
CAGGACCAAGAAGAAACTTATTGACCTCGATGGCAAGGAGCGTCGATCTAGTGAGAGCAGCATAGA
CGAGCACCCAGCTGCAGGAAGTTCCTCGCCTAGCCAAAATCAGGAAAACAACAGCAACAGCAGCAGT
CAGAGCGGAGCTAGTGGGGTTAGCACACCCAAAAAGAAGAGCAGCAAGGAGGACTCCAAGATAGA
TATGCGGAAAATGGTTCTGGAGCACAAATATAAGCTGACCAAAGCAGCGGTAACTGAAACCCCTGAG
AAGGAGAAGCTCAAAAAGGATAAGGACAAGGAAAAGTCTAGGGGGAAGGATAGAGATAAGGAAA
AGGAACGAGGCAAGGAAAAGGAAAAGGAAAAGAAAAGAAAAACGATACTCCTAAAAAACATGCA
ACGCCAGCAAAGAGCTCTTTGAGTGAGCCCTCGCCCGCCAAAATTCCCAAATTAAATAATGACACAGC
TCCTGCAACTGCACTTCCACTGAAATCTACACCCTCCTCTGCGCCACCTACGTGCACTCTTCCTGTAGCT
GCGCCGCCTATCCTGTCAGGACTGCCACCGCCACTCGTGCCTCAGACGCCTTTCATTGATCGTCTAAAT
CAAAAGGAGCCCGAGATCAAGATCGAAGTAAAGATTGTCTCCGACAGTAAAAGTGATATCAGTACCT
CCAATCCAGTTGTGTCAACAACAGCCCCACCACCGCCGCCACCTCCGGATCCGGTTGAAGATGTGGGT
AACGAGCTAATTGTGGAGAGCACACCCCAATTGATCGTTGACCACGAAGAGGAGGTCTCTGAAGTGT
GTTCCGAGGAGATTGTCGAAATGGATATACCGATACCCAAAGAACCAGTCCCAGCTGCTGCTTTTACA
TCATTGCCTAATACTCCGTCACCACTTATTAGTGCTCACTCCACTCTGGCTACTGCGCCCATACCAAAAT
CTTTTCTCCATCCACCAGTACCAATTGTAGCTGCACCCGCTCCTCCGCCACCCAGTCAAATTGTGAAAG
TGGTACTCCCAGGCTCTTTGACGCCTCTTACCAGAATGTCAGTGGTTTTGCCACCCTTGCCGCCATTAC
CACCATCACGGACATCTGCACCGCCACCGCCGCCGCCTCCTCCGAGCACCAAAAGCAGTGCTATCACA
ATGCCCATATCAGCCCCAGTTTCCACGTTGAAACCCACCTACAAAACCATTAATTTGCCCTCCCACAAG
ATTATCATCGCAGGCTCCGCGGCGGGAAGTCGAGTCTCGATGGCAAAGGGCACGTCGAAGAGACCT
GATTTATTGGGATCCATCATAGCCAGCATGGACAAGCCAGCGGGAAGTGGACCCAATAGCAACTCCT
CATCATCGACATCCTCGGCAACAACGAATAGTTCGCTCTCCTCTTACGCCGGTTCCAACAATAGCTTTT
GAAACTGGGAAAAGAAATGGACGGAAAGCGAGAGTAGAGGCTCCTATTTGGACTTTATGTGAAACT
ATATATTTTGTAAATTTAAAAGCACTCGTTATTACGTTAAACGGATTTCGAGTAGCTGAATCCTTTAAA
GGAAGATCTCAACTCTGAAAGGGATTTCAATACTTACTGAAGGTTGGCCATACAATTATATATTTTTAA
GTGGTGCGCCAAACTTCCGTATATCATGTTTTACATTCTGCAGGAGTCGCTATTATACCATTTATTTTG
AAAATTGTTATTCCTTGTTCTGCGATATTGAGCAGCCACAGTTTTGTGCGAACTGAGTGAAATCTCATC
CAGTGCCCAGAGATCTTCCTTTGTATATAGCCAAAAGCTTATATGTAGTCCTAAACTTCACATACAAAC
ACCCTTATATTTGATAACGAAACATAGGTTTATAAGCTAAATATGCGGCACTTGATAGAAATGTAGTT
TTTTAATTTATTTACACCACCAACCACACCGCAAGCACTTTTTAGATACAACAATTGAAAGTTTTTTAAT
GCAGGCTGGTACACTGAGCTGTTATAGTGTGTATTTTTAAAAGGCATGCACCTATTGGCCAGAGACAC
TCGTTCTCCTTATAACAATACATTTAAAGGTCCCAGTTTGATTTCCTTCTCAACAACTTTTTTGAAAAGT
GTTTTAAAAGTGCGTAGAACAATACACACATTGATAATTACTATATATATATATACCATATATAACTTT
AGGACTAAGACAGCTTACACGTTGAAGAAACAAT
```

Appendix F: CG10600 amino acid sequence

Aligned CG10600 WT and CG10600 mutant amino acid sequence
Red amino acids indicate changes in the mutant polypeptide compared to WT

```
WT   MENTRGGENNVELANGGKESTGNVRKPEKKPLEGAVGASKQRETPEKGVNKASAVENNNANARG
*    MENTRGGENNVELANGGKESTGNVRKPEKKPLEGAVGASKQRETPEKGVNKASAVENNNANARG

WT   NASPTTTSAQREATTTITPVEPTAAATTTNGNTLPTQQQQQLQSGDGNANNPTTNARIMLAAAE
*    NASPTTTSAQREATTTITPVEPTAAATTTNGNTLPTQQQQQLQSGDGNANNPTTNARIMLAAAE

WT   EEQAVASSVATTAGAPTAATTTATTTALADTTKTSTSTDEKTERNHKRKKRDRERDEDKNRRSN
*    EEQAVASSVATTAGAPTAATTTATTTALADTTKTSTSTDEKTERNHKRKKRDRERDEDKNRRSN

WT   NHNSSSKDRDKDRDRDKERDKERDRERERDRNRESSSKSSKEAEHKSSSSRSKEKYHRSDRDRE
*    NHNSSSKDRDKDRDRERDRDRDRDRERDRNRNRRSKSSSDAEAESSSSSSRDKYHHDRDRERE

WT   REKERDREKIPSSSSSSSSHSNRRRSSDSSRSSQNSSFKSKDPGADPVLPQAVVAALEKQIKVE
*    KESDRDKVNIPSSSSSSSSHSNRRRSSDSSRSSQNSSFKSKDPGADPVLPQAVVAALEKQIKVE

WT   EVIKKEPVETKQEVDTEPTPVLSNLIKTEAEIKTESTEKPNKPLVNGEAIAKTRDEVSRQLNFG
*    EVIKKEPVETKQEVDTEPTPVLSNLIKTEAEIKTESTEKPNKPLVNGEAIAKTRDEVSRQLNFG

WT   DKVSNSKLMPSPAPKSSSGSNSTASNHSSSQSSSSRKSYLSSSSSSSSSSNHHKSSSSSSSRHSS
*    DKVSNSKLMPSPAPKSSSGSNSTASNHSSSQSSSSRKSYLSSSSSSSSSSNHHKSSSSSSSRHSS

WT   STSRECSKCYKRSKIRRTSVGVQCLQHAPATGPWQTVQSLPPRPNRKPPAGLENLKYGCYFQVE
*    STSRECSKCYKRSKIRRTSVGVQCLQHAPATGPWQTVQSLPPRPNRKPPAGLENLKYGCYFQVE

WT   QYPNGGASIVHLYQHEIDALSPDEMEELVDEFFEVCFAEDEQGYAHHVMGIVHDAARYLPDLLE
*    QYPNGGASIVHLYQHEIDALSPDEMEELVDEFFEVCFAEDEQGYAHHVMGIVHDAARYLPDLLE

WT   HMAENYSTLTVKAGVLGRNSDIETCTMSQYNEQVVRNYCQGTFRYGPLHQISLVGKVHEEVGGY
*    HMAENYSTLTVKAGVLGRNSDIETCTMSQYNEQVVRNYCQGTFRYGPLHQISLVGKVHEEVGGY

WT   FPDLLGRIEMNPFLRKTMPWACNSILQTDPRQSNDGPILWIRAGEQLVPTAELNSKTPLKRQRT
*    FPDLLGRIEMNPFLRKTMPWACNSILQTDPRQSNDGPILWIRAGEQLVPTAELNSKTPLKRQRT

WT   RINELRNLQYLPRLSEARETMIEDRTKAHADHVGHGHERITTAAVGILKAVHCGQTYSQNRITK
*    RINELRNLQYLPRLSEARETMIEDRTKAHADHVGHGHERITTAAVGILKAVHCGQTYSQNRITK

WT   DVVAFAAQDFNTMVEMLQLDLHEPPISQCVQWIEDAKLNQLRREGVRYARIQLCDNDIYFLPRN
*    DVVAFAAQDFNTMVEMLQLDLHEPPISQCVQWIEDAKLNQLRREGVRYARIQLCDNDIYFLPRN

WT   IIHQFRTVTAVTSVAWHLRLRQYYPGQEVINEKNNPVLAETPHYKEKQTILPNPISHDECGKKT
*    IIHQFRTVTAVTSVAWHLRLRQYYPGQEVINEKNNPVLAETPHYKEKQTILPNPISHDECGKKT

WT   PSKRAHDGRTKKKLIDLDGKERRSSESSIDEHPAAGSSSPSQNQENNSNSSSQSGASGVSTPKK
*    PSKRAHDGRTKKKLIDLDGKERRSSESSIDEHPAAGSSSPSQNQENNSNSSSQSGASGVSTPKK

WT   KSSKEDSKIDMRKMVLEHKYKLTKAAVTETPEKEKLKKDKDKEKSRGKDRDKEKERGKEKEKEK
*    KSSKEDSKIDMRKMVLEHKYKLTKAAVTETPEKEKLKKDKDKEKSRGKDRDKEKERGKEKEKEK

WT   EKKNDTPKKHATPAKSSLSEPSPAKIPKLNNDTAPATALPLKSTPSSAPPTCTLPVAAPPILSG
*    EKKNDTPKKHATPAKSSLSEPSPAKIPKLNNDTAPATALPLKSTPSSAPPTCTLPVAAPPILSG

WT   LPPPLVPQTPFIDRLNQKEPEIKIEVKIVSDSKSDISTSNPVVSTTAPPPPPPPDPVEDVGNEL
*    LPPPLVPQTPFIDRLNQKEPEIKIEVKIVSDSKSDISTSNPVVSTTAPPPPPPPDPVEDVGNEL
```

```
WT    IVESTPQLIVDHEEEVSEVCSEEIVEMDIPIPKEPVPAAAFTSLPNTPSPLISAHSTLATAPI
*     IVESTPQLIVDHEEEVSEVCSEEIVEMDIPIPKEPVPAAAFTSLPNTPSPLISAHSTLATAPI

WT    PKSFLHPPVPIVAAPAPPPPSQIVKVVLPGSLTPLTRMSVVLPPLPPLPPSRTSAPPPPPPP
*     PKSFLHPPVPIVAAPAPPPPSQIVKVVLPGSLTPLTRMSVVLPPLPPLPPSRTSAPPPPPPP

WT    STKSSAITMPISAPVSTLKPTYKTINLPSHKIIIAGSAAGSRVSMAKGTSKRPDLLGSIIASM
*     STKSSAITMPISAPVSTLKPTYKTINLPSHKIIIAGSAAGSRVSMAKGTSKRPDLLGSIIASM

WT    DKPAGSGPNSNSSSSTSSATTNSSLSSYAGSNNSF
*     DKPAGSGPNSNSSSSTSSATTNSSLSSYAGSNNSF
```