An *in vivo* loss of function genetic screen in *Drosophila* identifies novel roles for the Cohesin complex in single and collective cell invasion

Brenda Ivonne Canales Coutiño
BSc (Hons), MSc

Thesis submitted to The University of Nottingham
for the degree of
Doctor of Philosophy

March 2019

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: 

Brenda Canales

Date: 29 March 2019

Abstract

The process of invasion and metastasis is the hallmark of cancer that accounts for approximately 90% of all cancer related deaths. Despite several years of research, this process is still poorly understood. To address this, we have recently completed a large-scale genetic screen to identify novel genes that either promote or inhibit tumour progression in the living fly. A total of 294 genes that, when knocked down, enhance tumour progression phenotypes were identified, 119 of these genes promote invasion of mutant cells.

From the genetic screen, I identified SA1 as a potential invasion suppressor. SA1 is the smallest sub-unit of the cohesin complex; this complex is highly involved in keeping sister chromatids together during cell division and in regulating gene expression, yet the specific role that SA1 plays in cancer cell invasion is yet to be understood. Here I present the further characterization of SA1, and its isoform SA2. My results show that i) the loss of function of either of these genes reduces the junctional localization of adherens junctions proteins; ii) the invading cells are not caused by the activation of apoptosis; iii) loss of the human orthologues STAG1 or STAG2 induces invasion in a breast cancer cell line, as observed in an in-vitro invasion assay; iv) Loss of STAG2 induces gene expression changes to MCF7 breast cancer cells and 85

percent of the differentially expressed genes in STAG2 KD interact with members of the adherens junctions pathway. Together my results strongly suggest that *SA1* and *SA2* are tumour suppressor genes that prevent cancer cell invasion by regulation the expression of genes involved in maintaining cell-cell junction integrity.

Additionally, loss of all cohesin, namely simultaneous loss of STAG1 and STAG2 cohesins, induced drastic cytoskeleton changes including multicellular actin ring formation, tissue folding and clonal extrusion of the mutant cells, basally. My results suggest that the Dpp signalling effector MAD, which is involved in the process of apical constriction, is activated in cohesin mutants. Altogether these results indicate that cohesin and apical constriction are very promising targets to study for their potential role in the porly understood process of collective cell migration. Additionally, these results outline the complexity of cohesin mediated gene expression regulation and the multiple effects cohesin mis-regulation could have in malignant cancer invasion.

Acknowledgements

Firstly, I want to express my deepest gratitude to my supervisor Dr Marios Georgiou, for providing me the opportunity to work on this exciting project and for all his advice and support during my research in his lab. I would also like to thank the lab members who worked alongside me on the screen project: Dr Zoe Cornhill, Dr Africa Couto, Dr Natalie Mack and Alexandra Rusu. It was a pleasure working with all of you for the past four years.

Additionally, I would like to thank our lab technician Zsuzsa Markus, for keeping the fly stocks healthy, co-ordinating the fly food cooking and all her kind motivating words. I want to also thank the project students who worked temporarily with me: Louisa Taylor, Simon Johnson, Rasar Mongkolkul, Annabel Browne, Juned Ahmed and Thin Kyi.

I am also grateful to Dr Ruman Rahman for giving me the opportunity to work in his lab for my cell culture experiments. I want to thank Dr Anbarasu Lourdusamy for his help with the screen data analysis. Dr Marcos Catellanos for his help with the transcriptomics experiments. I want to also mention Ian Ward, Tim Self and Chris Gell from the School of Life Sciences imaging facility, for all their help with specialized microsopy techniques.

I want to also thank my sponsors: CONACyT and the University of Nottingham, without their financial support this work wouldn't have been possible.

Special thanks to my brother Gerardo for always taking care of me, for encouraging me to be strong and for all his constant support. I want to thank my boyfriend Braulio, for his feedback and for providing uplifting words during the stressful moments. Additionally, I want to thank the rest of my family, all my aunts and uncles, for all their support.

Finally, I want to dedicate this work to my parents, Imelda and Gerardo. I will always carry your loving memory and all the lessons you taught me. I am truly blessed to be your daughter.

Contents

An *in vivo* loss of function genetic screen in *Drosophila* identifies novel roles for the Cohesin complex in single and collective cell invasion 1

Declaration 2

Abstract 3

Acknowledgements 5

Contents 7

Figures 10

Tables 12

Abbreviations 13

Chapter 1 - General introduction 16

1.1 The multistep process of carcinogenesis 18

1.2 Biological processes involved in tumour progression 23

1.2.1 Proliferation and growth 23

1.2.2 Survival 33

1.2.3 The immune system and inflammation 40

1.2.4 Genomic instability and mutation 42

1.3 Activating invasion and metastasis 44

1.3.1 Invasion 44

1.3.2 Metastasis 56

1.4 *Drosophila* as a model for cancer research 61

1.5 Research aims 64

Chapter 2 - Materials and Methods 66

2.1 *Drosophila* materials and methods 67

2.1.1 *Drosophila* materials 67

2.1.2 *Drosophila* methods 71

2.2 Cell culture materials and methods 90

2.2.1 Cell culture materials 90

2.2.2 Cell culture methods 93

Chapter 3 - A genetic screen to identify modifiers of tumour progression 100

3.1 Introduction 101

3.1.1 Preliminary experiments 102
3.1.2 Pilot screen 103
3.2 Results 107
3.2.1 A genetic screen to identify modifiers of tumour progression 107
3.2.2 Phenotypes relevant to tumour progression and scoring system 110
3.2.3 Quality control 114
3.2.4 Identification of genes that modify tumour progression 116
3.2.5 Characterization of invasion suppressors 122
3.2.6 Identification of the Cohesin complex subunits as modifiers of tumour progression. 128
3.3 Discussion 133
Chapter 4 - Investigating the role of SA1 and SA2 Cohesin subunits in single cell invasion 141
4.1 Introduction 142
4.1.1 Canonical function of Cohesin 142
4.1.2 Roles of Cohesin besides sister chromatid cohesion 149
4.1.3 Cohesin defects and disease 151
4.2 Results 153
4.2.1 SA1 and SA2 KD promotes single cell invasion in the living epithelium. 153
4.2.2 Investigating the migration speed and directionality of SA1 and SA2 KD invading cells 156
4.2.3 *SA1* and *SA2* invading cells are not apoptotic 160
4.2.4 Ectopic localization of adherens and septate junction proteins in *SA1* and *SA2* KD clones. 163
4.2.5 STAG1 and STAG2 KD promotes migration and invasion in a human adenocarcinoma cell line 167
4.2.5 *STAG2* KD induces changes in the transcriptome of MCF7 cells 171
4.2.6 Biological function and interactions of differentially expressed genes in MCF7 cells by STAG2 KD 179
4.3 Discussion 184
Chapter 5 - The role of the Cohesin complex in apical constriction and clonal extrusion 193
5.1 Introduction 194
5.1.1 Collective cell migration 195

5.1.2 Role of apical constriction in actin ring formation and collective cell migration 196
5.2 Results 200
5.2.1 Cohesin loss of function induces cytoskeletal changes and actin ring formation 200
5.2.2 SMC3-/- actin rings are supracellular and induce tissue invagination 202
5.2.3 Myosin II is enriched at the SMC3-/- actin rings. 204
5.2.4 *SMC3-/-* clones with rings extrude basally from the epithelium 208
5.2.5 Rho1 and myosin II have significant effects on SMC3-/- actomyosin ring formation. 213
5.2.6 Screen of apical constriction regulators 217
5.2.8 Loss of Cohesin increases phosphorylated Mad levels 222
5.2.7 *smc3* mutant clones do not affect the number of Cut positive cells 224
5.3 Discussion 227
Chapter 6 Discussion 237
6.1 Discussion 238
6.1.1 Genetic screen 238
6.1.2 Cohesin regulates single and collective cell invasion 241
6.1.3 Apical constriction and its potential role in collective cell invasion. 248
6.2 Future perspectives 250
6.3 Limitations 252
6.4 Conclusions 254
Chapter 7 – Bibliography 255
Appendixes 286
Appendix 1 – Screen Phenotypes and scoring system 287
Appendix 2- Screen genes 289
Appendix 3 - Screen hit genes 311
Top suppressor hits 311
Top enhancer hits 318
Lethal genes 332
Appendix 4 – Hierarchical clustering GO terms (BP) 333
Appendix 5 – Differentially expressed genes in STAG2 KD 338
STAG2 vs Untreated MCF7 338
STAG2 vs Non-targeting siRNA MCF7 339

Non-targeting siRNA MCF7 vs Untreated MCF7 340

Figures

Figure 1.1 Multistep process of carcinogenesis.....22
Figure 1.2 The hallmarks of cancer 25
Figure 1.3 Cell cycle.....30
Figure 1.4 Apoptotic pathways 37
Figure 1.5 Epithelial cell apical-basal polarity.....47
Figure 1.6 Cell adhesion proteins.....51
Figure 1.7 Types of cell invasion 56
Figure 1.8 Invasion-metastasis cascade.....60
Figure 2.1 Gal4/UAS system.....73
Figure 2.2 UAS-RNAi system 74
Figure 2.3 FLP/FRT system 76
Figure 2.4 MARCM system.....78
Figure 2.5 siRNA cut sites in Cohesin subunits mRNAs.....94
Figure 2.6 Trans-well migration and invasion assay diagram 97
Figure 3.1 Preliminary experiments and pilot study.....104
Figure 3.2 Outline of genetic screen 109
Figure 3.3 Scoring system specifically designed to identify modifiers of tumour progression phenotypes 113
Figure 3.4 Quality control 115
Figure 3.5 Clustering analyses identify ten RNAi line clusters.....119
Figure 3.6 Clustering analyses identify three distinct phenotypic subgroups.....121
Figure 3.7 An interaction network of invasion suppressors 124
Figure 3.8 Characterisation of selected invasion suppressors.....127
Figure 3.9 Effect of Cohesin subunit KD on lgl-/- clones.....129
Figure 3.10 Effect of Cohesin subunit KD in lgl mutant tumours 131
Figure 4.1 Somatic and germ cells Cohesin complexes 144

Figure 4.2 Conserved protein domains between Drosophila SA and human STAG proteins 145
Figure 4.3 Cohesin interactions with the chromosomes throughout the cell cycle 148
Figure 4.4 Effects of Cohesin subunit KD on invasion, apex defects and multilayering 155
Figure 4.5 Investigating the migration speed and directionality of SA1 KD invading cells.....158
Figure 4.6 SA1 and SA2 invading cells are not apoptotic.....162
Figure 4.7 Adherens junction protein localization is disrupted in SA1 and SA2 KD clones 165
Figure 4.8 FasIII localization is disrupted in SA1 and SA2 KD clones but no effect is observed on other septate junction and polarity proteins 166
Figure 4.9 Cohesin subunit KD in MCF7 cells.....169
Figure 4.10 STAG1 and STAG2 are invasion and migration suppressors in MCF7 cells 170
Figure 4.11 Data distribution from microarray analysis of MCF7 transcriptome 173
Figure 4.12 Volcano Plot of differentially expressed genes.....175
Figure 4.13 Hierarchical clustering of differentially expressed genes.....176
Figure 4.14 RT-qPCR validation of a selection of differentially expressed genes....178
Figure 4.15 Gene ontology (GO) Biological Process enrichment.....181
Figure 4.16 Interaction network of differentially expressed genes by STAG2 KD and adherens junction regulators.....183
Figure 5.1 Effects of apical constriction in epithelial polarized cells 199
Figure 5.2 Cohesin loss of function induces actin ring formation 201
Figure 5.3 SMC3 actin rings are multicellular and induce tissue invagination 203
Figure 5.4 Increased myosin and E-Cadherin localization in smc3 mutant clones .207
Figure 5.5 smc3$^{-/-}$ actomyosin rings contract over time leading to clonal extrusion 210
Figure 5.6 Clonal delamination does not trigger apoptosis 212
Figure 5.7 Rho1 activity affects ring formation and clonal extrusion 214
Figure 5.8 Myosin2 expression affects ring formation and clonal extrusion 216
Figure 5.9 An enhancer/suppressor screen of apical constrictor regulators 221

Figure 5.10 Cohesin loss of function increases pMad nuclear localization 223
Figure 5.11 Loss of smc3 has no effect on Cut expression 226
Figure 6.1 Proposed model of the role of cohesin in single and collective cell invasion 247

Tables

Table 2.1 General reagents……67
Table 2.2 Drosophila stocks……68
Table 2.3 UAS-transgenes Chapter 4 and 5……69
Table 2.4 Mutant alleles……70
Table 2.5 Primary Antibodies……70
Table 2.6 Secondary Antibodies……70
Table 2.7 Balancer chromosomes and genetic markers……80
Table 2.8 Genetic constructs……81
Table 2.9 Fiji plug-ins and macros installed and used……87
Table 2.10 General reagents……90
Table 2.11. Primer list for rt-qPCR……91
Table 2.12 siRNA sequences……92
Table 4.1 Gene Ontology (GO) terms……182
Table 5.1 Apical constriction genes studied in 5.2.6……219

Abbreviations

4EBP	eIF4E Binding Protein
AJ	Adherens junction
AKT	AKT Serine/Threonine Kinase 1
ALT	alternative lengthening of telomeres
AMPK	AMP-activated protein kinase
APAF-1	Apoptotic Peptidase Activating Factor 1
APC/C	anaphase-promoting complex
APF	after pupa formation
ATP	adenosine triphosphate
Bak	Bcl-2 homologous antagonist/killer
Bax	BCL2 Associated X
bcl-2	B cell leukemia/lymphoma 2
Bsk	Basket
CAF	cancer associated fibroblast
CDCA5	Cell Division Cycle Associated 5
Cdk	cyclin-dependent kinases
CO2	Carbon Dioxide
Cor	Coracle
CSF1	colony stimulating factor 1
CTCF	CCCTC binding factor
Dlg	discs large
DMBA	7,12-Dimethylbenz[a]anthracene
DMEM	Dulbeccos modified eagle medium
DMSO	Dimethylsulphide
DNA	Deoxyribonucleic Acid
Ecad	E-Cadherin
ECM	extracellular matrix
ECM	extracellular matrix
EGF	epidermal growth factor
EMT	Epithelial to mesenchymal transition
ESCO1	Establishment Of Sister Chromatid Cohesion N-Acetyltransferase 1
ESCO2	Establishment Of Sister Chromatid Cohesion N-Acetyltransferase 2
FADD	as Associated Via Death Domain
FLP	flippase
FRT	flippase recognition target
Gal4	Galactosidase 4
GDP	guanosine diphosphate
GFP	Green fluorescent protein
GTP	guanosine triphosphate
HER2	human epidermal growth factor receptor 2
His	histone

HR	homologous recombination
IL	interleukin
JNK	c-Jun N-terminal kinase
KD	knock down
lgl	lethal giant larvae
MARCM	mosaic analysis with a repressible cell marker
MCC	mitotic checkpoint complex
MET	mesenchymal to epithelial transition
MMP	matrix metalloproteinases
MOMP	mitochondrial outer membrane permeabilization
mTOR	mammalian target of Rapamycin
NipB	Nipped B
NK	natural killer cells
non-T	Non targeting siRNA (scrambled)
PARP	Poly [ADP-ribose] polymerase
PCNA	Proliferating Cell Nuclear Antigen
PDGF	Platelet-derived growth factor
PDS5	precocious dissociation of sisters 5
PI3K	Phosphoinositide 3-kinase
PIP3	phosphatidylinositol-3,4,5-trisphosphate
PKC-α	protein kinase C-alpha
pnr	pannier
Rb	retinoblastoma
RFP	Red fluorescent protein
RhoGEFs	Rho Guanine Exchange Factors
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
ROCK	Rho kinases
rt-qPCR	quantitative retro transcriptase PCR
S6K1	p70S6 Kinase 1
SA1	stromalin 1
SA2	stromalin 2
scc	sister chromatid cohesion
Scrib	Scribble
shg	shotgun
SJ	Septate junctions
SMC1	Structural maintenance of chromosomes protein 1
SMC3	Structural maintenance of chromosomes protein 3
sqh	
ssDNA	single stranded DNA
STAG1	Stromal Antigen 1
STAG2	Stromal Antigen 2
TGF-β	transforming growth factor-β
TNFR	tumour necrosis factor receptor
TPA	tetradecanoyl phorbol acetate

TSC1	Tuberous sclerosis 1
TSG	Tumour suppressor gene
UAS	upstream activation sequence
Ubx	Ultrabithorax
ULK1	Unc-51 Like Autophagy Activating Kinase 1
unt	untreated
VEGF	Vascular endothelial growth factor
VEGFR1	vascular endothelial growth factor receptor 1
WAPL	Wings apart-like protein
XPC	Xeroderma pigmentosum, complementation group C,

Chapter 1 - General introduction

General Introduction

Carcinogenesis is a multistep process where normal cells transform into malignant cancer cells through the accumulation of genetic abnormalities in key parts of the genome (Farber, 1984). These alterations provide cancer cells with the ability to grow and divide uncontrollably, creating over proliferation of malignant cells with invasive capabilities, and eventually some of them will spread to distant sites of the body to form secondary tumours through the process of metastasis (Poste et al., 1980). Signalling with neighbouring normal cells is affected, as is altered communication with blood vessels (Baluk et al., 2005) and the immune system (de Visser et al., 2006) providing a unique microenvironment that will contribute to the overall behaviour of the tumour.

There are over 200 types of cancer (Lodish, 2013) and they can result from abnormalities in different cell types, each with its own behaviour and response to treatment. The majority of human tumours are carcinomas, derived from epithelial cells. Despite years of cancer research, the current therapeutic options are not fully successful and carcinomas account for more than 80 percent of cancer related deaths in the Western world (Weinberg, 2014), making carcinomas one of the top public health diseases.

1.1 The multistep process of carcinogenesis

The multistep process of carcinogenesis involves a single cell that acquires and accumulates advantageous genetic alterations until it becomes malignant. The genes that must be targeted in tumorigenesis are categorized as tumour suppressor genes (TSGs) and oncogenes. TSGs, as their name suggests, functions to suppress tumour related phenotypes; therefore, a homozygous loss of function mutation in a TSG will provide cancer cells with survival and proliferative advantages (Weinberg, 1991). Oncogenes provide the same advantages to the cancer cells, yet unlike TSGs, they must be activated or overexpressed; before oncogenes are mutated they are referred to as proto-oncogenes (Cantley *et al.*, 1991). TSGs and oncogenes have opposite functions, since they control essential processes for tissue homeostasis. In a normal cell, TSGs and oncogenes are present at a balance; during oncogenesis, the balance shifts towards increased oncogene and decreased TSG function.

In 1990, Vogelstein *et al* described a linear process for colorectal cancer where they showed that the most significant difference between benign and aggressive tumours was the number of mutations in TSGs and oncogenes. A benign tumour has one or two mutations, while an aggressive tumour must

have accumulated at least four mutations (Fearon *et al.*, 1990). The Vogelstein model accurately described the positive relationship between mutation accumulation and tumour aggressiveness, yet the model has been updated, since a linear model was an oversimplification of tumour formation. Current models suggest that instead of linear progression, many branches can be formed; as the initial cell divides, and the population grows, many subclones will form due to new accumulated mutations (Figure 1.1a). The acquired mutations will confer some cells with a selective advantage and, much like the Darwinian model of natural selection, only the fittest will survive. This process is called clonal selection and some of the advantages the cells obtain include rapid growth, better survival mechanisms and invasive capabilities (Greaves *et al.*, 2012; Wang *et al.*, 2014)

Efforts to replicate epithelial cancer *in vivo* in model organisms has generated extensive evidence to support the multistep model (Okarmoto *et al.*, 1989; Bruckheimer *et al.*, 2000; H. Watson *et al.*, 2001). For example, the carcinogen dimethylbenz(a)anthracene (DMBA) was applied to mouse skin, which induced irreversible mutations, yet no visible papilloma were observed. Application of additional tetradecanoyl phorbol acetate (TPA), a phorbol ester that induces proliferation by the inactivation of the TSG protein kinase C-alpha (PKC), was sufficient to create papilloma after 3 months

(DiGiovanni, 1992). Papilloma were only formed when DMBA and TPA were applied in the same area, and in the same order. Additionally, if following the application of DMBA and TPA, an additional carcinogen was applied, this could generate a carcinoma instead of a papilloma after 3 months (DiGiovanni, 1992; Abel *et al.*, 2009) (Figure 1.1b).

Three stages have been described as part of this multistep model, namely: tumour initiation, tumour promotion and tumour progression (Farber, 1984; Vincent *et al.*, 2008). Tumour initiation begins with a single cell acquiring one or more irreversible mutations in genes with specific functions (O'Brien *et al.*, 2006; Goldstein *et al.*, 2010), which will provide the cells with different advantages, such as survival, proliferation, movement, amongst others. DMBA is an example of an initiator agent. An initiated cell is not a cancer cell and does not necessarily become one, but the acquired mutations are mitotically hereditable, and these clones might gain further genomic alterations in the future (Figure1.1b). Tumour promotion is the second step in the multistep carcinogenesis model, the DNA sequence is not affected but other types of reversible instabilities arise, such as epigenetic events leading to changes in gene expression (Feinberg *et al.*, 2006). Promoter agents, such as TPA, can induce alterations in the cell shape, growth rate, and degree of differentiation, amongst others. The presence of a promoting agent does not

make the cell carcinogenic; at the most it creates a benign tumour, but while in contact with the initiated cell it increases the likelihood of further mutations, mostly by increasing the population of initiated cells (Figure1.1b). Tumour progression is characterized by further mutations in tumour suppressor genes and oncogenes of an already initiated and promoted cell. The process of progression is irreversible, with clonal selection occurring continuously until the tumour becomes increasingly malignant and capable of invasion and metastasis (Figure 1.1b).

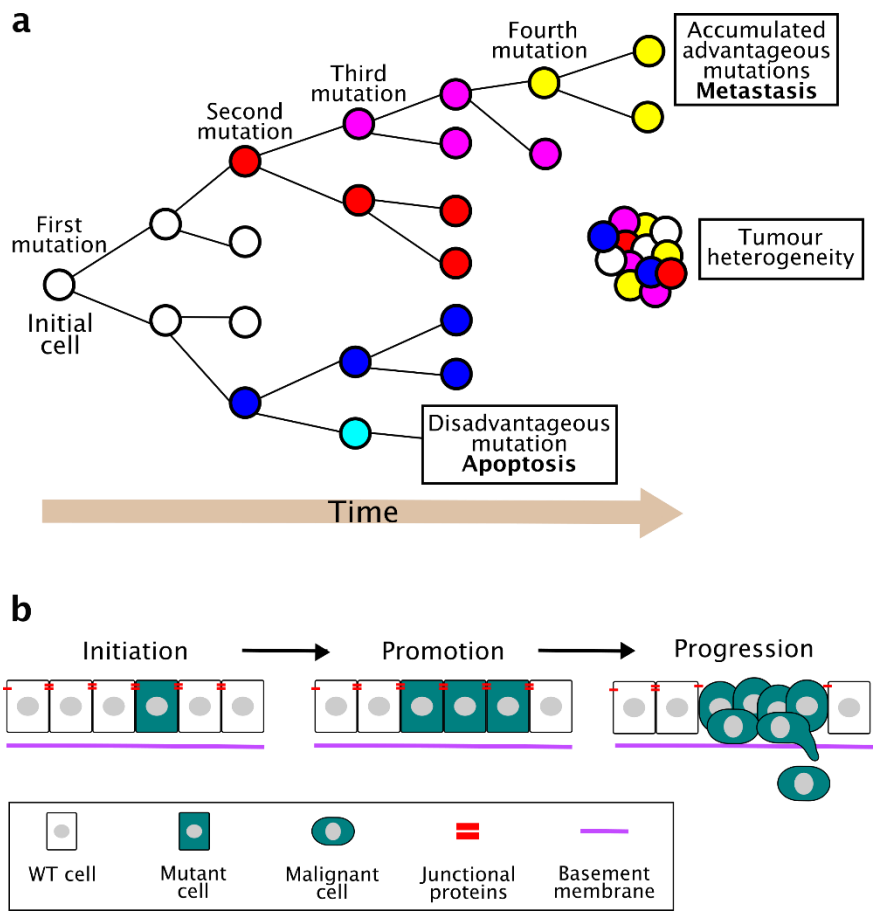


Figure 1.1 Multistep process of carcinogenesis

a) Illustration of clonal selection. A single cell acquires mutations in tumour suppressor genes and oncogenes. As the cell divides, the new populations accumulate novel mutations, creating a variety of clones with different genotypes and phenotypes. Most mutations will either confer no advantage or be deleterious to cell survival; while others will be advantageous and will provide the clones with proliferative, invasive and survival capabilities. Tumour heterogeneity refers to the different cells within a tumour **b)** The multistep process of carcinogenesis, as described in Farber, 1984. Initiation starts with a single cell acquiring somatic mutations in tumour suppressor genes and/or oncogenes. Promotion involves epigenetic changes that lead to proliferation and an increased number of initiated cells. Progression describes the accumulation of further mutations leading to malignancy through invasion and metastasis. Adapted from Greaves *et al.*, 2012.

1.2 Biological processes involved in tumour progression

As mentioned in section 1.1, tumour formation requires tumour cells to accumulate multiple mutations and epigenetic alterations. Specific biological processes must be altered to provide the cells with clonal advantages. The hallmarks of cancer, described by Hanahan and Weinberg, successfully classified the biological processes that must be altered to favour malignancy (Hanahan *et al.*, 2000, 2011) (Figure 1.2); genes that regulate these processes are very likely to be TSGs and oncogenes. This section will briefly describe the different biological processes involved in tumour progression, with specific emphasis on proliferation and on the activation of invasion and metastasis, since these processes are relevant to the results presented in Chapters 3, 4 and 5.

1.2.1 Proliferation and growth

Tissue homeostasis requires a tight control of cell division, where one cell replicates into two identical daughter cells; this should happen according to extracellular signals (growth factors) to meet the requirements of tissue homeostasis (Grinstein *et al.*, 1989; Rodrigues *et al.*, 2010) and only when there is no DNA damage (Zhou *et al.*, 2000). Naturally highly proliferative tissues in the body include the skin and the epithelium of the gastrointestinal tract (Feitelson *et al.*, 2015). Proliferation is also high in specific situations,

such as in certain development processes and also during epithelial wound healing (Landén *et al.*, 2016). Historically, uncontrolled proliferation has been the most iconic trait of cancer, with traditional chemotherapy treatment exclusively targeting hyper-proliferative cells; therefore, understanding the mechanisms regulating this process is vital to understanding tumour formation.

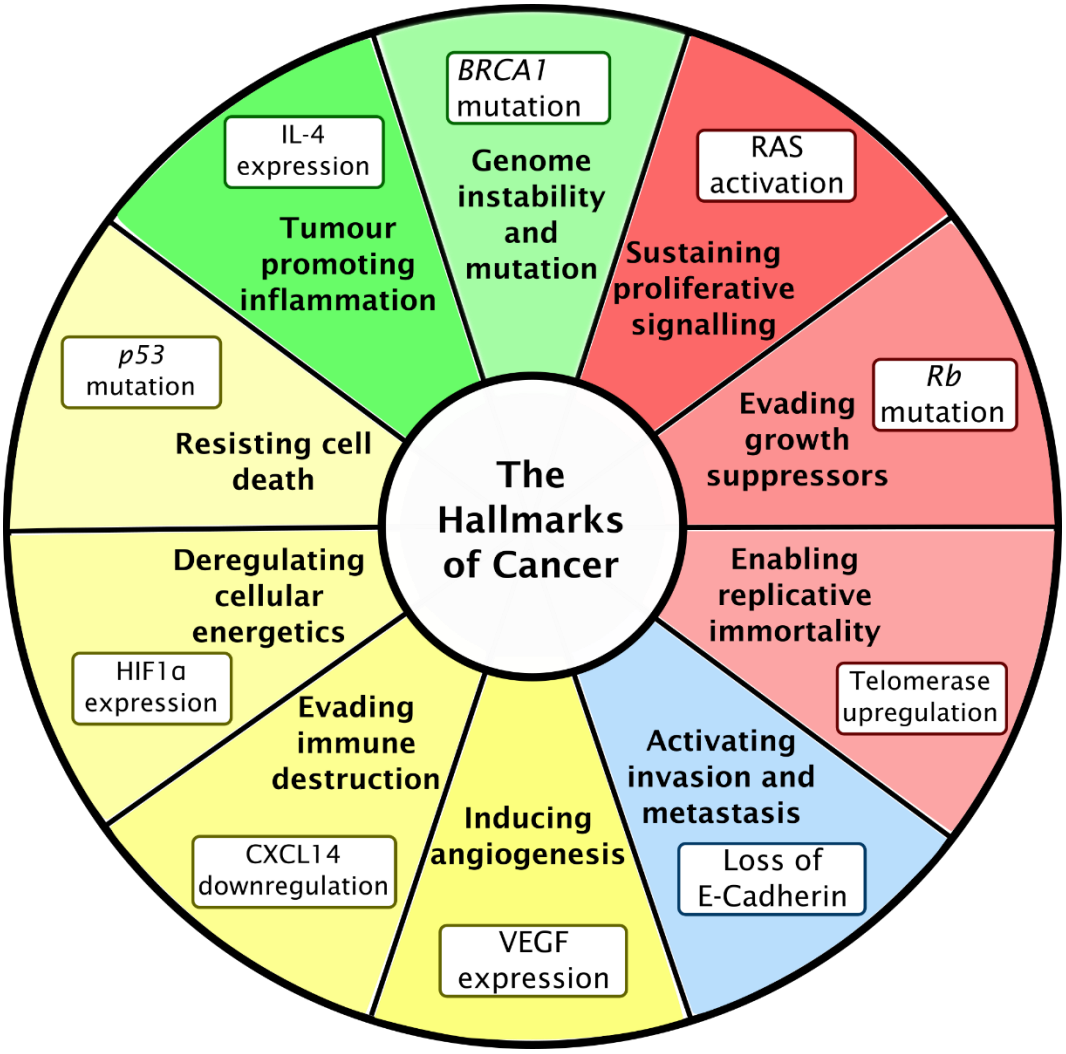


Figure 1.2 The hallmarks of cancer

An illustration of the ten capabilities, named hallmarks of cancer, which cancerous cells gain during the process of malignancy (Hanahan *et al.*, 2011). Proliferation related: Sustaining proliferative signal, evading growth suppressors and enabling replicative immortality. Survival related: Inducing angiogenesis, evading immune destruction, deregulating cellular energetics and resisting cell death. Others: Activating invasion and metastasis, tumour promoting inflammation, genome instability and mutation. An example of a gene or process, which is commonly disrupted in each hallmark is specified in the text boxes.

1.2.1.1 Cell proliferation

The cell cycle comprises two main stages: interphase, where cell contents will duplicate; and mitosis, where the cell divides forming two identical daughter cells (Figure 1.2). Interphase consists of 3 periods, the first being gap 1 (G1). After a cell has divided it enters this first gap stage that is highly responsive to extracellular signals which may promote or inhibit mitosis. Here, non-genomic cellular contents duplicate, and the cell begins to grow. During this phase the cell may commit to enter the next stage that leads to cell division or enter a resting phase, called G0, where the cells will stay dormant. Eventually, depending on extracellular signals and genetic material integrity, the G0 cells might re-enter the cell cycle, commit to programmed cell death; alternatively, highly differentiated and senescent cells (see section 1.2.1.3) permanently stay in G0 (Alberts *et al.*, 2002).

A somatic human cell contains 23 pairs of chromosomes and an exact copy of all of them is made during the second interphase period, named synthesis or S phase. The final interphase stage, gap 2 (G2), involves extensive protein synthesis and verification of the DNA synthesis process; if no errors are detected the cell commits to enter mitosis (Alberts *et al.*, 2002). Mitosis has 4 different stages: i) prophase, where chromosomes condense; ii) metaphase, duplicated chromosomes align on the spindle; iii) anaphase, sister

chromatids separate and iv) telophase, chromosomes decondense, and this final stage is followed by cytokinesis, where the mother cell is separated into two (Alberts *et al.*, 2002).

Transition between each of the stages of the cell cycle is controlled by 4 specific checkpoints (Hartwell *et al.*, 1989), regulated by tumour suppressor genes (Stewart *et al.*, 1995) or oncogenes (Ho *et al.*, 2002) and two highly conserved types of proteins: cyclins, that have a regulatory function; and cyclin-dependent kinases (cdks), that have a catalytic function (Peter *et al.*, 1994). Checkpoint one is very important, since at this point the cell commits to a round of cell division that will only be stopped if critical errors occur. This is regulated by cyclins D and E, which are expressed during G0 and G1. The TSG *Rb* is also involved in this stage; the hypophosphorylated form of the RB protein binds and sequesters the E2F transcription factor, which leads to E2F transcription mediated blockage and growth inhibition. Cyclin D-cdk4 and cyclin D-cdk6 phosphorylate RB on specific residues, releasing E2F from RB-mediated inhibition and thereby driving the transition to S phase (Weinberg, 1995). A homozygous loss-of-function mutation in *Rb* has the same effect as a hyperphosphorylated RB: since RB is not present to sequester E2F, and the cell goes through the G1/S checkpoint freely, causing uncontrolled cell division. The gene encoding cyclin D acts as an oncogene in

different human cancers (Hinds *et al.*, 1994; Sauter *et al.*, 2002). Therefore, cyclin D and its regulators are important therapeutic targets.

The other cyclins are less abundant; the cyclin E-cdk2 complex has histone H1 kinase activity and is thought to be required for the initiation and progression of DNA replication (Koff *et al.*, 1992). Cyclin A is synthesised during S phase and its main function is the inactivation of the no longer required E2F, necessary for progression to G2 (Krek *et al.*, 1995). Cyclin B is synthesised during late S phase; it binds to cdk1 forming a complex that is required to enter mitosis, with its degradation required for progression past anaphase (Brito *et al.*, 2006).

The tumour suppressor protein P53 plays an important role in the cell cycle checkpoint controls, it induces transcription of *p21*, a potent inhibitor of all the cdks and of DNA polymerase (Xiong *et al.*, 1993). Loss of P53 prevents *p21* transcription and increases the rate of checkpoint control failure. Additionally, P53 is involved in DNA repair, through the nuclear excision repair proteins proliferating cell nuclear antigen (PCNA) and xeroderma pigmentosum (XPC) (Wang *et al.*, 2003); block of angiogenesis, by positive regulation thrombospondin, an inhibitor of blood vessel formation (Dameron *et al.*, 1994); and induction of apoptosis, via BAX, FAS and PUMA,

a member of the Bcl-2 family (Wu *et al.*, 2007; Dogu *et al.*, 2009; Muller *et al.*, 2014) (see section 1.2.2). Overall, deprivation of P53 has an important effect on genetic instability, with its various functions coining it the nickname "the guardian of the genome". It is one of the most mutated genes in cancer, and loss of *p53* is found in over 50% of all tumours.

The members of the Ras family of proto-oncogenes are found localized to the plasma membrane and have important roles in signal transduction. In a normal cell, RAS is frequently found inactive, in its GDP-bound form. Once bound to GTP, it becomes activated and promotes growth and proliferation (Mulcahy *et al.*, 1985; Johnson *et al.*, 2001). Mutations that lead RAS to its constitutively active form are very important to the process of tumour progression, and over the years they have been linked to various cancers, such as lung, colon and pancreatic carcinomas (Forrester *et al.*, 1987; Malumbres *et al.*, 2003). Many other proto-oncogenes regulate proliferation, including HER2 (c-erbB-2), which is a member of the ERBB receptor tyrosine kinase superfamily that is overexpressed in several cancers. Its expression increases growth and proliferation by stimulating RAS and PI3K signalling (Wright *et al.*, 1989; Wong *et al.*, 2010) (see section 1.2.1.2). The Myc family, that are nucleus associated, are also highly implicated in cell proliferation through transcriptional activation and cell competition (Little *et al.*, 1983).

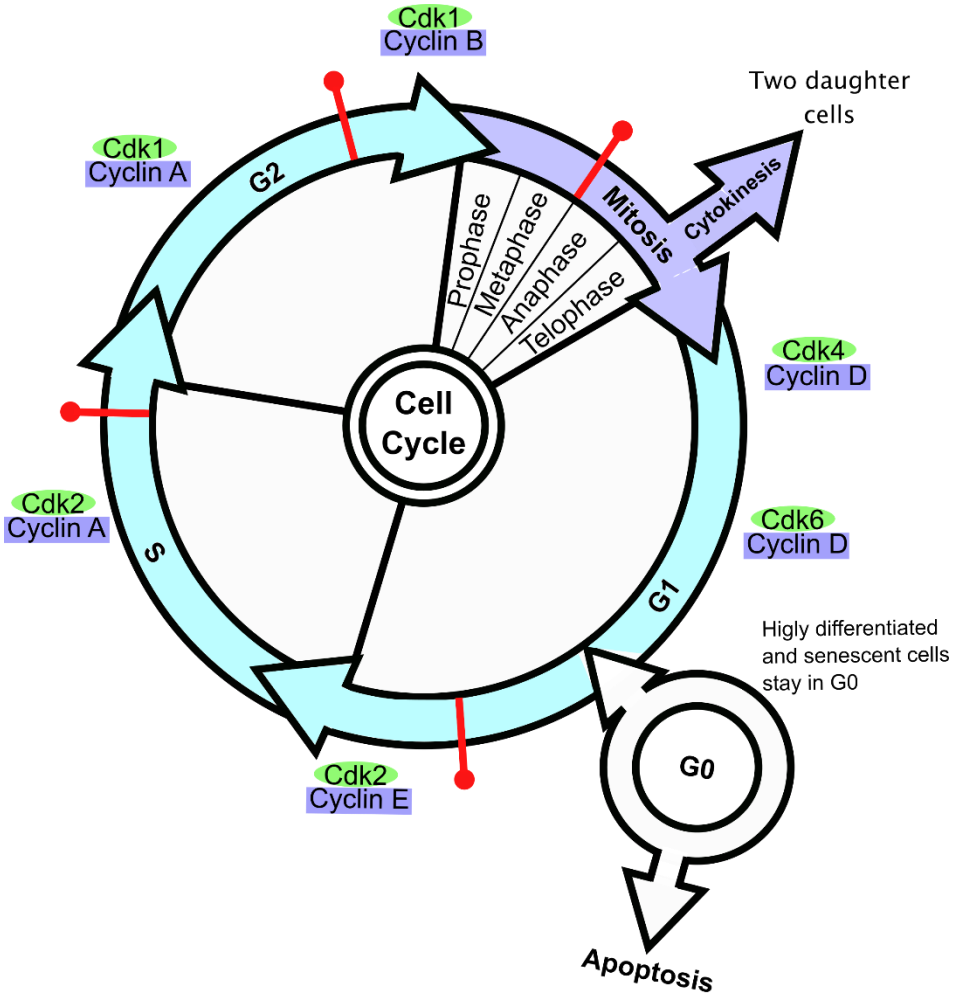


Figure 1.3 Cell cycle

A schematic diagram of the cell cycle showing interphase (with its three stages: G1, S and G2) and mitosis (and its five stages: prophase, metaphase, anaphase, telophase and cytokinesis). During mitosis, the parent cell divides into two daughter cells. Cells that have entered G0, might re-enter the cell cycle, commit to programmed cell death or permanently exit the cell cycle (highly differentiated and senescent cells). Red pins point the four cell cycle checkpoints that control the cell entry to the next phase, critical aspects that might block or allow progression through the cell cycle include cell growth and DNA integrity. The checkpoints are mediated by cyclins, cyclin dependent kinases (cdks), tumour suppressor genes and proto-oncogenes. Adapted from Lodish, 2013.

1.2.1.2 Cell growth

As mentioned in section 1.2.1.1, cell growth is an essential process during the G1 and G2 phases of the cell cycle. There are specific checkpoints at both gap stages that rely specifically on cell size and contents. Studies have shown that the cell cycle is delayed or accelerated depending on cell size (Prescott, 1956; Killander *et al.*, 1965; Dolznig *et al.*, 2004). Therefore, cancer therapeutics that target growth signalling are very important, as this process is directly related to proliferation rate.

The mammalian target of Rapamycin (mTOR) signalling is an important mechanism that regulates cell growth in various ways. mTOR can promote protein synthesis through the phosphorylation of two key effectors: p70S6 Kinase 1 (S6K1), that phosphorylates and activates several substrates that promote mRNA translation (Holz *et al.*, 2005); and eIF4E Binding Protein (4EBP), that also induces mRNA translation through the release of the translational regulator eIF4E (Fingar *et al.*, 2002). Additionally, mTOR blocks autophagy by preventing the AMPK mediated activation of ULK1, a kinase that drives autophagosome formation. mTOR has been shown to phosphorylate ULK1 in Ser757, this prevents AMPK from recognising and interacting with ULK1, therefore blocking autophagy and allowing macromolecule accumulation (Kim *et al.*, 2011).

mTOR signalling has an intricate interaction with the serine-threonine kinase (AKT). AKT is recruited to the cellular membrane by the phosphatidylinositol 3–kinase (PI3K) pathway; once localized to the membrane, AKT can be phosphorylated and activated by mTORC2 (Hagiwara *et al.*, 2012). Activated AKT can negatively regulate the tumour suppressor protein TSC1 to increase mTORC1 levels (Zhang *et al.*, 2003), and therefore increase cell growth and decrease autophagy. Activated AKT additionally regulates apoptosis, by the negative regulation of P53 (Gottlieb *et al.*, 2002). Furthermore, the tumour suppressor protein pTEN negatively regulates the PI3K pathway by dephosphorylating the AKT recruiter: phosphatidylinositol-3,4,5-trisphosphate (PIP3) into PIP2 (Cantley *et al.*, 1999). Therefore, loss of pTEN leads to increased AKT and mTOR activity.

1.2.1.3 Enabling replicative immortality

The "replicative age" of a cell can be measured by how many rounds of cell division it has been through. This is strictly monitored, since each cell division involves the duplication of every nucleotide, and errors could be overlooked and accumulated in cells that have been replicated extensively. To overcome this problem, every time the cell goes through the S phase, the outer regions of the chromosomes, called telomeres, are not fully replicated, and the last 50-100bp are lost (Levy *et al.*, 1992; Verdun *et al.*, 2007). The

consequence of this is telomere shortening with each cell cycle, until a limit is reached, called Hayflick limit (in honour of the discoverer of this process) (Hayflick *et al.*, 1961), which requires that the cells enter senescence (Williams, 1957; Collado *et al.*, 2005).

A hallmark of cancer is the ability of the mutant cells to overcome the Hayflick limit and continue to go through rounds of cell division even if their replicative age has reached the limit (Hanahan *et al.*, 2011). Two main mechanism are known: i) through the expression of the enzyme telomerase, that catalyses the addition of telomeric DNA repeats, lengthening of the telomeres (Kim *et al.*, 1994). ii) alternative lengthening of telomeres (ALT) based on homologous recombination (HR) dependent synthesis of telomeric DNA repeats (Bryan *et al.*, 1995).

1.2.2 Survival

An essential characteristic of tissue homeostasis is the constant communication and tight control of cellular processes through intracellular and intercellular signals. Therefore, proliferation and growth alone are not sufficient to drive tumorigenesis, as there are several pathways that identify and prevent increased uncontrolled cellular proliferation. Pathways that affect cellular survival are highly mis-regulated during tumorigenesis.

1.2.2.1 Apoptosis

Programmed cell death, or apoptosis, is the main mechanism that counterbalances proliferation. Two main pathways that trigger the apoptotic signal are: i) intrinsic, where a cell receives internal apoptotic signals in response to DNA damage and ii) extrinsic, where extracellular signals induce programmed cell death. In developing tissues that are highly proliferative the activation of apoptosis allows the maintenance of controlled cell numbers and the maintenance of correct tissue size.

The mitochondrial or intrinsic pathway begins with mitochondrial outer membrane permeabilization (MOMP) followed by the release of cytochrome C (Green *et al.*, 1998; Wang *et al.*, 2009), this results in the oligomerization and activation of APAF-1 and the formation of the apoptosome that activates the initiator apoptotic protein procaspase 9 (Zou *et al.*, 1999), followed by the activation of effector caspases 3 and 7. The B cell leukemia/lymphoma 2 (bcl-2) pro- and anti-apoptotic family of proteins is very important for this process to occur as pro-apoptotic BAX and BAK oligomerize to form the pores necessary for MOMP formation and cytochrome C release (Korsmeyer *et al.*, 2000). In contrast, anti-apoptotic BCL-2 binds to BAX and BAK, sequestering them and preventing pore formation (Cheng *et al.*, 2001).

.

The extrinsic or death receptor pathway involves members of the tumour necrosis factor receptor (TNFR) superfamily, which include: FAS (or CD95), TRAIL receptors and TNFR1. When the FAS and TRAIL receptors interact with their ligands, the adapter protein FADD binds to pro-caspase-8, allowing it to form an active caspase 8 dimer that will trigger apoptosis through the effector caspases 3 and 7 (Kischkel *et al.*, 1995) (Figure 1.3). The TNFR1 receptors trigger a different response that involves crosstalk with the mitochondrial pathway. Upon ligand interaction, binding of the adapter protein TRADD, and a complex signalling cascade is induced that can lead to apoptosis or proliferation, through 3 different pathways. i) TRADD binds to FADD in the cytosol and the activation of apoptosis is induced through the same mechanism as the FAS and TRAIL receptors (Hsu *et al.*, 1996). ii) NFκB pathway activation that promotes cell survival (Twu *et al.*, 2011) and iii) JNK pathway activation that can either induce proliferation or apoptosis (Liu *et al.*, 1996; Liu *et al.*, 2005).

The TNFR1 apoptotic pathway is highly relevant to cancer through the process of cell competition, where "winner" cells have increased proliferative capabilities compared to "loser" cells (Moreno *et al.*, 2002). Frequently, wild type (WT) cells outcompete mutant cells; yet, if the mutation is advantageous, the mutant cells will over proliferate. An example

is the proto-oncogene Myc: when inactivated, the *Myc* mutant cells are replaced by the WT tissue; in contrast, activation of *myc* to its oncogenic form provides the mutant cells with competitive advantage (Trumpp *et al.*, 2001). Furthermore, the effect of MYC is long range and WT cells, up to an 8 cell diameter from the MYC expressing cells, activate death receptor mediated apoptosis (de la Cova *et al.*, 2004). *Drosophila* studies have found that MYC activity induces apoptosis in loser cells by activating the JNK pathway, and the expression of a dominant negative form of Basket (BSK), the JNK Drosophila ortholog, prevents cell competition (Liu *et al.*, 2009).

Overall, apoptosis is a complex process that is highly mis-regulated in cancer. Mutant cells that effectively block cell death, while inducing apoptosis to its neighboring cells, through cell competition, acquire a high proliferative and survival advantage.

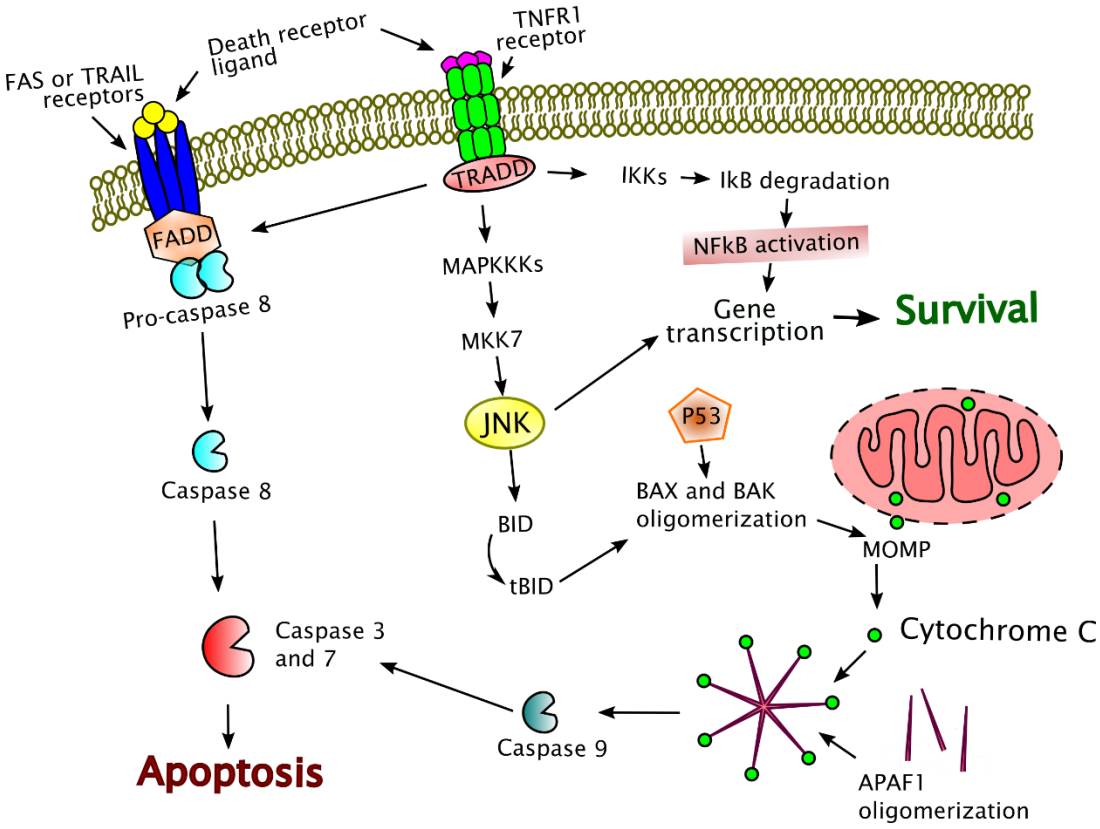


Figure 1.4 Apoptotic pathways

a) Extrinsic or death receptor pathway (left) is activated by extracellular signals. TNFR1 receptors binding to the TNFα ligand triggers three different possible responses. 1) Activation of NFκB that leads to survival. 2) Activation of JNK pathway that could lead to survival or BAX/BAK mediated cell death. 3) Interaction with FADD that leads to caspase 8 mediated apoptosis; FAS and TRAIL receptors use this mechanism of action to trigger apoptosis. **b)** Intrinsic or mitochondrial pathway (right) is triggered by intracellular signals, such as DNA damage. BAX/BAK oligomerize to form the pores leading to mitochondrial outer membrane permeabilization (MOMP), cytochrome C is released from the mitochondria, causing APAF1 oligomerization to form the apoptosome that will activate active caspases 3 and 7 through the initiator caspase 9. Adapted from Kalkavan *et al.*, 2017.

1.2.2.2 Inducing angiogenesis

Vasculature is essential for nutrient and oxygen distribution and is also important for the removal of cellular waste. Blood vessels are formed across the body in an organised pattern. Since cancer growth involves uncontrolled proliferation, accompanied by loss of contact inhibition and multilayering, the mutant cells grow further from the existing vasculature as they continue to proliferate. The acquired ability to form new blood vessels, called angiogenesis, provides these tumour cells with a survival advantage.

The angiogenic switch is normally turned off in adults, meaning that new blood vessels are not formed in developed tissues. Although, angiogenesis can be switched on by cytokine growth factors that activate endothelial cells (Margetts *et al.*, 2002). Adult tissues activate angiogenesis during wound healing and disease, i.e. vascular malformations (Chang *et al.*, 1999), rheumatoid arthritis (Koch, 1998), among others. Vascular endothelial growth factor (VEGF) has central importance in regulating angiogenesis. Studies in mice showed that heterozygous mutations of VEGF resulted in embryonic lethality, characterized by a reduction in vasculature and abnormal blood vessel morphology (Ferrara *et al.*, 1996). Additionally, the relevance of VEGF in tumour progression has been shown *in vivo*, where an overexpression of VEGF is sufficient to induce angiogenesis and positively

regulate *Bcl-2* expression (Nör *et al.*, 1999). Furthermore, VEGF expression and activity has been shown to be positively regulated by RAS (Rak *et al.*, 1995), HER2 (Kumar *et al.*, 2001) and a number of oncogenes and growth factors, such as Platelet-derived growth factor (PDGF) (Anan *et al.*, 1996) and transforming growth factor-β (TGF-β) (Breier *et al.*, 2002). VEGF is also negatively regulated by P53 (Ravi *et al.*, 2000), which, as mentioned previously, additionally regulates angiogenesis through thrombospondin.

1.2.2.3 Deregulating cellular energetics

The process of tumour angiogenesis is not triggered immediately and is not as organised and efficient as developmental angiogenesis; therefore, the mutant cells must additionally find a way to survive in hypoxic conditions. Glucose, the most important source of energy, is normally processed through an aerobic route involving glycolysis, followed by the Krebs cycle that efficiently produces 36 molecules of the biological energy molecule: adenosine triphosphate (ATP) (Alberts *et al.*, 2002). Since the first process, glycolysis, does not involve oxygen, cancer cells re-program their energetics to process glucose without entering the aerobic Krebs cycle (Baggetto, 1992). However, this process is highly inefficient with only 2 ATP molecules being produced.

To compensate, the mutant cells modify glucose uptake and processing mechanisms; for example, by upregulating glucose transporters, notably GLUT1, which substantially increases glucose import into the cytoplasm (Hoskin *et al.*, 2003; Godoy *et al.*, 2006). Additionally, RAS and hypoxia can increase the levels of the HIF1α and HIF2α transcription factors, which in turn upregulate glycolysis (Blum *et al.*, 2005; Kim *et al.*, 2009). Indeed, the adaptation of tumour cells to glycolysis is such that even when they are in the presence of oxygen, they continue to produce energy through the anaerobic pathway, in a process called "aerobic glycolysis" (Gatenby *et al.*, 2004; Lunt *et al.*, 2011).

1.2.3 The immune system and inflammation

The field of cancer immunology is very extensive and is commonly identified as a separate discipline, and therefore beyond the scope of this study to explain fully. For a more extensive introduction to the field, see Rees, 2014. In brief, it is currently accepted that immune surveillance has a negative effect on cancer growth. Studies in mice have shown that immunocompromised animals are more prone to develop aggressive tumours (Street *et al.*, 2002). Additionally, the transplantation of tumours between normal and immunocompromised mice has a critical effect in tumour survival (Smyth *et al.*, 2000). Therefore, in order for tumours to

prosper, tumour surveillance must be evaded. A mechanism that has been studied is the downregulation of the chemokine CXCL14, that attracts natural killer (NK) cells (Cicchini *et al.*, 2016).

A tumour's interaction with the immune system is a very complex one. While tumour cells must evade destruction by the immune system, inflammation and an enrichment of immune cells is often found in tumours, suggesting that the immune cells may play an additional advantageous role. Specifically, macrophages are recruited by IL-4 released by the tumour cells; once they are in the tumour, macrophages release matrix metalloproteinases (MMPs) that promote angiogenesis (Rundhaug, 2005) and catalyse the degradation of the basement membrane (Gocheva *et al.*, 2010), a key step in the process of invasion (see section 1.3.1). Additionally, in breast cancer, macrophages recruited by CSF-1 have been shown to release epidermal growth factor (EGF) that induces proliferation (Pollard, 2004).

While cancer cells are able to develop a collaborative relationship with macrophages; NK cells and the adaptive immune system are exploited for the development of currently successful targeted immune therapies including monoclonal antibodies (reviewed in Weiner *et al.*, 2010) and cancer vaccines (Palucka *et al.*, 2013).

1.2.4 Genomic instability and mutation

As mentioned in section 1.2.1.1, genes involved in cell cycle checkpoint controls are often mutated in cancer, thereby promoting cell cycle progression even if errors were made during DNA replication. An additional critical stage during cell division is anaphase. Even if the chromosomes were replicated with fidelity during S phase, failure to properly distribute them to the daughter cells leads to aneuploidy (changes in chromosome copy number). Aneuploidy is a double edged sword for tumour cells, since it can trigger a high alert and activate the apoptotic machinery through the P53 and/or JNK pathways (Li *et al.*, 2010; Dekanty *et al.*, 2012). In contrast, if the apoptotic pathways are disturbed, aneuploidy can immediately induce a high number of gene alterations. It has been shown that loss of P53 is sufficient for aneuploidy cell survival with high proliferative capabilities (Thompson *et al.*, 2010).

Additionally, the rate of mutation accumulation worsens when the function of DNA repair genes is impaired. Strand exchange protein RAD51 plays an important role in the process of homologous recombination (HR). It binds to the damaged single stranded DNA (ssDNA), and promotes strand invasion to a homologous non-damaged DNA sequence detected during homology identification (Baumann *et al.*, 1998). The DNA repair associated tumour

suppressor genes BRCA1 and BRCA2 have been found to regulate RAD51. Loss of BRCA1 and BRCA2 impairs HR and non-homologous end joining (NHEJ) is activated instead, which is highly prone to errors (Powell *et al.*, 2003). Loss of BRCA1 and BRCA2 is highly associated with breast and ovarian cancer (King *et al.*, 2003). PARP inhibitors, that target single strand break repair and base excision repair, are current treatments for BRCA deficient cancers; by targeting this additional mechanism for DNA repair, cell death is activated in these cells (Murai *et al.*, 2012). Paradoxically, RAD51 overexpression is found in many different types of cancers and it has been found to confer BRCA mutated cells with resistance to PARP treatment (Maacke *et al.*, 2000; Y. Liu *et al.*, 2017), outlining the very fine line between the differential effects of DNA repair gene misregulation and cancer.

1.3 Activating invasion and metastasis

The numerous different biological processes outlined above are essential for tumour growth and survival, yet a tumour is not considered a cancer until it colonizes further than its tissues of origin, through the process of invasion and metastasis.

1.3.1 Invasion

Invasion involves the detachment of cells from the primary tumour, movement through the basement membrane and extracellular matrix (ECM), until the cell reaches and infiltrates lymphatic ducts, blood vessels or cavities, which will become the route they will use to move to distant tissues. As previously mentioned, carcinomas are by far the most common tissues to develop tumours, and for that reason, epithelial cells have been extensively used as a model for invasion and metastasis.

1.3.1.1 Epithelial cell polarity

Epithelial tissues are widely distributed throughout the body; they form the lining of external and internal surfaces of organs and cavities (Alberts *et al.*, 2002). Structurally, their plasma membrane is comprised of three main domains: i) an apical domain that faces the external environment or lumen;

ii) a lateral domain, containing adherent (cadherin based) and tight junctions that form the structural link for epithelial cells in a tissue; iii) and a basal domain, that is in close contact with the basement membrane, that divides the epithelial and connective tissues (Rodriguez-Boulan *et al.*, 1989) (Figure 1.5).

Epithelia are highly organised and specialized tissues. In addition to providing a protective barrier, they are also involved in secretion, absorption and filtration. For their proper function, epithelial cells have a polarized structure and an asymmetrical internal distribution of molecules (Matter *et al.*, 1994). The maintenance of this cell polarity is given by the cross-talk between small GTPases and a network of polarity proteins. Small GTPases have been shown to be involved in the rearrangement of the actin cytoskeleton of the cell and in the maintenance of cell polarity. The best-characterized of these molecules are the members of the Rho family GTPases, which control the formation of stress fibres, cellular protrusions and focal adhesion (Hall, 1998). This includes RHO1, that regulates stress fibre formation, focal adhesions and Myosin II activation (Amano *et al.*, 1996, 1997); CDC42 that is related to filopodia formation (Krugmann *et al.*, 2001); RAC, which induces membrane ruffling (Ridley *et al.*, 1992) and its activity levels regulate protrusion form and position (Couto *et al.*, 2017).

Three protein complexes are directly involved in maintaining apical-basal epithelial cell polarity: The Scribble complex (SCRIB, DLG and LGL), located in the basolateral domain, are thought to behave as scaffold proteins and regulate protein–protein interactions between small GTPases and other polarity complexes (Enomoto *et al.*, 2011); the PAR complex (BAZ, PAR6 and aPKC) and the Crumbs complex (CRB, PALS1 and PATJ) are restricted to the apical domain (Margolis *et al.*, 2005). These three protein complexes maintain their localization on an epithelial cell through mutual exclusion of the basolateral Scribble complex with the apical PAR and Crumbs complexes (Yamanaka *et al.*, 2006); and the exclusion of the Crumbs complex by the basal Yrt complex (Laprise *et al.*, 2009) (Figure 1.5).

Modifications of epithelial cell polarity have been associated with tumour progression; the loss of apical-basal polarity and a gain of front-rear polarity has been found to be a key step in the process of epithelial to mesenchymal transition (see Section 1.3.1.3) (Nelson, 2009). An example of this is that the three components of the scribble complex have been identified to act as tumour suppressor genes in both invertebrates and vertebrates (Zhan *et al.*, 2008).

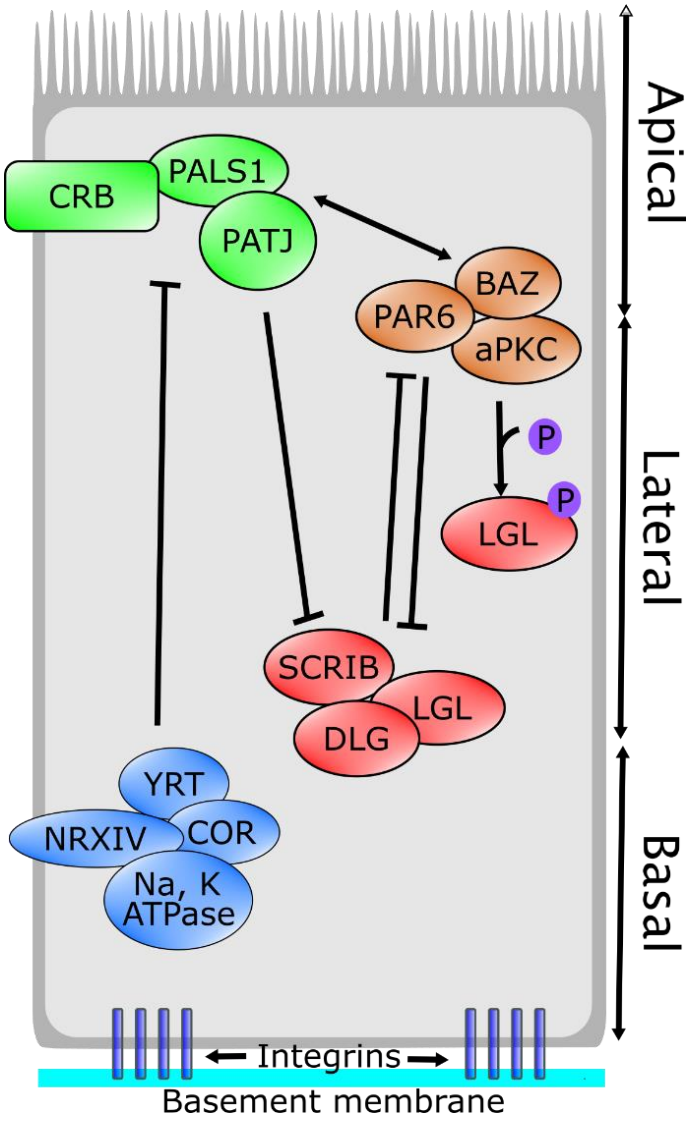


Figure 1.5 Epithelial cell apical-basal polarity

A polarized *Drosophila* epithelial cell containing three membrane domains: apical, lateral and basal. Four polarity complexes are shown: The two apical polarity complexes: Crumbs (CRB, PALS1 and PATJ) and PAR (PAR6, BAZ and aPKC); the basolateral complex, Scribble (SCRIB, DLG and LGL); and the basal Yrt complex (YRT, NRXIV, COR and Na,K-ATPase). The apical-basal polarity of the cell is mantained by mutual exclusion of the Scribble complex and the two apical complexes; and the exclusion of the Crumbs complex by the basal Yrt complex. Adapted from Martin-Belmonte *et al.*, 2011.

1.3.1.2 Cell adhesion

Adjacent cells are bound together, and to the extracellular matrix, by transmembrane and intracellular scaffolding proteins that regulate cell adhesion. Different protein complexes localized across the cell membrane, from the apex to the basal domain, mediate cell-cell adhesion: i) In vertebrates: tight junctions, adherens junctions, desmosomes and gap junctions (Kumar *et al.*, 1996; Garrod *et al.*, 2002; Niessen, 2007) (Figure 1.6) ii) In invertebrates: adherens junctions and septate junctions (the functional equivalent of tight junctions) (Peifer *et al.*, 1992; Willott *et al.*, 1993; Médina *et al.*, 2002) (Figure 1.6). Cell-extracellular matrix adhesion is regulated by focal adhesions (integrins). Integrin mediated adhesion is increased in various human cancers and positively correlated with invasion, as it provides traction for cell motility (Hood *et al.*, 2002). Additionally, some integrins promote growth factor and cytokine signals to promote growth and invasion (Arboleda *et al.*, 2003).

Adherens junctions, mediate intercellular adhesion via members of the cadherin/catenin family of proteins and their connections to the actin cytoskeleton (Figure 1.6b) (Mège *et al.*, 2006). Decreased E-cadherin expression is found in lung, breast, pancreatic, bladder, prostate, ovarian and other cancers (Umbas *et al.*, 1994; Guilford *et al.*, 1998; Hajra *et al.*, 2002; Witta

et al., 2006) and correlates with epithelial to mesenchymal transition (see section 1.3.1.3), reduced intercellular adhesion and increased tumour invasiveness (Cano *et al.*, 2000). Therefore, adherens junctions are widely studied in cancer research.

In addition to mutation, decreased E-cadherin expression and function is affected by other signalling pathways. For example, the zinc finger transcription factors TWIST, ZEB and SNAIL repress E-cadherin expression. SNAIL binds to the E-box near the promoter of E-Cadherin and recruits histone modificators that induce silencing methylation marks, preventing E-cadherin transcription (Batlle *et al.*, 2000). SNAIL can additionally induce histone acetylation and activate genes, such as the TWIST and ZEB transcription factors; matrix metalloproteinases (MMP) and N-cadherin; all important markers of epithelial to mesenchymal transition (Wang *et al.*, 2015) (see section 1.3.1.3). TWIST represses E-cadherin expression by inducing histone methylation independently of SNAIL. TWIST recruits SET8 that mediates the specific methylation of H4K20, this mark is also associated with N-cadherin activation (Vesuna *et al.*, 2008). ZEB acts similarly: it also binds to the E-box of E-cadherin and negatively regulates its expression through the C-terminal binding protein (CTBP) co-repressor (Romero-Pérez *et al.*, 2013).

Activation of EGFR promotes AKT mediated phosphorylation of β- catenin, which prevents β- catenin from interacting with E-cadherin, thereby promoting the disassembly of the adherens junction (Fang *et al.*, 2007; Lee *et al.*, 2010). Additionally, phosphorylated β- catenin re-localizes to the nucleus, where it interacts with the TCF family of transcription factors and can promote tumour survival, proliferation and progression towards malignancy (Behrens *et al.*, 1996; Schneider *et al.*, 1996; Moon *et al.*, 2004). Increased β-catenin nuclear localization is also observed when E-cadherin is lost (Alvarado *et al.*, 2011). Activation of the oncogene RAS has also been shown to disrupt adherens junctions by the phosphorylation of β-catenin and p120ctn, another member of the adherens junction complex (Figure 1.6) (Kinch *et al.*, 1995). Rho small GTPases also have a role in adhesion regulation. Since adherens junctions are bound to the actin cytoskeleton, a disruption or modification of the cytoskeleton will impact the integrity of the cadherin-catenin complex (Fukata *et al.*, 2001). RHO-C, RHO-A, RAC and CDC42 are frequently found upregulated in breast, testicular and other cancer tumours (van Golen *et al.*, 2000; Kamai *et al.*, 2004).

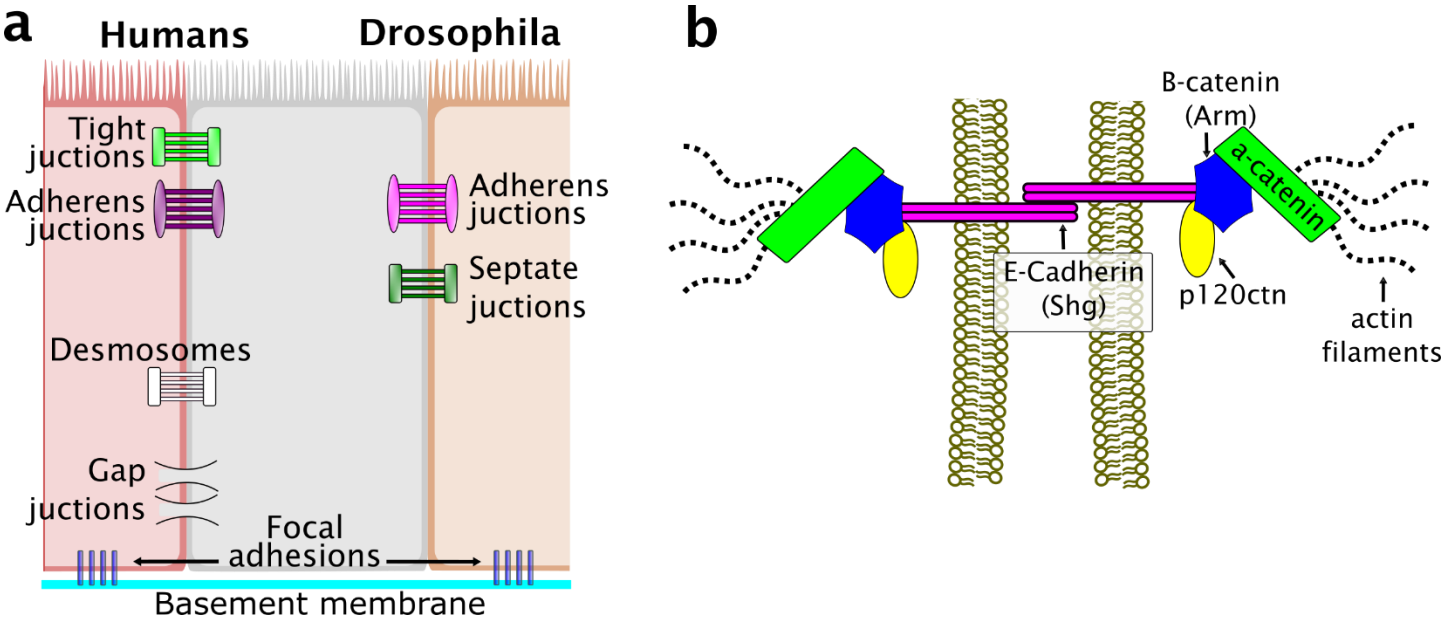


Figure 1.6 Cell adhesion proteins

a) Schematic of the protein complexes that mediate cell-cell and cell-matrix adhesion in humans and *Drosophila***.** Adhesion proteins in humans. Apical complexes: Tight and adherens junctions. Lateral: Desmosomes. Basal: Gap junctions. Cell-matrix complexes: Focal adhesions. Cell-cell Adhesion proteins in *Drosophila*. Apical: Adherens junctions. Lateral: Septate junctions, functional equivalent of tight junctions. **b)** Magnification of adherens junction proteins. Transmembrane E-cadherin mediated adhesion is linked to F-actin fibres by intracellular interactions with catenin proteins and to the microtubules through p120ctn. Adapted from Humbert *et al.*, 2003.

1.3.1.3 Single cell mesenchymal invasion

Epithelial to mesenchymal transition (EMT) is an intricate process resulting in the loss of epithelial characteristics and the gaining of a number of mesenchymal characteristics. This includes the loss of intercellular junctions, apical-basal polarity, epithelial cell morphology and the acquisition of front-rear polarity and cell motility (Huber *et al.*, 2005; Mani *et al.*, 2008) (Figure 1.7). EMT is a common feature during development and wound healing but when a cancerous cell undergoes EMT it also acquires an increased ability to invade (Yang *et al.*, 2008; Yan *et al.*, 2010).

The cytokine TGF-β has been described to be highly involved in EMT. TGF-β phosphorylates the polarity protein PAR6A and this causes loss of apical-basal polarity, a hallmark of EMT (Ozdamar *et al.*, 2005). Additionally, in its phosphorylated form, TGF-β can coordinate the expression of SNAIL, ZEB and TWIST transcription factors that act by repressing transcription of *Cdh1*, the coding gene for E-cadherin, contributing to the loss of adherens junctions and consequent cell-cell adhesion. (Medici *et al.*, 2006; Peinado *et al.*, 2007; Naber *et al.*, 2013). As mentioned in section 1.3.1.2, SNAIL is also involved in the transcriptional activation of N-cadherin; therefore, N-cadherin expression is found in mesenchymal cells instead of E-cadherin (Zhang *et al.*, 2013).

The extracellular matrix (ECM) is a key player in EMT, once the process of transition is initiated, the ECM releases proteins, such as type I collagen, that stimulate the mesenchymal phenotype (Gilles *et al.*, 1997; Medici *et al.*, 2010). Additionally, integrin-linked kinases co-operate with TGF-β to promote EMT (Serrano *et al.*, 2012). Furthermore, mesenchymal cells contribute to ECM remodelling through the production of matrix metalloproteinases (MMPs), enzymes that are well known for their matrix degrading activity and have been highly implicated in the process of invasion (Radisky *et al.*, 2010).

1.3.1.4 Single cell amoeboid invasion

Amoeboid invading cells are rounded, and their shape is governed by the activation of the RHO1 small GTPase and its downstream effectors RHO kinases (ROCKs), that induce cytoskeleton re-arrangements and activate the contractile protein non-muscular myosin II (Lawson *et al.*, 2018). Their movement is achieved in the same way as the amoeboid unicellular organism: by the formation of pseudopodia, a temporary protrusion that is established by the elongation of the membrane through RAC dependent actin filament formation (Yamazaki *et al.*, 2009; Tolde *et al.*, 2018). In addition to their shape, ameboid invading cells differ from mesenchymal cells in the absence of integrin expression and of contact with the ECM (Carragher *et al.*,

2006). Furthermore, amoeboid cells do not express MMPs or other proteolytic enzymes (Sabeh *et al.*, 2009) (Figure 1.7). Amoeboid cells are also capable of transitioning to the mesenchymal shape and vice versa, in response to signals from the microenvironment (Talkenberger *et al.*, 2017). Therefore, malignant invading cells are able to modify their gene expression and phenotype. Amoeboid invading cells have been observed in lymphoma, melanoma, small cell lung, prostate, and breast cancer (Carragher *et al.*, 2006; Gadea *et al.*, 2008; van Zijl *et al.*, 2011).

1.3.1.5 Collective invasion

Collective cell migration occurs when two or more cells migrate together, unlike single cell migration where the individual migrating cell breaks away from the tissue of origin. Cells within the invading cluster retain epithelial characteristics, the most distinctive are: the expression of junctional proteins and maintenance of cell-cell contact (Cheung *et al.*, 2013) (Figure 1.7).

Cell cluster movement is directed by actomyosin flows and protrusions, similarly to individual cell invasion. The most important difference is that this flow has to be orchestrated multicellularly (Alexander *et al.*, 2008; Shindo *et al.*, 2014). Cells within the invading cluster achieve the coordinated movement by expressing different proteins in the front 'leader' cells than in

the 'follower' cells. Leader cells have been shown to adopt a mesenchymal or cancer associated fibroblast (CAF) form (Gaggioli *et al.*, 2007). Regardless of the adopted cell type, leading cells are thought to have matrix remodelling capabilities, and integrin mediated traction to allow for directed movement and ECM degradation (Khalil *et al.*, 2010). Followers have a more tightly packed form and communicate with leaders to establish their directionality (Khalil *et al.*, 2010). Although, recent research suggests that follower cells do not simply obey instructions and are able to form their own tracks and have a more active role in the movement of the cell cluster (Theveneau *et al.*, 2017).

Collective cell invasion is found to be the most common mechanism for invasion in many different types of cancer including oral squamous cell carcinoma, colorectal carcinoma, melanoma and breast cancer; where circulating clusters of cells, that retain their cell-cell contacts are commonly found (Nabeshima *et al.*, 1999; Hegerfeldt *et al.*, 2002; Gaggioli *et al.*, 2007). This type of invasion is considered the most aggressive, since cancer cells gain all the phenotypes related to invasion (Figure 1.7). Furthermore, unlike single cell invasion, collective invading cells are not susceptible to anoikis (Kim *et al.*, 2017), the trigger of apoptosis in the absence of cell-cell contacts.

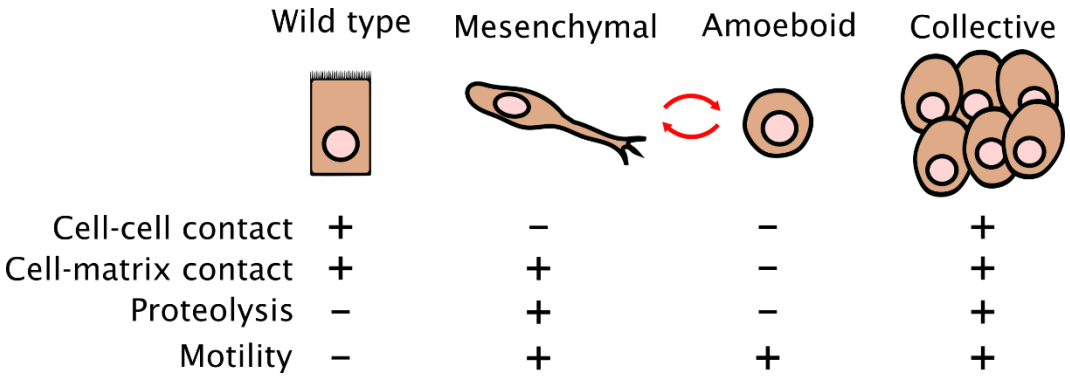


Figure 1.7 Types of cell invasion

Characteristics of the different types of invading cells and how they differ from WT non-invading cells. WT cells have strong cell-cell contact and cell-matrix contact (Alberts *et al.*, 2002). Mesenchymal single invading cells lose cell-cell contact and gain proteolytic and motility abilities (Huber *et al.*, 2005). Amoeboid single invading cells lose all epithelial characteristics, yet they gain high motility and are able to switch to mesenchymal type, and vice versa, in response to signals from the microenvironment (Talkenberger *et al.*, 2017). Collective invasion is considered the most aggressive type, cells retain all epithelial phenotypes and additionally gain motility and proteolytic capabilities (Cheung *et al.*, 2013).

1.3.2 Metastasis

The invasion-metastasis cascade model describes the events a cell or group of cells must follow to achieve the formation of secondary tumours (metastases). Invading cells, after successfully crossing the basal membrane, infiltrate the blood vessels through the process of intravasation (Zervantonakis *et al.*, 2012) (Figure 1.8). It is not fully clear how tumour cells achieve intravasation. Studies have shown that MMPs and other proteases

are required for tumour cells to enter the blood vessels (Kim *et al.*, 1998). Additionally, it is suggested that angiogenesis is a pre-requisite for intravasation, since newly formed tumour blood vessels are more permeable and might provide an easier route for invading cells (Deryugina *et al.*, 2015). Cell survival in the blood vessels is a difficult task to achieve. The first contact of invading cells with the vasculature are the capillaries, blood vessels with the smallest diameter, and most cells are torn apart by hydrostatic pressures (Wong *et al.*, 2014). Therefore, only cells that have high elasticity and cytoskeletal plasticity will survive and reach bigger vessels. Additionally, invading cells are exposed to the immune system. NK cells and monocytes eliminate the majority of them (van Zijl *et al.*, 2011). Some circulating tumour cells form contacts with blood cells, tumour cells within the vasculature have been associated with neutrophils, platelets and fibrin (Bekes *et al.*, 2011; Lou *et al.*, 2015). The invading cells that successfully form those interactions have a higher probability of survival (Figure 1.8).

Invading cells travelling in the blood eventually escape through the process of extravasation and travel to the parenchyma of a tissue (Kramer *et al.*, 1979; Bersini *et al.*, 2014) where they can begin the most difficult stage, namely colonization. A migrating cancer cell, once it has reached a secondary site, has to adapt significantly to the new microenvironment that it finds itself in.

It is believed that after extravasation, the mesenchymal-like migrating cancer cell does not have the stromal support that helped to stimulate the mesenchymal state at the primary tumour site. The absence of this stromal signalling leads to a mesenchymal to epithelial transition (MET) (Brabletz *et al.*, 2005). MET is an advantage, since it facilitates cancer invading cell insertion into an epithelial tissue at the secondary site and promotes the formation of a micrometastasis (Gunasinghe *et al.*, 2012). The process of micrometastasis formation is highly inefficient and tumour cells might stay dormant for years, as they may need to acquire new mutations to adapt to the secondary site. They also need to remodel the new local microenvironment to make it supportive of their survival and growth (Luzzi *et al.*, 1998; Wikman *et al.*, 2008). It has also been suggested that tumour cells have the ability to send signals, in the form of cytokines and chemokines, to the future site of metastatic growth (Peinado *et al.*, 2017). Therefore, tumour cells are able to prepare the microenvornment before their arrival. Specifically, upregulation of the extracellular protein, fibronectin (Kaplan *et al.*, 2006), has been observed. Fibronectin is associated with bone marrow progenitor cells, expressing vascular endothelial growth factor receptor 1 (VEGFR1) and VLA-4 integrin that are highly associated with the formation of tumour metastasis through matrix degradation and tumour associated adhesion, respectively (Kaplan *et al.*, 2005).

Once the cells have accumulated advantageous mutations, they will be able to survive and continue to proliferate in their secondary site, until they reach the size of macrometastasis (approximately 2mm in size). Cancer cells that have successfully adapted at the secondary site have a great survival advantage, dissemination is more efficient and new micrometastasis can rapidly form to colonize distant parts of the body (Chambers *et al.*, 2002). This process exponentially affects the aggressiveness of the disease, with metastasis accounting for 90% of all cancer related deaths (Seyfried *et al.*, 2013). Therefore, a better understanding of this process is essential to provide novel therapeutic targets.

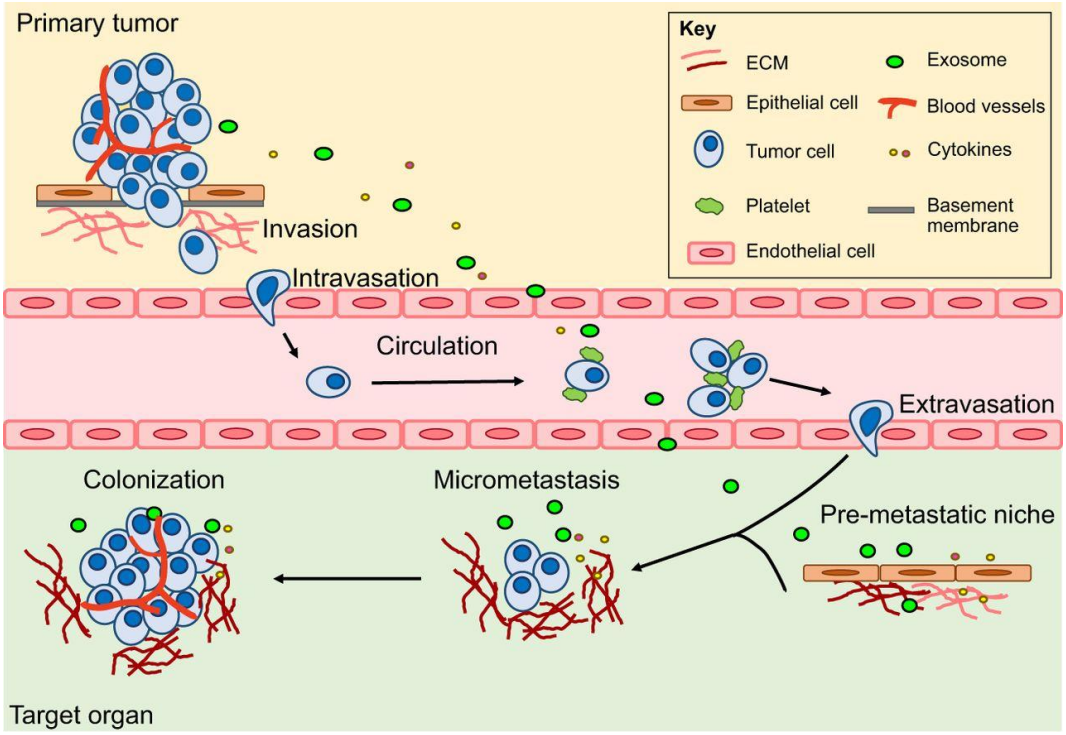


Figure 1.8 Invasion-metastasis cascade

The process of metastasis begins with a cell or a group of cells crossing the basement membrane (invasion). Once the cells enter the blood vessels, through intravasation, they can travel long distances. Association with blood cells (such as platelets) can sometimes provide survival advantage to the tumour cells through immune system evasion. It has also been proposed that the tumour cells can send cytokine and chemokine signals to the future secondary site, preparing the microenvironment and directing the invading cells. Eventually, some circulating invading cells will extravasate from the blood vessel to a new location. Invading cells start to grow in their new location, forming micrometastasis and eventually form a secondary tumour, or macrometastasis, through the process of colonization. Figure obtained from Gómez-Cuadrado *et al.*, 2017.

1.4 *Drosophila* as a model for cancer research

Use of the fruit fly as a model for cancer research is becoming increasingly popular. Observations from this model are of great interest since the effect of alterations in tumour suppressor genes and oncogenes can be studied in the living animal and with high spatial and temporal resolution. In addition to that, the *Drosophila* genome has high homology with more than 60% of cancer related genes (Potter *et al.*, 2000).

Drosophila has a life cycle of ten days, which means that it is possible to obtain many fly generations in a short period of time, allowing for complex genetic studies. Due to its small size, it is easy to keep large populations of flies in a limited space rendering *Drosophila* studies very cheap compared to other organisms. Currently, as the genome of *Drosophila* has been fully sequenced and due to its high conservation with the human genome (Adams *et al.*, 2000), a wide variety of research fields use *Drosophila* as a model, ranging from developmental and evolutionary biology to neurodegenerative disease and cancer research.

Many different precise techniques to genetically manipulate *Drosophila's* gene expression are available. For example, transposable elements, named P elements, are endogenously present in the fly's genome and have been used

for decades for the insertion of a wide range of transgenes (Rubin *et al.*, 1983; Robertson *et al.*, 1988). Targeted insertions are also available, such as the use of different vectors to direct insertions to specific loci (Horn *et al.*, 2005).

Sophisticated techniques adapted from yeast have been successfully introduced to *Drosophila*, such as the UAS/Gal4 system, that allows tissue specific transgene expression (Brand *et al.*, 1993); the Flp/FRT system, that induces site specific recombination in mitosis (Theodosiou *et al.*, 1998); and the mosaic analysis with repressible cell marker (MARCM) system, that combines the Flp/FRT with the UAS/Gal4 system, and Gal80 (repressor of Gal4) (Lee *et al.*, 1999). MARCM allows the creation of positively labelled homozygous mutant cells in an otherwise heterozygous *Drosophila* (see Chapter 2, section 1.2.2 for more detail). Recently, methods for site specific mutagenesis, such as CRISPR-Cas have been successfully introduced to *Drosophila*, increasing the available tools to manipulate the *Drosophila* genome (Gratz *et al.*, 2013).

Drosophila models have provided important contributions to the understanding of the mechanisms of tumour progression. The Notch signalling pathway was identified a century ago in the fly (Moohr et al., 1919). Crosstalk between Notch and Ras signalling pathways has been

described extensively in different species including *C. elegans* and humans (Carmena et al., 2002; Weijzen et al., 2002; Mittal et al., 2009). Currently, Notch is a well characterised oncogene involved in cell proliferation and survival (Go et al., 1998).

Additionally, our understanding of the co-operation between oncogenes and TSGs has been enriched by the use of *Drosophila* models. The co-operation between the oncogenic protein RAS^{v12} and loss of function mutations of the TSG *scrib* was studied in the *Drosophila* eye-antennal discs (Ming Wu *et al.*, 2010). Their findings show that clones expressing RAS^{v12} have increased proliferation, a well-known effect of increased Ras activity (Karim *et al.*, 1998). Clones mutant for $scrib^{-/-}$ lose apico-basal polarity and do not survive (Menéndez *et al.*, 2010). In contrast, $scrib^{-/-}$ clones simultaneously expressing RAS^{V12} grow into large tumours. Furthermore, $Ras^{V12}/scrib^{-/-}$ cells metastasised to the ventral nerve cord (Ming Wu *et al.*, 2010).

The Salvador/Warts/Hippo (SWH) signalling pathway is another important mechanism first identified in *Drosophila* (Xu *et al.*, 1995). The effector oncoprotein of this pathway is Yorkie (YKI), a transcriptional coactivator of genes involved in cell proliferation and survival (Huang *et al.*, 2005). The tumour suppressor protein HIPPO phosphorylates YKI leading to its

inactivation (Udan *et al.*, 2003). Studies in the Drosophila eye have shown that loss of the TSG *Lgl* leads to HIPPO mis-localization and un-phosphorylated YKI. Therefore, loss of *Lgl* has effects in tumour over proliferation, in addition to cellular apico-basal polarity defects (Grzeschik *et al.*, 2010). The human orthologue of *Lgl*, *Hugl1*, has been linked to a variety of cancers, including colorectal (Schimanski *et al.*, 2005), hepatocellular (Lu *et al.*, 2009), oesophageal (Song *et al.*, 2013), melanoma (Kuphal *et al.*, 2006), and others.

1.5 Research aims

The process of tumour progression is very complex and involves the alteration of several biological processes. It involves the mis-regulation of numerous tumour suppressor genes and oncogenes. Current sequencing techniques have allowed whole-genome sequencing of tumour biopsies from numerous cancer types and many genes have been found to be mutated in a significant number of patients with different carcinomas. However, many of these genes are uncharacterized and their potential role in tumour progression is unknown. To address this problem, we developed an *in vivo* model using *Drosophila melanogaster*, where we were able to generate positively labelled tumours with a known genotype on the back of the fly. We then studied the effect that the knock down (KD) of a specific

uncharacterised gene has on a variety of phenotypes related to tumour progression, such as tumour size, the amount of cell division, cell morphology, protrusion length, cell invasion, polarity defects, multilayering, and others.

The specific aims of this project are:

I. To screen through 500 individual genes and determine the effect that their RNAi mediated KD has on tumour progression related phenotypes.
II. To identify genes that promote or inhibit tumour progression phenotypes.
III. To select genes of interest for further characterization.
IV. To characterize the selected genes in more detail to better understand their mechanism of action and validate the genetic screen.

Chapter 2 - Materials and methods

2.1 *Drosophila* materials and methods

2.1.1 *Drosophila* materials

2.1.1.1 General reagents

Table 2.1 General reagents.

Reagent	Supplier
Agar pure powder	Thermofisher Scientific
Bovine serum albumin (BSA)	Sigma
BSA	Sigma
Cornmeal	ASDA
Cover slip 22mm x 22mm	Thermofisher Scientific
Cover slip 22mm x 50mm	Thermofisher Scientific
Dried yeast	DCL Yeast
Foetal bovine serum (FBS)	Thermofisher Scientific
Fluorogel mounting medium	Source Bioscience UK
Glycerol	Sigma
Glycerol	Sigma
Golden syrup	ASDA
Microscope slide plain	Thermofisher Scientific
Oil 10 S VOLTALEF	VWR
Paraformaldehyde (PFA)	Thermofisher Scientific
Paraformaldehyde (PFA)	Thermofisher Scientific
PBS	Sigma
PBS	Sigma
Propionic acid	Sigma
Soya flour	ASDA
Tween 20	Sigma
Tween 20	Sigma

2.1.1.2 *Drosophila* stocks

All *Drosophila* RNAi lines were obtained from Vienna *Drosophila* Stock Centre (VDRC) in Austria, National Institute of Genetics (NIG) in Japan or Bloomington *Drosophila* Stock Centre (BDSC) in USA. *Drosophila* RNAi lines used in the genetic screen (Chapter 3) are listed in Appendix 2.

Table 2.2 *Drosophila* stocks

Stocks used were either bought commercially from BDSC or previously made in the Georgiou lab (Couto *et al.*, 2017).

Stock	Genotype	Source
1744	y, w; FRT19A	BDSC
57065	w, FRT19a, SMC3a/ +/ +	BDSC
57066	w, FRT19a, SMC3b/ +/ +	BDSC
59024	w; +; sqh-mCherry	BDSC
64186	w; sna[Sco]/CyO; UAS-iCasper/TM6B	BDSC
M23	w; tubGal80, FRT40A; MKRS/TM6B	Georgiou Lab
M34	w; Tub-Gal80, FRT40A; +	Georgiou Lab
M34.5	w, Ubxflp, tub-Gal80, FRT19a; +; PnrGal4, UAS-GFP:Moe/ TM6B	Georgiou Lab
M413	w, UbxFlp; FRT40A; PnrGal4, UAS-GFP:Moe/TM6B	Georgiou Lab
M519	w, UbxFlp; CyO/Kr^{If-1}; TM6B/MKRS	Georgiou Lab
M563	w, FM7; +; TM6B/MKRS	Georgiou Lab
M58:10	w, UbxFlp; lgl4, FRT40A/CyO-GFP; PnrGal4, UAS-GFP:Moe/TM6B	Georgiou Lab
M826	w, FM7; CyO-GFP/ Kr^{If-1}; +	Georgiou Lab
M94	w; ubi-HisRFP/CyO; +	Georgiou Lab

Table 2.3 UAS-transgenes in Chapter 4 and 5

List of *Drosophila* UAS-RNAi lines (for gene KD) and UAS-CDS (for gene overexpression) used in Chapter 4 and 5 (UAS-transgene in section 2.2.3). For UAS-transgenes used in Chapter 3 refer to Appendix 2.

Target gene	RNAi or UAS-CDS	Stock	Source
SA1	RNAi	3423R-2	NIG
SA1	RNAi	3423R-3	NIG
SA1	RNAi	33395	BDSC
SA2	RNAi	108267	VDRC
SA2	RNAi	51569	VDRC
SMC1	RNAi	6532	VDRC
SMC1	RNAi	6057R-1	NIG
SMC3	RNAi	39207	VDRC
SMC3	RNAi	39205	VDRC
RAD21	RNAi	36786	BDSC
RAD21	RNAi	65229	BDSC
NipB	RNAi	32406	BDSC
Sqh	RNAi	32439	BDSC
Tkv	RNAi	31041	BDSC
Mad	RNAi	31316	BDSC
Cut	RNAi	29625	BDSC
Shrm	RNAi	40942	BDSC
hh	RNAi	31042	BDSC
Fog	RNAi	36790	BDSC
RhoV14	UAS	8140	BDSC
RhoN	UAS	7328	BDSC
SqhEE	UAS	M201	Georgiou Lab
Tkv	UAS	36537	BDSC
Mad	UAS	44256	BDSC
Cut	UAS	36496	BDSC
Shrm	UAS	59845	BDSC
p35	UAS	5072	BDSC

Table 2.4 Mutant alleles

Source and mutation type for the three mutant alleles used in this research.

Genotype	Abbreviation	Description	Citation
l(2)gl^{4}	*lgl*^{4}	Loss of function allele, Deletion of sequences from 0 to 22kb	(Mechler *et al.*, 1985)
sm3c^{a}	*smc3*^{a}	Loss of function allele, K575stop	(Haelterman *et al.*, 2014)
smc3^{b}	*smc3*^{b}	Loss of function allele, Q146stop	(Haelterman *et al.*, 2014)

2.1.1.3 Antibodies

Table 2.5 Primary Antibodies

Antibody	Dilution	Source	Catalogue number
rat anti-E-Cadherin	1:100	DSHB	DCAD2
mouse anti-Armadillo	1:100	DSHB	N27A1
rat anti- α-Catenin	1:100	DSHB	DCAT-1
mouse anti-Fasciclin III	1:100	DSHB	7G10
mouse anti-Coracle	1:100	DSHB	C566.9
mouse anti-Discs Large	1:100	DSHB	4F3
rabbit anti-aPKC	1:50	Santa Cruz	sc-216-G
mouse anti-Cut	1:100	DSHB	2B10
rabbit anti-pMad	1:250	Abcam	ab52903

Table 2.6 Secondary Antibodies

Antobody	Dilution	Source
Alexa Fluor ® 546 goat anti-mouse	1:2000	Molecular Probes
Alexa Flour ® 546 goat anti-rabbit	1:2000	Molecular Probes
Alexa Flour ® 546 goat anti-rat	1:2000	Molecular Probes
Alexa Fluor ® 647 goat anti-mouse	1:2000	Molecular Probes
Alexa Flour ® 647 goat anti-rabbit	1:2000	Molecular Probes
Alexa Flour ® 647 goat anti-rat	1:2000	Molecular Probes

2.1.2 *Drosophila* methods

2.1.2.1 Fly pushing

Flies were kept in vials containing standard cornmeal medium (7340ml H20, 565ml golden syrup, 127g yeast, 73g soya flour, 536g cornmeal, 42g agar and 35.4ml propionic acid) and stored in incubators at 25°C (for a development of approximately 10 days) or at 18°C (for a development of approximately 20 days) (Dillon *et al.*, 2007). To maximise progeny and evade mould/mite infestation vials were flipped (transferred into new vials) weekly for those at 18°C or twice a week when kept at 25°C.

For observation, flies were anaesthetised using a CO^2 porous polyethylene gas diffuser and viewed under a Leica M60 microscope or a Leica MZ10F fluorescence microscope for flies expressing GFP or RFP. Balancer Chromosomes (homozygous lethal alleles due to chromosome re-arrangement with a dominant visible phenotype) were used to suppress meiotic recombination that naturally occurs in females and to track genotypes through genetic crosses (see section 2.2.3).

2.1.2.2 Genetic techniques

Commercially bought lines, listed in section 2.1.1.2 and Appendix 2, carry different transgene insertions, such as UAS-RNAi lines, UAS-CDS, lines that

express fluorescently labelled fusion proteins, among others. Insertion of the transgenes to the *Drosophila* genome is done by genetic manipulation of P-elements, transposons naturally expressed in the fly (Rubin *et al.*, 1983; Dietzl *et al.*, 2007). Modified transposons are microinjected into the *Drosophila* embryo, their integration to the genome and successful expression is monitored by markers within the P element vector. The most common marker is the mini white gene that restores pigmentation to the eye in an otherwise white eyed mutant (Silicheva *et al.*, 2010).

Drosophila experiments in this research required the combination of different sophisticated genetic manipulation techniques. A brief description of these techniques is included in sections 2.2.2.1 – 2.2.2.4.

2.2.2.1 Gal4/UAS system

Galactosidase 4 (Gal4) was identified in the yeast *Saccharomyces cerevisiae* as a regulator of genes, it recognises and binds to an upstream activation sequence (UAS) driving transcription of the UAS-tagged genes (Brand *et al.*, 1993), this system has been widely used for the expression of RNAi lines for tissue specific KD and fluorescent reporters allowing the observation and analysis of many different phenotypes (Figure 2.1).

Gal4 can be engineered under the control of different promoters in *Drosophila*, such as pannier (pnr) which restricts Gal4 activity to the central region of the notum (Heitzler *et al.*, 1996); or neuralized (neu), that drives expression in the sensory organ precursors (SOP) (Venugopala Reddy *et al.*, 1999). Before committing to differentiation, SOP cells behave like epithelial cells, allowing transgene expression in single epithelial cells in the back of the fly. Through the use of many different promoters, the Gal4/UAS system enables transgene expression in specific cell types at a particular time point during development.

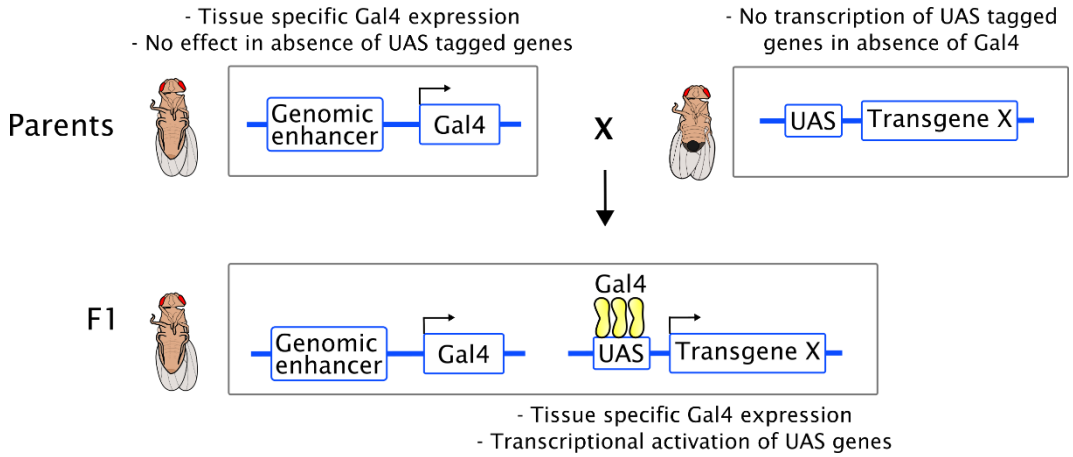


Figure 2.1 Gal4/UAS system

Parental flies expressing the Gal4 are crossed to flies with the upstream activating sequence (UAS) linked to the gene of interest. In the progeny of the cross, the Gal4 binds to the UAS sequence leading to the transcriptional activation of the transgene X (Brand *et al.*, 1993).

2.2.2.2 RNAi mediated KD

In vivo targeted RNAi has quickly become a powerful and widely used tool for loss of function studies. Double-stranded RNA (dsRNA) hairpins used in this study 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 and loss of function (Okamura *et al.*, 2008) (Figure 2.2).

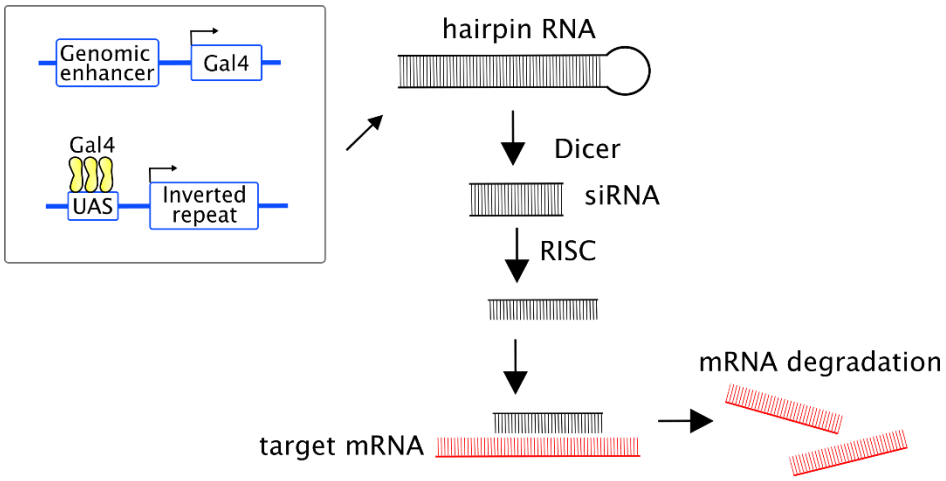
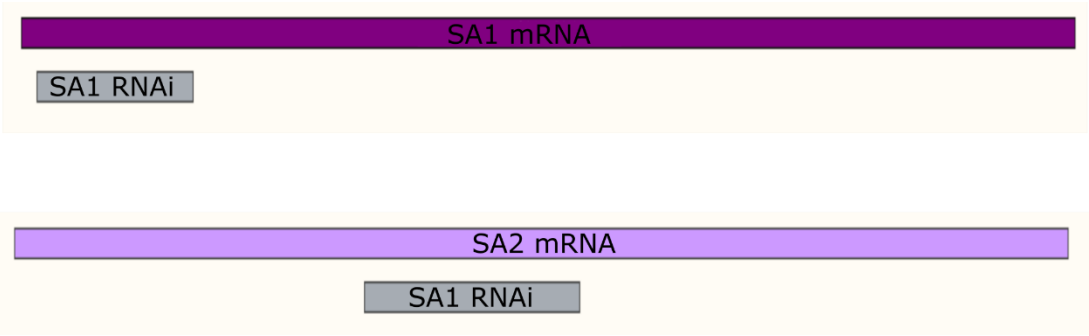


Figure 2.2 UAS-RNAi system

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 complementary strands results in the degradation of mRNA and subsequent gene silencing (Dietzl *et al.*, 2007).



siRNA cut sites in *Drosophila* Cohesin subunits mRNAs

Schematic representing the mRNA for the *Drosophila* Cohesin complex subunits and the specific target sites of the RNA interference sequences.

2.2.2.3 FLP/FRT system

The flippase (FLP)/ flippase recognition target (FRT) is a yeast site-directed recombination system that has been adapted for the use in *Drosophila*; the enzyme FLP recognizes identical FRT sites located on homologous chromosomes during the G2 phase of the cell cycle and induces recombination at those sites (Theodosiou *et al.*, 1998). The system allows the formation of clones by generating one homozygous mutant and one homozygous WT daughter cell as a mitotic product of a heterozygous cell (Figure 2.3). Flippase activity can be restricted by different promoters, such as Ultrabithorax (Ubx), which restricts the induced recombination process to the dorsal thorax of the fly (Bienz *et al.*, 1988).

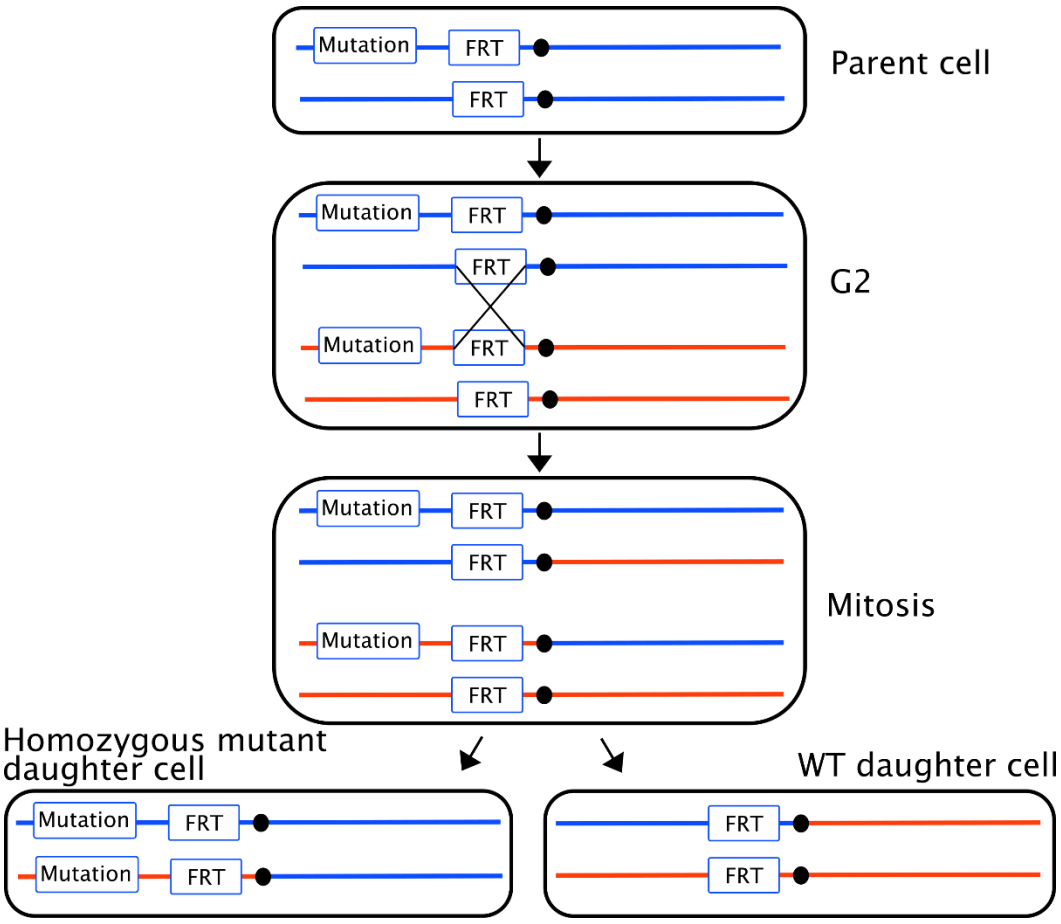


Figure 2.4 FLP/FRT system

In a heterozygous parental cell, one chromosome arm contains the mutant allele and FRT site whilst the other arm contains an identical FRT site. Following DNA replication, during the G2 phase, FLP drives mitotic recombination between the FRT sites of two chromosome arms, leading to the exchange of chromosome segments. Daughter cells will be different from each other, one will be homozygous for the mutant and the other WT (Theodosiou *et al.*, 1998).

2.2.2.4 MARCM system

The mosaic analysis with a repressible cell marker (MARCM) is a technique that combines the Gal4/UAS and the FLP/FRT systems, with the addition of a Gal4 repressor, Gal80. In this system, the products of FLP induced recombination will be one homozygous mutant and one homozygous Gal80 cell, the consequence of this is that any Gal4 activity will be repressed in the daughter cell containing Gal80, yet active in the homozygous mutant cells (Lee *et al.*, 2001). MARCM technique can be used to create homozygous mutant clones surrounded by WT tissue, and the additional expression of any UAS-tagged transgenes specifically within the mutant clones (Figure 2.4).

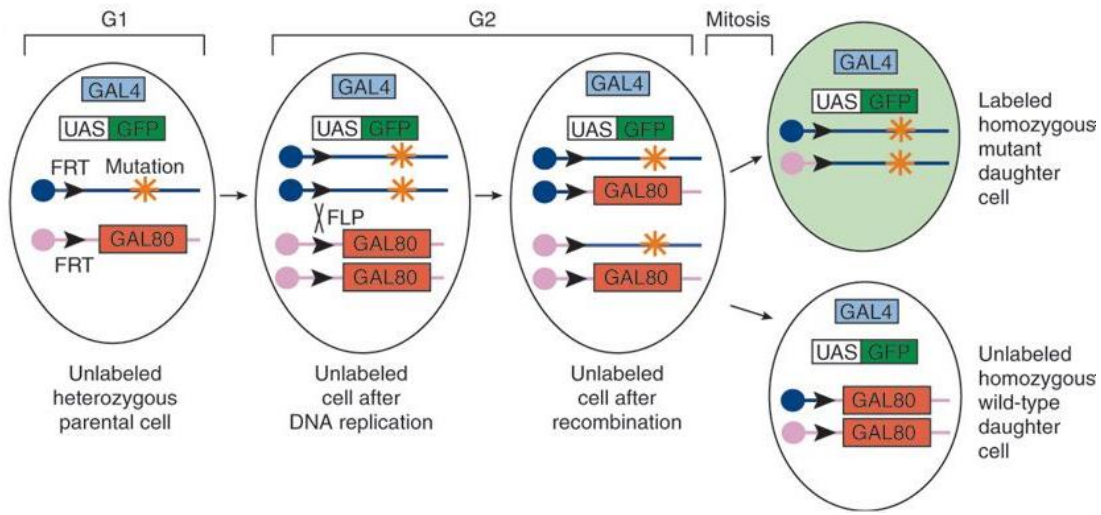


Figure 2.5 MARCM system

A cell containing the Gal80 protein will block the activity of Gal4, therefore blocking transgene expression. Combination of the Gal4/UAS system with FLP/FRT system and the addition of Gal80 can be designed so that a recombination event is induced between Gal80 and a specific gene mutation. After mitosis one daughter cell will have homozygous Gal80, which will block transgene expression, phenotypically a WT cell. The other daughter cell will be homozygous mutant and will additionally express transgene X. For example, if the transgene is a GFP reporter, mutant cells will be positively labelled whilst wild-type cells will be negatively labelled, allowing identification of mutant cells with high spatiotemporal resolution. Figure obtained from (Wu *et al.*, 2006).

2.1.2.3 Genetic crosses

To collect virgin females, flies were sexed based on the identification of ventrally located claspers on the posterior abdomens of males; this process was conducted every 8 or 16 hours, at 25°C or 18°C respectively, since females will not mate within the first 8 hours of adulthood (Ashburner, 1989). For all genetic crosses approximately 20 virgin females were crossed to 5-10 males and kept at 25°C. Progeny flies and pupae were then screened against appropriate genetic markers and balancers to ensure the inheritance of desired genetic constructs.

2.1.2.4 Genotypes analysis

Drosophila genome is composed of 4 chromosome pairs: 1 pair of sex chromosomes (X and Y) and 3 somatic chromosome pairs. The fourth and Y chromosomes are usually ignored in genetic studies, due to their small size and low genetic relevance. Males do not undergo recombination during meiosis. There are tools available, such as balancer chromosomes, to prevent recombination in females, facilitating the inheritance of traits and the study of genotypes. Additionally, genetic markers were used, homozygous lethal alleles with a dominant visible phenotype, yet they do not prevent recombination.

Table 2.7 Balancer chromosomes and genetic markers

Name	Chromosome	Type	Phenotype
CyO (Curley of Oster)	II	Balancer	Curly Wings
FM7 (ln(1)FM7)	X	Balancer	Kidney-shaped eyes
KR^{IF1} (Irregular facets)	II	Marker	Small 'rough' eyes with irregular ommatidia.
MKRS	III	Marker	Stubble
TM6B (In(3LR)TM6)	III	Balancer	Tubby (short, fat body) and Humeral (extra macrochaetes)

Table 2.8 Genetic constructs. Table outlining the function of genetic constructs incorporated into the *Drosophila* genome.

Genotype	Abbreviation	Description
P{Ubx-FLP}	ubxFlp	Expresses FLP in the pattern of the Ubx gene
P{neoFRT}40A	FRT40a	Carries FRT site at Chr 2, 40A3, 2L:21794705..21794705
P{UAS-GMA}	UAS-GFP:moe	Expresses the actin-binding domain of moesin tagged with GFP[S65T] under the control of UAS
P{tubP-GAL80}	tubGal80	Expresses GAL80 'ubiquitously' under the control of the alphaTub84B promoter
P{GawB}pnr[MD237]	pnrGal4	Expresses GAL4 in dorsal cells along the length of the fly under the control of pnr.
P{UAS-lhRNAi}	UAS-RNAi	Expresses dsRNA for RNAi under UAS control
PBac{UAS-iCasper-noGFP-T2A-HO1}	UAS-iCasper	Expresses a reporter protein under UAS control that shows infrared fluorescence upon caspase cleavage in the presence of heme oxygenase.
P{neoFRT}19A	FRT19a	Carries FRT site at Chr X, 19A2, X: 19910870..19910870.

A list of the genotypes analysed in this research is included in this section. To avoid listing redundant genotypes, "UAS-transgene" refers both to UAS-RNAi line (for gene KD) or UAS-CDS gene (for gene overexpression). The specific lines used are listed in Table 2.2 and Appendix 2. All other transgenes used are specified in each genotype.

Genotype nomenclature is presented based on standard guidelines: chromosomes are listed in order: X, 2 and 3 (chromosomes Y and 4 are omitted). Chromosomes are separated by semicolons (;), alleles are separated by solidus or slash (/). WT chromosomes are specified with a plus symbol (+), any other transgene or mutation is specifically written.

- **FRT40a WT MARCM clones**

w, ubxFLP/+; FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe/+

- ***lgl*^{4} MARCM clones**

w, ubxFLP/+; lgl^{4}*, FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe/+*

- ***lgl*^{4} MARCM clones plus UAS-transgene on the III chromosome.**

w, ubxFLP/+; lgl^{4}*, FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe/UAS-transgene*

- ***lgl*^{4} MARCM clones plus UAS-transgene on the X chromosome.**

w, UAS-transgene/ubxFLP; lgl^{4}*, FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe/+*

- ***lgl*^{4} MARCM clones plus UAS-transgene on the X chromosome and iCasper reporter.**

w, UAS-transgene/ubxFLP; lgl^{4}*, FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe /UAS-iCasper*

- **FRT19a WT MARCM clones**

w, FRT19a, tubGal80, ubxFLP /FRT19a; +; PnrGal4, UAS-GFP:moe/+

- **MARCM clones with UAS-transgene on the third chromosome.**

w, ubxFLP/+; FRT40a/ tubGal80, FRT40a; PnrGal4, UAS-GFP:moe/UAS-transgene

- **MARCM clones with UAS-transgene on the second chromosome.**

w, FRT19a, tubGal80, ubxFLP /FRT19a; UAS-RNAi/+; PnrGal4, UAS-GFP:moe /+

- **MARCM clones plus UAS-transgene on the II chromosome and iCasper reporter.**

w, FRT19a, tubGal80, ubxFLP /FRT19a; UAS-RNAi/+; PnrGal4, UAS-GFP:moe/UAS-iCasper

- ***smc3*$^{a\ or\ b}$ mutant MARCM clones.**

w, smc3$^{a\ or\ b}$*, FRT19a/FRT19a, tubGal80, ubxFLP; +; PnrGal4, UAS-GFP:moe/+*

- ***smc3*$^{a\ or\ b}$ mutant MARCM clones plus Histone-RFP.**

w, smc3$^{a\ or\ b}$*, FRT19a/FRT19a, tubGal80, ubxFLP; ubiHis-RFP/+; PnrGal4, UAS-GFP:moe/+*

- ***smc3$^{a\ or\ b}$*** **mutant MARCM clones plus mCherry:Sqh**

w, smc3$^{a\ or\ b}$, FRT19a/FRT19a, tubGal80, ubxFLP; +; PnrGal4, UAS-GFP:moe/ mCherry:Sqh

- ***smc3$^{a\ or\ b}$*** **mutant MARCM clones plus iCasper**

w, smc3$^{a\ or\ b}$, FRT19a/FRT19a, tubGal80, ubxFLP; +/+; PnrGal4, UAS-GFP:moe/UAS-iCasper

- ***smc3$^{a\ or\ b}$*** **mutant MARCM clones plus UAS-transgene on the III chromosome**

w, smc3$^{a\ or\ b}$, FRT19a/FRT19a, tubGal80, ubxFLP; +/+; PnrGal4, UAS-GFP:moe/UAS-transgene

- ***smc3$^{a\ or\ b}$*** **mutant MARCM clones plus UAS-transgene on the II chromosome**

w, smc3$^{a\ or\ b}$, FRT19a/FRT19a, tubGal80, ubxFLP; UAS-transgene/+; PnrGal4, UAS-GFP:moe/+

2.1.2.5 Pupa mounting and notum dissections

Progeny of interest were picked at the white prepupal stage (zero hours after pupae formation; 0 APF) and analysed under a Leica Fluorescence Microscope for the identification of genetic markers. Pupae with the genotype of interest were kept in a petri dish, with paper towel soaked in distilled water to maintain moisture. Pupae carrying a UAS-RNAi construct

were maintained at 29°C and pupae from any other experiment were kept at either 25°C or 18°C.

2.1.2.5.1 Mounting for live imaging

Mounting for live imaging was achieved by pupae immobilization, ventral side down, on a microscope slide using double sided sticky tape. The notum was then exposed by removing a region of the pupal case, covering the dorsal thorax, using tungsten forceps. Small glass coverslips (0.13-0.17mm) were fixed to the anterior and posterior ends of the microscope slide to generate a stage on which a long coverslip (24x50mm) covered in a thin layer of 10S voltalef injection oil was placed to create an interface between the coverslip and the exposed notum.

2.1.2.5.2 Notum dissections

Notum (dorsal thorax of the fly) dissections were carried as described in (Jauffred et al., 2012). In short, pupae were pinned, dorsal side down, in a silicon dish in cold PBS and the pupal casing was removed. The head and abdomen were then excised using dissecting scissors. Fat bodies and other organs were removed by gently pipetting PBS. Trachea was cut with scissors and the notum was cut in a rectangular shape, leaving the central Pnr region, where the MARCM clones are created. Dissected tissue was fixed for 20

minutes in 4% Paraformaldehyde solution and then washed 3 times in PBS-T (0.1% Tween in PBS). Tissue were then used for immunohistochemistry.

2.1.2.6 Immunohistochemistry

Dissected notum was stained following the protocol described in (Jauffred *et al.*, 2012). Each notum was isolated and transferred to a glass dish with 3% BSA and blocked for 1 hour at room temperature (RT) or overnight (O/N) at 4°C. Blocking solution was removed, and antibody solution added (for antibody concentrations see Table 2.5), and tissue was incubated O/N at 4°C. Antibody solution was removed and tissue were washed 3 times with PBS-T and incubated with secondary antibody for 40min at RT. Tissues were then washed 3 times with PBS-T and incubated 1 minute in DAPI nuclear stain solution. Tissues were washed 3 times with PBS-T and then moved to a solution of PBS (50%):Glycerol (50%) O/N. The glass dish was always kept in a humidified chamber to prevent tissue to dry. The tissues were mounted on a microscope slide in a drop of fluorogel mounting medium and covered with a coverslip.

2.1.2.7 Confocal imaging

Imaging was performed using an inverted Leica confocal TCS SP2, Zeiss LSM880 or LSM exciter confocal laser scanning microscopes with 40x oil

immersion objective. Lasers used: 405nm, 488nm, 543nm and 633nm. Clones from live imaging were imaged by generating z-stacks of 1µm planes from the cuticle to basal lamina at a resolution of 1024x1024. Time lapses were recorded using the parameters described, in the time intervals stated in each result section. Clones from dissected tissue were imaged by generating z-stacks of 0.5µm planes from the cuticle to basal lamina at a resolution of 2048x2048.

2.1.2.8 Image analysis

Prism (GraphPad) was used to generate graphs and carry out statistical analysis. Fiji (imageJ) software was used to analyse images. The following plug-ins were installed and used in addition to standard Fiji processing options:

Table 2.9 Fiji plug-ins and macros installed and used

Plugin or Macro	Use	Source
Random Sample and Measure	Used to identify random regions of interest for the measurement of fluorescence intensity.	Dr. Gabriel Landini. University of Birmingham
Manual tracking	For the 3D tracking of individual invading cell trajectories.	Dr. Fabrice Cordelières. French National Center for Scientific Research

2.1.2.9 Identification of genetic screen "hits" and clustering (done by Dr. Anbarasu Lourdusamy)

For each category, the scores for each RNAi line across five different animals were averaged. We calculated total within-cluster sum of square (wss) varying number of clusters using mean scores and estimated the optimal number of clusters (k = 10) using the Elbow method. We then used hierarchical clustering to cluster RNAi lines by applying Ward's method and dissimilarity based on Euclidean distance. To detect robust clusters of categories, we used hierarchical clustering with agglomerative average linkage as our method for consensus clustering (R package: ConsensusClusterPlus). Clustering was performed over 1000 iterations at a sub-sampling ratio of 0.8.

For each category, genes with a mean score above (positive) or below (negative) interquartile range (IQR) from the median were selected as target genes. In addition, the cut-off score for positive should be greater or negative should be less than the score at Basal condition was used to select the target genes. All statistical analyses were done in R (version 3.5.0).

2.1.2.10 Interaction map

The interaction network was generated by integrating publicly available data to obtain a unified map of genes involved in cancer invasion. The data resources used were DroID (www.droidb.org, version 2018_08), Flybase (version FB2018_03) and BioGRID (version 3.5.165). PubMed ID was used as a unique identifier for each interaction, to prevent the inclusion of redundant information. Interactions were included into further consideration only if both interactors were hits in the categories of 'invasion', 'multilayering', 'cell body rounding', 'lethals' or if the interactor was not part of the screen yet linked an otherwise not interacting gene from any of these categories. The resulting network was drawn using Cytoscape 3.6 (http://www.cytoscape.org). MCODE v1.5.1 algorithm was used to identify subnetwork complexes of highly interconnected nodes.

2.2 Cell culture materials and methods

2.2.1 Cell culture materials

2.2.1.1 General reagents

Table 2.10 General reagents

Reagent	Supplier
50x TAE	GeneFlow
5x siRNA buffer	Dharmacon
Chloroform	Sigma
Clarion Assay, human	Thermofisher Scientific
CyQUANT Cell Proliferation Assay Kit	Thermofisher Scientific
DEPC	Thermofisher Scientific
Dharmafect Reagent	Dharmacon
DMEM	Thermofisher Scientific
DMSO	Thermofisher Scientific
DNase 1	Invitrogen
dNTP Mix	Thermofisher Scientific
Ethanol	Sigma
Foetal bovine serum (FBS)	Thermofisher Scientific
GeneChip Hybridization, Wash, Stain kit	Thermofisher Scientific
GeneChip WT PLUS Reagent Kit	Thermofisher Scientific
Glycogen	Sigma
Hanks balanced salt solution (HBSS)	Thermofisher Scientific
iQ SYBR Master Mix	BioRad, UK
Isopropyl alcohol	Sigma
L-Glutamine	Thermofisher Scientific
Matrigel basement membrane	SLS
Penicillin/Streptomycin (P/S)	Thermofisher Scientific
Proteinase K	Thermofisher Scientific
RevertAid Reverse Transcriptase	Thermofisher Scientific
RiboLock RNase Inhibitor	Thermofisher Scientific
sybersafe	Invitrogen
Taq polymerase	Thermofisher Scientific
TRIzol Reagent Ambion	Life Technologies
Trypsin	Thermofisher Scientific
X48 Millicell Hanging Cell Culture Insert	Thermofisher Scientific

2.2.1.2 Primers

Primers were designed using the National Center for Biotechnology Information (NCBI) resource. Selected primers overlapped an exon-exon junction. Oligonucleotides were synthesised by Sigma Aldrich (UK) at 0.01μM scale and desalted.

Table 2.11. Primer list for RT-qPCR

Primer name	Forward (5′ – 3′)	Reverse (5′ – 3′)
STAG1	GGCCAGCCGAAGTTAGAAGA	CCATTAGACCCCGAACAGCA
STAG2	AGACATGCCTGAGCAGATTGT	AAACGCAGCAAGTCCTCCTT
SMC1	ATCCACAGCTACAAGCCAGC	TTGCAAGTCGACTGCTCCTT
SMC3	TGGAGTTGGAATTAGGGTGTCA	GTCCACCTGAAAGCTGTTGC
RAD21	TGTATCAATGGGTGGGCCTG	GGCTCCAATGCAAATGCTTCT
PCDH1	GTCATCCAGGTGAAGGCCAA	CGGTCCTTAGCAAGCACTGA
EHD2	GGTGCGAGTTCACGCTTACA	AGATGACGGGCAGTTTGAGG
AKR1B10	CAGCAACAGAGAGCAGGACG	TGCCAAGAGGAGACTTCCAA
GAPDH	ATGTTCGTCATGGGTGTGAA	GTCTTCTGGGTGGCAGTGAT
ING1	TACTGTCTGTGCAACCAGGT	TTCTCCAGGGCTTTGTCCAT
ARHGAP23	CCGATGAATGGAGTCGCCTT	ATGAAGTGGCGCAGAGTGAA
DST	TACTGCCCTGGTCACTCTCAT	TACATCTGCTACCGGCCAC
RPS6KA3	ACCACATCTAGTAAAGGGTGCC	TTCACAGGGCTGTTGAGGTG
RIMS2	ACAGCAAGAACAGAAGGGTGA	CACATAACCTTGTTTGAGCGT

2.2.1.3 siRNA

All the siRNA pools were obtained from the human ON-TARGET Dharmacon library.

Table 2.12 siRNA sequences

Target gene	siRNA
JAG1	GGGAUUUGGUUAAUGGUUA
	GGAACAUACUGCCAUGAAA
	UAACAUAGCCCGAAACAGU
	GAGCAACCCUUGUAGAAAC
ARHGAP23	CCACAGGCAUUUACCGAGU
	CCGUCAACCGCAAGCGCAA
	UCAAGUCCUUCUUCCGAAA
	CCCACAUGCCUGACCGCUA
DST	GGGAUUUGGUUAAUGGUUA
	GGAACAUACUGCCAUGAAA
	UAACAUAGCCCGAAACAGU
	GAGCAACCCUUGUAGAAAC
RPS6KA3	AUACAAUGCUUACCGGUUA
	CUACAUAGCCUGGGAAUAA
	CGUAAUCAGUCACCAGUUU
	GACAGCAAAUUAUGGAUGA
RIMS2	GAUCAUACGUCUUGGCAUA
	AGACCAGACUUGAGGCGUA
	CCAUUGUGAUGUCCAAUAAA
	GGUCAAAUUUCUCCUAAGUU

Target gene	siRNA
STAG1	GAAUAGAGAUGUUUCGAAA
	UCAAGAAAUCGGCGAGAGA
	GAUUAGAGCCAUUUGUAUU
	GAAUGUUGGUGAAUGUCAA
STAG2	GAAAUUUACUUGCAGCAUU
	GUAGAUGAUUGGAUAGAAU
	GGGAUUUAUUUGCUUGUAA
	CCACUGAUGUCUUACCGAA
SMC1	GACAGCUCUAUUUGAAGA
	GCUCGUAACUUCCUCGUUU
	GUACAAGGGUCGACAGAUU
	GAACUGGCCUCAAAGAACA
SMC3	CAACGUAGCUUACAGAGUU
	GGUGUAAAGUUCAGAAAUA
	GAGAGUAGAUGCACUGAAU
	GCAGUGCAACACAGAAUUA
RAD21	GAAUCUAUCACAGGAAAGCC
	GAGAUAACCAUGAGAGAAGA
	UAGAGUCUGAACAGAGCACC
	GAUCAAUAUAAGGAUGAUAA

2.2.2 Cell culture methods

2.2.2.1 Cell maintenance

MCF-7 human breast carcinoma cells were obtained from Dr Anna Grabowska (Faculty of Medicine & Health Sciences, University of Nottingham). Cells were cultured in RPMI medium without phenol red (Invitrogen) supplemented with 10% fetal bovine serum (Sigma-Aldrich) and grown in T75 culture flasks at 37°C in a 5% CO^2 atmosphere.

2.2.2.2 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 104/ml the following equation was used to count and seed appropriate number of cells. $Cells\ per\ ml$ = ($Total\ number\ of\ cells\ from$ 4 $corner\ squares$ 4) x 104

2.2.2.3 siRNA transfection

To knock-down gene expression, MCF-7 cells were incubated for 48 hours with transfection media containing DharmaFect1 (Dharmacon) transfection reagent and ON-TARGET plus SMARTpool siRNA (Dharmacon) against ArhGAP23, DST, RPS6KA3, RIMS2 and ING1. ON-TARGETplus Non-

targeting siRNA (Dharmacon) was used as a negative control and siGLO controls (Dharmacon) for transfection optimisation experiments.

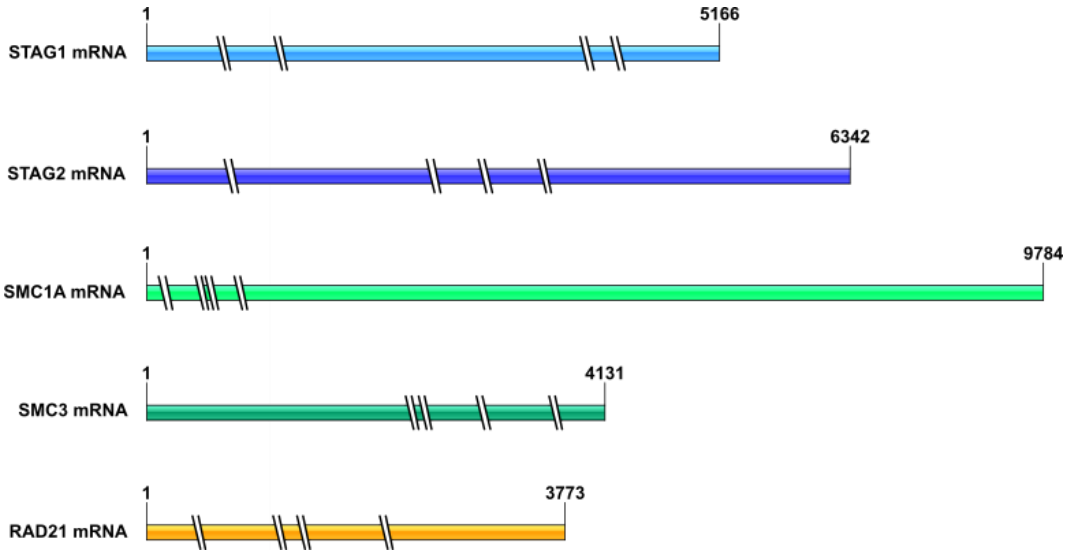


Figure 2.6 siRNA cut sites in Cohesin subunits mRNAs

Schematic representing the mRNA for the Cohesin complex subunits and the specific target sites of the Dharmacon siRNA sequences. For specific target sequences of all the genes analysed see section 2.2.1.3.

2.2.2.3 RT-qPCR

Total cellular RNA was extracted from un-transfected cells, cells featuring knock-down and cells transfected with ON-TARGETplus Non-targeting Control siRNAs (Dharmacon) using TRIzol (Thermo Fisher Scientific) following the manufacturer's instructions. The concentration and purity of RNA were determined by spectrophotometry, only samples with a 260/280 ratio higher than 1.8 and 260/230 ratio higher than 2.0 were used. 2µg of total

RNA were used to synthetize cDNA following the RevertAid First Strand cDNA Synthesis Kit [Thermo Fisher Scientific (k1621)] protocol. Quantitative real-time PCR (RT-PCR) was performed with A C1000 thermal cycler CFX96 RT system. Non-universal primers were designed using primer 3. Samples were labelled using IQ SYBR Green qPCR Master Mix (BioRad). Primer specificity and efficiency was confirmed by standard and melting curve analyses.

Relative transcript levels were calculated relative to the transcript level of the reference gene GAPDH using the fold change method:

$$\Delta Cq = \text{mean Cq (target gene)} - \text{mean Cq (housekeeping gene)}$$

$$E\Delta Cq = \frac{\text{power (PCR efficiency for the housekeeping gene,mean Cq (target gene))}}{\text{power (PCR efficiency for the target gene,mean Cq (housekeeping gene)}}, \text{FC } \frac{E\Delta Cq}{\text{mean } E\Delta Cq}.$$

2.2.2.4 Invasion Assay

After the transfection, cells with and without the knock-down were harvested. Appropriate cell density was optimized according to the Corning cell invasion assay protocol. $1x10^5$ cells, in a final volume of 250µL, were plated in each 24 well hanging insert. The base of the inserts had and 8µm pore membrane (Millipore). For invasion experiments, the inserts were covered with 75µg extracellular matrix (Corning). Cells were plated in 1% FBS RPMI and 10% FBS RPMI was used underneath the inserts to create a

chemoattractant gradient. The amount of migrating or invading cells was assessed after a 48 hours incubation. Migrating or invading cells were detached from the bottom of the inserts using accumax cell detachment solution (Merk Millipore). The fluorescence intensity of the cell suspension was measured using 480/520 nm filters and optimal gain on a BMG fluorescent plate reader after a 15 minutes incubation with CyQuant GR dye (Thermofisher Scientific). Images of the migrating or invading cells were acquired using a 10x/ 0.3 NA objective on a Zeiss Axioplan/Axiophot brightfield microscope after inserts were stained using a Shandon Kwik-Diff stain kit (Thermofisher scientific) following the manufacturer's instructions. A background correction was performed using Gimp Software.

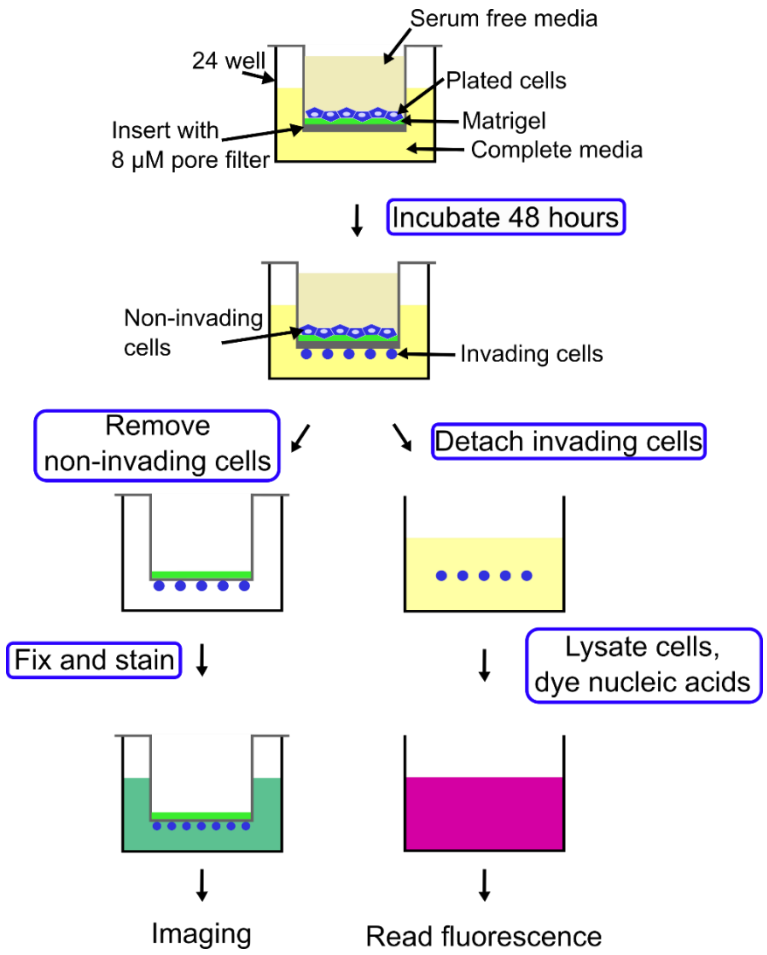


Figure 2.7 Trans-well migration and invasion assay diagram

A Boyden chamber, or insert, with an 8µM pore membrane at the bottom, is placed on a 24-cell culturing well containing complete media. For invasion assays, a layer of extracellular matrix proteins is added to coat the membrane. $4x10^4$ cells are plated in each insert with serum free media and incubated for 48 hours. After incubation, non-invading cells are removed with a cotton swab. For microscopy imaging, invading or migrating cells attached to the membrane are fixed and stained, number of invading or migrating cells can be counted by field of view (bottom left on the diagram). For fluorescent quantification, invading or migrating cells are detached from the insert, lysated and nucleic acids are stained with a fluorescent dye;

number of migrating or invading cells can be calculated with a standard curve (bottom right on the diagram).

2.2.2.5 Transcriptomics (done by Dr. Marcos Castellanos)

Whole-genome transcriptome analysis of transfected and un-transfected (both un-treated and treated with non-targeting siRNA) MCF-7 cells was conducted at the Nottingham Arabidopsis Stock Centre (NASC) of the School of Biosciences. The RNA concentration and quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technlogies Inc., Palo Alto, CA) and the RNA 600 Nano Kit (Caliper Life Sciences, Mountain View, CA). Samples with a minimum RNA concentration of 100 ng/µl and RNA Integrity Number (RIN) ≥ 8 were used for gene expression analysis. Single stranded complimentary DNA was prepared from 200 ng of total RNA as per the GeneChip™ WT PLUS Reagent Kit (Applied Biosystems and Affymetrix). Total RNA was first converted to cDNA, followed by *in vivo* transcription to make cRNA. Single stranded cDNA was synthesised, end labelled and hybridised for 16 h at 45°C to Clariom™ S Assay arrays (Thermofisher Scientific).

Gene expression data were analysed using Partek Genomics Suite 6.6 software (Partek Incorporated). The raw CEL files were normalised using the RMA background correction with quantile normalisation, log base 2

transformation and mean probe-set summarisation with adjustment for GC content. Differentially expressed genes (DEG) were identified by a two-way ANOVA. DEG were considered significant if p-value with FDR was ≤ 0.05 and fold change of >1.5 or <-1.5. Regulatory network prediction was accomplished by use of GeneMANIA plug-in for Cytoscape. Based on this database, an analysis of genes interacting with adherens junction protein coding genes was carried out.

2.2.2.6 GO term enrichment

Genes were analysed for GO biological process enrichment, using the online annotation database: DAVID Bioinformatics Resources 6. Functional annotation clustering with default settings was used; medium stringency and Benjamini–Hochberg correction was applied. The enriched GO terms with FDR <0.05 were selected and displayed in the interaction map.

Chapter 3 - A genetic screen to identify modifiers of tumour progression

3.1 Introduction

The process of carcinogenesis is very complex and requires the alteration of several signalling pathways. Advances in techniques such as whole genome sequencing have generated a large amount of data across tumours from different types of tissues, and while this has contributed largely to the advances in cancer research, many questions are still unanswered. A solid human tumour has an average of 30-60 mutations, and this number goes even higher in older patients (Kandoth *et al.*, 2013; Vogelstein *et al.*, 2013; Lawrence *et al.*, 2014). While several tumour suppressor genes (TSG) and oncogenes have been identified and characterized, many of the genes found significantly mutated in cancer have no direct link to their potential role in driving tumour progression. Furthermore, differentiating between driver and passenger mutations has not been achieved to date. To address this problem, a loss of function screen was done in the living epithelial tissue of the dorsal thorax (notum or back of the fly) of *Drosophila* pupa. 500 different uncharacterised genes, that are significantly found to be mutated in different epithelial cancers, were knocked down individually within positively marked tumours, with the aim of identifying a potential mechanism of action for these genes in tumour progression.

3.1.1 Preliminary experiments

An *in vivo* model, adapted from the method previously described by Georgiou and Baum (Georgiou *et al.*, 2008, 2010), was designed to study uncharacterized genes involved in tumour progression. The model combines the MARCM technique with RNA interference to create tumours with specific genotypes on the back of the living fly. Tumours are homozygous mutant for a tumour suppressor gene and simultaneously expressing an RNAi transgene to KD the expression of an additional gene of interest, specifically within the mutant tissue. Furthermore, expression of GFP:moe (the actin binding domain of moesin fused to GFP) labels the cytoskeleton and other actin-rich structures, again specifically within the mutant clones.

The first step in method development was to choose the tumour suppressor gene that was suitable as a background in this system. As mentioned in Chapter 1, cancer is a multistep process; therefore, a background mutation will allow the identification of additional genes that either promote or inhibit tumour progression in this model. The three members of the Scribble basolateral polarity complex: lethal 2 giant larvae (*lgl*), discs large (*dlg*) and scribble (*scrib*) were analysed during preliminary experiments. Besides their role in polarity, these genes are also involved in tumour proliferation and their orthologues have been established as tumour suppressor genes in

humans (Schimanski *et al.*, 2005; Massimi *et al.*, 2008). Homozygous mutant clones surrounded by WT tissue were generated. *dlg* and *scrib* mutant clones showed very low clonal coverage when compared to the WT control (Figure 3.1a-d). Additionally, *dlg* and *scrib* clones were highly multi-layered (Figure 3.1c'-d'). In contrast, *lgl* mutant clones had no significant effect on clonal size (Figure 1e) and showed a significant increase in the number of invading cells (Figure 1f) when compared to WT, making it an ideal background for the study of enhancers or suppressors of tumour progression.

3.1.2 Pilot screen

During an initial pilot screen, 50 well known tumour suppressor genes previously implicated in cancer were studied to verify that the system was working efficiently. Expected phenotypes were observed, including: increased clonal coverage following RNAi of Tsc1 (Figure 3.1h) (Potter *et al.*, 2001); increased invasion with pTEN KD (KD) (Figure 3.1i) (Tamura *et al.*, 1999); abnormal cell shape with ECad KD (Figure 3.1j) (Tepass *et al.*, 1996); increased blebbing and abnormal cell division in clones expressing RNAi for moe (Figure 3.1k) (Kunda *et al.*, 2008); and multilayering from pak KD (Figure 3.1l) (Aguilar-Aragon *et al.*, 2018). Additionally, negative controls failed to generate significant phenotypes.

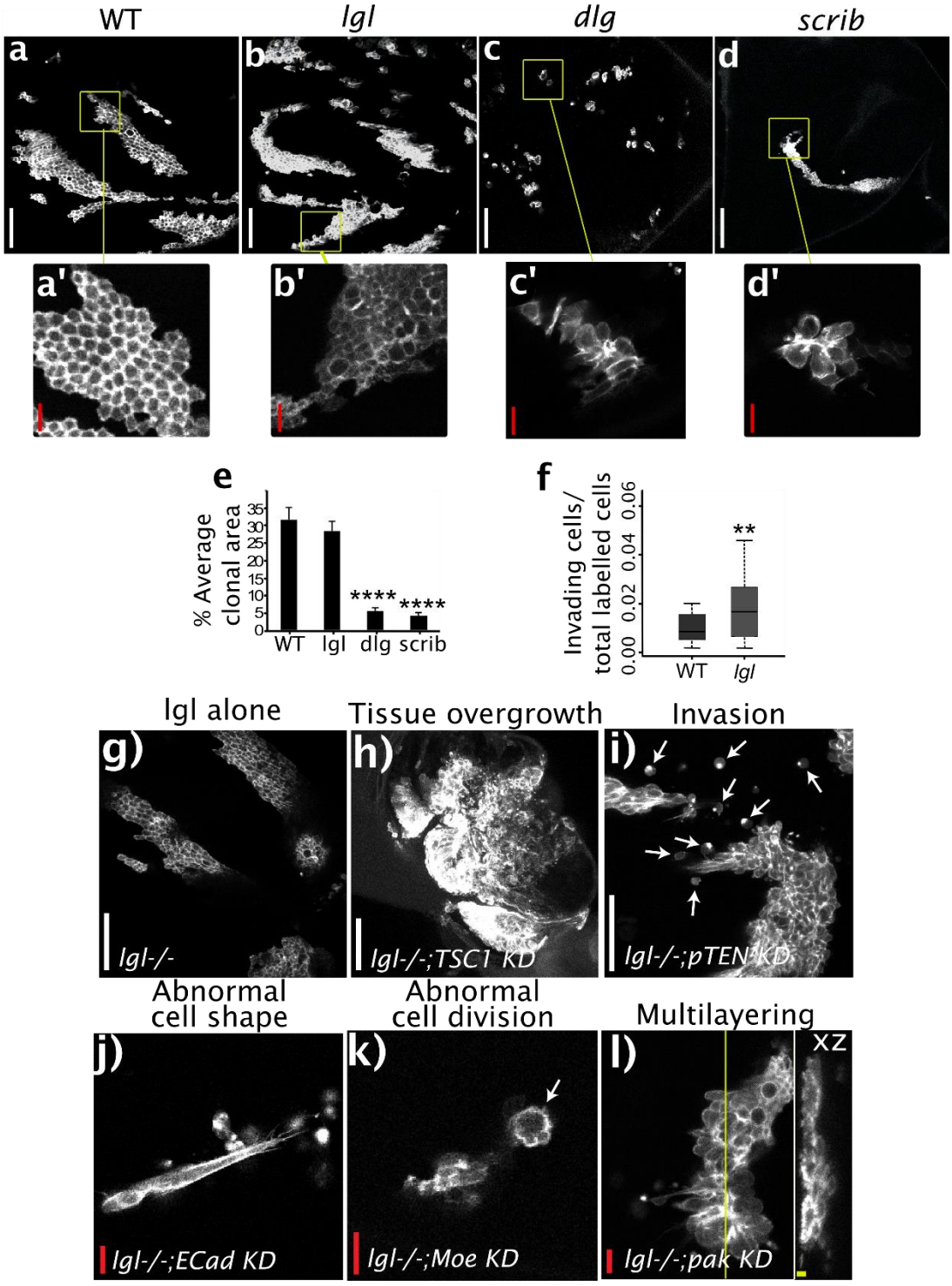


Figure changed

Figure 3.1 Preliminary experiments and pilot study

(a-d) GFP:Moe labelled genetic clones in the dorsal thorax epithelium of living fly pupae. Clones shown are wild-type (a-a') or homozygous mutant for the

neoplastic tumour suppressor genes *lgl* (b-b'), *dlg* (c-c'), or *scrib* (d-d'). *Dlg* and *scrib* clones are small and extremely multilayered. **(e-f)** Quantification of average clonal area and the number of invading cells over the total number of labelled cells. Quantification shows *lgl* mutant clones to be similar to wild-type clones in size, with a significant increase in the number of invading cells. Error bars represent ± s.e.m. Student's t-test (e) and Kruskall-Wallis test (f) were performed to determine statistical significance. **(g-l)** Examples of phenotypes observed within the pilot screen: (h) tissue overgrowth, (i) invasion, arrows point to invading cells, (j) abnormal cell shape, (k) abnormal cell division (arrow highlights abnormal blebbing), (l) multilayering. Genotypes indicated on bottom left of panels. White scale bar: 50μm; red scale bar: 10μm; yellow scale bar: 10μm in xz plane. Data generated by Cornhill Z, Couto A and Mack N.

Following the pilot screen, which confirmed that our *in vivo* system could identify modulators of tumour progression, the main genetic screen was performed. Section 3.2 will describe: i) a more detailed description of the rationale behind the screen method; ii) the results obtained and the identification of genes that promoted or inhibited 32 different phenotypes related to tumour progression; iii) a detailed analysis of the genes that enhanced invasive related phenotypes; iv) the identification of the cohesin complex subunits as tumour modifiers, which will be characterized in detail in Chapters 4 and 5.

Since this is a large-scale project, the results that are presented in the next section were generated and analysed over the course of approximately 7 years, involving the hard work of many people, not just my own, especially Dr Zoe Cornhill, Dr Africa Couto, Dr Natalie Mack and Alexandra Rusu. I have been involved, in collaboration with them, in all the aspects of the main genetic screen, except the preliminary results and pilot screen described in this introduction. Data analysis was done in collaboration with a bioinformatician, Dr Anbarasu Lourdusamy. The interaction network, verification experiments and Cohesin complex quantification, presented in Figures 3.7 - 3.11, were performed by myself exclusively.

3.2 Results

3.2.1 A genetic screen to identify modifiers of tumour progression

A large-scale screen to identify novel genes involved in the process of tumour progression was done to identify a potential mechanism of action for currently uncharacterized genes. Tumour-like clones were created through a combination of the MARCM technique, UAS/pnrGal4 system controlling the expression of UAS-GFP:moe (labelling the cytoskeleton of the cells) and UAS-RNA interference (targeting a specific gene of interest). ubxFLP induces site specific recombination in homologous FRT sites during mitosis; the flies analysed in the screen are designed so that the FRT sites are located in the second chromosome, and recombination is induced between the *lgl* mutation and the tubGal80 transgene, generating two different daughter cells in mitosis: i) a daughter call that is homozygous mutant for the *lgl* mutation and has lost Gal80 ii) a daughter cell that has two copies of tubGal80 and no *lgl* mutation. In WT cells, tubGal80 will block pnrGal4 mediated expression, and therefore UAS transgene expression, thereby blocking the expression of both the RNAi transgene and the GFP:Moe reporter (Figure 3.1 a) (see Chapter 2 for more details). Therefore only homozygous mutant *lgl* clones will be labelled with GFP:Moe and express the RNAi transgene. Additionally, since FRT and Gal4 are under specific promoters, their expression is restricted to the notum, namely the epithelial cells of the dorsal

thorax of the fly (Figure 3.1 b, area in the magenta rectangle), meaning that mitotic recombination and clone formation is restricted to that area (Figure 3.1 b).

A total of 497 individual genes were studied, as previously mentioned. These genes were selected from different studies where they have been found to be mutated or downregulated in a significant number of patients with carcinomas, yet their potential role in tumour progression is unknown (Beroukhim *et al.*, 2010; Consortium *et al.*, 2010; Lawrence *et al.*, 2014). UAS-RNAi lines were obtained from three different near genome wide libraries: VDRC in Austria, NIG in Japan, and Bloomington in the USA; 754 different RNAi lines were used. When available, KD of one gene was done by two different RNAi lines. If animals from two independent RNAi lines targeting the same gene showed very similar phenotypes, we could be more confident that the phenotype observed was a consequence of the KD of the gene of interest rather than off target effects, or an RNAi line that does not efficiently reduce the expression of the candidate gene.

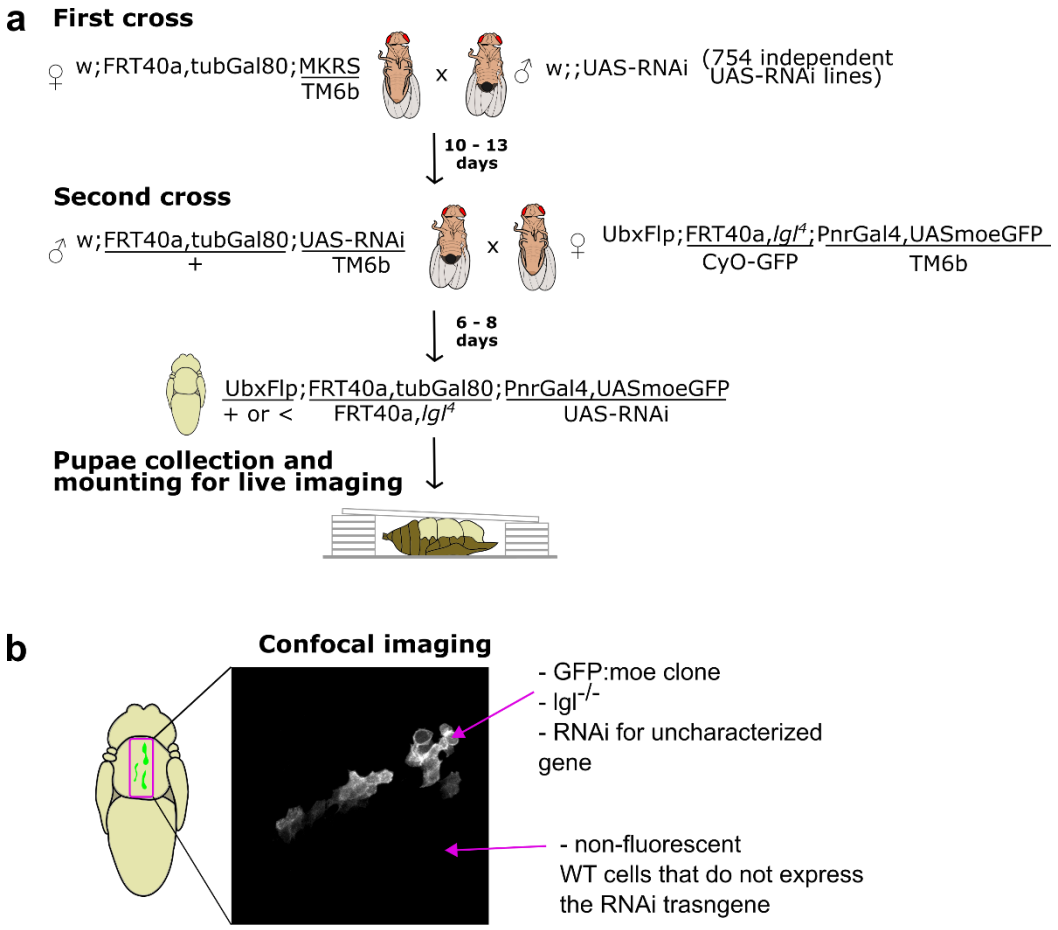


Figure 3.2 Outline of genetic screen

a) Screen cross scheme. First cross, females containing FRT site and tubGal80 were crossed with UAS-RNAi males. Second cross, male F1 progeny of the correct genotype was identified by the absence of short hairs (MKRS) and crossed with females containing FRT site, pnrGal4 driver and UAS-GFP:moe. F2 progeny was collected at the pre-pupal stage (0h APF), based on pupal shape (TM6b, absence) and fluorescent markers, i.e. tubGal80 absence (salivary glands not fluorescent) and CyO-GFP absence (See methods for more details). Selected pupae were aged at 29C and mounted for live imaging at 20-24h APF. **b)** Magnification of the pupal area imaged. The notum, or epithelial tissue of the dorsal thorax of the fly, is outlined in the magenta rectangle. Image on the right is a sample confocal image of the mosaic tissue. GFP:moe positive cells are homozygous mutant for the TSG *lgl* and express an RNAi transgene targeting an uncharacterised gene. Non-fluorescent cells that appear as black background are the WT cells that act as the tumour microenvironment.

3.2.2 Phenotypes relevant to tumour progression and scoring system

There are numerous phenotypes relevant to tumour progression that we can observe in this system (see Appendix 1 for full list). By labelling the mutant cells with GFP:moe we have access to the actin rich structures of the cell, including the cytoskeleton and other accessory components such as protrusions. One of the most relevant phenotypes includes clonal tissue size. As shown in Figure 3.1, *lgl* clonal area has no significant difference with WT clones, yet the additional KD of certain genes either increases or decreases tumour size (Figure 3.3a).

From a basal view of the clones, individual cells that detached from the epithelium can be analysed. We found, in all genotypes studied, that invasive cells would round up and decrease size after delamination. Additionally, detaching cells form a characteristic actin-rich spot at one side of the cell prior to invasion (Figure 3.1i). The number of cells presenting this polarised phenotype within the epithelial sheet is therefore an indicator of invasive potential. Figure 3.1 shows that lgl subtly, yet significantly, enhances the number of invading cells compared to WT tissue; therefore, genes that increase *lgl* invasion rate, towards malignancy; or genes that reduce *lgl* invasion rate, towards WT; are studied in the invasion category (Figure 3.3b).

Similarly, several other effects on tumour behaviour can be studied, including: multilayering, (cells no longer form a monolayered epithelial sheet; Figure 3.3c); apex size (the area at the top of the cell can be constricted or expanded; Figure 3.3d); and changes in the basal protrusion length (Figure 3.3e).

An enormous amount of data was generated in the large-scale screen, therefore, an efficient system for data analysis was necessary. A scoring system was developed, where all the phenotypes related to tumour progression are identified and compared to the phenotypes of the *lgl* mutant alone (Figure 3.3, middle column). An enhancement of the $lgl^{-/-}$ phenotype was scored as either plus one or plus two, depending on the severity of the phenotype (Figure 3.3, right column); and equivalently an inhibition of the $lgl^{-/-}$ phenotype was scored as a minus one or minus two (Figure 3.3, left column). Binary categories were also included, where the scoring is done as either a zero (no change from lgl) or one (enhancement of lgl) (Appendix 1). There are a total of 32 different phenotypic categories in the screen database. As the analysis performed for the phenotypes of the screen is qualitative, the scoring process involves two different scientists that analyse each animal blind (at least 5 animals were imaged per genotype), and the averages of the scores provide a less biased representation of the real phenotype.

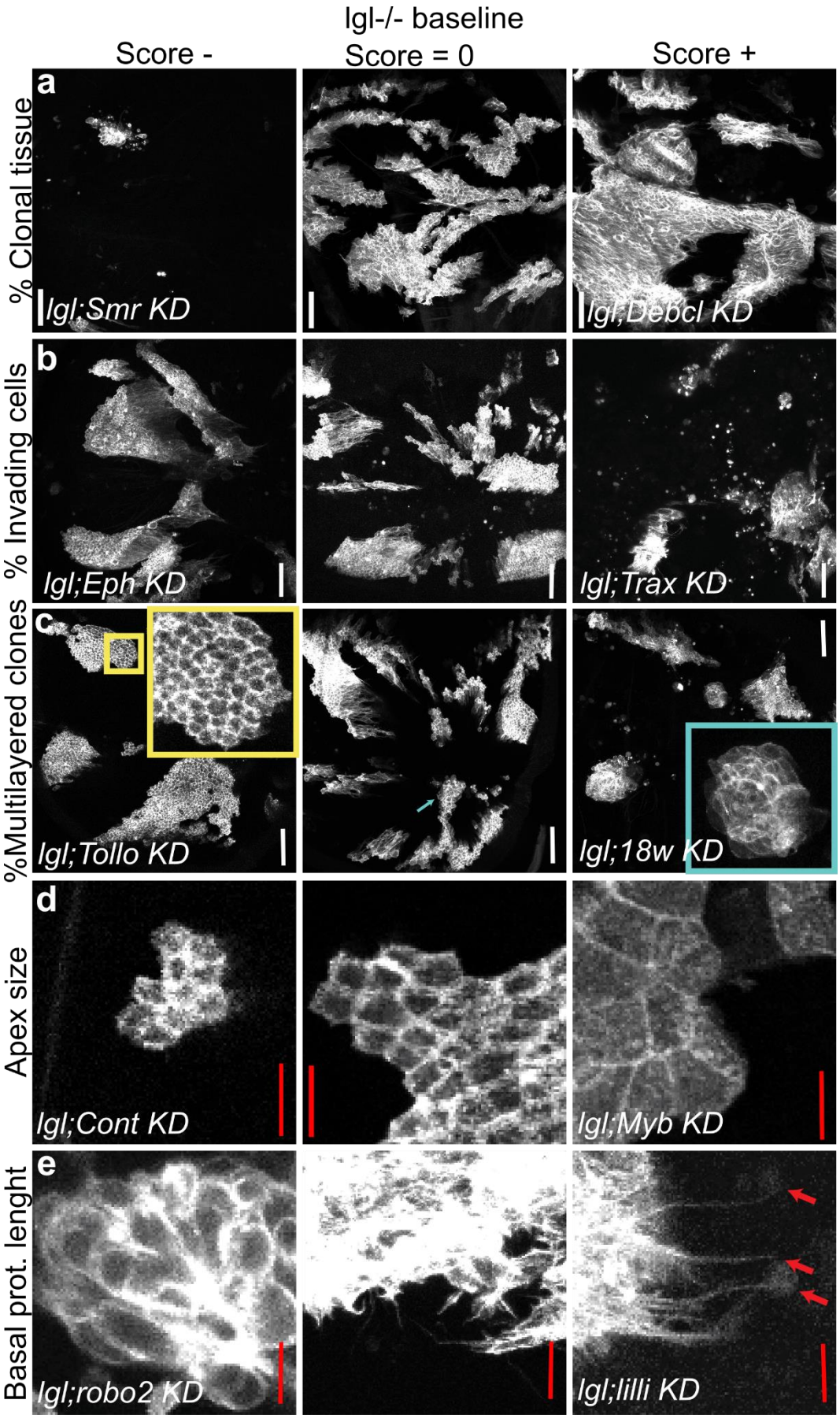

lgl-/- baseline
Score -
Score = 0
Score +
% Clonal tissue
% Invading cells
%Multilayered clones
Apex size
Basal prot. lenght
a
lgl;Smr KD
lgl;Debcl KD
b
lgl;Eph KD
lgl;Trax KD
c
lgl;Tollo KD
lgl;18w KD
d
lgl;Cont KD
lgl;Myb KD
e
lgl;robo2 KD
lgl;lilli KD

Figure 3.3 Scoring system specifically designed to identify modifiers of tumour progression phenotypes

Sample images from 5 relevant phenotypes of tumour progression that were studied in the genetic screen. There are a total of 32 phenotypic categories (see Appendix 1). The *lgl*$^{-/-}$ alone phenotype is assigned a score of 0 (middle column), an enhancement of a phenotype is scored in positive numbers, +1 or +2 (right column) and a suppression of the phenotype is scored negative numbers, -1 or -2 (left column). **a)** Clonal tissue, percentage of clonal tissue within the fly pupal notum (i.e. GFP-positive area versus GFP-negative area). **b)** Invading cells, percentage of invading cells against total GFP-positive tissue. c) Multilayering, percentage of multilayered clones against total GFP-positive tissue; d) Apex size, average apex size of epithelial cells; e) Basal protrusions, average length of basal protrusions. Panel names on right and left columns refer to the specific gene knocked down in an *lgl* mutant background. Yellow box indicates a magnified monolayered clone. Turquoise box indicates a magnified multilayered clone; turquoise arrows highlight multilayered clones within the notum (others are monolayered). Red arrows point to basal protrusions. White scale bar: 50μm; red scale bar: 10μm.

3.2.3 Quality control

Quality control was performed for a selection of categories, to verify that our scoring system had a high correlation with quantitative measurements. Randomly selected genes were quantified for four phenotypes that were easy to quantify: clonal tissue, invading cells, basal protrusions and apical area. A strong correlation was observed between the qualitative scores and the quantitative measures (0.91 – 0.97, Spearman correlation test) (Figure 3.4 a-e), validating the scoring system and therefore indicating that there is a high accuracy in the database results.

Additional validation was performed for the 256 genes that were targeted by two independent RNAi lines. A comparison between the scores given for each RNAi pair targeting the same gene was done, where correlation between the average scores across the 32 phenotypic categories was determined. We found that 224 genes (87.5%) gave statistically similar phenotypes; therefore, there is a stronger confidence that the phenotypes observed were due to a loss of gene expression instead of error. An example of two RNAi lines targeting the TSG pTEN is shown in Figure 3.4 e-j.

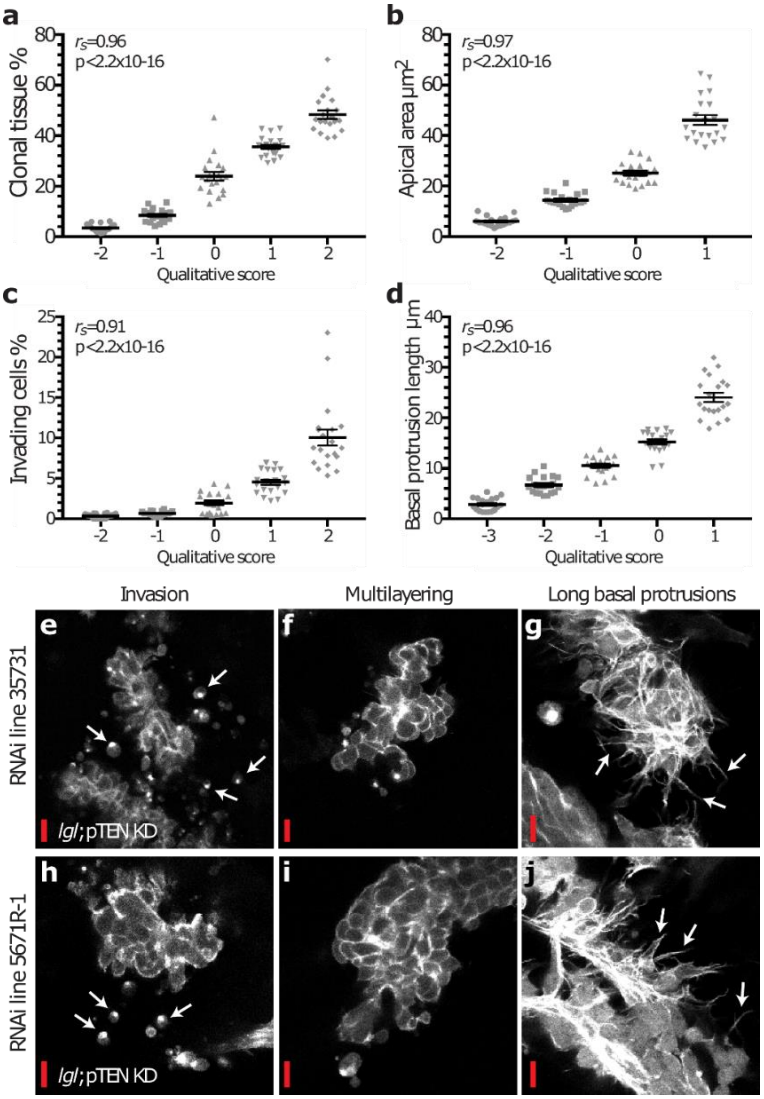


Figure 3.4 Quality control

(a-d) Graphs comparing qualitative scores given to individual animals against the quantitative measure for a range of phenotypes, including: (a) percentage clonal tissue; (b) apical area; (c) percentage invading cells (d) basal protrusion length (n= 20 animals for each category of qualitative score). (e-j) A comparison of the phenotypes observed when two independent RNAi lines (35731-GD and 5671R-1) were used to target the same gene (pTEN). Both showed strong effects on invasion (e and h), multilayering (f and i), and basal protrusion length (h and k). Arrows highlight: invading cells (f and i); long basal protrusions (g and j). Scale bars = 10 µm. Error bars represent ± s.e.m.

3.2.4 Identification of genes that modify tumour progression

To identify the genes that induce a change in the *lgl* phenotype, average scores were calculated for the 5 animals analysed and the scores given by the two different scientists, i.e. a minimum of 10 different scores for each RNAi line; to have one average score for each of the 32 phenotypes for each RNAi line. The distribution of the scores for each phenotype was then determined; for categories with a two-tail distribution, enhancers and inhibitors were identified; for categories with one-tail distribution, only enhancers were identified. Genes with a mean score above or below the interquartile range from the median were selected as genes of interest (these genes are referred to as hits in the next sections). The full list of genes that modified each of the phenotypes studied is included in Appendix 3.

Once the hit genes were identified, different clustering methods were performed as part of the data analysis and to render the information easier to interpret. Hierarchical clustering was done first: all the RNAi lines studied were clustered across all the phenotypes. This resulted in the identification of ten phenotypic clusters (Figure 3.5), the complete list of genes from each cluster and the GO biological processes significantly enriched in each cluster is included in Appendix 4. Cluster 10 includes the genes that did not induce a change in any of the phenotypes studied, namely they were not hits for any

category; accordingly, the lgl mutant and the negative controls (*kirre* and *otu*) were clustered together. Additionally, a number of genes that could not be analysed, because no clones were observed (6 RNAi lines), or the clonal area was less than 5% of the notum (32 RNAi lines), are also clustered together, since their average scores are 0 for every category.

Cluster 1 includes genes that reduced proliferation phenotypes, including clonal area and cell division. This cluster is highly enriched in the biological process gene expression (GO:0006357, *Marcal1, Tkv, Salr, Babo, Hsp83, Blimp-1, Brm, Snr1, Taf1, Osa, Cic*) and morphogenesis (GO:0008587, *Brm, Snr1, Osa, Mys*). Consistent with the phenotype observed, loss of the Dpp signalling members TKV and SALR, has been shown to reduce proliferation in the *Drosopihila* wing disc (Martín-Castellanos *et al.*, 2002; Organista *et al.*, 2013). Additionally, Cluster 1 includes the completely uncharacterised genes *CG5645* and *CG6143*. Cluster 8 is also enriched in genes that reduced clonal tissue; yet, unlike Cluster1, alteration of genes in this cluster also promote multilayering and cell body rounding; the most relevant biological processes overrepresented here are mRNA splicing (GO:0000381, *U2af38, Rm62, How, Dom*), cell differentiation (GO:0048863, *Arm, Vn, Scrib, Robo2*) and cell adhesion (GO:0019991, *Nrg, Dlg1, Cont, Scrib, Arm, Robo2*).

In addition to cluster 8, clusters 2, 4, 5 and 6 are also enriched for multilayering and cell body rounding. Cluster 6, additionally has increased invasion. Interestingly, the significant GO terms in this cluster include CNS development (GO:0007417, *Robo3, Bun, Pdm2, Cnn, Abi, Mbc, Ephrin, Trio, Robo2, Apc*), oxidation-reduction and immune response (GO:0006909 and GO:0055114, *Pp2C1, Fbp2, CG30268, Klu, Idh*). Additionally, cluster 6 has the uncharacterized CG genes: *CG13185, CG14372, CG30268, CG32066, CG32669, CG34371, CG4096, CG45263, CG8197*, and *CG9572*.

Cluster 2 has increased apex defects, in addition to multilayering and invasion; GO terms enriched are adhesion (GO:0007156, *Fw, Ft, Dscam3, Cad99C, Hbs),* transport (GO:0016192, *Cac, Rab40, Cdc42, Ex, SCAR, Dscam1, Cirl, Rim, Cad99C, Unc-13*), and cell polarity (GO:0001736, *Fw, Ft, Pten, Jbug, Patj, Hbs*). This cluster includes the well-known TSG pTEN; additionally, both adhesion and polarity defects are highly involved in invasion and cell changes that lead to multilayering. Cluster 2 also includes completely uncharacterized genes *CG3430, CG11836* and *CG31626*. Overall, hierarchical cluster analysis is a powerful tool to identify genes that induced a change across the 32 different phenotypes studied. The genes that are part of the same clusters are likely to have similar mechanisms of action and could provide a starting point for future identification of novel genes involved in a cellular process relevant to tumour progression.

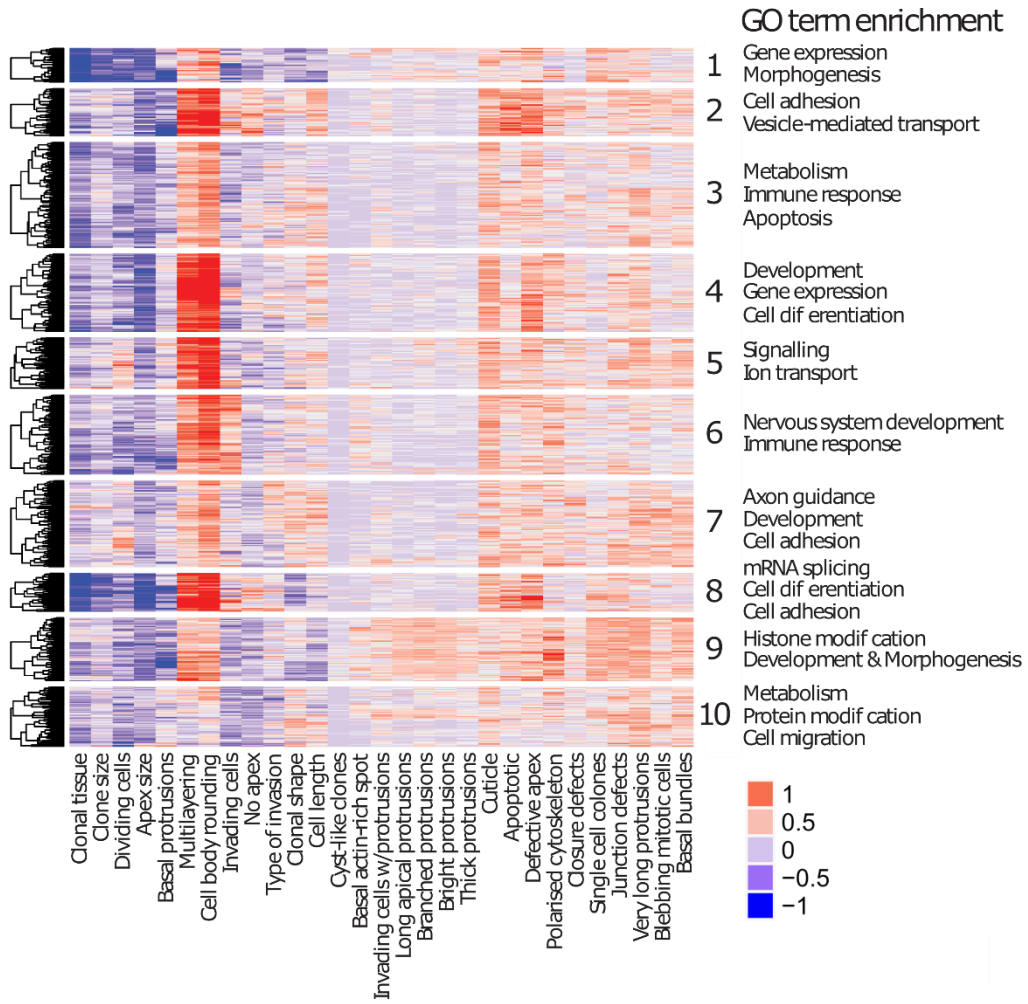


Figure 3.5 Clustering analyses identify ten RNAi line clusters

Heat-map representation of supervised clustering of 754 RNAi lines with average phenotype scores. Each row represents an RNAi line; each column represents a phenotype category. *A priori*, the model-based optimal number of $K = 10$ (phenotypic clusters) was determined. The clustering of rows and columns are based on Euclidean distance. Map colours represent row-scaled average scores: blue indicates the lowest score, light blue indicates an intermediate score, and red indicates the highest score. Each cluster was analysed with regard to their biological function by GO enrichment analysis. The most overrepresented GO categories are shown on the right-hand side of each cluster. Data generated by Canales B, Cornhill Z, Couto A, Mack M and Rusu A. Map generated by Lourdusamy A.

Consensus clustering was additionally used with the aim of identifying similar categories based on phenotypes presented across all RNAi lines; namely, a gene that had an effect on a category has high probabilities of having the same effect in other categories from the same cluster. Three distinct category clusters were identified (Figure 3.6 a). Categories that clustered together included those related to (a) actin cytoskeleton regulation, (b) invasion and multilayering, and (c) cell proliferation and cell and tissue morphology. This cluster method is also a useful tool to reduce redundant categories and identify the genes that affect a cellular process. In total, 294 genes had an enhancement effect across the three phenotypic clusters, with 114 increasing actin related changes, 33 increasing proliferation and 80 increasing invasion (Figure 3.6 b).

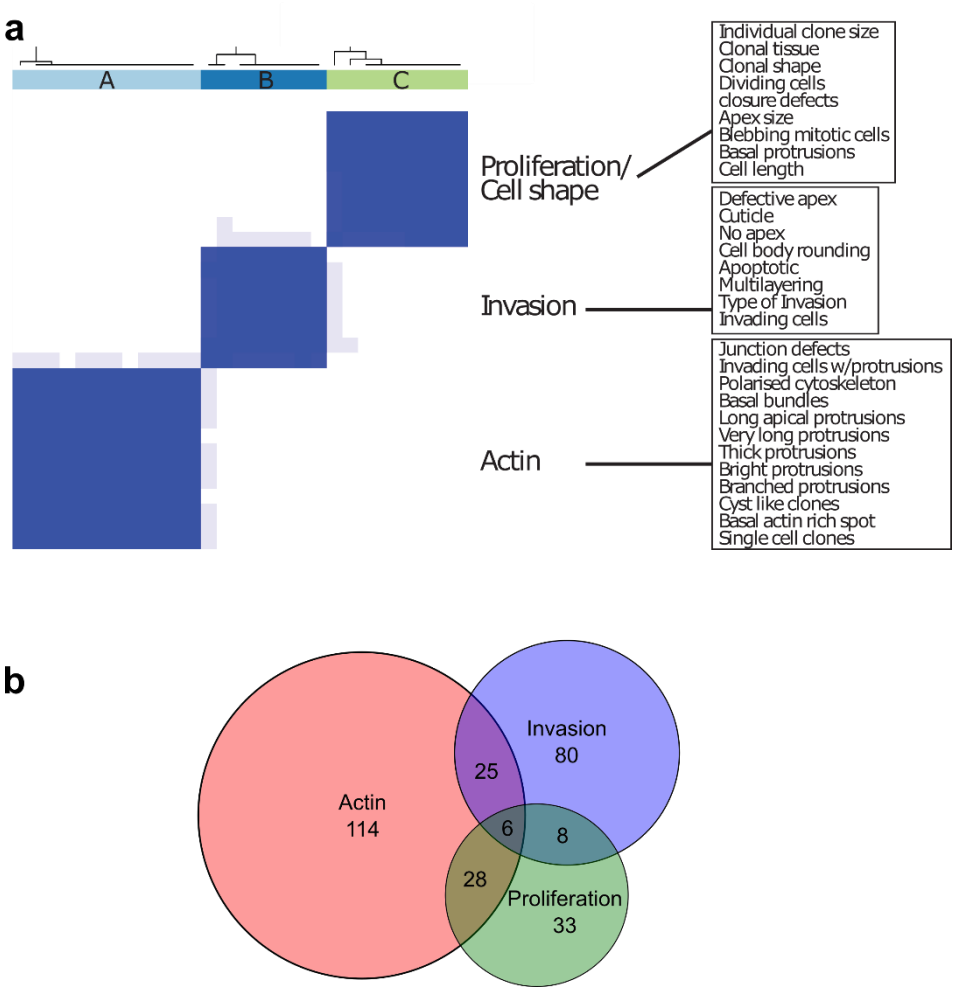


Figure 3.6 Clustering analyses identify three distinct phenotypic subgroups

a) Consensus clustering of average scores of 29 phenotypic categories reveals three distinct subgroups. The categories: "not developed", "no clones" and "malformed" were not included in the clustering since these animals were not analysed and we categorised them as "lethals". Each column represents one phenotype. Heat-maps display consensus values between pairs of phenotypes by blue shading. High consensus corresponds to phenotypes that always occur in the same cluster and is shaded dark blue. b) Venn diagram indicating the number of genes that were identified as "hits" for each cluster; genes that overlap were hits for more than one phenotypic cluster. i.e. the total of hit genes for actin is 173, for invasion: 119 and for proliferation: 75.

3.2.5 Characterization of invasion suppressors

Our live model of tumour progression provides many advantages on the study of cell invasion, since this process is difficult to replicate *in vitro*, due to the complex interaction required between the invading cells and the stroma. Therefore, this category became of special interest to us and a more detailed analysis of the phenotypes related to invasion was performed. Interaction networks have become a powerful tool to identify novel disease-associated genes. To generate a functionally validated interaction map of invasive genes, we combined all hits in three categories that clustered strongly together (Figure 3.6a): invasion, multilayering and cell body rounding. For each gene, we searched for physical or genetic interactions, validated by experimental data, including yeast two-hybrid, co-immunoprecipitation, and other interaction data from various databases (see Methods). We maintained interactions only between hit genes from these categories, together with lethals and 'linker genes', which linked hit genes from our screen by one interaction (Figure 3.7). The resulting network includes 321 interactions between 140 genes, 99 of which have not been previously implicated in cancer cell invasion or migration and including 9 genes that are completely uncharacterised.

To identify putative complexes, we used the Molecular Complex Detection (MCODE) software and found seven clusters of highly interconnected nodes (Figure 3.7). MCODE clusters often represent signalling cascades and/or molecular complexes. Complex 1 comprises core proteins involved in cytoskeleton organisation, including RAC2, SCAR, WASp, ARP2 and MBC. Adhesion proteins highly involved in cancer invasion are present in Complex 6: SHOT, FAT, p120ctn and GUG. Complex 5 is enriched in axon guidance molecules, whilst other identified complexes have an increased number of proteins that have not been previously linked to cancer cell invasion, such as Complexes 4 and 7. By integrating hits in invasive categories from our screen, together with protein and genetic interaction data, we have therefore identified a large number of novel genes that are now implicated in cancer cell invasion.

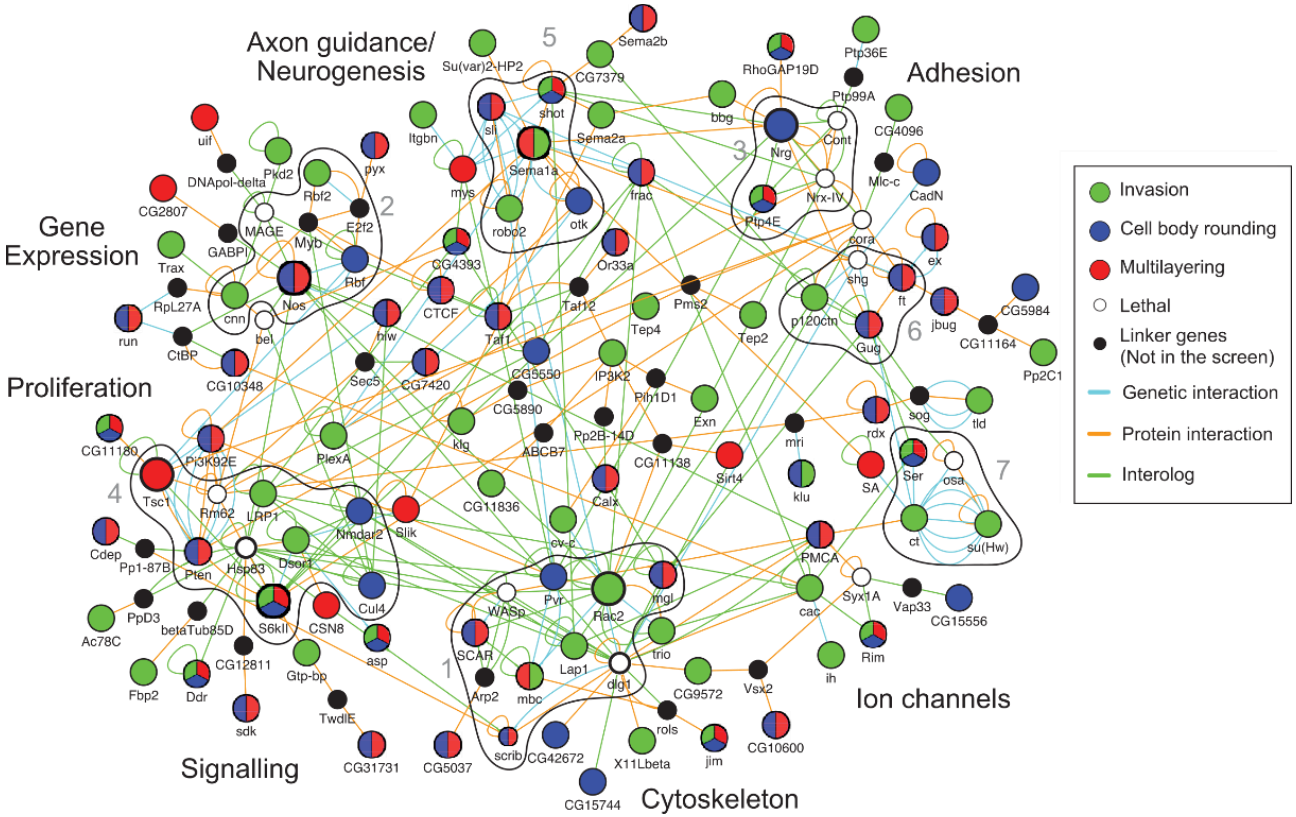


Figure 3.7 An interaction network of invasion suppressors

Interactions between genes for which KD enhanced the categories 'invasion', 'multilayering' and 'cell body rounding' are shown. Each circle node represents a gene. Node colour indicates phenotype observed in the screen: green = invasion; blue = cell-body rounding; red = multilayering; multi-coloured nodes = genes that were hits for more than one phenotype; white = lethal; black = 'linker genes', i.e. genes that were not part of the screen, but which connect screen hit genes by one interaction; nodes with a bold outline = hub genes in this network. Lines represent interactions: cyan = genetic; orange = protein-protein; green = interolog. MCODE complexes of highly interconnected genes are outlined in black. Significantly enriched GO terms are indicated.

It has previously been shown that WT epithelial cells delaminate from the pupal notum at early pupal stages, but this delamination is concentrated at the midline region and is rapidly followed by cell death (Marinari *et al.*, 2012). To specifically test for the viability of invading cells within highly invasive tumours, we used the genetically encoded apoptosis reporter iCasper (To *et al.*, 2015) (see Chapter 2 for more detail). We expressed iCasper within wild-type clones, *lgl* mutant clones, and in clones for five strong hits for invasion from our screen, namely: *lgl*; CG12268 KD, *lgl*; RHOGAP19D KD, *lgl*; SEMA1a KD, *lgl*; CG10931 KD, *lgl*; CAC KD. We observed that in four of the five invasive genotypes tested, a high proportion (~70%) of invading cells were iCasper negative. Only WT, *lgl* alone and *lgl*; CG12268KD mutant clones showed a high proportion of invading cells that were positive for apoptosis (~64%; Figure 3.8 a-b).

Having identified a number of invasion suppressors in our screen, we wanted to test whether human orthologues of the fly genes within this category would also act in a similar way. We took a panel of five fly genes that (1) strongly promote invasion when their expression is knocked down, and (2) have high-confidence, high-scoring best match human orthologues. Genes included were *RhoGAP19D, Rim, S6kII*, *CG7379*, and *Shot* (their closest human orthologues are *ArhGap23, RimS2, Rps6KA3, Ing1* and *Dst*). We

designed siRNAs against these human genes to see if their loss would lead to similar effects in the MCF7 breast cancer cell line. We used an *in vitro* invasion assay to test whether gene KD would promote MCF7 invasion and/or migration. We found a significant increase in both invasion and migration following gene KD of RPS6KA3, ING1 and DST, and a significant increase in migration alone with gene KD of RIMS2 (Figure 3.8c-i). This suggests that the invasion hits identified in our genetic screen are likely to have relevance to human disease.

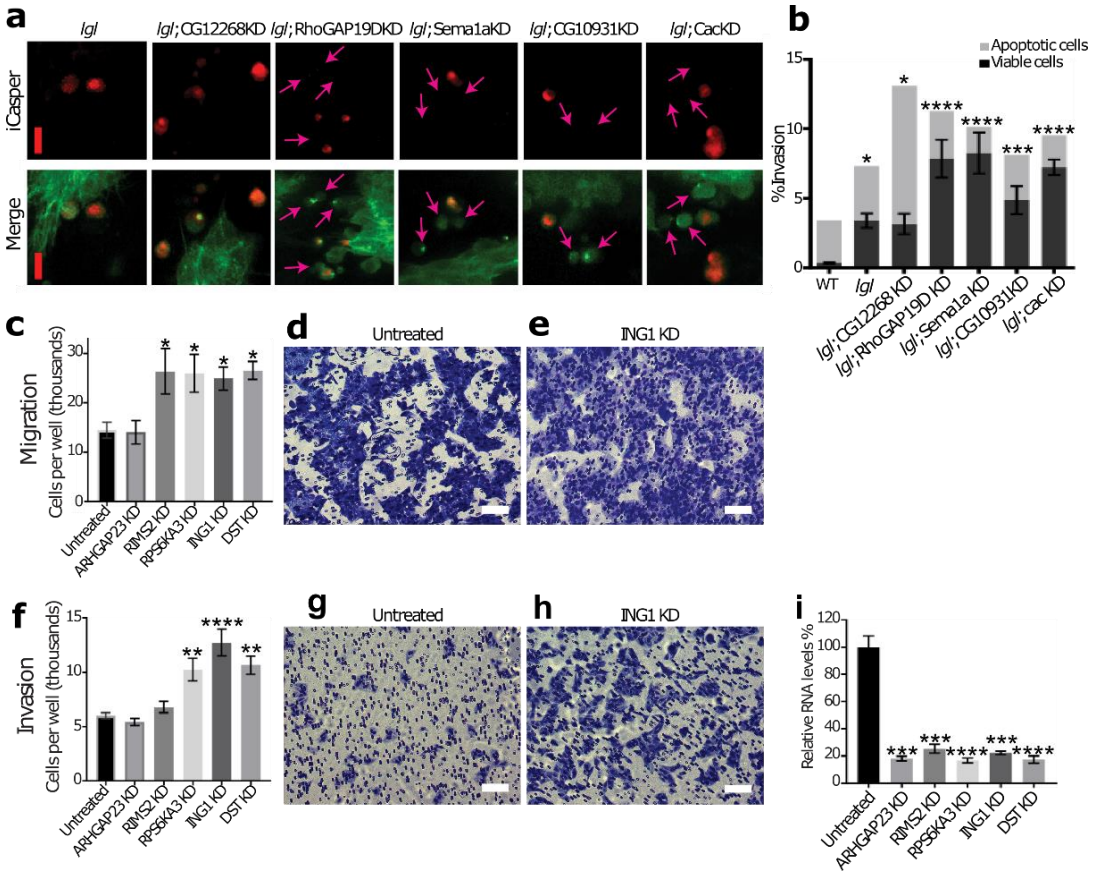


Figure 3.8 Characterisation of selected invasion suppressors

(a) i Casper (red) and GFP:Moe (green) labelled mutant clones (genotypes specified above panels). Arrows highlight invading cells that are iCasper negative. Four out of the five invasive genotypes tested showed a high proportion of invading cells that were iCasper negative (quantified in **b**; n=10 animals for each genotype). **(c-i)** *In vitro* transwell migration and invasion assays for human orthologues of a selection of highly invasive fly genotypes. Quantification of migration and invasion (**c** and **f**), showing the number of MCF7 cells, untreated or transfected with siRNA targeting the indicated gene, which migrated through an 8μm pore membrane. For the invasion assay the membrane was coated with a layer of Matrigel (n=3 for each assay). Representative images of ING1, the most invasive gene of the 5 genes analysed, are shown for migration (**d-e**) and invasion (**g-h**). Relative mRNA levels, determined by qRT-PCR, are shown in (m). Error bars represent ± s.e.m. One-way ANOVA with Dunnett's post hoc test for multiple comparisons was performed to determine statistical significance. Red scale bar: 10μm; white scale bar: 100μm.

3.2.6 Identification of the cohesin complex subunits as modifiers of tumour progression.

After the screen data analysis was finished, the next step in this research was to identify a strong candidate, or group of candidates, and characterise the phenotypes identified in the screen in more detail. SA1, a member of the cohesin complex mainly known for its role in sister chromatid cohesion (SSC) and gene expression regulation, enhanced various phenotypes related to tumour progression and was part of the invasion network (Figure 3.7). This brought attention to the other cohesin subunits, since SMC1, SMC3 and RAD21 were part of the genetic screen as well.

Hierarchical clustering of the qualitative scores from the genetic screen shows that loss of any of the cohesin subunits in combination with *lgl*, increases multilayering, cell-body rounding, apex defects and decreases clonal area (Figure 3.9, bottom block). Size and shape of the basal protrusions are unaffected by the KD of any cohesin subunit (Figure 3.9, middle block). For the categories 'invading cells', 'polarized invasion' and 'dividing cells'; different and opposing phenotypes were observed between the cohesin subunits (Figure 3.9, top block). Phenotypes where all the cohesin subunits scored zero are not included in the heat map, namely cohesin had no effect on those phenotypes. Phenotypes where any of the cohesin subunits were

ranked with a drastic enhancement score (≥1.5) or an inhibition score (-1), were picked for a quantitative analysis, to determine if the effect was statistically significant. The phenotypes are: multilayering, apex defects, clonal size and invading cells.

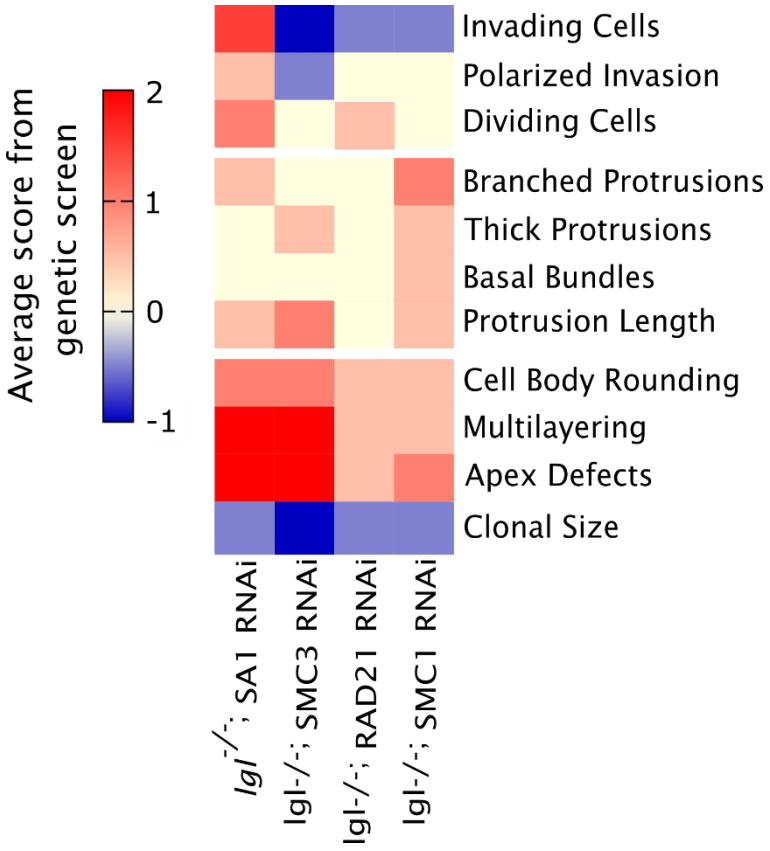


Figure 3.9 Effect of cohesin subunit KD on *lgl*$^{-/-}$ clones

Heat map illustrating the qualitative scores given to the cohesin subunits in the genetic screen. Positive score, or red, indicates an enhancement of a phenotype. No phenotype change is shown in yellow, or zero. A negative score, or blue, represents an inhibition of a phenotype. All the cohesin subunits scored zero in the categories not included in the heathmap. Detailed description of the scoring system can be found in Appendix 1.

Apex defects were significantly enhanced by SA1 (p=0.0353) and SMC3 (p=0.0416) KD (Figure 3.10a). Around 60% of mutant cells in SA1 (n=10 animals; 2,375 cells) and SMC3 (n=10 animals; 1,709 cells) KD clones had defective apices (Figure 3.10a'). The affected cells lost their distinct geometrical apical shape and appeared rounded instead (Figure 3.10a). SA1 and SMC3 KD also had a significant effect on multi-layering. 82% of SA1 KD (n=10 animals, 47 clones. p=0.0349) and 69% of SMC3 KD clones (n=10 animals, 47 clones. p=0.035) were multi-layered. Compared to $lgl^{-/-}$ alone, where only 29% of the clones showed multi-layering (n=10 animals, 28 clones,) (Figure 3.10b, b'). Clonal area was significantly reduced by the KD of any cohesin subunit (Figure 3.10d).

Invasion was only affected by SA1 KD, showing a significant increase in the number of invading cells (Figure 3.10c). Collectively, these results indicate that the cohesin complex subunits cooperate with *lgl* to maintain cytoskeletal organization. Additionally, downregulation of the expression of *Sa1* or *Smc3* in *lgl* depleted cells induces cancer phenotypes, such as multi-layering and invasion. Since the cohesin complex has not been previously linked to the phenotypes detected in the screen, a further characterization of their effect on the cytoskeleton could reveal a potential novel role for the cohesin complex in tumour progression.

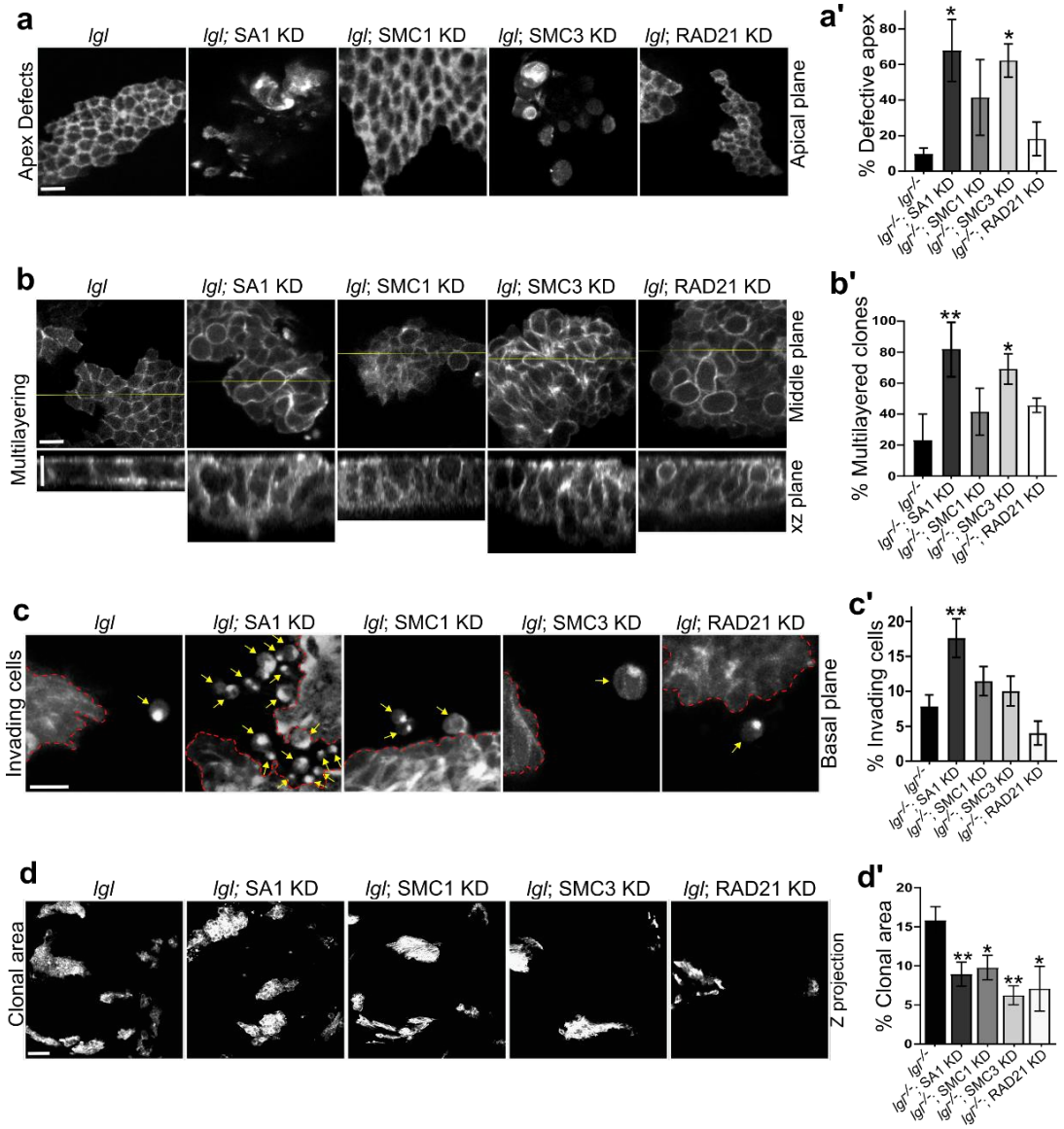


Figure 3.10 Effect of cohesin subunit KD in *lgl* mutant tumours

(a,b,c,d) GFP:moe positively labelled *lgl* mutant clones with additional KD of a specific cohesin complex subunit, specified at the top of each image. (**a)** Example images showing the apices of mutant cells. SA1 and SMC3 KD clones lost normal geometrical apical shape. (**a')** Quantification of clonal area with defective apex, represented as a percentage of area with defective apex against total clonal area. (**b)** Example images at an intermediate confocal plane (top) and an orthogonal view of the z-stack (bottom). Multilayered clones were frequently observed with cohesin subunit KD. (**b')** Quantification of the multilayered clonal area, represented as a percentage of

multilayered area against total clonal area. (**c)** Example images at a basal confocal plane. SA1 KD increased the number of invading cells. (**c')** Quantification of invasion, represented as a percentage of invasion against total clonal area. (**d)** Sample images of Z projection from the apex to the base of the clones, highlighting clonal area. (**d')** Quantification of clonal area, represented as a percentage of clonal area against total notum area. Scale bars = 10μM. Statistical analysis: Student's t-test. *p<0.05. **p<0.01. Error bars represent ± s.e.m.

3.3 Discussion

Our *in vivo* system in *Drosophila melanogaster* provides a powerful model where we can study a great variety of phenotypes relevant to tumorigenesis. By generating *lgl* homozygous mutant clones, with the simultaneous KD of a specific uncharacterised gene of interest, surrounded by WT tissue, we are able to phenotypically replicate a tumour with a controlled known genotype that can be studied in a high spatiotemporal resolution, as we specifically label the mutant cells with GFP:moe. Additionally, since transgene expression is mediated genetically, including the RNAi for the uncharacterised genes, variability is reduced compared to *in vitro* models, where sample manipulation is a great source of error and variation. We show that in this system we can study the effect that the loss of a single gene of interest has in several phenotypes related to cancer, such as invasion, cell division, multilayering, clonal coverage, polarity, among others. This provides valuable information since the high heterogeneity of human tumours renders the identification of the role that individual genes have in the progression of the tumour impossible. The identification of a single gene's contribution to malignancy is essential to develop biomarkers and targeted therapies.

One of the most challenging stages of the screen was image analysis. Around 4,000 confocal z-stacks were generated during the genetic screen, and due to the many categories measured from each animal, time of analysis was crucial to obtain the results within a reasonable amount of time, since the novelty of the findings is a key aspect to consider. Attempts to automate quantification of the phenotypes were tried in collaboration with the mathematics department of the University of Nottingham, yet artificial intelligence analysis was not a possibility for this study, due to the complexity of tumour cell morphology (especially with phenotypes such as multilayering), the multiple slices within a Z-stack, tissue folding and others. Instead, we found that the best approach to analyse the screen data was using a qualitative analysis; a trained eye proving to be much more sensitive than any computer software. Our scoring system proved to be very accurate, a high correlation was found between qualitative and quantitative scores in four important categories: clonal tissue, apical area, invading cells and basal protrusion length. Yet, with the field of deep learning advancing quickly (Litjens *et al.*, 2016), automatic analysis could be successfully optimized for future studies that involve a large number of microscopy images, to reduce the bias and time involved in image analysis and to share the findings with the scientific community faster.

During our analysis, a selection of genes had phenotypes that were categorized as non-analysable, since the clonal area was extremely small or no clones at all were observed. One of the genes that had this effect was shotgun (shg), the fly ortholog of the adherens junctions member E-cadherin. This gene is a very well-established tumour suppressor gene, and its loss has been widely implicated in EMT and invasion (reviewed in Pećina-Slaus, 2003). Loss of E-cadherin has also been shown to be involved in polarity defects (Desai *et al.*, 2009) and loss of the CDC42-aPKC-PAR6 polarity proteins leads to ectopic apical structures in the fly notum epithelium (Georgiou *et al.*, 2008). While both of these studies did not find a direct link with *Lgl*, it is a possibility that SHG KD had drastic effects on the *lgl* phenotype in this study, causing either an extreme rate of cell invasion where the cells are not visible anymore at the time of imaging, or cell death is triggered in the mutant clones due to essential cell mechanisms being disrupted. This opens a question for the other uncharacterised genes that had very low to no clonal area, when combined with an *lgl* mutation. Therefore, a recommendation for future experiments is the study of the KD of the genes that induced very low, or no clonal area, on their own, i.e. without the lgl background mutation, to better understand their potential role in tumour progression.

Consensus clustering identified three different phenotypic clusters across all the categories studied: proliferation, actin and invasion. This means that the genes that had a significant effect in a phenotype was likely to be a hit in the other phenotypes from the same category identified in the consensus. Around 15 percent of the total number of genes studied had an effect on increasing proliferation or tumour size. Increased proliferation was, however, the most suppressed phenotype, with most of the genes inducing smaller clonal size. This category may be difficult to assess and quantify since clonal area was assessed in the XY plane and many genes that induced multilayering caused aberrant growth in the Z plane; therefore, it is a possibility that this phenotype could have been underrepresented for this reason. Hierarchical clustering shows that the genes that increased proliferation are mostly enriched in clusters 2 ,5 and 7. These groups have significant increase of the biological processes: cell proliferation, biosynthetic processes and morphogenesis, respectively, indicating that very specific pathways must be lost to induce tumour growth. This is in accordance with previous studies that show the relevance of cell competition in cancer development, where in the majority of the cases, WT cells overgrow the mutant tissue, unless specific pathways are altered, the most important of them being alterations of Myc signalling (Di Giacomo *et al.*, 2017; Patel *et al.*, 2017).

The most represented phenotypic category cluster was modification of actin processes. Around 35 percent of all the genes analysed had an effect on actin related phenotypes. Unlike the genes involved in increasing proliferation, the hits from actin related phenotypes were not particularly enriched in a specific hierarchical cluster, indicating that this process is less specific. Since LGL, as an important member of the SCRIB/DLG/LGL basolateral polarity complex, is involved in actin cytoskeleton regulation, a possibility for the over-representation of this category could be an enhancement of the *lgl* phenotype. Additionally, since the 500 genes studied in this screen are frequently found mutated in cancer, another possibility could be that these genes are involved in cytoskeleton regulation, since this is a commonly altered process in tumour progression.

The consensus cluster of migration enhancers includes 119 genes, around 25% of all the genes studied. These genes were predominantly found in the hierarchical clusters: 2, 4, 5, 6, 7 and 8 which include the GO biological processes: Cell adhesion, biosynthetic process, calcium ion transport, regulation of transcription, axon guidance, histone modifications and chromatin remodelling. Cell adhesion is an important biological process that is highly disrupted in cancer invasion (Umbas *et al.*, 1994), and this was an expected result. Additionally, calcium is also linked to cell adhesion (Hirano

et al., 1987; Grunwald, 1993); therefore, calcium ion transport might have a role in regulating cancer invasion through adhesion. Genes that regulate gene expression are very interesting targets, since they might regulate well known tumour suppressor genes. Additionally, the enrichment of axon guidance genes in cluster 7 is an interesting finding that will be discussed in more detail in the next paragraph.

To determine how the genes from different biological processes might link to regulate invasion, an interaction network was generated, this revealed that 99 of the genes that enhanced invasion interact with each other, through genetic and physical interactions. Most of the genes from this network have not been linked to invasion before, and this opens an interesting line for future studies. Several of these genes are regulators of gene expression; how they might regulate genes involved in invasive related processes such adhesion or cytoskeleton, as shown in the interaction network, could be studied next. Furthermore, there is a high number of genes that are involved in axon guidance. The Slit/Robo pathway has been shown to inhibit invasion and metastasis in pancreatic and gastric epithelial cancers (Tie *et al.*, 2010; Göhrig *et al.*, 2014); therefore, our study provides additional interesting potential candidates that might be involved in epithelial invasion. In addition to the invasive hits, 12 genes from the screen, that were identified as

lethal for inducing very small or no clones, interacted highly with the invasive genes. This suggests that these genes might be implicated in cancer cell invasion, but their phenotype in combination with the *lgl* mutation was too severe to detect in our screen. Indeed, the two members of the LGL basolateral polarity complex, DLG and SCRIB, are part of the list of lethal interactors in the invasive network, indicating that other lethal uncharacterized genes are an interesting line for further characterisation.

Moreover, we found that loss of the closest human orthologues for a selection of invasive hits significantly increased migration and invasion in a human adenocarcinoma cell line. RIMS2 is a regulator of membrane exocytosis that is highly uncharacterized in cancer. RPS6KA3 is a protein kinase whose KD has been previously linked to increased malignancy in liver cancer (Fujimoto *et al.*, 2016). DST, whose main function in epithelial tissue is anchoring keratin-containing intermediate filaments to hemidesmosomes (Young *et al.*, 2007) and it's decreased expression has been found in metastatic breast cancer (Casey *et al.*, 2009). ING1 is a transcriptional regulator mainly linked to regulating proliferation related genes (Coles *et al.*, 2007). Indeed, ING1 was further characterised by other members of the Georgiou lab (Canales Coutino *et al.*, 2019), and a severe disruption of the adherens junctions was found; additionally, transcriptomic

analysis of ING1 KD identified altered gene expression of two genes that are involved in the promotion of cancer invasion WASF3 (Teng *et al.*, 2016) and RHOC (Mei Wu *et al.*, 2010). Therefore, our study also provides evidence that the identified genes in our *in vivo* model are very likely to be potential tumour suppressor genes in human cancer.

Overall, the results from the screen provide an extensive list of potential invasion regulators, and open up a wide line of research. Due to the advantages and novelty of our live model, we focused on invasive genes, yet a larger list of genes that target other cancer related processes such as proliferation, cell protrusions, cell shape, among others, could be analysed in more detail in future studies, since the images and scores from our qualitative analysis will be made open access. Furthermore, this research focused on phenotype enhancers, yet interesting candidates that suppress tumour progression in this model can also be easily found in the database results. Particularly in this study, the effects that the cohesin complex members had in cytoskeletal related phenotypes such as multilayering, apex defects and invasion is an interesting finding, since the cohesin complex has not been linked to these processes before. This line of study was picked for further characterization and the role of the cohesin complex in cytoskeletal related phenotypes and its impact in tumour progression will be described in the following chapters.

Chapter 4 - Investigating the role of SA1 and SA2 cohesin subunits in single cell invasion

4.1 Introduction

Four subunits of the cohesin complex were studied in the genetic screen described in Chapter 3: SMC1, SMC3, RAD21 and SA1. The KD of these subunits induced significant cytoskeletal changes to *lgl* mutant tumours in the epithelial tissue of the back of the living fly, including increased multi-layering, cell body rounding and apex defects. Additionally, SA1 KD significantly enhanced the $lgl^{-/-}$ invasive phenotype, while the other cohesin subunits had no effect on invasion. The aims of this chapter are *i)* to provide further evidence in support that SA1, and its isoform SA2, act as invasion suppressors in the living fly and in human adenocarcinoma in vitro and *ii)* to show preliminary results of the potential molecular mechanisms of action of SA1 and SA2 in the regulation of cell invasion. This brief introduction provides a background of what is currently known about the cohesin complex.

4.1.1 Canonical function of cohesin

Cohesin is a multi-protein complex mainly known for its role in sister chromatid cohesion (SCC) (Michaelis *et al.*, 1997). The cohesin complex is formed by the kleisin superfamily protein (Schleiffer *et al.*, 2003), RAD21 and two Structural Maintenance Chromosome proteins, SMC1 and SMC3; these proteins bind with each other to form a tripartite ring (Gruber *et al.*, 2003).

Additionally, RAD21 binds to a stromalin protein: SA in the fly or STAG in humans; from this point, the *Drosophila* stromalin will be addressed as SA1/2 and the human orthologue as STAG1/2. RAD21 being the only subunit that interacts with all the components of the cohesin complex (Losada *et al.*, 2000; N. Zhang *et al.*, 2013) (Figure 4.1). Somatic cells simultaneously express two different SA/STAG proteins, SA1/STAG1 and SA2/STAG2, and therefore two cohesin complexes; the subcellular localization of the ring will depend on the SA/STAG protein that binds to the tripartite ring. STAG1 cohesin is found in the chromosome telomeres; and STAG2 cohesin localizes in the centromeres (Canudas *et al.*, 2009). In addition to that, germ cells of higher organisms have specific cohesin subunit isoforms, such as SMC1β (Revenkova *et al.*, 2004), RAD21L (Ishiguro *et al.*, 2011) and STAG3 (Prieto *et al.*, 2001) (Figure 4.1b).

Cohesin is evolutionary conserved, functional cohesin complexes are found ubiquitously in all Eukaryotic organisms, from yeast to humans (Michaelis *et al.*, 1997; Losada *et al.*, 2000). Relevant to this study, the cohesin subunit orthologues of humans and *Drosophila* SA1/STAG1 and SA2/STAG2 have high homology, and the functional protein domains are strictly conserved (Figure 4.2). Additionally, loss of the cohesin complex shows SCC defects in chicken (Sonoda *et al.*, 2001), *Drosophila* (Vass *et al.*, 2003), mice (Hodges *et al.*,

2005), humans (Hauf, 2001; Hoque *et al.*, 2002), *C. elegans* (Pasierbek *et al.*, 2001), and many others; denoting that not only the structure of the proteins but also the function is evolutionary conserved.

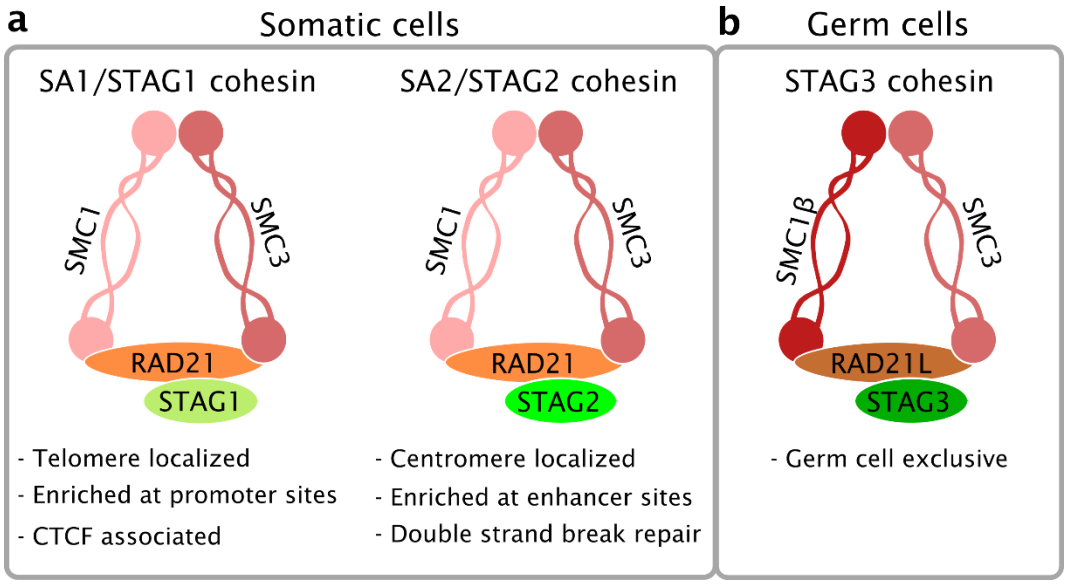


Figure 4.1 Somatic and germ cells cohesin complexes

a) The core of a mitotic cohesin ring is formed by the subunits SMC1, SMC3 and RAD21. Somatic cells simultaneously express two different cohesin rings, differentiated by the presence of either SA1/STAG1 or SA2/STAG2. STAG1 cohesin is localized to the centromeres during mitosis and is associated to CTCF (a transcriptional regulator known for inducing DNA loop formation) and promoter sites in non-dividing cells. STAG2 cohesin is the most abundant in humans, it localizes to the centromeres during mitosis; in non-dividing cells it is highly enriched at enhancer sites and provides stability to double strand break sites. **b)** Germ cells of higher organisms have specific isoforms of all cohesin subunits but SMC3. Loss of meiotic exclusive subunits leads to gender specific defects and sterility.

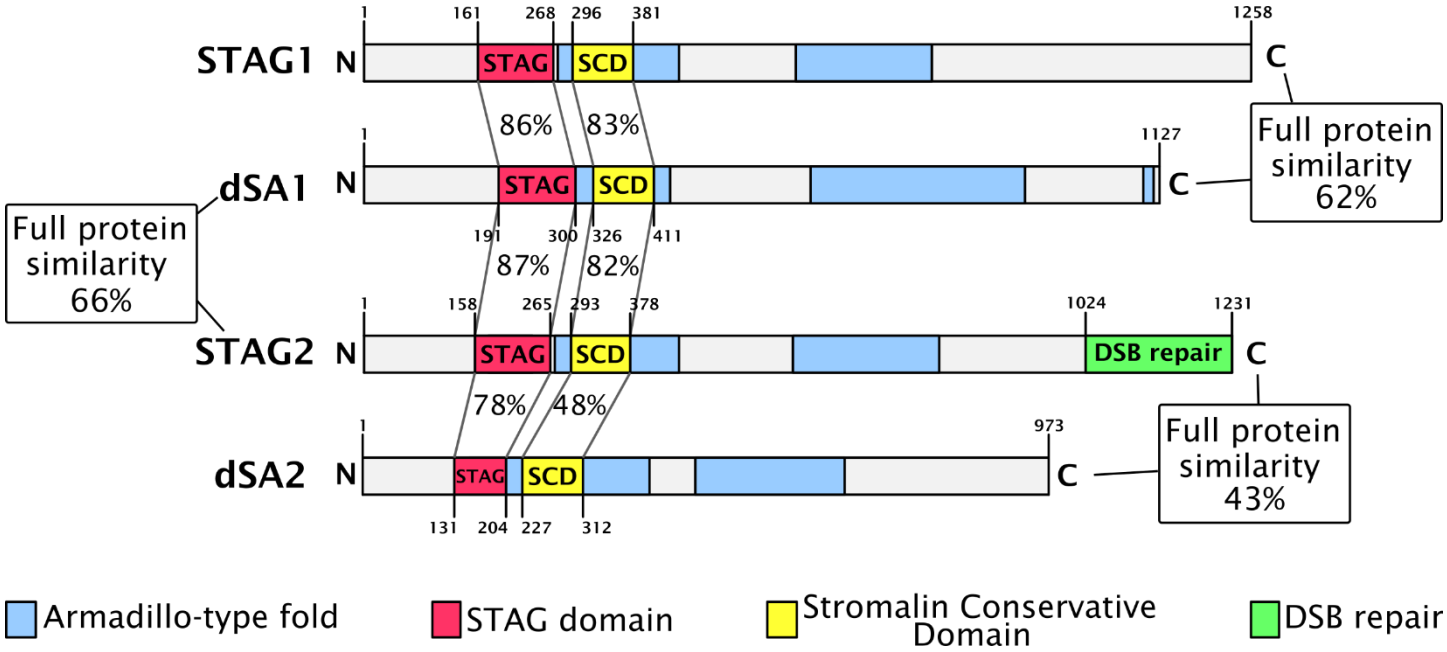


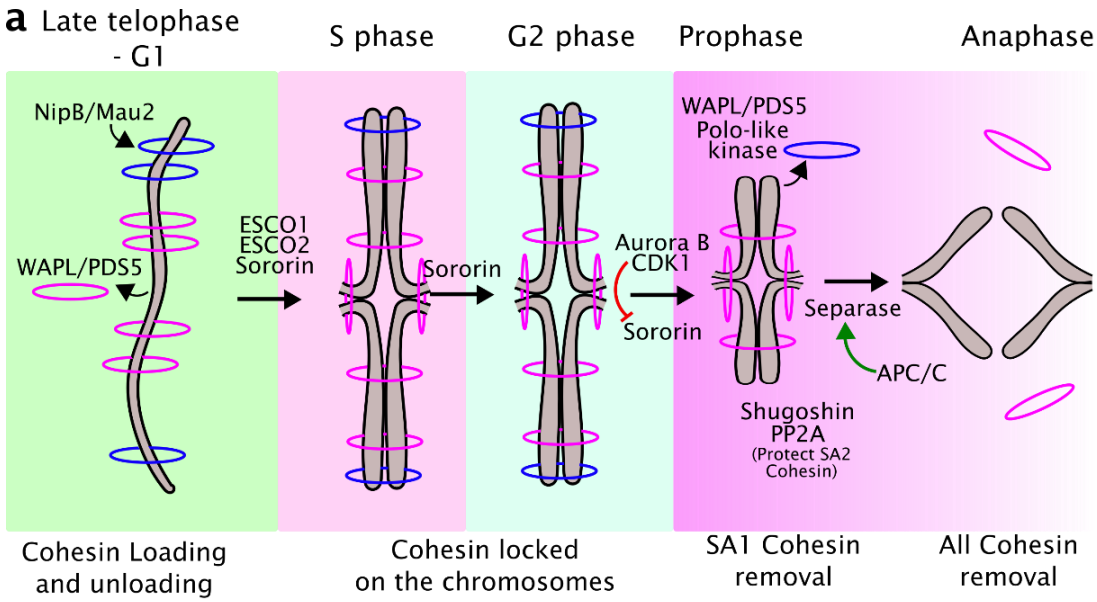
Figure 4.2 Conserved protein domains between Drosophila SA and human STAG proteins

Linear protein domain models of *Drosophila* SA1 and SA2 and human orthologs STAG1 and STAG2. Domains were identified by Uniprot and Interpro databases. Percentages of similarity were identified by Interpro. STAG and stromalin conservative domain (SCD) domains bind directly to RAD21 and are tightly conserved. Human STAG2 has a specific domain linked to the localization of the cohesin complex to double strand break repair sites.

As mentioned before, the canonical role of the cohesin complex is SCC during cell division, this requires the coordinated loading and unloading of the cohesin ring throughout the cell cycle. In short, the ring encircles the DNA as soon as the chromatids are segregated to each daughter cell, in late telophase. During S phase this interaction is strengthened and cohesin provides stability to the newly formed strands during DNA replication (Gerlich *et al.*, 2006). Both sister chromatids are kept together by cohesin until the onset of anaphase, where timely disassociation of the ring, in combination with the microtubule machinery (Wong *et al.*, 2008), anaphase-promoting complex (APC/C) (Waizenegger *et al.*, 2000) and mitotic checkpoint complex (MCC) (Chao *et al.*, 2012) allow the distribution of an exact chromatid copy to the newly formed daughter cells (Figure 4.3a).

Interaction of cohesin with the chromosomes is established by key regulators. The NipB-Mau2 heterodimer loads the ring to the DNA (Klein *et al.*, 2005). ESCO1, ESCO2 and CDCA5 (or Sororin) provide stability to the core of the ring during DNA replication (Hou *et al.*, 2005; Zhang *et al.*, 2012). WAPL and PDS5 unload STAG1 cohesin from the telomeres in prophase (Losada, 2005; Kueng *et al.*, 2006), while STAG2 cohesin is protected from unloading by SGO1 (shugosin) (Kitajima *et al.*, 2006). Therefore, STAG1 and STAG2 cohesin provide SCC during the S and G2 phases, while STAG2

cohesin continues to bind the sister chromatids together during mitosis; at the onset of anaphase, the APC/C complex activates the cysteine protease, separase, which binds to RAD21, releasing the remaining cohesin from the centromeres (Hauf, 2001) (Figure 4.3a).



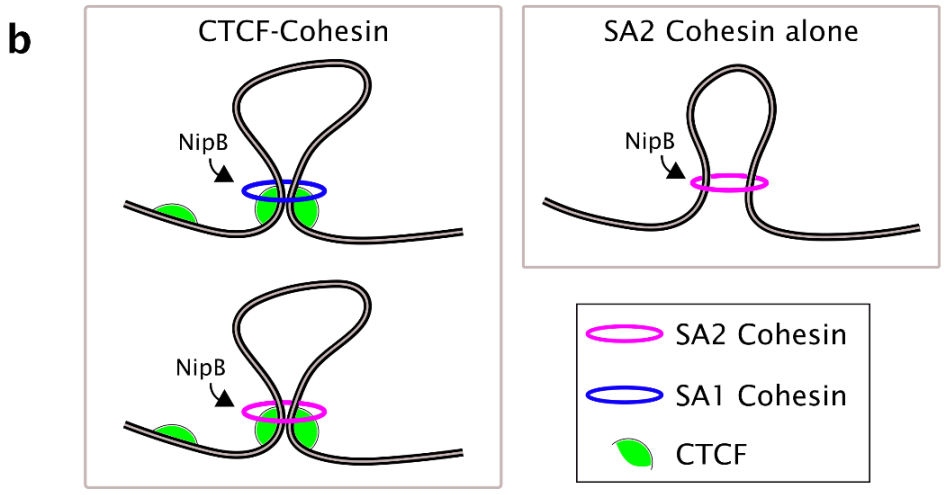


Figure 4.3 Cohesin interactions with the chromosomes throughout the cell cycle

a) Cohesin loading and unloading throughout the cell cycle. Cohesin is loaded to the chromosomes after sister chromatid segregation, this interaction is highly dynamic with constant loading (NipB/Mau2) and unloading (WAPL/PDS5) of the cohesin rings to allow chromatid re-arrangement. During S phase cohesin is locked on the chromosomes, ESCO1/2 acetylate SMC3, stabilizing the cohesin-DNA interaction; Sororin binds to PDS5, blocking the unloading complex. As the cell enters mitosis, Aurora kinase B and CDK1 remove Sororin from the cohesin unloading complex. Polo-like kinase phosphorylates SA1, facilitating SA1 cohesin removal from the telomeres. SA2 cohesin is protected from removal by Shugoshin and PP2A binding. At the onset of anaphase, the anaphase-promoting complex (APC/C) activated separase, which cleaves RAD21, releasing all cohesin and allowing sister chromatid segregation. **b)** Cohesin mediated chromatin re-arrangement. SA1 and SA2 cohesin are recruited to CTCF sites and induce chromatin loop formation that allows long range interactions. Cell type specific chromatin loops formed by SA2 cohesin without CTCF are also commonly found near gene enhancers. Adapted from (Pratap Singh *et al.*, 2015).

4.1.2 Roles of cohesin besides sister chromatid cohesion

Historically, cohesin research has focussed to its role during cell division. However, in the past decade, interest has been raised to the potential roles of cohesin in post mitotic cells, uncovering a very important role for the cohesin ring in transcriptional regulation (Lengronne *et al.*, 2004; Wendt *et al.*, 2008). Genome wide mapping studies in *S. cerevisiae* found cohesin enrichment at transcriptionally active sites (Glynn *et al.*, 2004). Additionally, in mice and humans cohesin is also found accompanying the CCCTC binding factor (CTCF), a transcriptional regulator known for inducing DNA loop formation (Parelho *et al.*, 2008; Rubio *et al.*, 2008). The current model suggests that CTCF recruits cohesin to transcriptionally active sites where cohesin mediated DNA loops are formed (Figure 4.3b). The loop formation allows the interaction between long range genetic regions. Depending on whether a silencer or an enhancer is either moved to proximity or separated from a gene promoter, cohesin in combination with CTCF might silence or induce gene transcription. In humans, both STAG1 and STAG2 cohesin have been shown to be associated with CTCF, yet STAG2 cohesin is also found at transcriptionally active sites in the absence of CTCF, only accompanied by the cohesin loader NipB (Kojic *et al.*, 2018) (Figure 4.3b). It is still unclear how cohesin might regulate gene expression in non CTCF sites, it probably acts in

a similar manner, by creating DNA loops, yet how the ring is recruited to these sites is currently unknown.

An additional role for cohesin was discovered when double strand break (DSB) repair defects were observed in *rad21* mutants in *S. pombe* (Birkenbihl *et al.*, 1992). The current mechanistic role describes cohesin binding to the damaged DNA, providing stability to the broken strands, and allowing the repair machinery to perform their role efficiently. Unal et al have proposed that histone H2AX posttranslational modifications near the DSB site form a conformation that allows cohesin recruitment and binding (Unal *et al.*, 2004), yet more research is needed to better understand this process. Interestingly, only STAG2 cohesin is involved in this process. Kong et al have identified the C terminal region of the STAG2 protein to be directly involved in the recruitment of the complex to the damaged site (Kong *et al.*, 2014). The C and N terminal regions of STAG1 and STAG2 share no homology between each other, while the central part of the protein, where the cohesin interacting domains are located are highly similar (Figure 4.2). Tt has been proposed that it is the C and N domains of the STAG proteins that determine the differences in localization and function between the two somatic cohesins. Furthermore, chromatin condensation, mediated by loop formation (Schalbetter *et al.*, 2017) and preventing telomere replication (Daniloski *et al.*,

2017) are other very recent novel roles that have been attributed to cohesin, indicating that there might be many other unknown processes influenced by cohesin activity.

4.1.3 Cohesin defects and disease

In humans, germ mutations that reduce cohesin function lead to developmental diseases collectively called "cohesinopathies", including:
i) Cornelia de Lange syndrome (CdLS), caused by heterozygous mutations in *nipB*, *smc1*, *smc3* and *stag3*; patients show developmental delays, microcephaly, craniofacial and upper limb abnormalities (Krantz *et al.*, 2004).
ii) Roberts syndrome, a very rare disease, caused by recessive mutations in *esco2*; the phenotypes are very similar to CdLS (Vega *et al.*, 2005). Patients from both cohesinopathies have effective SCC, and therefore it is hypothesised that these syndromes are caused due to defects in cohesin mediated gene expression rather than genomic instability.

A full loss of function of the cohesin ring leads to defective chromosome segregation; therefore, it is no surprise that somatic homozygous mutations of cohesin subunits have been highly implicated in cancer through aneuploidy (Cucco *et al.*, 2014). Mutations in n*ipB*, *stag1*/2, *smc1*/3, and *rad21* are found in colorectal cancer and acute myeloid leukaemia (AML) (Barber *et*

al., 2008; Rocquain *et al.*, 2010). Furthermore, *Stag2* is the most mutated Cohesin subunit in human cancers, inactivating mutations of *stag2* have been additionally found in glioblastoma multiforme, bladder and breast cancer (Kandoth *et al.*, 2013; Leiserson *et al.*, 2015). Hypothesis for this include that, since *Stag2* is located in the X chromosome, only one mutation is needed for its inactivation (Soardi *et al.*, 2017); another explanation could be that, since STAG2 is involved in tissue specific gene expression, its loss might have a higher impact in certain tissues (Faure *et al.*, 2012). Specifically in breast cancer, *stag2* has been found significantly mutated in metastatic patients compared to lower tumour grades (Repo et al., 2016). Surprisingly, loss of STAG2 in AML, colorectal and other cancers is accompanied by a normal karyotype (Walter et al., 2012; Kim et al., 2016). It has been shown that loss of STAG2 triggers compensation by STAG1 Cohesin, and vice versa (Benedetti et al., 2017). Therefore, it is likely that STAG1/STAG2 compensation is the reason why no aneuploidy is observed. Although Stag2 has been shown to be potentially acting as a tumour suppressor gene in different cancer types, little to no research has been done to understand the molecular mechanism of action. Furthermore, there is little understanding of the mechanisms of loss of cohesin driven tumour progression when no aneuploidy is observed, making the cohesin subunits interesting genes to study.

4.2 Results

4.2.1 SA1 and SA2 KD promotes single cell invasion in the living epithelium.

In the genetic screen, SA1 KD promoted cell invasion in *lgl* mutant tumours in the notum of the living fly (Chapter 3). It is therefore important to determine whether the invasive phenotype was a consequence of a co-operation between the loss of SA1 and LGL function, or if the KD of SA1 alone would be sufficient to promote invasion. GFP:Moe positive clones that specifically express an RNAi hairpin construct, designed to knockdown the expression of each cohesin subunit, were expressed in the epithelial cells of the fly notum using MARCM (Lee *et al.*, 2001). As previously mentioned, a mitotic cohesin complex is formed with either SA1 or SA2 (Figure 4.1). Since both SA proteins have been very poorly studied to date, there are limited tools available to use. Particularly for SA2, only two UAS-RNAi lines are available commercially and they are both located on the 2nd chromosome; since *lgl* is also located on chromosome 2, it was not possible to simultaneously KD SA2 expression within *lgl* homozygous mutant clones. However, the effect of the SA2 KD alone was studied and included in this section. Significantly increased invasion was observed in both SA1 KD (8.82% of the mutant cells invaded, n=10 animals, $p<0.0001$) and SA2 KD (7.27%, n=10, $p<0.005$), compared to WT animals (3.41%, n=10). Conversely,

SMC1 KD (2.75%, n=10, p=0.95), SMC3 KD (3.8%, n=10, p=0.99) and RAD21 KD (2.61%, n=10, p=0.91) each had no effect on the number of invading cells when compared to WT animals (Figure 4.4 a-d). These results, though less severe, show the same trend and significance difference to cohesin subunits KD in combination with the *lgl* mutation (Figure 3.10), suggesting that SA has a specific role in the regulation of single cell invasion, which is largely independent of *lgl* function.

Additionally, multilayering and apex defects were also measured in cohesin KD clones without the *lgl* mutant background. A significant increase in multilayered tissue was observed for all cohesin subunit KD experiments, when compared to WT clones (Figure 4.4e): SA1 KD (31.67% of clonal tissue was multilayered, n=10 animals, $p<0.01$); SA2 KD (23.24%, n=10, $p<0.05$); SMC1 KD (24.26%, n=10, $p<0.05$); SMC3 KD (35.47%, n=10, $p<0.001$); and RAD21 KD (31.48%, n=10, $p<0.01$). WT (0%, n=10).

Mild apex defects were also observed in: SA1 KD (4.21% of clonal tissue had apex defects, n=10 animals, $p<0.05$); SA2 KD (3.09%, n=10, $p<0.05$); SMC1 KD (3.40%, n=10, $p<0.05$); SMC3 KD (7.18%, n=10, $p<0.05$); and RAD21 KD (7.32%, n=10, $p<0.0001$) when compared to WT tissue (0%, n=10) (Figure 4.4 f). Overall these results suggest that the cohesin complex might regulate

cytoskeletal structure and/or apicobasal polarity, yet only SA1 and SA2 act as invasion suppressors.

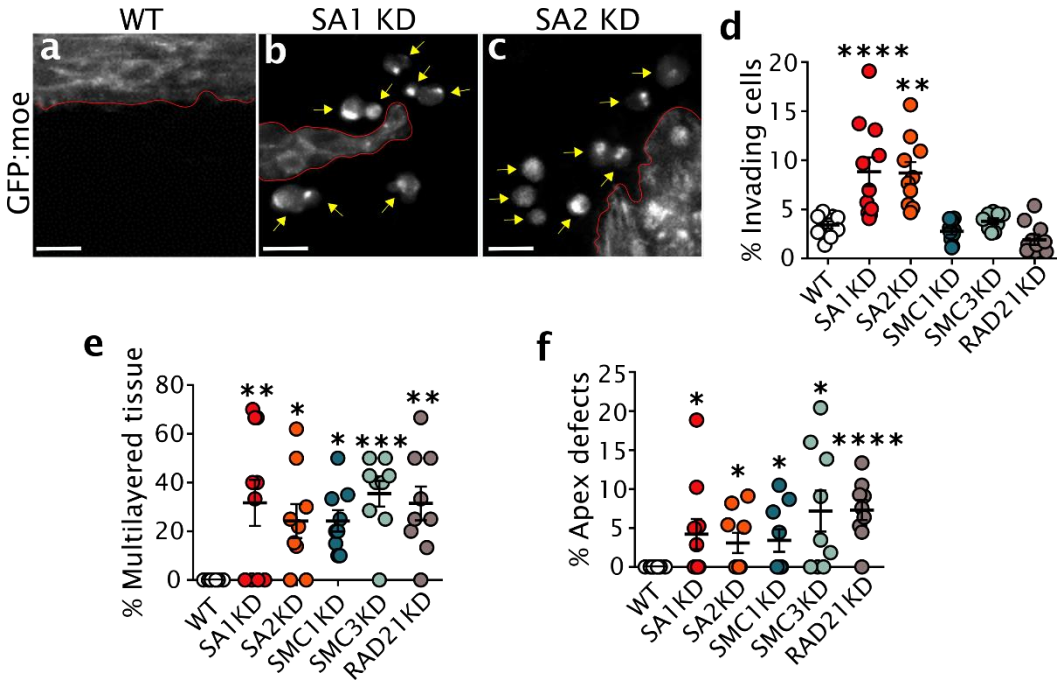


Figure 4.4 Effects of cohesin subunit KD on invasion, apex defects and multilayering

(a-c) Confocal images of the basal surface of moe-GFP labelled WT **(a)**, SA1 KD **(b)** or SA2 KD **(c)** clones in the living epithelium of the dorsal thorax of *Drosophila* pupae. Yellow arrows highlight invading cells. Red dashed line outlines clonal tissue. **(d)** Quantification of **(a-c)**, shown as percentage of mutant invading cells against total number of mutant cells. **(e)** Quantification of the percentage of multi-layered tissue, calculated by measuring the area in μm^2 of multi-layered tissue against total clonal area. **(f)** Quantification of apex defects calculated by measuring the area in μm^2 of tissue with apical defects (described in detail in Chapter 3) against total clonal area. Statistics performed: One-way ANOVA with Dunnett's post hoc test for multiple comparisons. **** $p<0.0001$. *** $p<0.001$, ** $p<0.01$, *$p<0.05$. Scale bars: 10μm.

4.2.2 Investigating the migration speed and directionality of SA1 and SA2 KD invading cells

The uniqueness of the invasive phenotype observed by the KD of SA1 and SA2 compared to the other cohesin subunits was a very interesting finding that I decided to investigate further. SA1 KD invading cells appear to be very motile, which is unlike many hits for invasion in our screen, many of which appear to have no directionality to their migration and randomly move about over a number of hours. One of the biggest advantages of a live model is that directionality and speed of the invading cells can be studied and quantified in real time (Figure 4.5a). Single cell tracking of *lgl*$^{-/-}$ and SA1 KD invading cells was performed to determine the X, Y and Z trajectories and calculate their speed and directionality. An illustration of representative trajectories measured is shown for 5 cells in Figure 4.5b for X and Y planes, and for the same cells in X and Z in Figure 4.5b'. To determine directionality, the trajectory of each cell was measured every 3 minutes for a total of 30 minutes (10 readings for each cell), the total number of micrometres travelled was documented (Length in Figure 4.5 c). Additionally, the distance an invading cell would have travelled if following a straight line from 0min to 30min was measured, (Displacement in Figure 4.5 c).

If an invading cell has directionality, their length and displacement should be very similar. Quantification of the length and displacement of 25 cells from 5 animals is shown in Figure 4.5d, there is a significant increase in the length and displacement for SA1 KD cells (41.55μm length, $p<0.01$. 26.55μM displacement, $p<0.05$) compared to *lgl* mutant cells (16.07μm length. 4.16μm displacement). Furthermore, there is no significant difference between length and displacement in SA1 KD cells, indicating that their trajectories are directional. Furthermore, the speed of migration was higher by 2.7 fold in SA1 KD (1.46μm/min, $p<0.01$) when compared to the tumour suppressor gene *lgl* (0.53μm/min) (Figure 4.5e).

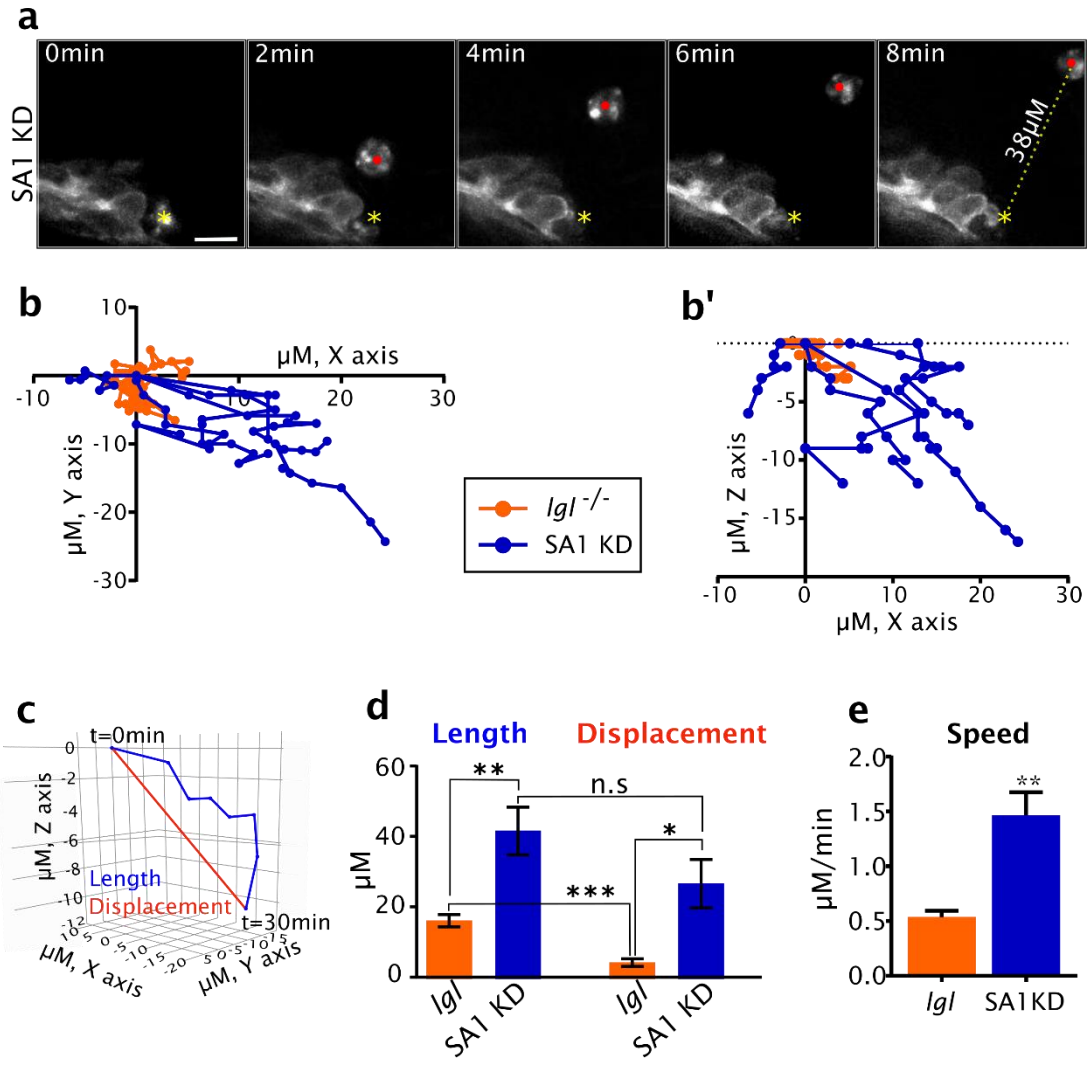


Figure 4.5 Investigating the migration speed and directionality of SA1 KD invading cells

(a) Stills from a time lapse of the basal surface of a GFP:moe labelled SA1 KD clone. Yellow star marks the initial location of an invading cell. Red dot shows the location of the invading cell at the indicated time. The cell shown has moved 38μm in 8 minutes. Scale bar=10μm **(b-b')** Representative single cell trajectories from *lgl*$^{-/-}$ invading cells (orange) and SA1 KD invading cells (blue). The position of each cell was measured every five minutes and each cell was followed for 30 minutes. To better display the 3-D trajectory for each cell, both their X/Y and X/Z movement is shown (e-e'). **(c)** Illustration showing the two trajectories measured for each invading cell in order to determine directionality. Length follows the full trajectory of an invading

cell. Displacement measures a straight line from the initial point to the final point. **(d-e)** 25 cells from 5 animals were used for quantification. **(d)** Quantification of length and displacement from SA1 KD cells compared to *lgl*$^{-/-}$ cells. Cells that have directionality had no significant difference between length and displacement. **(e)** Quantification of speed of migration, showing average of μm travelled per minute. Statistics performed: Student's t-test. *** $p<0.001$, ** $p<0.01$, *$p<0.05$. Scale bars: 10μM.

4.2.3 SA1 and SA2 invading cells are not apoptotic

Live analysis of the levels of apoptosis was performed to discard the possibility that the cell delamination observed in SA1 and SA2 KD clones was a consequence of apoptotic cell extrusion or that delamination was followed by anoikis, as observed in developing WT animals (Marinari *et al.*, 2012; Levayer *et al.*, 2016). The in vivo caspase sensor, iCasper (To *et al.*, 2015), was used to measure the levels of apoptosis in invading cells. The iCasper construct was inserted in the 3rd chromosome of the *Drosophila* genome; since every RNAi line available for SA1 is also in the 3rd chromosome, this experiment was only done for SA2. Young WT pupae were used as a control (12-15h APF), since older WT animals have little to no invading cells (Figure 4.4a,d). In accordance with Marinari and Levayer publications, nearly all the invading cells from WT animals were iCasper positive (94.26%, n=50 cells from 5 animals), while around half of the SA2 KD invading cells were iCasper positive (56.25%, n=50 cells from 5 animals, $p<0.001$) (Figure 4.6a-c). Additionally, there is a 4-fold increase in the proportion of iCasper negative invading cells in SA2 KD compared to WT (Figure 4.6c).

An additional indirect study was also performed; the baculovirus p35 was overexpressed specifically within WT, SA1 and SA2 KD clones. p35 binds to

the effector caspase, caspase-8 in the fly, and changes its conformation, therefore rendering it inactive and preventing apoptosis (Hay *et al.*, 1994). If the promotion of delamination by SA1 or SA2 KD was a consequence of apoptosis, the overexpression of p35 should block this delamination. If the SA1 or SA2 KD promotes delamination by different mechanisms, increased invasion would still be observed even when apoptosis is blocked. WT clones overexpressing p35 had nearly no invading cells (0.25%, n=8) while there was a significant increase in SA1 KD +p35 (3.02%, n=8, $p<0.0001$) and SA2 KD + p35 (3.63%, n=8, $p<0.0001$) invading cells compared to p35 alone (Figure 4.6d). Overall these results indicate that the effect that the loss of SA1 and SA2 have on promoting delamination is independent of apoptosis and therefore other mechanisms must be altered.

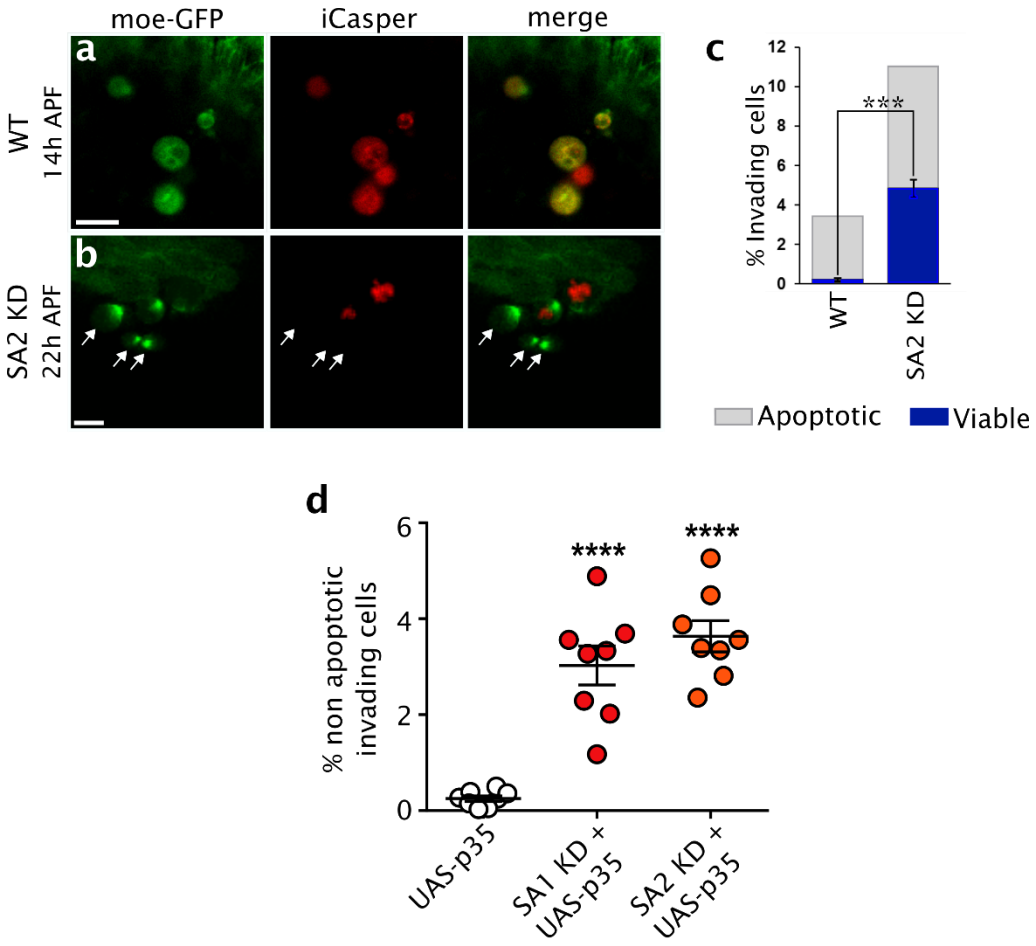


Figure 4.6 SA1 and SA2 invading cells are not apoptotic

(a-b) Confocal images of the basal surface of iCasper (red) and GFP:moe (green) labelled (a) WT clones and (b) SA2 KD clones. Arrows highlight invading cells that are iCasper negative. Scale bars = 10μm. **(c)** Quantification of (a-b). Invading cells categorized as viable (blue) are iCasper negative. Apoptotic cells (light grey) are iCasper positive. 50 cells from 10 animals were measured for each genotype. **(d)** The baculovirus inhibitor of apoptosis, p35, was overexpressed in WT, SA1 KD and SA2 KD tissue. n=10 animals. Statistical analysis: Student's t-test. **** $p<0.0001$. *** $p<0.001$.

4.2.4 Ectopic localization of adherens and septate junction proteins in SA1 and SA2 KD clones.

Adhesion and polarity have been shown to be crucial in the maintenance of cell shape, cell-cell connections and tissue integrity. Since loss of SA1 or SA2 promotes drastic cell shape changes that lead to invasion, which was not induced by cell death, the next steps followed in the characterization of these genes was to study polarity and junctional integrity. SA1 and SA2 KD positively marked clones with GFP:moe were generated, surrounded by WT tissue. This allowed the direct comparison of the localization of polarity and junctional proteins between mutant and WT cells within the same tissue. The mosaic tissue was then stained for adherens junction (AJ) proteins: E-Cadherin, Armadillo (β-catenin), and α-catenin; and the fly equivalent of tight junctions, septate junction (SJ) proteins: Fasciclin III (FASIII), Coracle (COR) and Discs Large (DLG). A significant reduction in the localization of E-cadherin (Figure 4.7a-a''), alpha-catenin (Figure 4.7b-b''), β-catenin (Figure 4.7c-c'') and FAS III (Figure 4.8a-a'') was observed at the junctional level in both SA1 and SA2 KD clones, when compared to the surrounding WT tissue, with some levels of ectopic localization of AJ and SJ in the cytoplasm of the mutant cells (Figure 4.7 and 4.8). Additionally, there was no significant change in the localization of the septate junction proteins DLG and COR; and the polarity protein aPKC in SA1 and SA2 KD clones when compared to WT

tissue (Figure 4.8b-d). These results suggest that SA1 and SA2 act as invasion suppressors by maintaining cell-cell junction integrity.

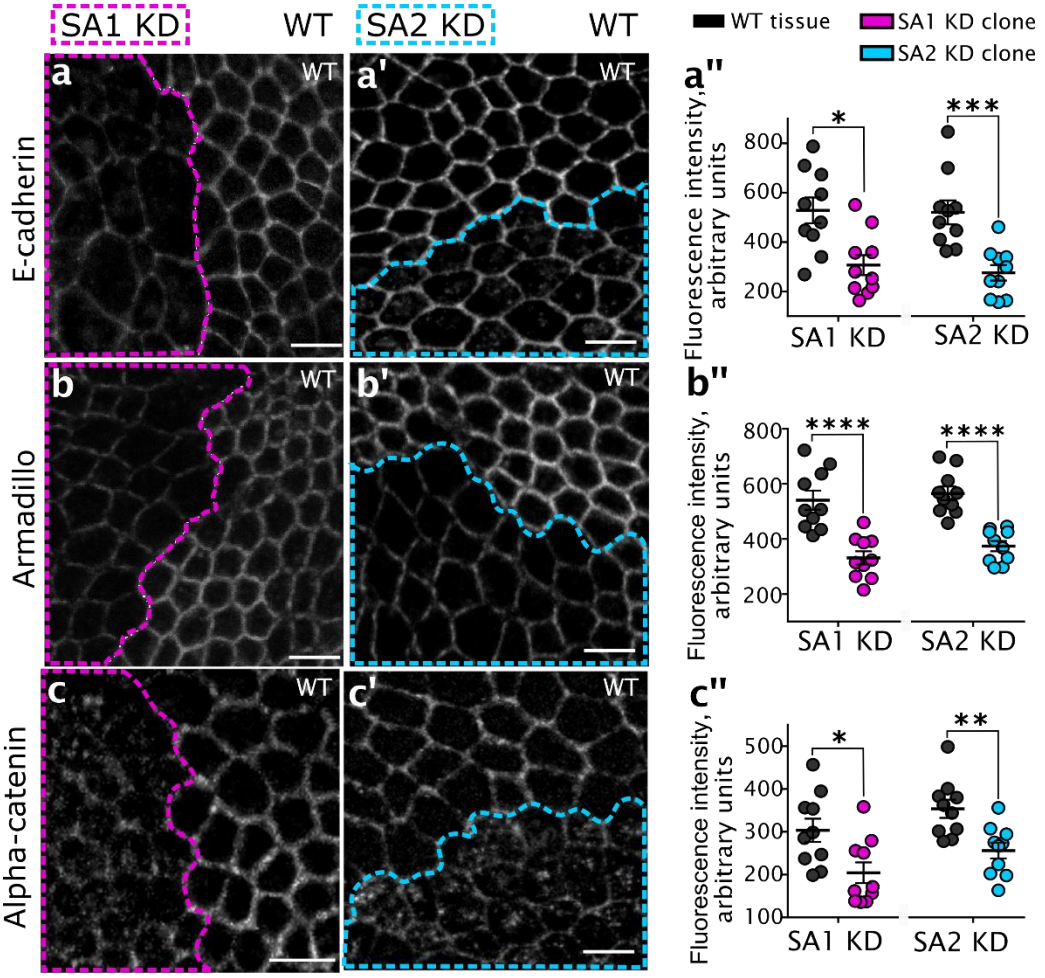


Figure 4.7 Adherens junction protein localization is disrupted in SA1 and SA2 KD clones

20-24h APF *Drosophila* nota containing positively marked clones were dissected and stained for proteins essential to the establishment of the adherens junction: **(a-a'')** E-Cadherin; **(b-b'')** Armadillo, (the fly orthologue of β-catenin); **(c-c'')** alpha catenin. Mutant clones are highlighted: SA1 KD tissue is outlined with the magenta dashed line; SA2 KD tissue is outlined with the cyan dashed line. Tissue surrounding these cones are WT and serve as an internal control. (a''-c'') show the quantification of the accumulation of the adherens junction proteins at the junction (n=100 junctions from 10 animals for each genotype). Error bars represent s.e.m. Student's t-test was performed to determine statistical significance. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. Scale bars=10μM.

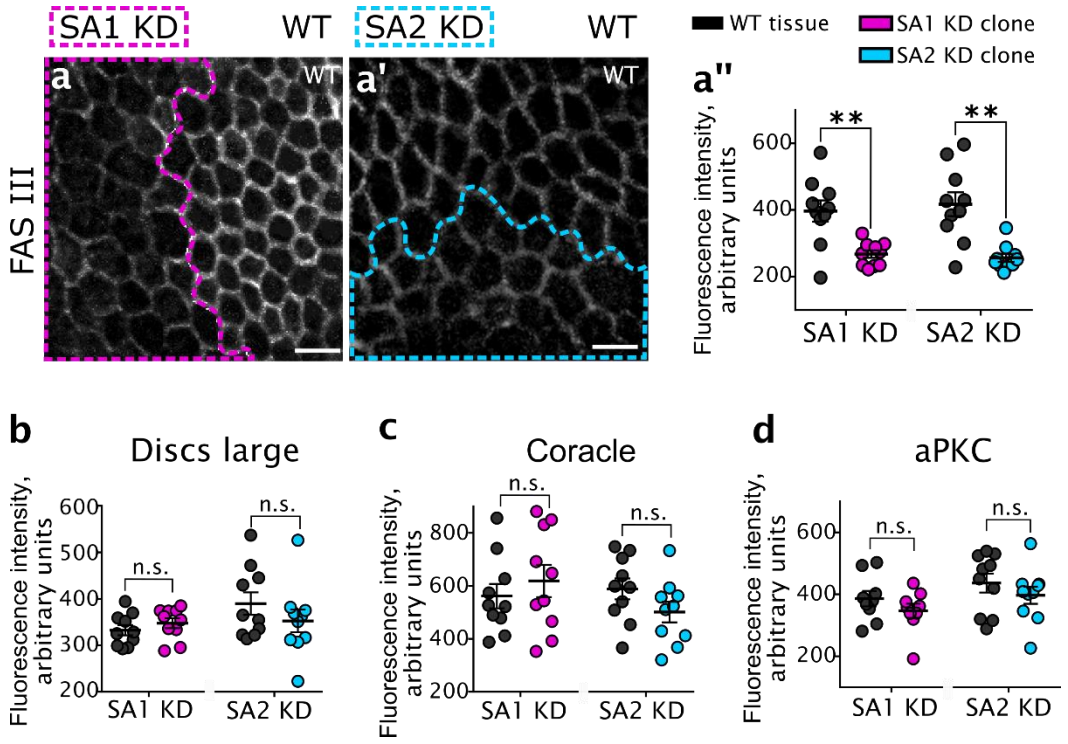


Figure 4.8 FasIII localization is disrupted in SA1 and SA2 KD clones but no effect is observed on other septate junction and polarity proteins

20-24h APF *Drosophila* nota containing positively marked clones were dissected and stained for proteins essential to the establishment of the septate junction (equivalent to the mammalian tight junction): **(a-a'')** FASIII; **(b)** discs large; **(c)** coracle and **(d)** the polarity protein aPKC. Mutant clones are highlighted: SA1 KD tissue is outlined with the magenta dashed line (a); SA2 KD tissue is outlined with the cyan dashed line (a'). Tissue surrounding these cones are WT and serve as an internal control. (a''-d) Graphs show the quantification of protein levels at the junction for the indicated proteins (n=100 junctions from 10 animals for each genotype). Error bars represent s.e.m. Student's t-test was performed to determine statistical significance. **p<0.01, n.s. not significant. Scale bars=10μM.

4.2.5 STAG1 and STAG2 KD promotes migration and invasion in a human adenocarcinoma cell line

To determine if the role of SA1 and SA2 as invasion suppressors is conserved in different species, I next studied the effect that the loss of their human orthologues, STAG1 and STAG2, would have in human cell invasion in-vitro. Loss of function mutations of STAG2 are found significantly in metastatic breast cancer tumours compared to lower grades (Repo *et al.*, 2016), suggesting that STAG2 has a role in preventing the transition of a benign tumour to malignancy. I therefore decided to study the effect of the cohesin subunit KD in breast cancer. The MCF7 adenoarcinoma cell line was a gift from Dr. Anna Grabowska; this line is not invasive and expresses all the subunits of the cohesin complex, making it a good model to study the effect of the cohesin subunit KD in single cell invasion.

A pool of siRNA designed specifically for each human cohesin subunit was transiently transfected in MCF7 cells. KD levels for each subunit were verified by RT-qPCR (Figure 4.9a). Further optimization was performed when knocking down STAG2, so that KD levels were similar between all the cohesin subunits (Figure 4.9b). Next, the cells were plated into Boyden chambers, these 3D culture structures allow the movement of the cells basally, through an 8μM polycarbonate membrane, and migratory cells are

collected and quantified after 48 hours. An additional barrier of extracellular matrix proteins is added to coat the polycarbonate membrane to study invasion. A significant increase of both migration and invasion was found in STAG1 and STAG2 KD cells, compared to untreated MCF7 cells, and no significant change in invasion was observed by the KD of hSMC3, hSMC1 or hRAD21 (Figure 4.10); these results corroborate that the subunits STAG1 and STAG2 have a specific role in regulating cell invasion and that their role is conserved in human breast cancer cells.

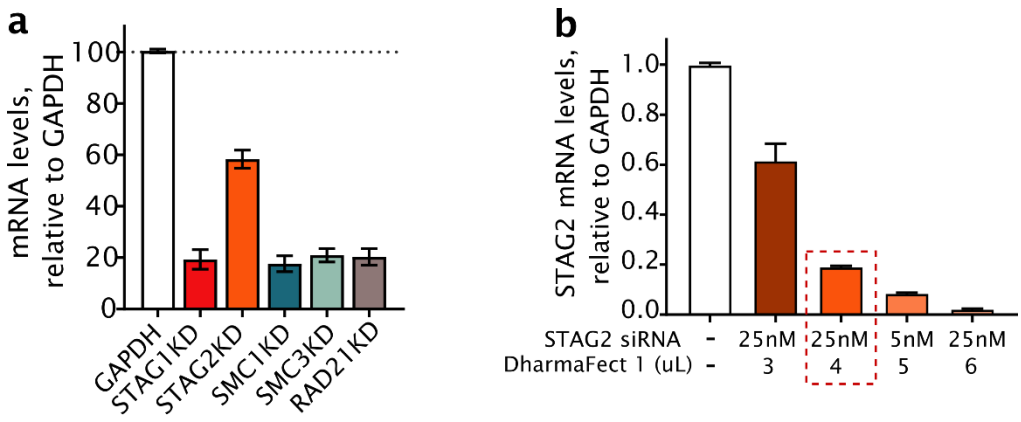


Figure 4.9 Cohesin subunit KD in MCF7 cells

(a) RT-qPCR measuring mRNA levels after siRNA treatment for each of the genes specified on the Y axis of the graph was used. Conditions for each of the genes were siRNA 25nM and Dharmafect1 3µL. Expression is shown as percentage of expression after KD compared to untreated cells, levels were normalized to GAPDH. **(b)** Optimization of STAG2 KD with different Dharmafect1 concentrations. 25nM siRNA and 4µL of Dharmafect1 was used for subsequent experiments.

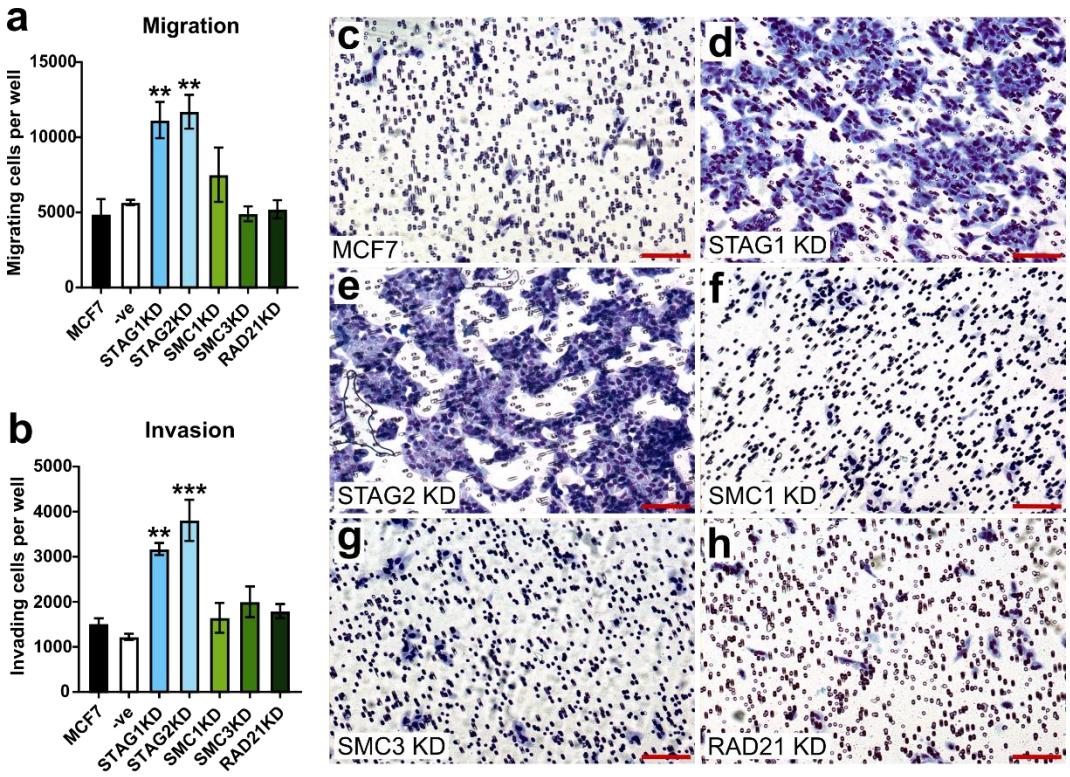


Figure 4.10 STAG1 and STAG2 are invasion and migration suppressors in MCF7 cells

In vitro transwell migration and invasion assays with the human adenocarcinoma cell line, MCF7, with or without the KD of human orthologues of the somatic cohesin complex subunits. Quantification of **(a)** migration and **(b)** invasion, showing the number of MCF7 cells, untreated or transfected with siRNA targeting the indicated gene, which migrated through an 8µm pore membrane. For the invasion assay, the membrane was coated with a layer of Matrigel. n=3 for each assay. Error bars represent ± s.e.m. One-way ANOVA with Dunnett's post hoc test for multiple comparisons was performed to determine statistical significance. **(c-h)** Representative images of the invasion assay, showing increased invasion specifically with STAG1 or STAG2 KD. Red scale bar: 100µm.

4.2.5 STAG2 KD induces changes in the transcriptome of MCF7 cells

The cohesin complex has been linked to the regulation of gene expression, therefore I hypothesised that STAG1 and STAG2 might regulate the expression of genes that are involved in cell-cell junction or cytoskeletal integrity. Since loss of STAG1 and STAG2 promotes invasion in MCF7 cells (Figure 4.10), and STAG2 is the most abundant and most mutated cohesin gene in human cancers, the effect that the KD of STAG2 has on gene expression in MCF7 cells was studied by genome wide transcriptomic analysis through microarrays. Three different total RNA samples were prepared in triplicate: untreated MCF7 (unt), MCF7 non-targeting siRNA (non-T) and MCF7 STAG2 siRNA. Samples were then sent to the Nottingham Arabidopsis Stock Centre (NASC) for processing.

The data was first analysed to determine the variability and validity of the results. Principal component analysis (PCA) was used to study the distribution of the data. PCA uses linear transformation to convert multidimensional data into linear values called principal components. In Figure 4.10a, each circle represents the principal component of the genome wide gene expression for each sample, this shows that the three samples from STAG2 KD (red) cluster together and are separated from unt (yellow) and non-T (blue) samples. Furthermore, there is little variability within

STAG2 KD samples and unt samples, while non-T samples had larger variability, yet this variability is distant to STAG2 KD (Figure 4.11a). Additionally, the F ratio was calculated, this compares the variability within the same group (i.e. within STAG2 KD triplicates) and compares it with the variability between different groups (i.e. STAG2 KD against unt triplicates). Figure 4.11b shows that the variability between different samples (attribute) is bigger than the variability within the same samples (error), therefore there is more confidence that the results obtained are a consequence of STAG2 KD rather than error.

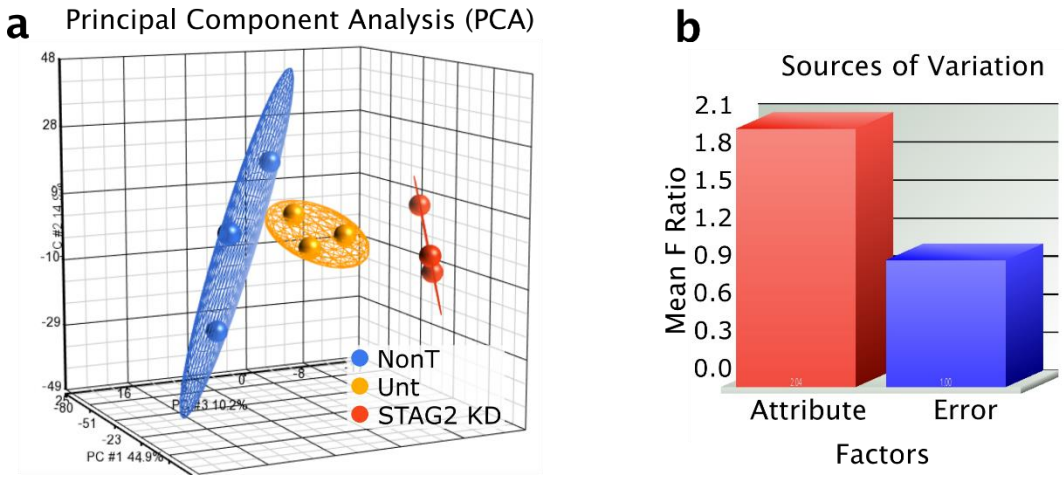


Figure 4.11 Data distribution from microarray analysis of MCF7 transcriptome

(a) Principal component analysis plot of untreated MCF7 cells (yellow), MCF7 transfected with non-targeting siRNA (blue) and MCF7 transfected with STAG2 siRNA. Each individual dot represents the distribution of one individual sample. (b) Bar graph showing the sources of variation in the genome-wide mRNA expression analysis. The *X*-axis of the plot represents the factors or interactions in the ANOVA model and the *Y*-axis represents the *F* ratio of the factors or interactions. *F* ratio describes the signal to noise ratio and is the average signal to noise ratio of all the computed variables for each factor.

Next, volcano plots were made where each gene is plotted individually based on the mean value from the three samples of two-way ANOVA analysis of fold change of expression against p-value. A cut-off of <-1.5 and >1.5 was established for fold change and $p<0.01$ for significance. As negative control, the expression of genes between non-T samples were plotted against unt samples, four genes were differentially expressed significantly between these samples (Figure 4.12a), therefore these genes were a consequence of error. A total of 33 genes were differentially expressed between unt and STAG2 KD (Figure 4.12b) and 32 genes between non-T and STAG2 KD (Figure 4.12c). Then, to reduce the possibility of false positives, only the genes that were differentially expressed in STAG2 KD against unt and in *STAG2* KD against non-T were selected. A total of 23 genes were selected to be differentially expressed in MCF7 cells by the KD of STAG2.

Unsupervised hierarchical clustering was applied to the fold change of the 23 differentially expressed genes, two main clusters were identified from the samples, one containing non-T (orange) and unt (yellow) samples and the other STAG2 KD samples (green) (Figure 4.13, vertical clustering), confirming that these genes are differentially expressed in STAG2 KD. Additionally, two other main clusters were identified, one includes the genes that are downregulated by STAG2 KD and the other upregulated genes (Figure 4.13, horizontal clustering).

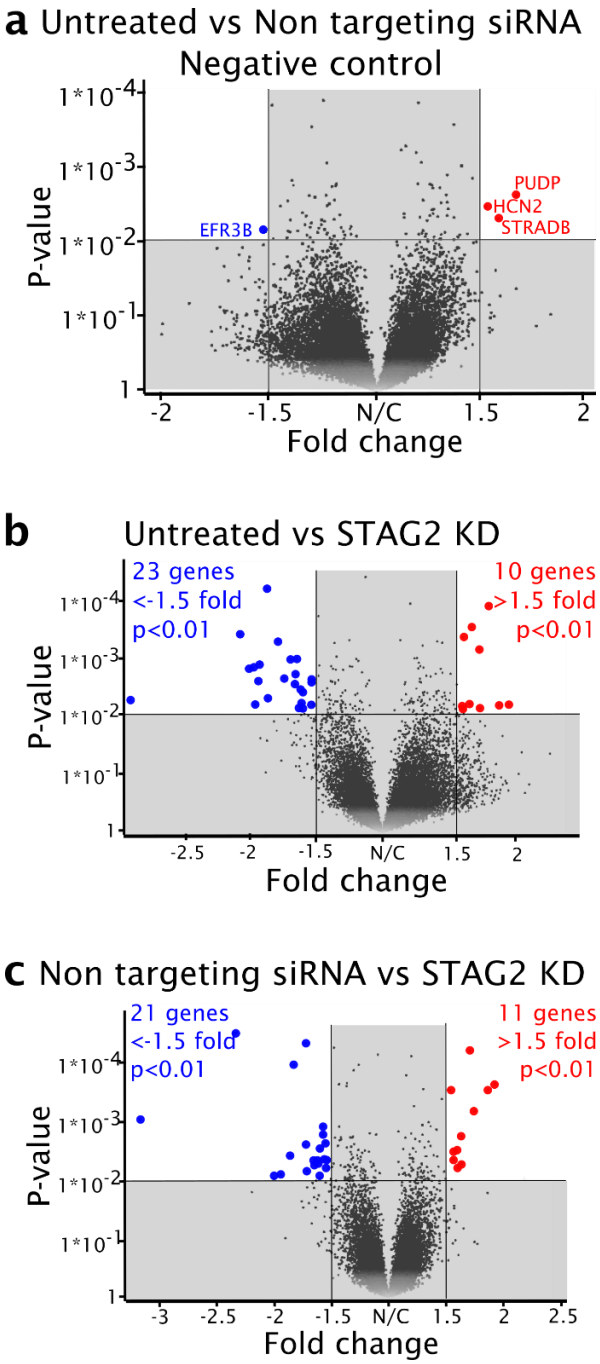


Figure 4.12 Volcano Plot of differentially expressed genes

Volcano plot of significance level (p value determined by one-way ANOVA) versus fold change in genetic expression from treated and untreated MCF7 cells. Each grey dot represents a gene that had no significant change in expression i.e. that had a p value greater than 0.01 and an expression fold change between -1.5 and 1.5. Each blue dot represents a gene that was downregulated compared to the control (Untreated in **a** and **b** and Non- targeting siRNA in **c**), these genes had a p value smaller than 0.01 and an expression fold change of less than -1.5. Red dots mark genes that were upregulated compared to the control, these genes had a p value smaller than 0.01 and an expression fold change greater than 1.5. Graph in **(a)** is the negative control, where genes from samples treated with non-targeting siRNA are compared to genes from untreated cells. Differentially expressed genes between STAG2 KD cells compared to untreated MCF7 cells and MCF7 cells treated with non-targeting siRNA are shown in **(b)** and **(c)** respectively.

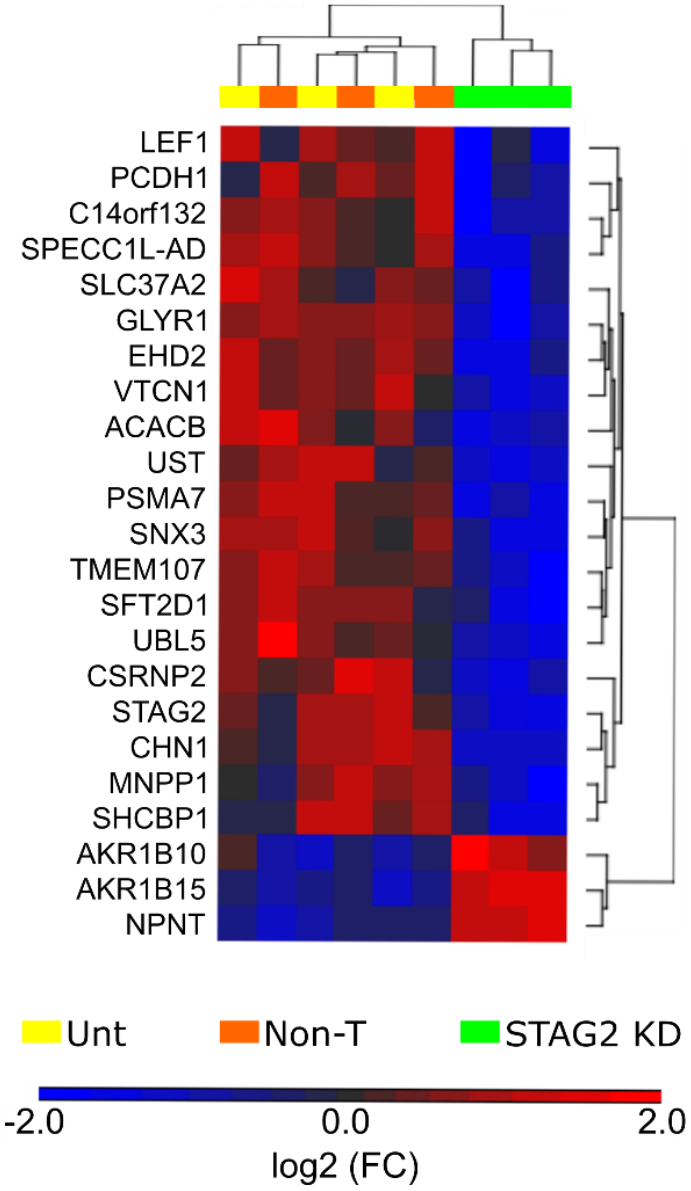


Figure 4.13 Hierarchical clustering of differentially expressed genes

Heat-map representation of unsupervised clustering of the 23 differentially expressed genes by STAG2 KD in MCF7 cells. Each row represents a gene. Each column represents a sample: yellow, untreated MCF7; orange, MCF7 non-targeting siRNA treated; green, MCF7 STAG2 siRNA treated. Colour code within the graph represents log2 of the fold change of expression: blue, downregulated; red, upregulated. Horizontal and vertical clusters were created based in Euclidean distance. The full list of differentially expressed genes can be found in Appendix 5.

RT-qPCR was performed for a selection of genes to verify the results obtained by microarray analysis. *Stag2, Pcdh1, Ehd2* and *Akr1b10* had significant difference in their expression in *STAG2* KD MCF7 cells when compared to untreated MCF7 (Figure 4.14 a-d). Comparison of the fold change levels as determined by microarray and qPCR were analysed; the same expression change trend was observed in the four genes analysed using either technique, with a stronger change detected by RT-qPCR in EHD2 and AKR1B10 (Figure 4.14e). Overall, both microarrays and RT-qPCR techniques showed a differential expression for the four genes analysed in STAG2 KD compared to untreated MCF7 cells.

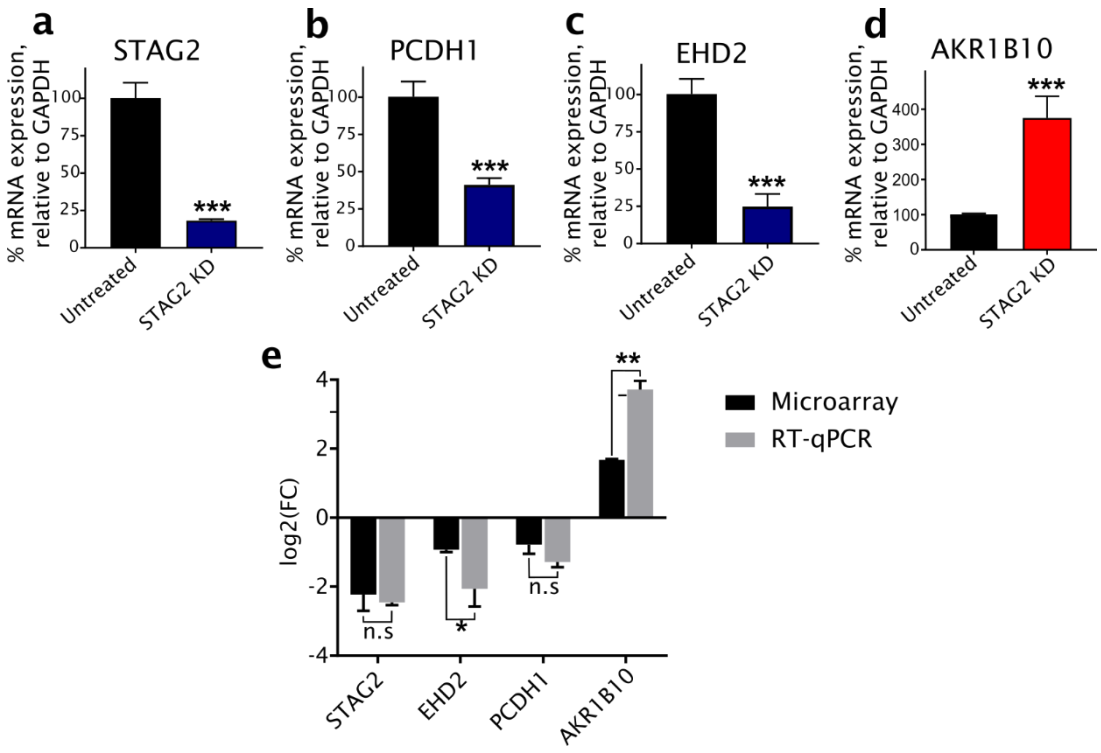


Figure 4.14 RT-qPCR validation of a selection of differentially expressed genes

(a-d) mRNA expression levels of four genes with differential expression by STAG2 KD in MCF7 cells were confirmed by RT-qPCR analysis. Data are expressed as percentage of mRNA levels, adjusted to GAPDH. **(e)** Bar chart comparing the results obtained from the microarray analysis and RT-qPCR validation expressed as log2 of fold change. Error bars represent s.e.m. Student's t-test was performed to determine statistical significance. n.s=not significant, *p<0.05, **p<0.01, ***p<0.001.

4.2.6 Biological function and interactions of differentially expressed genes in MCF7 cells by STAG2 KD

To determine the biological function of the differentially expressed genes by STAG2 KD, gene ontology (GO) analysis was performed to the 23 significantly mis-regulated genes, using the DAVID analytical tool. Six biological processes were significantly enriched within the differentially expressed genes: Transport (SFT2D1, VTCN1, SLC37A2, SNX3, ACACB, EHD2, UBL5), protein localization (SFT2D1, VTCN1, NPNT, SNX3, TMEM107, EHD2, UBL5), regulation of cell differentiation (NPNT, UST, CHN1, LEF1, SNX3, EHD2), cell projection organization (UST, CHN1, SNX3, TMEM107), cell-cell adhesion (PCDH1, VTCN1, NPNT, LEF1) and cellular ketone metabolic process (AKR1B10, ACACB, PSMA7) (Figure 4.15, Table 1).

Additionally, an interaction network was generated, using the Cytoscape plugin GeneMania, to display any genetic and physical interactions. Since SA2 KD significantly affected the localization of AJ proteins in the *Drosophila* notum (Figure 4.7), the members of the AJ KEGG pathway were also included in the network, to analyse whether they interact with the STAG2 KD differentially expressed genes. Only genetic and physical interactions that were backed up by experimental data were included in the interaction

map, this resulted in 95 interactions, displayed as edges (cyan: genetic, orange: physical), between 20 differentially expressed genes, displayed as nodes (blue, downregulated and red, upregulated) and 20 members of the AJ KEGG pathway (white) (Figure 4.16). Overall these results indicate that the differentially expressed genes in STAG2 KD interact with members of the adherens junction pathway; furthermore, the AJ pathway member LEF1 was differentially expressed by STAG2 KD, indicating that the role of STAG2 in cell invasion might be through gene expression alterations of junctional integrity regulators.

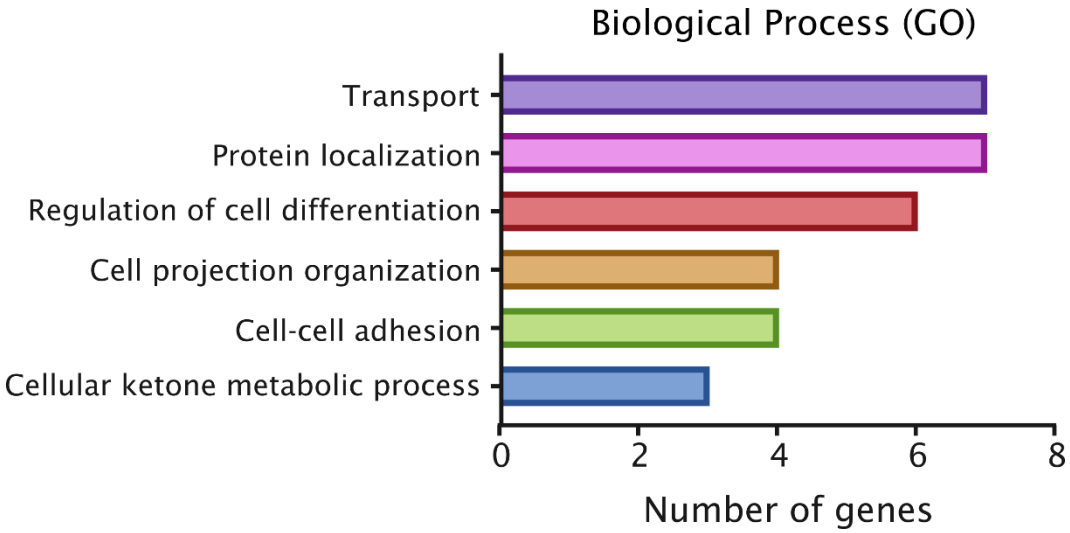


Figure 4.15 Gene ontology (GO) Biological Process enrichment

Genes that were differentially expressed in STAG2 KD MCF7 cells vs untreated or treated with non-targeting siRNA MCF7 cells (genes shown in Figure 4.12) were analysed for GO biological processes enrichment, using the online annotation database: DAVID Bioinformatics Resources 6.8 (Huang et al, 2009). Functional annotation clustering with default settings was used; medium stringency and Benjamini–Hochberg correction was applied. Only the enriched GO terms with FDR <0.05 were selected and displayed in the bar chart. The number of genes refers to the number of differentially expressed genes described by that annotation. Full table describing which genes are clustered in each GO term is found in Table1.

GO ID	GO Name	Genes	p-value (Bonferroni Corrected)
GO:0008104	protein localization	SFT2D1, VTCN1, NPNT, SNX3, TMEM107, EHD2, UBL5	0.0049
GO:0098609	cell-cell adhesion	PCDH1, VTCN1, NPNT, LEF1	0.0143
GO:0042180	cellular ketone metabolic process	AKR1B10, ACACB, PSMA7	0.0028
GO:0045595	regulation of cell differentiation	NPNT, UST, CHN1, LEF1, SNX3, EHD2	0.0026
GO:0030030	cell projection organization	UST, CHN1, SNX3, TMEM107	0.0181
GO:0006810	transport	SFT2D1, VTCN1, SLC37A2, SNX3, ACACB, EHD2, UBL5	0.0448

Table 4.1 Gene Ontology (GO) terms reaching Bonferroni corrected significance for enrichment amongst the 20 genes that had significantly altered their expression by STAG2 KD. Results were obtained using DAVID Bioinformatics Resources 6.8.

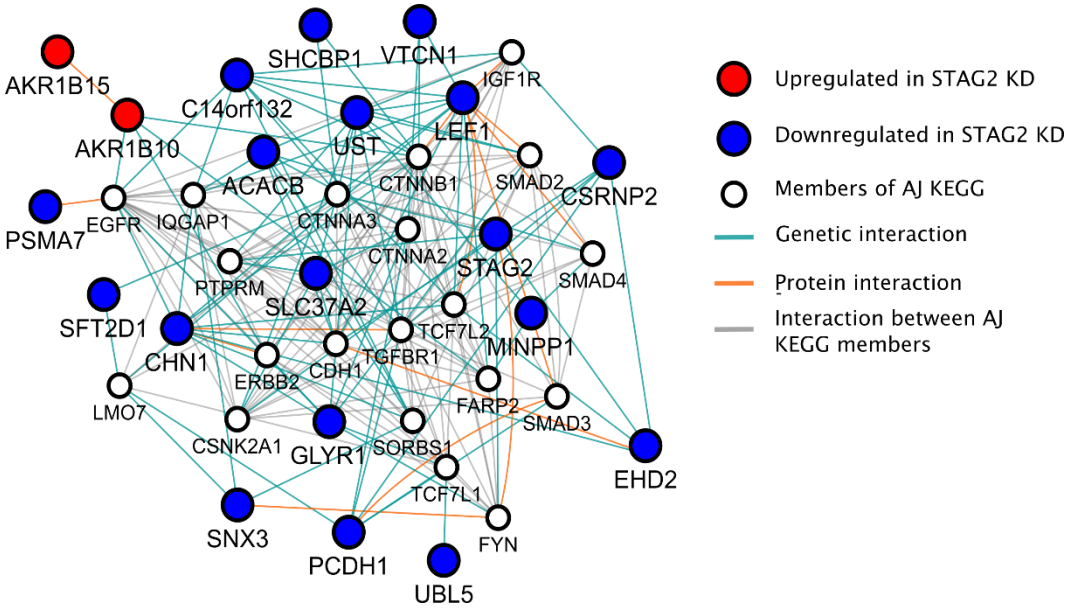


Figure 4.16 Interaction network of differentially expressed genes by STAG2 KD and adherens junction regulators

Physical and genetic interactions between genes misregulated by STAG2 KD and proteins of the adherens junction (AJ) KEGG pathway. The GeneMANIA plug-in for Cytoscape was used to generate an interaction network based on previously documented interactions. White nodes mark AJ components; red nodes: genes that were upregulated by STAG2 KD; blue nodes: genes that were upregulated by STAG2 KD. Orange lines: physical interactions; cyan lines: genetic interactions; grey lines: interactions between AJ genes.

4.3 Discussion

Stag2 is the most mutated subunit of the Cohesin complex in human cancers, truncating mutations are found in glioblastoma, AML, bladder and breast cancer (Kandoth *et al.*, 2013; Leiserson *et al.*, 2015). Several studies to date have analysed the mutational profile of *Stag2* in different tumours, establishing it as a tumour suppressor gene, yet the specific molecular mechanism that STAG2 plays in cancer is currently unknown. A *Drosophila* orthologue of STAG2, SA1, was found in the genetic screen (described in Chapter 3), since its KD induced phenotypes relevant to tumour progression such as multilayering, apex defects and invasion to mutant tumours in the back of the living fly. Following from the screen, further research in this study shows that loss of SA1, and its isoform SA2, induce a significant increase in the number of invading cells compared to WT tissue even without the *lgl* mutation. This indicates that SA1 and SA2 have a specific role as invasion inhibitors, independent of *lgl*. Indeed, KD of either STAG1 or STAG2 in the human breast carcinoma cell line, MCF7, significantly increased migrating and invading cells compared to untreated cells, suggesting that the role that STAG2 plays in tumour progression is by regulating single cell invasion, yet a morphological analysis of the cells is required to determine the effect of STAG2 in cell size and cell proliferation.

SA1 and SA2 KD invading cells observed in the notum of the fly were not apoptotic, unlike delaminating cells observed in the notum of WT developing pupae (Marinari *et al.*, 2012), indicating that the SA1 and SA2 KD cells detach from the epithelium by mechanisms independent of apoptotic cell extrusion. Loss of STAG2 has been reported to give resistance against apoptosis in tumour biopsies from metastatic melanoma patients, by regulating the ERK signalling pathway (Shen *et al.*, 2016). Additionally, Wang *et al.* show that bladder cancer cells, T24 and UMUC-3, treated with shRNAi for STAG2 are resistant to apoptosis compared with untransfected cells, determined by Annexin V staining and western blots of active caspases, yet the exact role STAG2 might play in the regulation of apoptosis in bladder cells is unknown (Wang *et al.*, 2017). Transcriptomic analysis of MCF7 cells with STAG2 KD did not reveal any differentially expressed genes involved in either proliferation or the ERK signalling pathway, and thus it is not clear if STAG2 has a role in apoptosis or proliferation in breast epithelial cells. Since this study is limited to one cancer cell line, and the baseline expression of genes within these breast cancer epithelial cells is already altered, the use of additional cell lines is recommended for future studies to determine if the role of STAG2 regulating apoptosis through the ERK signalling pathway, or through different mechanisms, is also conserved in different tissues. More specific assays to detect apoptosis could be performed first, such as the Annexin V assay, TUNEL staining or caspase detection assays, to determine

if STAG2 has a role regulating apoptosis in different breast cancer lines. Additionally, this could be studied in more depth *in vivo* in *Drosophila* since the proportion of iCasper SA2 KD positive cells was significantly smaller when compared to WT tissue.

SA1 KD cells migrated significantly faster and with directionality, compared to *lgl* mutant cells. GO biological analysis of differentially expressed genes by STAG2 KD in MCF7 cells showed an enrichment of cell projection organization, which directly affects cell motility (Yamaguchi *et al.*, 2007), including CHN1 a Rac interacting protein previously involved in dendritic cell projection (Miyake *et al.*, 2008) and TMEM107 that regulates cilium composition (Lambacher *et al.*, 2015). Additionally, STAG2 KD MCF7 cells had overexpression of both ROCK1 and ROCK2, that are critical to actin cytoskeleton regulation (Vigil *et al.*, 2013); yet while their fold increase was 1.87 – 2.3, the change was not significant. Since both ROCKs were differentially expressed, I suggest that further RT-qPCR confirmation could be performed to determine if STAG2 might directly or indirectly regulate their expression. This recommendation is also due to the variability observed between the gene expression of the three negative controls analysed. While this variation was separate from the distribution of the STAG2 KD samples, as seen in the principal component analysis, the variation could have had an

effect on the significance of the differentially expressed genes, increasing the amount of false negative results.

E-cadherin, beta-catenin and alpha-catenin localisation was misregulated in SA1 and SA2 KD clones in the fly. Gene ontology analysis of the differentially expressed genes after STAG2 KD in MCF7 cells show a high enrichment of genes involved in protein localization and transport. *Ehd2*, a gene that was significantly downregulated in STAG2 KD cells, is a plasma membrane protein that has been linked to E-Cadherin localization and expression. Additionally, lower EHD2 is associated with increased invasion and metastatic tumours (Shi *et al.*, 2015; Yang *et al.*, 2015). Since the expression of E-cadherin was not altered in STAG2 KD MCF7 cells, based on the results obtained from the adherens junction immunostainings in *Drosophila* and the transcriptomic analysis of MCF7 cells, a possible mechanism of action is that loss of STAG2 reduces the expression of genes important for E-cadherin localization to the plasma membrane, such as EHD2. Since the localization of both beta and alpha catenin are dependent to the proper localization of E-cadherin, the mislocalisation of both proteins is expected. Further experiments to confirm the role of EHD2 in the localization of E-Cadherin and the invasive phenotype induced by STAG2 KD could be performed as future work, such as an overexpression of EHD2 in STAG2 KD

cells to test whether the invasive phenotype is rescued. The localization of E-cadherin in MCF7 cells can also be studied to determine if there is mislocalisation of adherens junction proteins in human breast cancer cells in the absence of STAG2.

The localisation of FAS III, a septate junction protein, was also affected in SA1 and SA2 KD clones in the fly. However other members of the septate junction (coracle and dlg) and also the polarity protein aPKC were not affected by the KD of SA1 or SA2. Additionally, in the transcriptomic analysis, there was no change in the expression of genes that regulate the septate junction equivalent in humans, the tight junctions. While FASIII is a member of the septate junction, and it forms transmembrane adhesion between adjacent cells, it is regulated differently from the other members of the septate junctions, furthermore there is no clear ortholog of FASIII in humans (Snow *et al.*, 1989; Knust *et al.*, 2002). Therefore, it seems that *SA1* and SA2 specifically regulate *FASIII* in an unknown manner in the epithelial tissue of the thorax of the fly, yet overall there is no evidence to suggest that SA1 or SA2 are involved in septate junction regulation, nor *STAG2* in tight junction regulation.

The expression levels of the adherens junction member LEF1 was significantly reduced in STAG2 KD cells, yet LEF1 is well established as an oncogene in many different cancers, including breast cancer (Gebeshuber *et al.*, 2007; Santiago *et al.*, 2017). LEF1 binds directly to beta-catenin and together they regulate gene expression. Overexpression of LEF1 has two major effects in cancer cells: i) it sequesters beta-catenin to the nucleus, weakening the adherens junction, leaving the cell more vulnerable to have adhesion defects. ii) LEF1 positively regulates the expression of several genes involved in EMT, including N-Cadherin, Vimentin, and Snail (Grumolato *et al.*, 2013). An explanation for the severe invasive phenotype of STAG2 KD MCF7 cells, even when a reduced expression of LEF1 is observed, might be that E-cadheren mislocalisation still prevents beta-catenin localization to the plasma membrane, regardless of the levels of the beta-catenin nuclear partner. Additionally, there was no alteration in the levels of expression of beta-catenin/LEF1 target genes in MCF7 STAG2 KD cells.

Network analysis shows that several genes mis-regulated in STAG2 KD MCF7 cells interact with members of the adherens junction KEGG pathway. These genes have been shown to be involved in cancer invasion in different ways. ACACB, a gene involved in lipid metabolism, has been found to be downregulated in invasive ductal breast carcinomas (Castellana *et al.*, 2012),

and overall lipid metabolism is currently being studied for its potential role in metastasis, since alterations in their function affect the structure of the lipid-rich plasma membrane (Luo *et al.*, 2017). Additionally, the adhesion protein proto-cadherin 1 (PCDH1) was significantly downregulated in STAG2 KD; in addition to its role in the junction of different tissues, downregulation of PCDH1 has been shown to upregulate fibronectin type II containing 3 (FNDC3) expression (Tellez *et al.*, 2016); FNDC3B overexpression has been highly involved with tumour metastasis in hepatocellular carcinoma (Lin *et al.*, 2016). The transcriptomic analysis from this study shows that both FNDC3A and FNDC3B had a high fold increase in expression in MCF7 cells with STAG2 KD, compared to untreated cells, yet the change was not significant; in a similar situation to the ROCK proteins, validation of their expression in STAG2 KD cells through RT-qPCR could be performed to discard the possibility of false negatives due to variability of the samples.

Two interacting members of the aldo-keto reductase family: AKR1B10 and AKR1B15 were significantly overexpressed in STAG2 KD cells, these genes have been previously established as oncogenes, specifically involved in breast cancer cell invasion by the activation of the ERK signalling pathway (Ma *et al.*, 2012; Li *et al.*, 2017). These results make the study of STAG2 and

the aldo-keto reductase family an interesting line of continuing research, since the significant increase of AKR1B10 mRNA levels in STAG2 KD MCF7 cells was confirmed in this study by RT-qPCR. Additionally, as mentioned previously, Shen et al also suggested a role for STAG2 in the regulation of the ERK signalling pathway in metastatic melanoma patients (Shen *et al.*, 2016). Therefore, further investigation to determine if these mechanisms are connected, by studying ERK signalling in SA2/STAG2 KD cells, either in the *Drosophila* live model or in different breast cancer cell lines; could provide valuable evidence to better understand the role of STAG2 in tumour progression.

Overall these results show that SA1 and SA2, and STAG1 and STAG2, act as invasion suppressors in the epithelial cells of the *Drosophila* notum, and human adenocarcinoma cells, respectively. Loss of STAG2 induces significant changes in the expression of several genes that regulate adherens junction proteins and are relevant to tumour progression. Furthermore, results from the transcriptomic analysis indicate that the role that STAG2 has in tumour progression could be beyond the localization of adherens junction proteins, by affecting the expression of several genes involved in single cell invasion through different molecular mechanisms. The transcriptomic analysis showed the complexity of the changes induced by the loss of STAG2

in MCF7 cells and how these affected genes interact physically and genetically in a network; this must be considered if a specific pathway is to be studied in more detail in the future. It is also intriguing as to whether STAG1 might be involved in single cell invasion by regulating the same genes as STAG2, as both Cohesin complexes are found in different transcriptionally active sites, the effect that the loss of STAG1 has in MCF7 cells could be another exciting line to follow.

Chapter 5 - The role of the Cohesin complex in apical constriction and clonal extrusion

5.1 Introduction

SA1 and SA2 Cohesin complex subunits were identified in the genetic screen (Chapter 3) for their role in regulating single cell invasion, which was characterized in detail in Chapter 4. Interestingly, the other Cohesin subunits: SMC3, SMC1 and RAD21, did not have any effect on invasion; yet they did have significant effects in enhancing multilayering and apex defects, raising the question of whether the Cohesin complex might be involved in regulating epithelial cell polarity and/or cell shape. As mentioned in section 4.1, loss of SA1 triggers compensation by SA2 and vice versa (Benedetti *et al.*, 2017); while loss of a core subunit (SMC3, SMC1 and RAD21*)* leads to disruption of all Cohesin (Carvalhal *et al.*, 2018). This chapter will focus on the effects that loss of function of all somatic Cohesin complexes, namely simultaneous loss of SA1 Cohesin and SA2 Cohesin, has on tumour progression. All Cohesin loss of function induced drastic cytoskeletal changes including the formation of a supracellular actin ring, eventually followed by clonal extrusion. This extrusion of a cluster of mutant cells from the epithelium is likely to have relevance in the poorly understood process of collective cell migration in cancer.

5.1.1 Collective cell migration

Collective cell migration occurs when two or more cells migrate together, retaining their cell-cell junctions, unlike single cell migration where the individual migrating cell breaks away from the tissue of origin. Two-dimensional collective cell migration is observed in wound repair (Antunes *et al.*, 2013) and in development, i.e. dorsal closure in *Drosophila* (Kiehart *et al.*, 2000). In both processes two sheets of tissue are mobilized through apical constriction and the formation of an actomyosin ring at the leading edges of the migrating tissue, until they merge, closing the epithelial gap (Figure 5.1 a). Three-dimensional collective cell migration is also observed during development. For example, in zebrafish and *Drosophila*, primordial cells migrate collectively in the embryo, through actin polymerization and myosin contraction (Jaglarz *et al.*, 1995; Blaser *et al.*, 2006). In vertebrates, neural crest cells follow the same mechanism (Gurniak *et al.*, 2005; Berndt *et al.*, 2008).

3D collective cell migration has been found to be the most common mechanism for invasion in many different types of cancer including: oral squamous cell carcinoma, colorectal carcinoma, melanoma and breast cancer; where circulating clusters of cells, that retain their cell-cell contacts are commonly found (Nabeshima *et al.*, 1999; Hegerfeldt *et al.*, 2002; Gaggioli *et al.*, 2007). Additionally, spheroid injection of mutant xenografts in mouse *in*

vivo models have allowed the study of cell cluster movement, where similar to single cells, migration is directed by actomyosin flows through the leading and rear ends of the moving bodies. However, in collective cell migration, this flow has to be orchestrated multicellularly (Alexander *et al.*, 2008; Shindo *et al.*, 2014). Overall, the process of collective cell migration in cancer is poorly understood. It is not clear how the cells detach as a group from the epithelium. Some studies have reported that a level of de-differentiation might be required (Friedl *et al.*, 2009), or that a polyclonal group of cells is necessary to develop the front-rear polarity required for movement (Cheung *et al.*, 2013). Additionally, how the cluster of cells travel through the stroma is poorly understood.

5.1.2 Role of apical constriction in actin ring formation and collective cell migration

Additional 3D collective cell migration mechanisms during development involve apical constriction. In the process of tube formation, apical constriction is induced in a circle of cells, causing an invagination; eventually the leading cells establish junctions between each other, forming a three-dimensional circle or tube of cells, where the lumen is facing internally, in a process known as wrapping (Andrew *et al.*, 2010) (Figure 5.1 c). This is observed in neural tube (Haigo *et al.*, 2003) and retina formation (Plageman

et al., 2010). In contrast, during gastrulation (Behrndt *et al.*, 2012), apical constriction is induced in a stripe of cells across a tissue, causing tissue folding and invagination, in a process known as budding (Kim *et al.*, 2013) (Figure 5.1 b).

Apical constriction is induced by actin polymerization molecules and active myosin II localized at the apex of the cells. Formins induce actin nucleation and unbranched actin filament formation. Loss of the formin family member, Diaphanous (Dia), inhibits gastrulation in *Drosophila* (Lai *et al.*, 2008). Additionally, the Arp2/3 complex (actin-related protein 2/3) that induces branched actin filament formation, and one of its main activators, the small GTPase Rac, have been shown to be essential to actin ring formation in zebrafish epiboly during gastrulation (Hurst *et al.*, 2004; Leskow *et al.*, 2006). Overall, the Rho family of small GTPases (RAC, CDC42 and RHOA) are very important for actin polymerization and therefore ring formation. Specifically, RHOA has a critical role as additionally, its downstream effectors, Rho kinases (ROCK), phosphorylate Myosin II light chain to activate it (Gallo, 2006). Rho Guanine Exchange Factors (RhoGEFs) are also very important, as they recruit RHOA and therefore determine the position of the actomyosin ring (Padash Barmchi *et al.*, 2005).

In non-developing tissues, the mechanisms that induce budding are maintained (i.e. intestine folds) yet wrapping is inactivated. Results obtained from this study indicate that the loss of function of the cohesin complex induces aberrant induction of apical constriction in the epithelial tissue of *Drosophila*, followed by the basal extrusion of the actin ring containing cohesin mutant clones. The potential role that the cohesin complex plays in apical constriction and the implications that the extruded group of mutant cells could have in tumour progression in the context of collective cell migration is studied in this chapter.

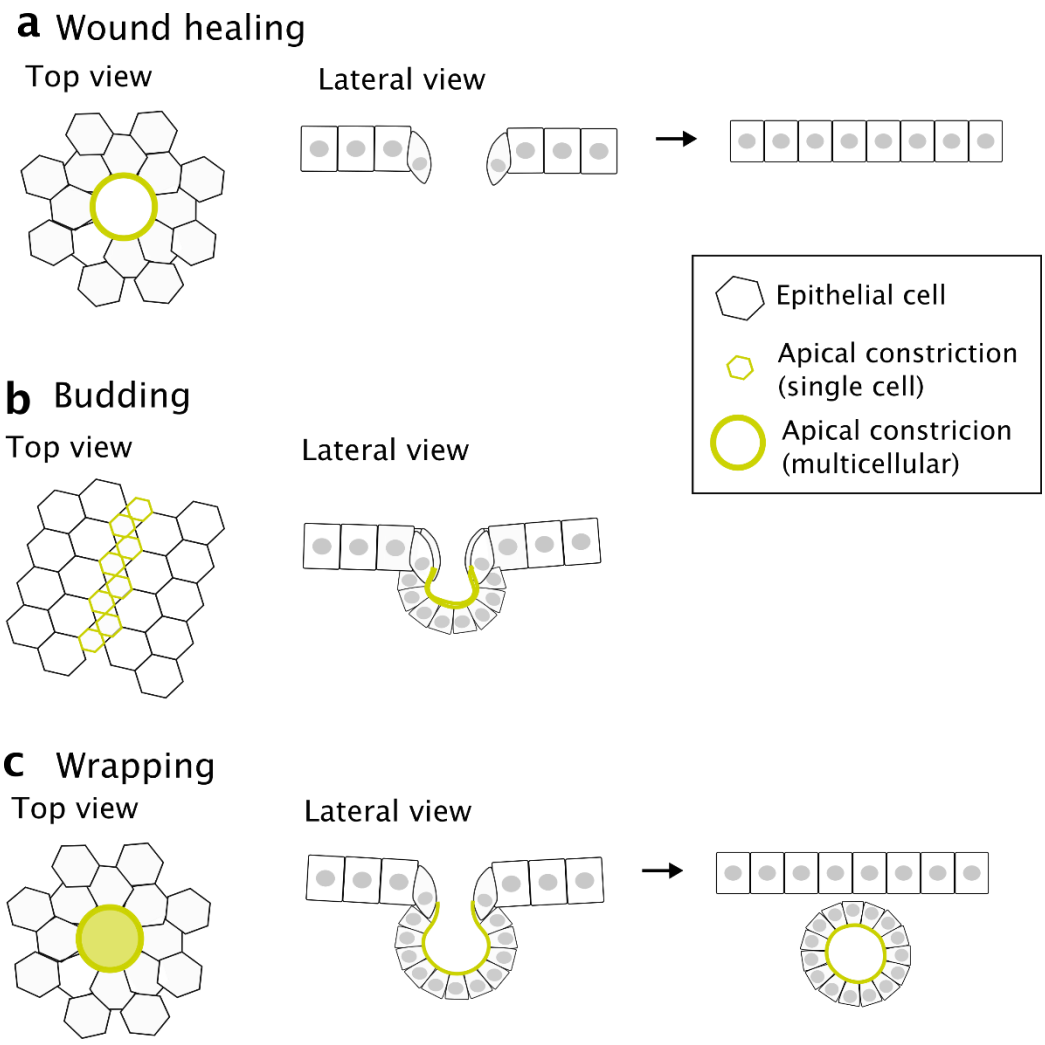


Figure 5.1 Effects of apical constriction in epithelial polarized cells

a) Wound healing requires apical constriction and actomyosin ring formation at the wound edge, mobilising the tissues to close the gap (Antunes *et al.*, 2013) **b)** Budding is induced when a stripe of cells becomes apically constricted, this is the mechanism that drives folding during gastrulation (Behrndt *et al.*, 2012) **c)** Wrapping is induced when a group of cells in a circle become apically constricted, an invagination is formed, and the group of cells is extruded from the epithelium. This is the mechanism for neural tube formation (Haigo *et al.*, 2003).

5.2 Results

5.2.1 Cohesin loss of function induces cytoskeletal changes and actin ring formation

Loss of function of both somatic cohesin complexes in genetically generated clones of epithelial cells on the fly pupal notum was achieved using three different approaches: i) the simultaneous KD of SA1 and SA2, to prevent the compensation of either cohesin ring when the other one is lacking. ii) a truncating loss of function mutation of *smc3* (Haelterman *et al.*, 2014) (see methods). Since it has been shown that a KD of SMC3 by RNAi, where at least 20% of the protein is functional, will still maintain the Cohesin sister chromatid cohesion (SCC) function (Carvalhal *et al.*, 2018). iii) NippedB (NipB) KD, as this will prevent Cohesin from interacting with the DNA.

WT clones, SA1 KD, SA2 KD and SMC3 KD had a very similar apical phenotype, namely an organized pentagonal shape (Figure 5.2 a-b and d-e). In contrast, the three all cohesin loss of function genotypes: $smc3^{-/-}$, SA1 + SA2 KD and NipB KD, consistently formed clones with a round clonal shape and frequently presented an apical actin rich ring (Figure 5.2 c,f,g). Quantification of the occurrence of these ring structures per mutant tissue confirm a significant increase in frequency: SA1+SA2 KD (0.23 rings/mm^2, n=8 animals, $p<0.0001$), $smc3^{-/-}$ (0.39 rings/mm^2, n=8, $p<0.0001$) and NipB KD (0.052 rings/mm^2, n=8, $p<0.01$), WT tissue (0 rings/mm^2, n=8) (Figure 5.2 h).

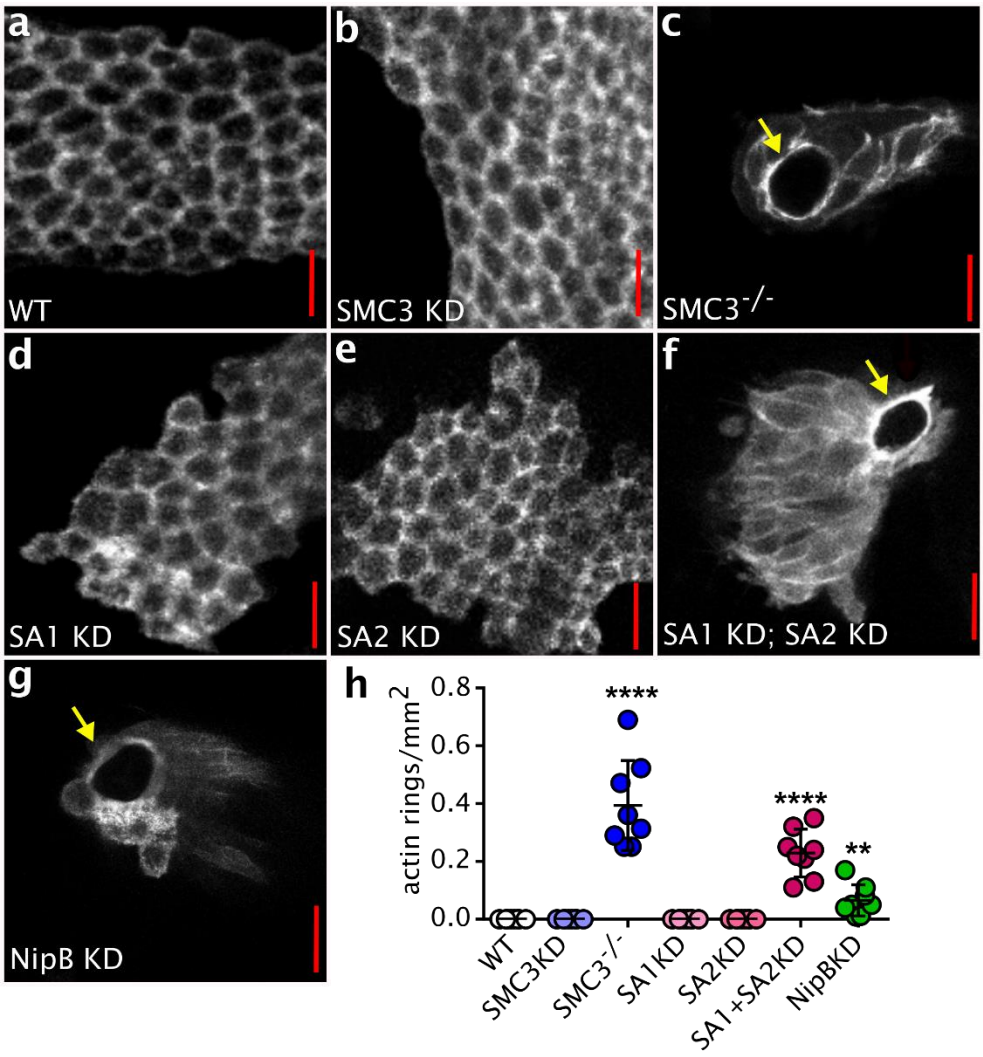


Figure 5.2 Cohesin loss of function induces actin ring formation

a-g) Confocal images of GFP:moe positive clones with different genotypes, as indicated on the bottom left of each panel. Actin rich rings were observed in *smc3*$^{-/-}$mutants, SA1 and SA2 simultaneous KD, and NipB KD clones. The actin ring phenotype is indicated with yellow arrows. Scale bars = 10μM. **h)** Quantification of the number of actin rings per mm^2 of clonal tissue. Eight animals were analysed for each genotype. Statistical analysis: Student's t-test. **** $p<0.0001$. ** $p<0.01$.

5.2.2 *smc3*$^{-/-}$ actin rings are supracellular and induce tissue invagination

I generated *smc3*$^{-/-}$ mutant clones to further characterise the actin ring phenotype. I specifically chose *smc3* clones for two reasons: 1) *smc3* mutant clones presented with actin rings most frequently (see Figure 5.2h), and 2) the generation of *smc3* mutant clones, compared to SA1 and SA2 double KD, was much simpler genetically.

To better understand the morphology of the actin rings induced by loss of cohesin function, positively marked *smc3*$^{-/-}$ clones were generated in the notum of the living fly. The entire coding region of His2A, which is tagged with mRFP1 (His-RFP), was ubiquitously expressed to identify whether there was any genetic material inside the actin ring structure and to differentiate unicellular (as observed in cell division and apoptotic cell extrusion) from multicellular rings (as observed in many developmental processes). No nuclear material was observed inside the actin ring in *smc3*$^{-/-}$ clones (Figure 5.3b). Orthogonal views of *smc3*$^{-/-}$ clones show that the actin ring induces tissue invagination, forming a multicellular bowl-like structure. Additionally, mutant cells that contribute to the supracellular ring have a basal nuclear localization (Figure 5.3b bottom panel), which is unlike the symmetric and organized morphology observed in WT clones that form a monolayer of cells whose nuclei are localized centrally (Figure 5.3 a).

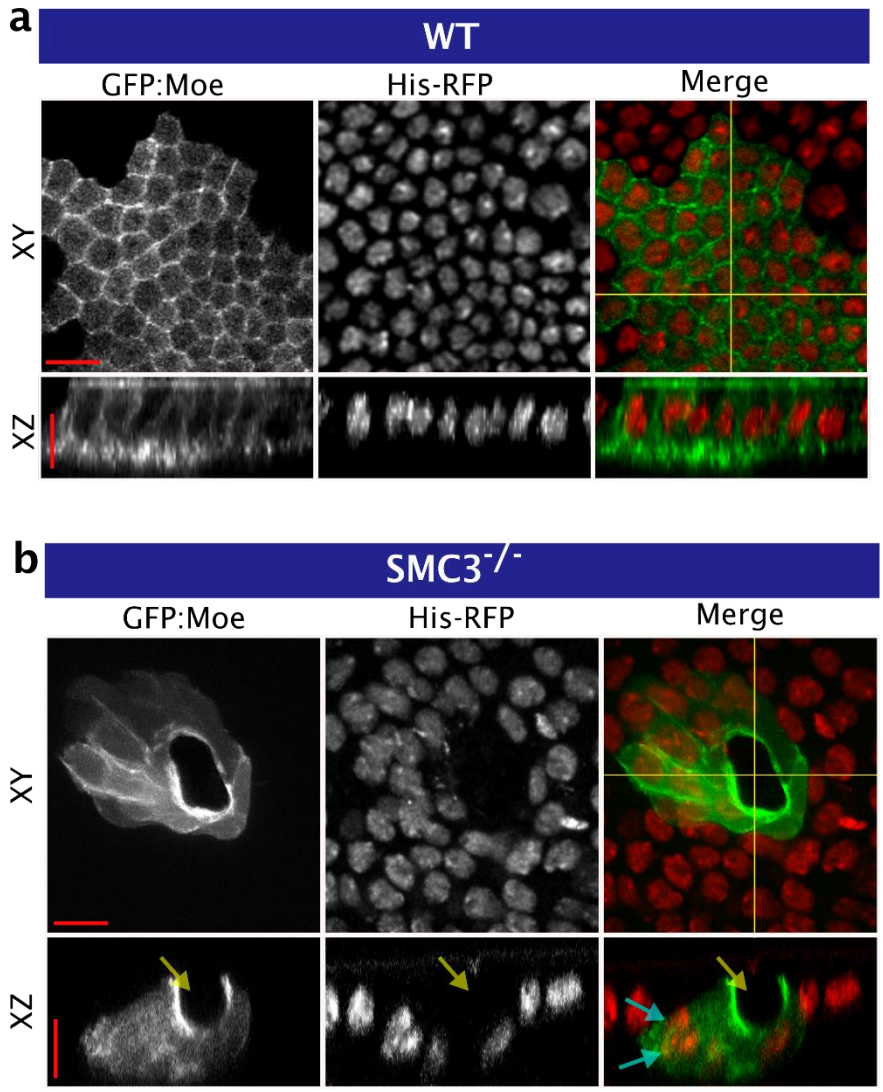


Figure 5.3 *smc3*$^{-/-}$ actin rings are multicellular and induce tissue invagination

a-b) Confocal images of GFP:moe positively marked clones (green) and ubiquitously expressed His-RFP (red) which labels the nuclei of all cells. **a)** WT tissue organization is uniform and the orthogonal (xz) view shows a monolayer of cells. **b)** A *smc3*$^{-/-}$ mutant clone containing an actin ring. No His-RFP signal is observed inside the ring in apical confocal sections of the mutant clone (xy). The orthogonal view (xz) shows that the actin ring induces invagination, as indicated by yellow arrows, and the nuclei of the *smc3*$^{-/-}$ cells are localized basally, as indicated by cyan arrows. Scale bars = 10μM.

5.2.3 Myosin II is enriched at the *smc3*$^{-/-}$ actin rings.

The supracellular actin ring formed in *smc3*$^{-/-}$ clones induce an invagination of the mutant tissue (Figure 5.3b) and promote a basal nuclear localization (Figure 5.3b, bottom); these phenotypes resemble the actin rings that form during development through apical constriction (Andrew *et al.*, 2010). An additional, crucial phenotype of apical contraction is the increased localization of myosin II at the apex, anchoring to the adherens junctions (Martin *et al.*, 2008).

To study Myosin localization, tissue containing GFP:moe positively labelled *smc3*$^{-/-}$ mutant clones and mCherry:spaghetti squash (sqh; the fly ortholog of the regulatory light chain of non-muscle myosin II) was dissected, fixed and subsequently stained for E-Cadherin, together with DAPI counterstain. Confocal imaging at the apex of the clones show that E-cadherin and myosin II are localized at the edge of the ring structure, indicating a rearrangement of apical morphology (Figure 5.4a). *smc3*$^{-/-}$ nuclei are localized in different planes from WT tissue and there is no DAPI signal from the inside of the ring, further confirming that the ring is supracellular (Figure 5.4a). E-cadherin staining from a more basal plane of the tissue shows that beneath the ring there are junctions, indicating the apex of cells that have invaginated

from the tissue, which become visible at a more basal confocal plane (Figure 5.4a bottom).

Additionally, mCherry:Sqh mean fluorescence intensity was measured at the junctions of WT clones, *SMC3*$^{-/-}$ clones with no actin rings and *smc3*$^{-/-}$ clones with actin rings. No significant change was observed between *smc3*$^{-/-}$ clones without actin rings (200.3 arbitrary units (AU), +/- 58.32, n=5 animals, p=0.4701) and WT tissue (177.3 AU, +/- 34.78, n=5); yet there is a significant increase in the localization of myosin II at the apex of the actin ring forming cells of *smc3*$^{-/-}$ clones (328.6 AU, +/- 81.38, n=5, $p<0.01$) (Figure 5.4). The levels of junctional E-Cadherin were also measured, and a significant increase in E-Cadherin mean fluorescent intensity was also observed in *smc3*$^{-/-}$ clones within the actin ring (708 AU, +/- 62.06, n=5 animals, $p<0.01$) and within non actin rings (687.5 AU, +/- 105.3, n=5, $p<0.05$) when compared to WT tissue (481.1 AU, +/- 93.79, n=5). Overall these results provide further evidence that *smc3*$^{-/-}$ actin rings are supracellular structures and suggest that non-muscle myosin II and E-cadherin might play essential roles in either the formation or maintenance of these rings.

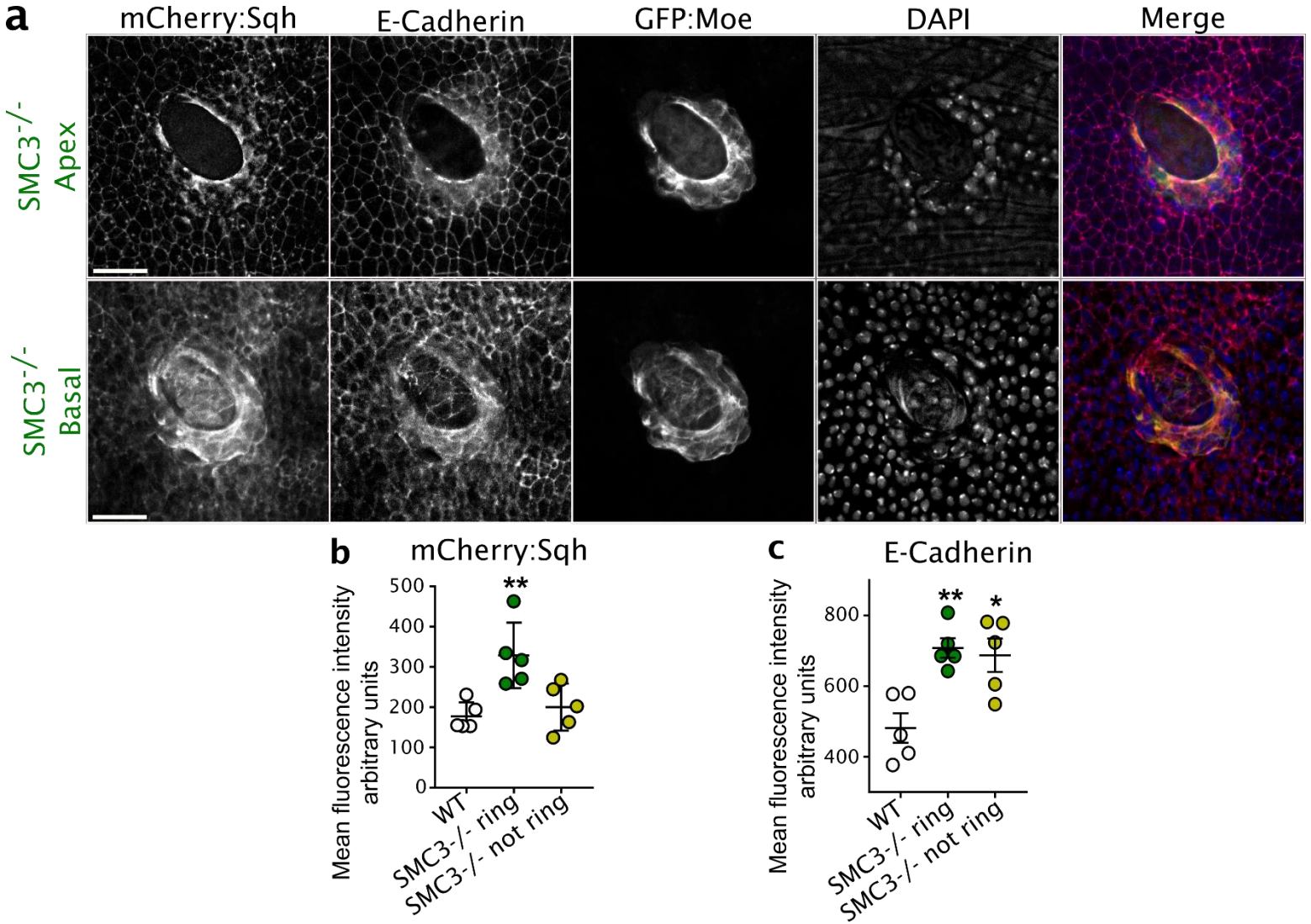

a
mCherry:Sqh
E-Cadherin
GFP:Moe
DAPI
Merge
SMC3-/- Apex
SMC3-/- Basal
b
mCherry:Sqh
Mean fluorescence intensity arbitrary units
500
400
300
200
100
0
**
WT
SMC3-/- ring
SMC3-/- not ring
c
E-Cadherin
Mean fluorescence intensity arbitrary units
800
600
400
**
*
WT
SMC3-/- ring
SMC3-/- not ring

Figure 5.4 Increased myosin and E-Cadherin localization in *smc3*$^{-/-}$ mutant clones

A dissected *Drosophila* notum highlighting a GFP:moe *smc3*$^{-/-}$ clone with mCherry:Sqh to label myosin II. The tissue is additionally stained for E-cadherin and DAPI. An apical view of the clone shows no protein localization or signal from the DAPI nuclear stain inside the actin-rich ring. A more basal confocal section shows E-cadherin and DAPI staining, marking the adherens junctions and nuclear material of the invaginated cells. **b-c)** Quantification of the mean fluorescence intensity from **b)** mCherry:Sqh and **c)** E-Cadherin at the cellular apex of WT, and *smc3*$^{-/-}$ clones within the actin ring cells or in mutant cells without actin rings. 50 measurements from 5 animals were quantified, each dot in the graph represents one animal. Statistical analysis: Student's t-test. ** $p<0.01$.

5.2.4 *smc3*$^{-/-}$ clones with rings extrude basally from the epithelium

During development, invagination is induced when a circular group of cells becomes apically constricted, and this is followed by clonal extrusion (Figure 5.1c). Since *smc3*$^{-/-}$ clones with actin rings induce an invagination, similar to that observed in apical constriction and there is an increased localization of myosin at the ring, contraction and clonal extrusion might be additional phenotypes. To test whether *smc3*$^{-/-}$ clones with actomyosin rings extrude, positively marked *smc3*$^{-/-}$ clones were generated, together with the mCherry:Sqh reporter for ubiquitous Sqh labelling, thereby providing a reference label for WT tissue. Gibson *et al* studied the effect of Dpp signalling in apical constriction and clonal extrusion in the wing disks of *Drosophila* (Gibson *et al.*, 2005). They reported the process of clonal extrusion to be slow, with a duration of approximately 10 hours. Therefore, in this study, animals were imaged live, hourly for approximately 17 hours, to determine if the actin ring containing clones extrude from the epithelium.

From a lateral view, as seen previously (Figure 5.3 b), the rings from young animals (12-20h APF) induce an invagination (Figure 5.5a). The apical actin-rich ring contracts over time (21-23h APF) until the apical opening is closed and the invagination generates a ring that can be observed laterally, generating a ball of cells with a central lumen (Figure 5.5a'). The internal

multicellular ring continues to contract, and the entire clone is extruded basally, and the neighbouring WT cells form new junctions with each other, thereby restoring the epithelial monolayer (Figure 5.5a''-a''''). An additional 3D reconstruction is shown in Figure 5.6b, with the aim of showing a clearer view of the structure of the clones. At 14h APF, the *smc3*$^{-/-}$ multilayered clone (GFP:moe), is still part of the epithelium (mCherry:sqh); yet at 27h APF, the clone is completely detached from the WT tissue.

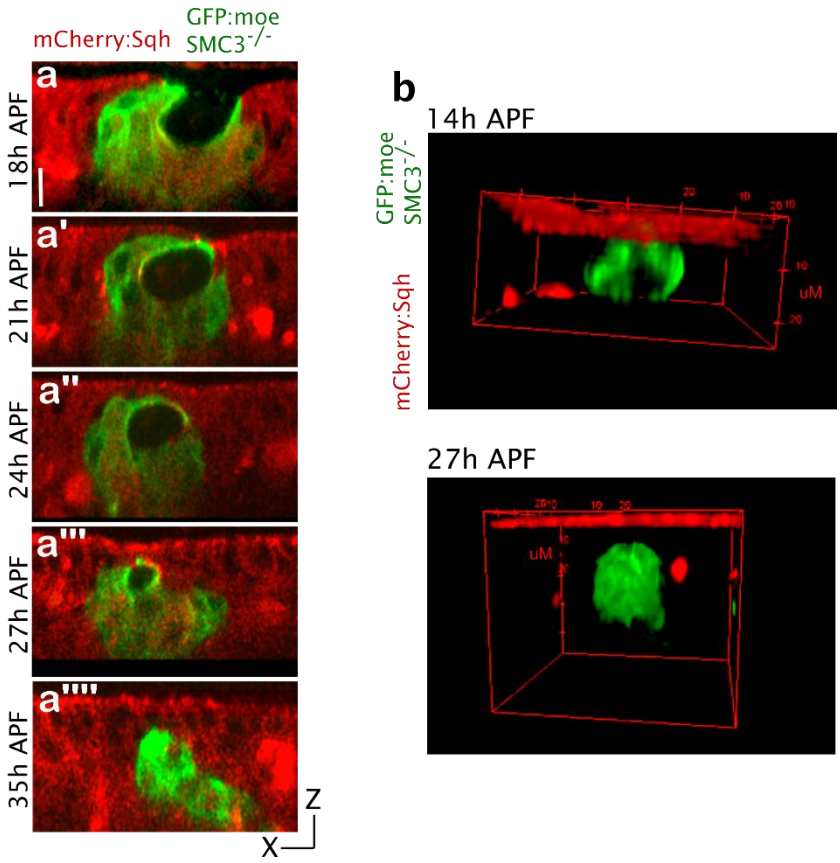


Figure 5.5 *smc3*$^{-/-}$ actomyosin rings contract over time leading to clonal extrusion

a-a''') Lateral view of GFP:moe positively labelled *smc3*$^{-/-}$ clones and mCherry:Sqh to label the epithelial tissue. The same animals were imaged live for approximately 15 hours, to study the morphology and dynamics of the actin ring clones. **a)** At ~18h APF, the actin ring induces an invagination of the tissue, as previously observed (Figure 5.3). **a')** At 21h APF, the actin ring closes, forming a central lumen. **a''-a'''')** The mutant clone is extruded from the epithelium, and the actin ring contracts over time, closing the internal lumen (35h APF). **b)** 3D reconstruction of confocal Z stacks, top image shows an *smc3*$^{-/-}$ clone containing an actin ring at 14h APF, labelled with GFP:moe, the epithelial tissue is labelled with mCherry:Sqh. The mutant clone is multilayered but is still part of the epithelium. The bottom image shows the same animal at 27h APF, after the mutant clone has been extruded from the epithelium.

To test whether, in this system, the extrusion of the *smc3*$^{-/-}$ clones acts as a way to remove mutant cells from the epithelium, or if this could be a mechanism applicable to collective cell migration, a live study of apoptosis was done next. The caspase sensor, iCasper, was used to determine the levels of active caspases in *smc3*$^{-/-}$ clones. Levels of apoptosis were compared between WT clones, *smc3*$^{-/-}$ clones that are still connected to the epithelial sheet, and extruded *smc3*$^{-/-}$ clones. An increase in the number of iCasper positive cells in *smc3*$^{-/-}$ clones compared to WT tissue was observed, yet this was not significant (Figure 5.6a). Additionally, there is no significant difference between *smc3*$^{-/-}$ clones that are part of the epithelium and extruded clones (Figure 5.6a). Furthermore, time lapse imaging was performed in the delaminated clones, and no increase in iCasper signal was observed in the 1-hour slots imaged (Figure 5.6b). This indicates that the basal extrusion of the *smc3*$^{-/-}$ clones does not trigger cell death.

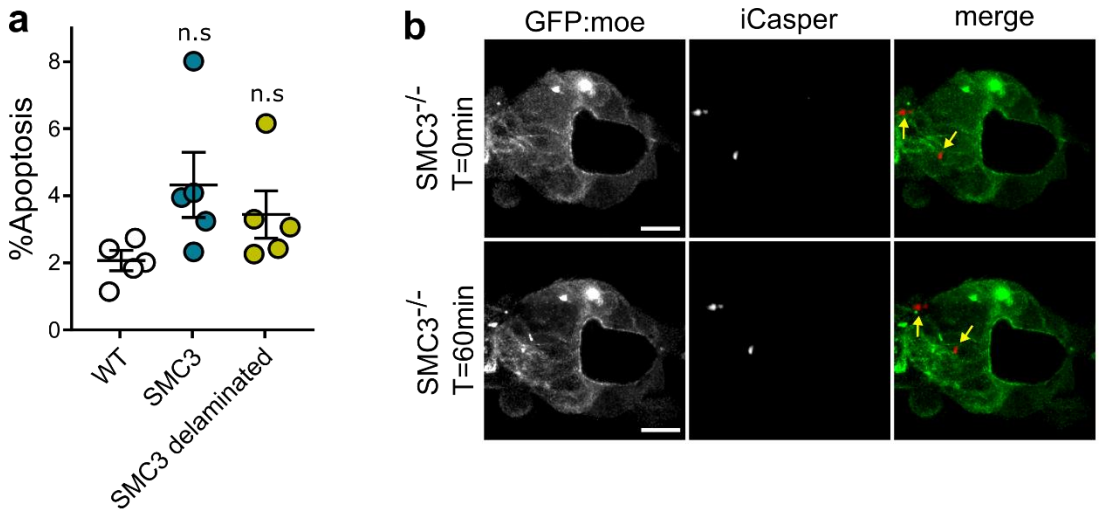


Figure 5.6 Clonal delamination does not trigger apoptosis

a) Sample images of a time lapse from a delaminated *smc3*$^{-/-}$ mutant clone imaged every 10minutes for 1 hour. Two iCasper positive cells, highlighted with yellow arrows, were observed at time 0. The same two cells changed their position within the clone at time 60. Yet, no new iCasper positive cells were detected. Scale bars= 10μM **b)** Quantification of iCasper positive cells within WT clones, *smc3*$^{-/-}$ clones that are still connected to the epithelial sheet, and extruded *smc3*$^{-/-}$ clones. n=5 animals, all the labelled cells within each animal were quantified, each dot represents the average of one animal. Statistical analysis: Student's t-test, n.s. = not significant

5.2.5 Rho1 and myosin II have significant effects on *smc3*$^{-/-}$ actomyosin ring formation.

As mentioned in section 5.1.2, the crucial mechanisms that trigger apical constriction and clonal extrusion are the apical localization of activated RHO1 that will in turn recruit and activate myosin II (Padash Barmchi *et al.*, 2005). Therefore, I next studied whether an increase or decrease in RHO1 or myosin II activity would have any effect on ring formation and clonal extrusion within *smc3*$^{-/-}$ clones.

Positively marked clones expressing a constitutively active allele of RHO (RhoV14) or dominant-negative allele (RhoN) were generated in the notum of the fly. Neither RhoN nor RhoV14 expression induced actin ring formation (Figure 5.7). Cell shape in RhoV14 clones was highly heterogenic, and although no actin rings were observed, a multicellular rosette pattern was induced instead (Figure 5.7a). When RhoN or RhoV14 were expressed within the *smc3*$^{-/-}$ mutant tissue, the frequency of ring formation was significantly affected. RhoN decreased the amount of actin rings or delaminated clones (0.126 actin rings + delaminated clones/mm^2, +/- 0.06, n=8 animals, $p<0.001$) compared to *smc3*$^{-/-}$ mutant clones alone (0.393, +/- 0.16, n=8) (Figure 5.7b-b'). The expression of RhoV14 in combination with the *smc3*$^{-/-}$ mutation resulted in very small unicellular clones, or no clones at all

(Figure 5.7b-b'), and therefore it was not possible to fully quantify and compare this genotype.

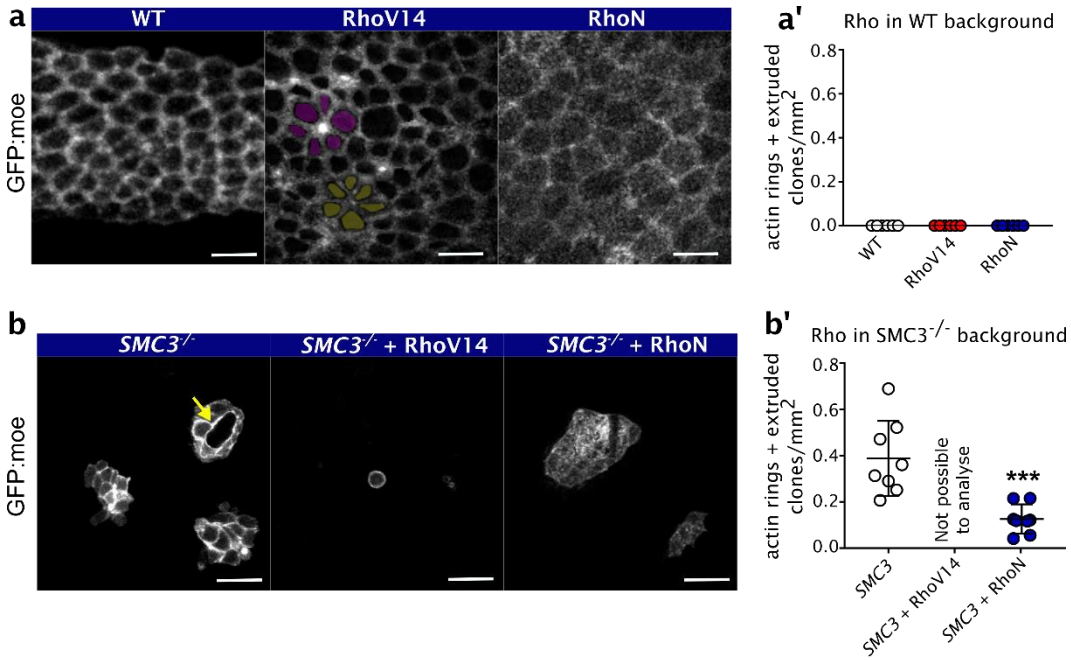


Figure 5.7 RHO1 activity affects ring formation and clonal extrusion

(a,b) Live imaging of the apex of GFP:moe positively marked clones in the epithelial tissue of the dorsal thorax of *Drosophila* pupae. Dominant negative RHO1 (RhoN) and constitutively active RHO1 (RhoV14) were expressed either alone (a-a') or in combination with the *smc3*$^{-/-}$ mutation (b-b'). Scale bars = 10μM. **(a',b')** Quantification of the number of actin rings or delaminated clones per mm^2. Each dot represents an animal. Statistical analysis: Student's t-test. ***p<0.001.

I next asked whether phosphomimetic SQH (Sqh-EE) or Sqh KD (Sqh-RNAi) could affect actin ring formation, either alone or in combination with the *smc3*$^{-/-}$ mutation, in positively marked clones in the *Drosophila* notum. Expression of Sqh-EE alone was sufficient to induce actin ring formation (0.19, n=8, $p<0.0001$) while SQH KD had no effect in actin ring formation (0.0 actin rings + delaminated clones/mm^2, n=8) when compared to WT tissue (0.0, n=8) (Figure 5.8a-a'). The expression of Sqh-EE within *smc3*$^{-/-}$ clones significantly enhanced the frequency of actin rings and delaminated clones (1.062, n=8, $p<0.05$) (Figure 5.8b-b'). In contrast, SQH KD significantly reduced actin ring formation (0.153 actin rings + delaminated clones/mm^2, n=8, $p<0.01$) when compared to *smc3*$^{-/-}$ alone (0.393, n=8). Overall these results indicate that RHO1 and non-muscle myosin 2 are essential for the determination of *smc3*$^{-/-}$ cell morphology and actin ring formation.

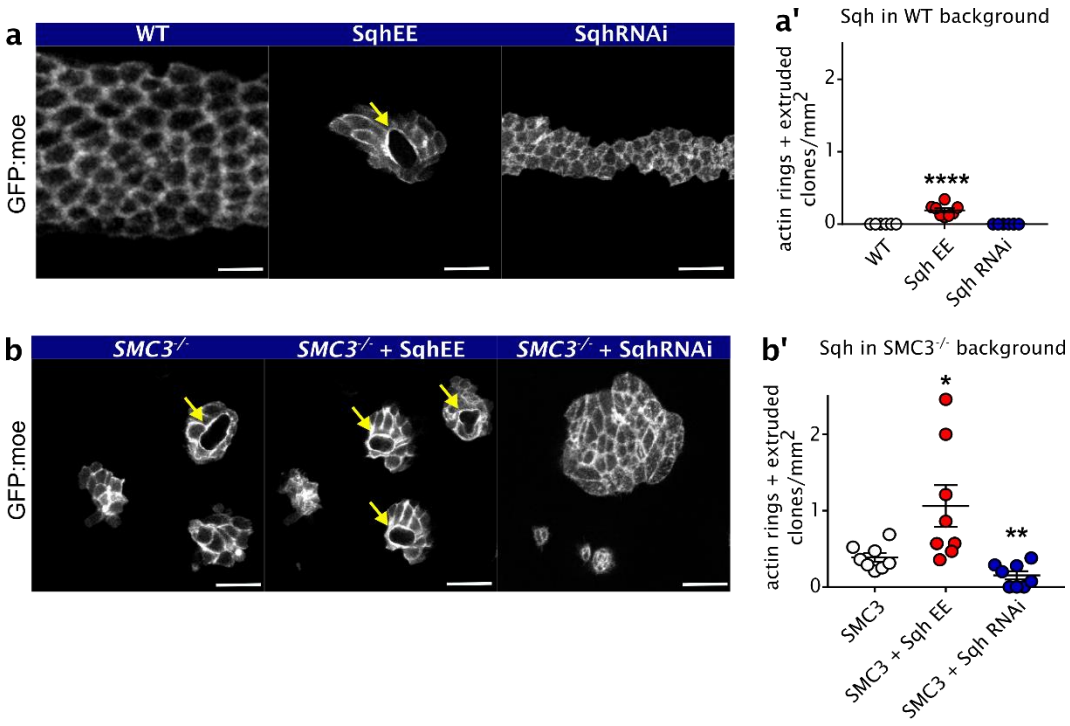


Figure 5.8 Myosin2 expression affects ring formation and clonal extrusion

(a,b) Live imaging of the apex of GFP:moe positively marked clones in the epithelial tissue of the dorsal thorax of *Drosophila* pupae. Phosphomimetic spaghetti squash (Sqh EE) and Sqh-RNAi were expressed either alone (a-a') or in combination with the *smc3*$^{-/-}$ mutation (b-b'). Scale bars = 10µM. **(a',b')** Quantification of the number of actin rings or delaminated clones per clonal area. Each dot represents an animal. Statistical analysis: Student's t-test. *p<0.05, **p<0.01, ****p<0.0001.

5.2.6 Screen of apical constriction regulators

To better understand the potential mechanism of action of SMC3 in apical constriction and actin ring formation, an enhancer/suppressor screen of genes involved in regulating the localization of myosin II and RHO1 to the apex of the cell, thereby inducing apical constriction was performed. A selection of 6 genes were either KDed and where possible overexpressed; due to a limitation of transgenic constructs not every gene studied was overexpressed. The KD or overexpression of each gene was done in positively marked clones on the notum of the fly, both alone and in combination with the *smc3*$^{-/-}$ mutation, to determine if these genes enhance or rescue the actin ring and clonal extrusion phenotype. A summary of each of the genes studied is included in Table 5.1.

Mad is the main effector of the *Drosophila* Dpp signalling pathway (the BMP pathway in humans). In *Drosophila*, activation of MAD (hSMAD) is governed by its phosphorylation from the receptor kinases TKV (hBMPR1/2) and PUNT (hACVR1/2) (Wiersdorff *et al.*, 1996). An increase in Dpp signalling, and therefore of phosphorylated MAD (pMAD), has been directly implicated in apical constriction and actin ring formation (Jidigam *et al.*, 2015). The overexpression of either MAD (0.1448, n=8, $p<0.01$) or TKV (0.173, n=8, $p<0.01$) significantly induced actin ring formation when compared to WT

animals (0.0, n=8), while the clones expressing RNAi of MAD (0.0, n=8) or TKV(0.0, n=8) did not have any ring formation (Figure 5.9). Mad overexpression within *smc3*$^{-/-}$ clones significantly increased the number of actin rings and delaminated clones (1.196, n=8, $p<0.05$) when compared to *smc3*$^{-/-}$ alone (0.393, n=8). Furthermore, MAD KD *smc3*$^{-/-}$ tissue, also significantly decreased actin ring formation and clonal extrusion (0.196, n=8, $p<0.01$). In contrast, there was no significant effect from either TKV KD (0.587, n=8, p=0.13) or overexpression (0.467, n=8, p=0.42) in *smc3*$^{-/-}$ clones when compared to the *smc3*$^{-/-}$ mutant alone. Another important pathway for apical constriction specification in different tissues is Hehdgehog (Hh), yet a KD of Hh did not have any effect on the formation of actin rings or clonal delamination, either alone (0.0, n=8) or in combination with *smc3*$^{-/-}$ (0.574 n=8, p=0.23).

Table 5.1 Apical constriction genes studied in 5.2.6

Gene target	Construct	Description
Fog	KD	Fog signalling stimulates localisation of myosin to the apical surface, this then provides the contractile force for apical constriction (Sweeton *et al.*, 1991)
Tkv	KD + O/E	Thick veins is a type I receptor for decapentaplegic (Dpp) (homolog of BMP). Dpp is thought to be implicated in dorsal closure of Drosophila.(Wiersdorff *et al.*, 1996)
Cut	KD + O/E	A high level of Cut gene expression is implicated in apical constriction of anterior follicular cells during oogenesis within *Drosophila.* (Levine *et al.*, 2010)
Hh	KD	During *Drosophila* eye development, Hh promotes apical constriction along with the DPP pathway across the epithelium. (Curtiss *et al.*, 2000)
Shrm	KD + O/E	Shrm is able to independently promote apical constriction in WT epithelial cells that are transcriptionally quiescent, but not in non-polarised cells. (Nishimura *et al.*, 2008)
Mad	KD + O/E	In the developing *Drosophila* retina, Mothers against DPP (Mad), the DPP signalling effector, is required to maintain the apical microtubules and is sufficient to drive apical constriction. (Jidigam *et al.*, 2015)

Pathways that are tissue specific were also studied. CUT specifically induces actin ring formation in sensory organ precursors (SOP) cells (Levine *et al.*, 2010). Overexpression of cut alone induced ring formation (0.28, n=8, $p<0.01$), yet it did not have an effect in combination with *smc3*$^{-/-}$ (0.361, n=8, p=0.7). Additionally, the KD of CUT failed to rescue the *smc3*$^{-/-}$ phenotype (0.22, n=8, p=0.078). FOG is the main inducer of apical constriction during gastrulation (Sweeton *et al.*, 1991), yet there was no significant change in the actin ring phenotype when knocked down within *smc3*$^{-/-}$ clones (0.189, n=8, p=0.058). Finally, Shrm is another well-established pathway for apical constriction in *Drosophila* (Nishimura *et al.*, 2008), and while overexpression of SHRM slightly induced actin ring formation in a WT background (0.066, n=8, $p<0.0001$), neither the KD (0.163, n=8, p=0.08) nor overexpression of SHRM within *smc3*$^{-/-}$ clones (0.25, n=8, p=0.42) had any significant effect in the formation of actin rings (Figure 5.9).

Overall, the results from the screen of apical constriction enhancers and suppressors suggests that the role of cohesin in apical constriction could be due to the regulation of MAD or other members of the Dpp pathway that regulate MADactivity.

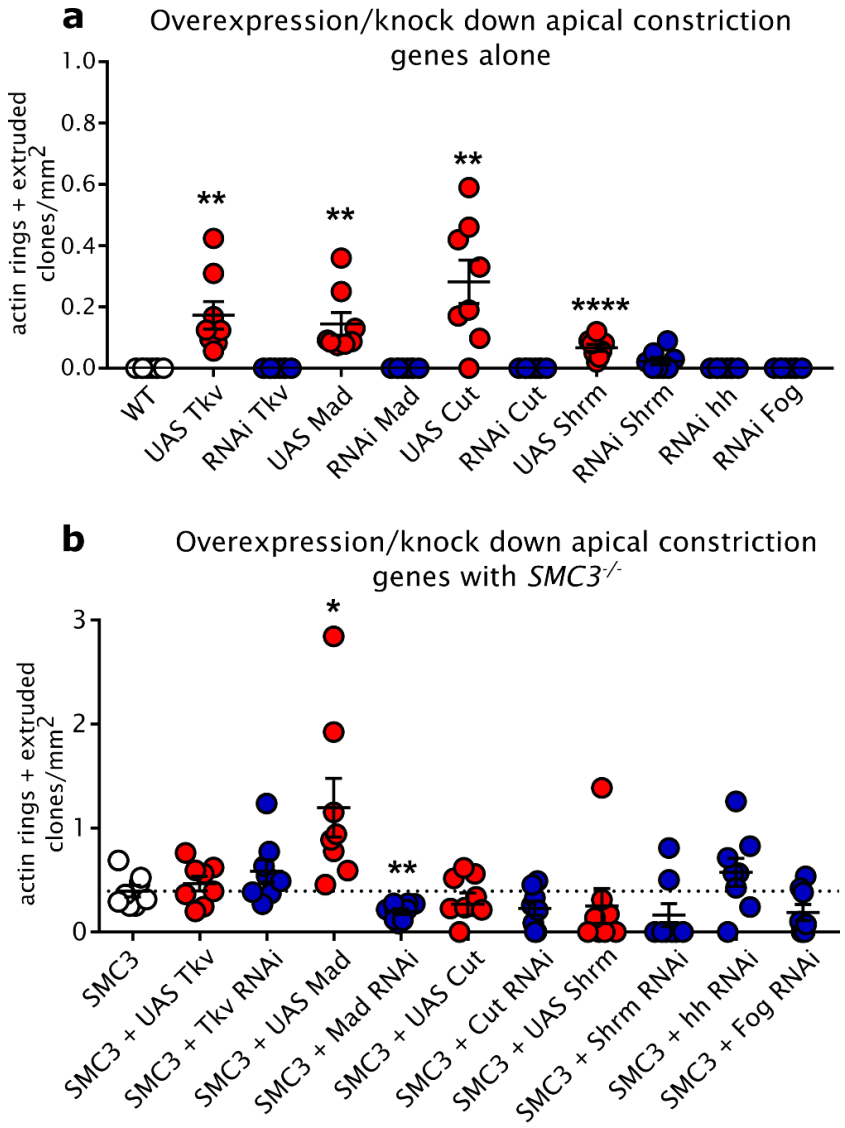


Figure 5.9 An enhancer/suppressor screen of apical constrictor regulators

Genes involved in apical constriction were either knocked down or overexpressed in the epithelial tissue of the dorsal thorax of the fly, either on their own **(a)** or within *smc3*$^{-/-}$ clones **(b)**. The number of multicellular actin rings and delaminated clones were quantified and represented as occurrence within every mm^2 of mutant tissue. Each dot represents 1 animal. Red dots are transgenes that were hypothesised to induce actin ring formation, based on the literature. Blue dots are transgenes that were hypothesised inhibit actin ring formation. Statistical analysis: Student's t-test. *p<0.05. ** p<0.01. ****p<0.0001.

5.2.8 Loss of Cohesin increases phosphorylated MAD levels

The Dpp signalling effector, MAD, had significant effects on actin ring formation in *smc3*$^{-/-}$ mutant clones, suggesting that the role of cohesin in apical constriction could be through the regulation of MAD. To better understand whether cohesin influences MAD activity, *smc3*$^{-/-}$ mutant and SA1 + SA2 KD positively marked clones were generated in the epithelial tissue of the back of the fly. The tissue was then dissected and stained with a phosphorylated MAD (pMAD) antibody (ab52903, see methods), which is its active form.

A significant increase in pMAD levels was detected in *smc3*$^{-/-}$ clones (391.7 AU, n=7 animals, $p<0.0001$) and SA1 + SA2 KD clones (396.9, n=7, $p<0.01$) specifically when these clones contained actin rings (Figure 5.10). There was no significant difference in pMAD mean intensity between WT cells (~218, n=7) and *smc3*$^{-/-}$ clones (281.1, n=7, $p=0.668$) or SA1 + SA2 KD clones without an actin ring (269.8, n=8, $p<0.0633$) (Figure 5.10), suggesting that MAD activity is necessary to induce apical constriction in cohesin LOF clones.

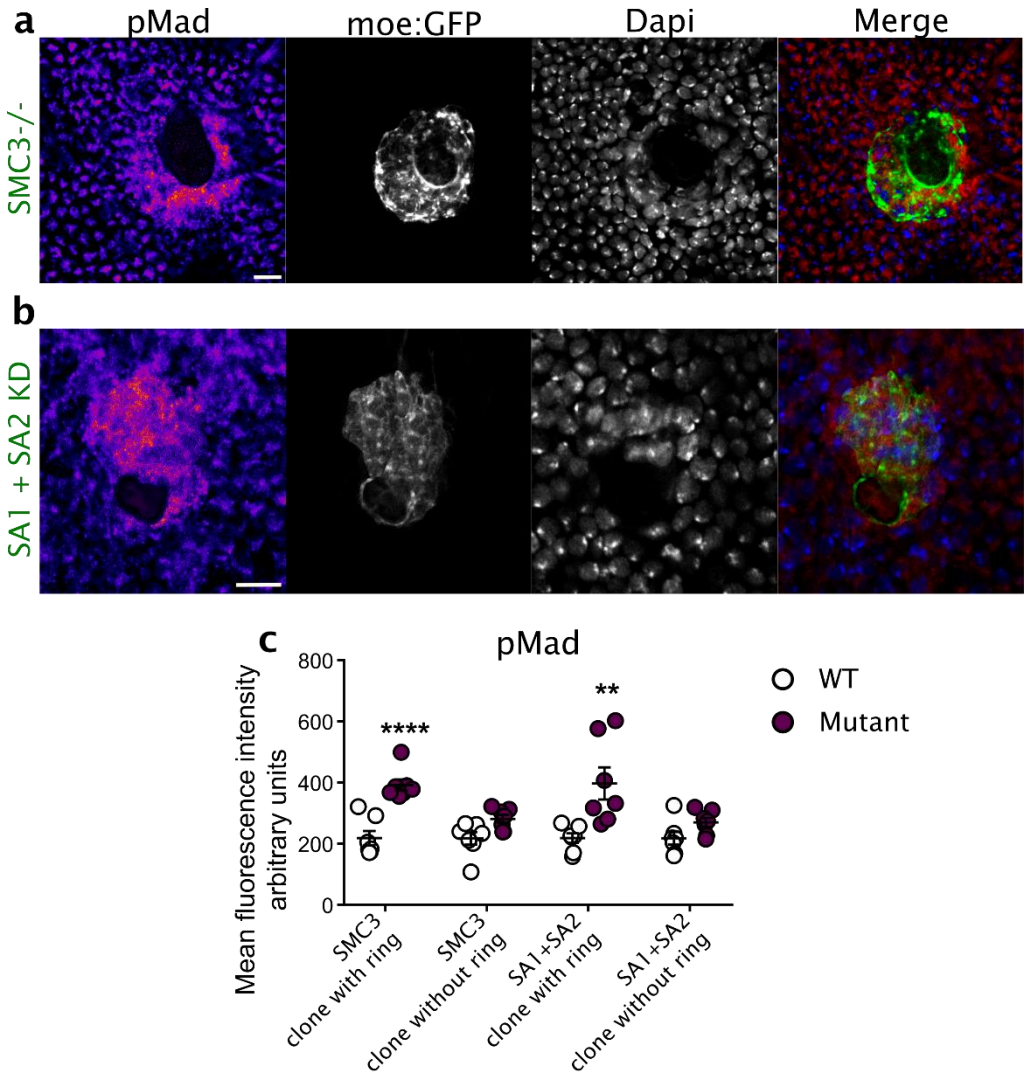


Figure 5.10 Cohesin loss of function increases pMAD nuclear localization

(a-b) Example images from dissected *Drosophila* nota containing GFP:moe labelled *smc3*$^{-/-}$ (**a)** and SA1 + SA2 KD (**b)** clones stained for the active form of the Dpp signalling effector, phosphorylated MAD (pMAD). Scale bars = 10μM. (c) Quantification of mean fluorescence intensity from the nuclei of *smc3*$^{-/-}$ and SA1 + SA2 KD clones with and without actin rings, compared to WT tissue within the same animal. 35 nuclei from 7 animals were measured. Each dot represents one animal. Statistical analysis: Student's t-test. ****P< 0.0001. **p<0.01.

5.2.7 *smc3*$^{-/-}$ mutant clones do not affect the number of CUT positive cells

It has been reported that LOF of cohesin regulates CUT expression, yet this regulation is very sensitive and depending on the cohesin mutation, this might lead to an overexpression or downregulation of CUT in the wing disc of *Drosophila* (Rollins *et al.*, 2004; Dorsett *et al.*, 2005). In turn, an overexpression of CUT induces actin ring formation. No significant changes in actin ring formation were observed by either the upregulation or KD of CUT within *smc3*$^{-/-}$ tissue. Since cohesin has been reported to regulate CUTexpression, further validation was performed to verify if CUT was affected in *smc3*$^{-/-}$ mutant clones.

smc3$^{-/-}$ mutant positively marked clones were generated in the epithelial tissue of the back of the fly. The tissue was then dissected and stained using a CUT antibody (2B10 DSHB, see methods). Two different pupal ages were analysed. Since CUT is only expressed in the SOP cells of the notum of *Drosophila* (Blochlinger *et al.*, 1993), a time point before SOP specification (12h APF) and after SOP specification (25h APF) was used. In accordance with the literature, there was no CUT staining in the notum of young pupae, either from WT tissue or *smc3*$^{-/-}$ tissue, while actin ring formation still occurred (Figure 5.11a). In older animals, the nuclei of cells that follow the pattern of SOP cells were stained with CUT antibody (Figure 5.1b), while the rest of the

tissue did not show any staining. Additionally, the *smc3*$^{-/-}$ mutant tissue had CUT positive cells, following the pattern of SOP wild tissue, without any increase in the number of CUT positive cells. Although there is a reduction in CUTpositive cells in *smc3*$^{-/-}$ mutant clones, quantification shows no significant difference between WT (1.8, n=5) and *smc3*$^{-/-}$ CUT positive cells (1.3, n=5, p=0.065) (Figure 5.11c).

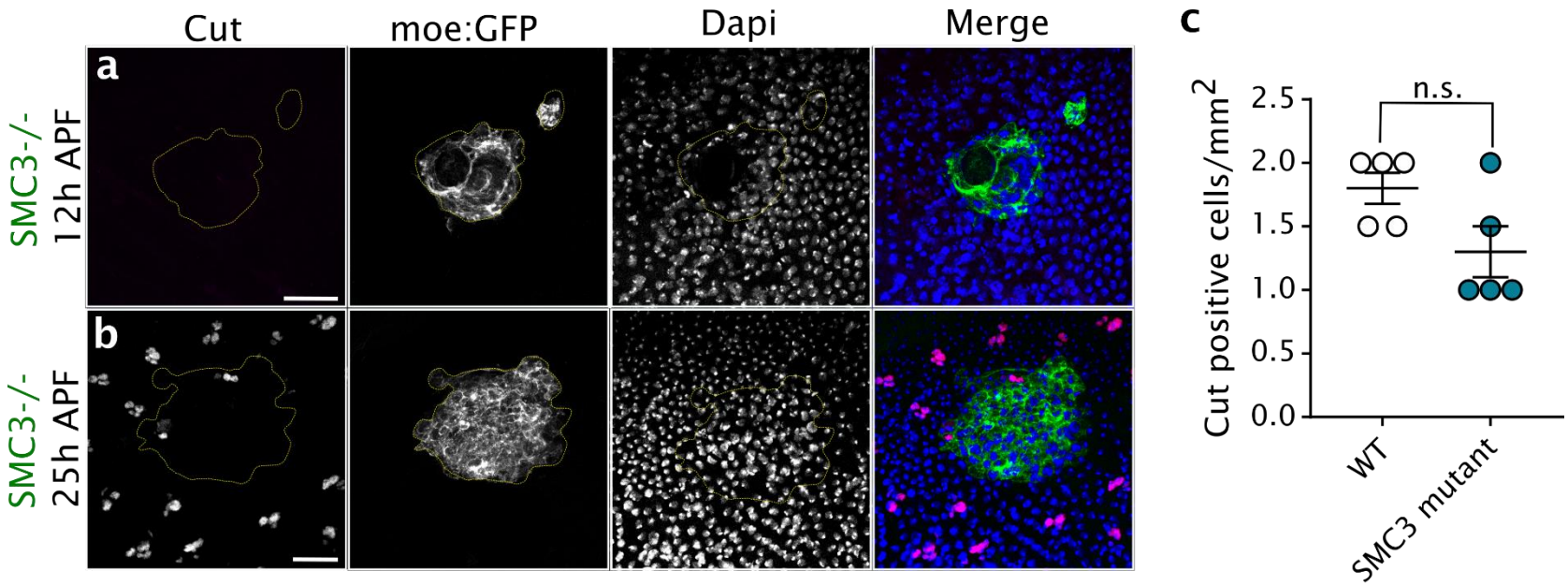


Figure 5.11 Loss of *smc3*$^{-/-}$ has no effect on CUT expression

a-b) Example images of dissected *Drosophila* nota containing GFP:moe *smc3*$^{-/-}$ clones stained with a CUT antibody. **a)** Young pupa (12h APF) where CUT expression is still not activated. *b)* Pupa of 25h APF where SOP specification has occurred, and CUT is expressed in the sensory organs. Scale bars=10μM. **c)** Quantification of the number of CUT positive cells per mm^2 area of either WT or mutant tissue. Each dot represents an animal. Statistical analysis: Student's t-test, n.s. = not significant.

5.3 Discussion

Apical constriction and clonal extrusion are crucial events during development, these processes induce the formation of compartments in the embryo that will later form different organs (Larsen *et al.*, 2003; Prin *et al.*, 2014) and they also play an important role in the specification of tissue morphology in events such as gastrulation (Oda *et al.*, 2003), neural tube (Haigo *et al.*, 2003) and retina formation (Plageman *et al.*, 2010). In adult epithelial tissue though, multicellular apical constriction is normally silenced, and activated in events such as wound healing (Martin *et al.*, 2014). Loss of cohesin function in the *Drosophila* notum induced apical constriction, actomyosin ring formation and clonal extrusion. Interestingly, the KD of individual cohesin subunits did not induce apical constriction; only the LOF mutation of $smc3^{-/-}$, the simultaneous KD of SA1 and SA2, and and KD of the cohesin loader NipB could induce supracellular actin rich rings and clonal extrusion. It has previously been shown that a small dose of SMC3 (as little as 30 percent of protein levels) are sufficient for the cohesin ring to successfully perform its role in SCC (Laugsch *et al.*, 2013). Additionally, it has been shown that the loss of a specific SA protein can be compensated for by the other isoform. The KD of one SA gene has been shown to have effects on the regulation of gene expression, yet SCC is performed normally (Benedetti *et al.*, 2017), although this seems to be cell type specific and STAG2 KD has

been shown to cause aneuploidy in normal bladder cells (Li *et al.*, 2015). In contrast, cells are more sensitive to changes in the levels of NipB. It has been reported that as little as a 15% reduction of NipB transcript is sufficient to cause the most common cohesinopathy CdL (Krantz *et al.*, 2004; Zuin *et al.*, 2017). These results provide more evidence to support the theory that cohesin levels have very important effects on overall cohesin function and tightly determine the phenotype induced. Therefore, the results from this study indicate that the cellular levels of cohesin are crucial to whether supracellular apical constriction will be activated, suggesting that only very low levels of cohesin will trigger this phenotype.

The main effectors of apical constriction, myosin2 and RHO1, induced drastic apical shape changes when overexpressed in the epithelial cells of the notum of *Drosophila*. The overexpression of an activated form of myosin II was sufficient for the formation of actin rings. This is consistent with other studies that have shown that even though active myosin II is required at the apex of the cell for apical constriction to occur, an ubiquitous overexpression of myosin across the cell induces apical constriction and ring formation (Escudero *et al.*, 2007). Additionally, the expression of the constitutively active form of RHO1 did not induce actin ring formation. Instead a multicellular rosette structure was observed. Rosette formation involves a

combination of apical constriction and changes in planar polarity (Blankenship *et al.*, 2006) and are observed during development in the retina formation (Rauzi *et al.*, 2008), epithelial elongation (Zallen *et al.*, 2008), pancreatic branching (Yan *et al.*, 2019), among others. Therefore, these results suggest that although increased RHO1 activity has effects on cell shape and induces apical constriction, an overexpression of Myosin2 is required for a stronger phenotype, including supracellular ring formation and clonal extrusion. Furthermore, the activity of both Myosin2 and RHO1 within *smc3* mutant tissue had significant effects on the *smc3* mutant phenotype. The KD of either SQH or RHO1 significantly reduced the number of actin rings and extruded clones within $smc3^{-/-}$ tissue, indicating that the two apical constriction effectors have relevant roles in the induction of supracellular actin ring formation in *smc3* mutant background. In accordance with that, the overexpression of activated SQH significantly enhanced the actin ring phenotype in $smc3^{-/-}$ mutant clones. In contrast, constitutively active RHO1 in combination with $smc3^{-/-}$ was not possible to analyse, since the mutant tissue was too small and only circular unicellular bodies were observed. The very small clonal area could have been a consequence of RHO1 having a too severe effect in $smc3^{-/-}$ mutants, inducing a higher or faster rate of clonal extrusion and therefore the mutant clones could have already been removed from the epithelium at the time of imaging. Additionally, the smaller clonal size could be explained by the activation of apoptosis, since increased RHO1

activity has been linked to JNK activity (Neisch *et al.*, 2010). Tools to measure RHO1 activity within *smc3*$^{-/-}$ mutant clones could be used in the future, such as FRET (Yoshizaki *et al.*, 2003) or live fluorescent RHO1 activity sensors (Munjal *et al.*, 2015).

RHO1 and Myosin2 activity were important for the actin ring formation in *smc3*$^{-/-}$ mutant clones. To better understand the role of *smc3*$^{-/-}$ in actin ring formation, an enhancer/suppressor screen of different apical constriction regulators was done in the notum of the living fly by either knocking down or overexpressing genes from apical constriction pathways both alone and in combination with *smc3*$^{-/-}$. The overexpression of four well known regulators of apical constriction: TKV, MAD, CUT and SHRM significantly induced actin ring formation, compared to WT tissue. Additionally, in accordance with the literature, the KD of these genes and two other inducers of apical constriction: FOG and Hh did not induce actin ring formation. When combined with the *smc3*$^{-/-}$ mutant, Mad had a significant effect, both by enhancing the *smc3*$^{-/-}$ phenotype when overexpressed and by suppressing the *smc3*$^{-/-}$ phenotype when knocked down. While the KD or overexpression of the other apical constriction genes within *smc3*$^{-/-}$ mutant tissue had no significant effect in actin ring formation or clonal extrusion.

The difference in the effects observed in the apical constriction enhancer/suppressor screen might be explained by analysing each of the five pathways studied: i) Dpp, including TKV and MAD, ii) Hh, iii) CUT, iv) FOGand v) SHRM. Some of these pathways are redundant and others are tissue specific. For example, Folded gastrulation (FOG) is known for being the main inducer of apical constriction during *Drosophila* gastrulation by recruiting RhoGEF2, a RHO1 activator, to the apex of the cell. Yet, while FOG activity is essential for midgut invagination, FOG has no significant effect in the ventral furrow (Sweeton *et al.*, 1991). Since the effect of FOGoutside the midgut is not essential and the KD of Fog did not influence *smc3*$^{-/-}$ actin ring formation, it seems unlikely that the role of cohesin in apical constriction is through the regulation of *Fog*.

Shroom (SHRM), named thus as the invaginations induced by this protein resemble mushrooms, directly recruits ROCKs, the downstream effectors of RHO1, to the apical junctions (Nishimura *et al.*, 2008). Dominant negative expression of RHO1 within activated SHRM tissue has been shown to have no effect on apical constriction in the *Xenopus* neural tube (Haigo *et al.*, 2003), confirming ROCKs functionality even without RHO1 activity, in SHRM tissue. Overexpression of SHRM had no effect on *smc3*$^{-/-}$ actin ring formation. Potentially, if ROCK is already being recruited by RHO1, then SHRM

activity might be redundant and therefore no enhancement of *smc3*$^{-/-}$ actin ring formation was observed. Additionally, the KD of SHRM did not suppress the *smc3*$^{-/-}$ actin ring phenotype. This could be as well due to ROCK being activated independently of SHRM. Furthermore, the expression of dominant negative RHO1 in *smc3*$^{-/-}$ mutant clones was sufficient to rescue the apical constriction phenotype. If *smc3*$^{-/-}$ was inducing actin ring formation through SHRM, the activity of RHO1 would not interfere with the induction of apical constriction. Therefore, these results indicate that *smc3*$^{-/-}$ function in apical constriction is likely to be independent of SHRM.

High levels of CUT are sufficient to induce apical constriction in *Drosophila* border cells (Levine *et al.*, 2010). Additionally, cohesin has been linked to the negative regulation of CUT expression in the *Drosophila* wing disc (Dorsett *et al.*, 2005). Therefore, it was surprising that the overexpression or KD of CUT did not have a significant effect on the actin ring formation rate within *smc3*$^{-/-}$ mutant clones. CUT expression is tightly regulated during fly development. It is highly expressed in embryos, yet silenced in later stages of development. Specifically, within the pupal notum, CUT is again activated within the SOP cells (Blochlinger *et al.*, 1990, 1993). No increased CUT levels were observed in *smc3*$^{-/-}$ clones, both before and after SOP specification. Indicating that in this tissue, loss of cohesin does not induce CUTexpression. While no

significant change was observed in the number of CUT positive SOP cells, between WT and *smc3*$^{-/-}$ tissue, a negative trend in the number of CUT positive cells was observed in *smc3*$^{-/-}$ tissue. As sample size was small, a higher number of animals could be studied in the future. While Dorsett et al. found cohesin to suppress CUT expression, the same research group published that, within the same tissue, NipB had opposite effects to SA1 and two different *rad21* mutations had opposite effects in *Cut* expression, their model proposed that this difference was due to chromatin domain boundaries (Rollins *et al.*, 2004). Therefore, a decrease in *Cut* expression in *smc3*$^{-/-}$ mutants could also be explained through the same model. Overall, while more studies are needed to determine if loss of cohesin decreases *Cut* expression in the notum and the link between cohesin and *Cut* is an interesting interaction to study. Actin rings were formed in *smc3*$^{-/-}$ mutants, even when no CUT signal was observed in both young and older pupae. Additionally, since the overexpression and KD of CUT had no effect on ring formation in *smc3*$^{-/-}$ tissue, it is highly unlikely that CUT activity is involved in apical constriction and actin ring formation following cohesin loss of function.

The Dpp pathway is an inducer of apical constriction in *Drosophila* dorsal closure (Fernández *et al.*, 2007), eye development (Heberlein *et al.*, 1993) and others (Schlichting *et al.*, 2008; Jidigam *et al.*, 2015). Overexpression of the

Dpp effector, MAD, and its upstream activator, the receptor TKV, was sufficient to induce actin ring formation. Additionally, MAD overexpression within *smc3*$^{-/-}$ mutant clones, induced a significant increase in the number of actin rings and a rescue of the phenotype was observed by MAD KD. Furthermore, phosphorylated MAD was increased in *smc3*$^{-/-}$ mutant and SA1 + SA2 KD clones with actin rings, indicating that the most plausible mechanism of action for cohesin actin ring formation is through increased Dpp signalling. Interestingly, TKV had no effect on the Cohesin actin ring phenotype. Dpp signalling is tightly regulated. While TKV is a main kinase of MAD, the PUNT receptor is a parallel and redundant MAD kinase (Wiersdorff *et al.*, 1996). Studies have shown that individual KD of TKV and PUNT had no effect on the invagination needed for lens formation, yet the simultaneous KD of both prevented invagination (Rajagopal *et al.*, 2009). Therefore, a more detailed study of both TKV and PUNT kinases is required to better understand the mechanism of how cohesin might regulate Dpp signalling. Additionally, the Hh pathway, that acts redundantly to Dpp in the context of apical constriction (Curtiss *et al.*, 2000; Corrigall *et al.*, 2007), was studied. While no effect was observed by knocking down the Hh receptor within *smc3*$^{-/-}$ mutant clones, the study of the MAD equivalent in Hh signalling, Cubitus interruptus (Ci), is a recommendation for future experiments. It is a possible that this signalling pathway is also affected, yet

by only studying Hh and not Ci, the effect could be masked by the complex regulation of the Hh signalling pathway.

Limitations from this study include the time window for when it's possible to image *Drosophila* pupae. Since 12h APF is the youngest age for pupal mounting, it is not possible to determine whether animals with lower clonal coverage had a higher rate of clonal delamination and actin ring formation before 12h APF or if the smaller clones were observed due to increased apoptosis and cell competition. Further experiments could include measuring the levels of apoptosis with tools like the live construct iCasper, or immunodetection with the TUNEL assay or Dcp-1 staining, to have a better understanding of which mechanism is being triggered.

Additionally, the induction of apical constriction and clonal extrusion is an interesting process to be studied in the context of collective cell migration, since clonal delamination did not increase apoptotic levels. Furthermore, the levels of junctional E-cadherin were significantly increased within *smc3*$^{-/-}$ mutant clones. A hallmark of collective invasion is the retention of cell-cell junctions, therefore additional studies could be done in order to identify other markers that the cohesin LOF clones express. For example, levels of cytokeratin-14 (K14) and p63 epithelial proteins could be studied, as they

have been shown to be essential to breast cancer collective cell migration (Cheung *et al.*, 2013). Overall these results show an exciting and novel potential role for the cohesin complex in the regulation of the Dpp pathway, which may additionally have relevant links to the mechanism of action of the developmental cohesinopaties, such as Cornelia de Lange syndrome. Furthermore, inducers of apical constriction and clonal extrusion could be studied for their potential role in collective cell migration during carcinogenesis.

Chapter 6 Discussion

6.1 Discussion

The idea that cancer is a genetic disorder (Knudson, 1971, 1973) revolutionised the field in the 1970s, triggering the identification of several genes and signalling pathways that affect tumour progression. While great advancements have been made over the past 5 decades, there are still many challenges to overcome in cancer treatment and detection. Unlike other genetic diseases, cancer is not systemic, it starts with a single cell accumulating advantageous mutations (Farber, 1984; Beroukhim *et al.*, 2010). Current diagnostic tools fail to detect tumour cells at such early stages, leading to tumour evolution, cancer development, clonal selection and diversification of the mutant population, providing significant challenges to therapy. Furthermore, since tumour cells accumulate several mutations, many of the genes mutated are not fully characterized and identification of the driver genes that are involved in tumour progression is necessary to have a better understanding of the disease and to design a wider selection of biomarkers and treatment targets.

6.1.1 Genetic screen

A large-scale genetic screen was designed with the aim of providing preliminary characterization of genes that are frequently found mutated in human carcinomas, yet their mechanism of action is unknown. *Drosophila*

provide a model that is especially advantageous for large-scale studies, due to the fact that they have a short fly life cycle, they are economic to maintain, and there are extensive and sophisticated genetic techniques available. We screened through 500 genes, by knocking them down individually within positively labelled tumours on the back of the living fly and determining the effect that this has on tumour progression related phenotypes. A total of 294 genes were identified as potential tumour suppressor genes, since their KD enhanced phenotypes such as cell division, invasion, polarity defects, protrusion length, among others.

The screen is a powerful versatile tool. Several phenotypes were analysed which allow different approaches for the identification and selection of target genes for future study based on specific research interests. For example, to study a broad process, such as invasion, proliferation or actin related defects; the genes identified in consensus clustering can be selected. Additionally, individual categories can be picked for more specific processes, i.e. within actin defects: protrusion length, actin spots, polarized cytoskeleton, among others. The same rationale applies to the invasion and proliferation related categories.

If instead of a process, a gene of interest is in mind, results from hierarchical clustering can be used to identify genes that induce a similar effect across all the categories studied in the screen. For example, if there is a particular research interest in the well-known TSG p53 (Cluster 3), other genes from cluster 3 will be interesting hits. Indeed, p53 well-known partners are part of the same cluster, such as E2F (Wu *et al.*, 1994; Polager *et al.*, 2009), FAS (Bennett *et al.*, 1998; Müller *et al.*, 1998), SCOX (Martínez-Morentin *et al.*, 2015) and others. Additionally, completely uncharacterized genes that are predicted to bind DNA or chromatin are part of cluster 3: CG10623, CG10669, CG4360, CG8290 (Stampfel *et al.*, 2015; Sarov *et al.*, 2016; M. Liu *et al.*, 2017). Therefore, they are interesting genes to study and they might regulate P53 related processes.

Network analysis is an additional tool to select genes based on previously identified interactions (Chen *et al.*, 2009). Network analysis, in combination with the screen results provide strong evidence of a potential unknown function for uncharacterized genes. Therefore, genes that interact in the network of invasion suppressors are strong candidates for future study. The interaction network of invasion suppressor hits from the genetic screen generated a complex net of interactions between well-known genes that are involved in cancer invasion, i.e. the adherens junction members E-cadherin

and p120ctn (Jeanes *et al.*, 2008; Peglion *et al.*, 2013); together with 99 genes that have not been previously linked to invasion. Additionally, network analysis identified 6 "hub genes", highly interactive genes within this network: Sema1a, Nrg, S6KII, Nos, Rac2 and Tsc1; these genes are of special interest since they affect several signalling pathways and their alteration will have greater effects when compared to genes that are less interactive (Zaman *et al.*, 2013).

Overall, the results obtained in the genetic screen provide a wealth of candidate genes with great potential for future targeted therapy and molecular diagnosis. Several genes that are currently only linked to cancer by its mutational frequency, or reported interactions with cancer related genes, might be overlooked for more detailed research. Therefore, the screen contributes by providing additional evidence on the effect these genes have in tumour progression in a living epithelium.

6.1.2 Cohesin regulates single and collective cell invasion

Cohesin, a nuclear multiprotein complex known for its role in sister chromatid cohesion, gene expression regulation and DNA repair (Hoque *et al.*, 2002; Unal *et al.*, 2004; Rubio *et al.*, 2008), was identified in the screen for inducing different phenotypes related to actin cytoskeleton re-arrangement;

roles that have not been previously linked to cohesin. Interestingly, individual KD of each subunit led to different phenotypes. KD of one subcellular localization subunit, SA1 or SA2, increased invasion, multilayering and apex defects. Reduced expression of the core subunits, SM1, RAD21 and SMC3, increased multilayering and apex defects, yet had no effect on invasion. Loss of all Cohesin function: *smc3*$^{-/-}$ mutation, SA1 + SA2 simultaneous KD or NipB KD, induced multicellular actin ring formation.

The different functions of the Cohesin subunits (SA1 and SA2 provide subcellular localization; SMC1, SMC3 and RAD21 form the core of the ring) (Gruber *et al.*, 2003), the redundancy of the different isoforms (SA1/SA2, SMC1A/SMC1B) (Losada *et al.*, 2000; Revenkova *et al.*, 2004), in combination with the specific dose required for each subunit to efficiently perform its role in either gene expression or SCC (Laugsch *et al.*, 2013), could be key to understanding the different effects observed in this study. Several recent studies have shown that individual loss of SA1 or SA2 has different effects compared to loss of all cohesin (Rao *et al.*, 2017; Schwarzer *et al.*, 2017; Wutz *et al.*, 2017). Loss of SA1 or SA2 leads to chromatin rearrangement with no signs of aneuploidy (Kim *et al.*, 2012; Djos *et al.*, 2013), yet this is cell type dependent, since aneuploidy has been observed by the loss of SA2 in

glioblastoma and bladder cancer cell lines (Solomon *et al.*, 2011, 2013). Additionally, a study in which a partial KD of SMC3, where protein expression is reduced by 70%, has shown low levels of *SMC3* to be sufficient for sister chromatid cohesion, with no signs of aneuploidy (Laugsch *et al.*, 2013). In contrast, loss of all cohesin leads to chromosome instability and aneuploidy (Losada, 2014).

The model I propose for the different phenotypes observed is that individual KD of SA1 or SA2 will trigger compensation by the other subunit, depending on the subunit that was knocked down, and either SA1-Cohesin or SA2-cohesin will be upregulated to compensate the loss of the other subunit (Benedetti *et al.*, 2017). As mentioned before, the two SA cohesins are not fully interchangeable; at interphase, SA1 cohesin is found to co-localize with CTCF to induce DNA loop formation, while SA2 cohesin is also found to be interacting with gene promoters on its own (Kojic *et al.*, 2018). Therefore, loss of one specific SA subunit will have drastic effects on how cohesin interacts with chromatin and on the genes whose expression is regulated by cohesin. Specifically, in the cell line MCF7, STAG2 KD significantly affected the expression of 23 genes, 20 of which interact with adherens junction genes. Therefore, it is likely that loss of SA1 or SA2 induces single cell invasion by affecting cohesin mediated gene expression during interphase (Figure 6.1).

KD of the core cohesin subunits, SMC1, SMC3 and RAD21, while significantly reducing the expression of each subunit, does not completely deplete their expression (Mouri *et al.*, 2012). As mentioned previously, it has been shown that 30 percent of SMC3 expression is sufficient for SCC to occur normally (Laugsch *et al.*, 2013). Therefore, although effects on the cytoskeleton are observed with subunit KD, a more severe phenotype is observed with a genetic mutation (Figure 6.1).

Full loss of cohesin function, as observed with SA1+SA2 KD, *smc3*$^{-/-}$ mutant and NipB KD, has drastic effects. In addition to the cytoskeletal phenotypes observed by the KD of the individual core subunits, an apical actin ring is additionally induced. My experiments suggest that this phenotype is induced by a mis-regulation of Dpp signalling that leads to Myosin II accumulation at the apex of the cell. Potentially, cohesin loss could induce Dpp upregulation through chromosome instability (Figure 6.1).

Important future experiments include the quantification of cohesin levels in each of the cohesin fly lines used, to determine the levels of protein expressed in the cohesin KD and mutant, as this will provide a better

understanding of the effects cohesin expression levels may have on the severity of phenotypes observed. Additionally, transcriptomic analysis of cells with loss of all cohesin function can be done and compared to the profile of genetic expression of STAG2 KD cells, this will provide a better understanding of the potential mechanisms of action for the actin ring formation. More economical alternatives could be performed as well, since the DPP signalling pathway has been identified as a strong candidate for cohesin mediated actin ring formation, RT-qPCR analysis of Dpp signalling members, especially the SMAD effectors, will provide valuable information. If the proposed model in Figure 6.1 is correct, SMAD expression is not affected in STAG2 KD, in accordance with the results observed in the MCF7 STAG2 KD transcriptomic analysis. In contrast, a Dpp signalling upregulation in STAG1 + STAG2 KD might be observed. Fluorescence in situ hybridization (FISH) could also be done to determine if SMAD or other Dpp regulators mis-segregate in Cohesin induced aneuploidy. Additionally, Chromatin immunoprecipitation (ChIP) can be used to identify cohesin interacting domains.

The most striking difference between individual loss of SA1 or SA2, compared to all Cohesin loss of function was E-cadherin localization, that was reduced in SA1 or SA2 KD. Conversely, E-Cadherin was significantly

increased in *smc3*$^{-/-}$ mutant clones. Phenotypically, E-Cadherin expression correlates with what is currently known about single and collective cell invasion; since SA1/SA2 invading cells are circular and resemble amoeboid invading cells, and one of their characteristics is the loss of cell-cell junctions (Ly *et al.*, 2009). In contrast, the most iconic hallmark of collective cell migration is the retention of cell-cell junctions (Friedl *et al.*, 2009), and the *smc3*$^{-/-}$ mutant clones extrude from the epithelium collectively. However, how cohesin differentially affects gene expression that leads to E-Cadherin reduction in localization by individual SA1 or SA2 KD, as opposed to E-Cadherin increased junctional localization in *smc3*$^{-/-}$ mutants is unknown. The previously recommended experiment, of the study of differentially expressed genes following loss of all cohesin will be very informative to generate a more specific hypothesis for cohesin mediated gene expression that leads to cell-cell adhesion defects.

This research outlines the complexity of cohesin function regulation and cohesin mediated gene expression and provides further evidence of a novel tumour suppressive role of SA1 and SA2 in the regulation of single cell invasion. Additionally, Cohesin mediated actin cytoskeleton rearrangement, are exciting novel roles for cohesin that will provide a better understanding of the cohesin role in tumour progression. Moreover, actin ring formation

and apical constriction are exciting potential mechanisms of action for the poorly understood process of collective cell invasion.

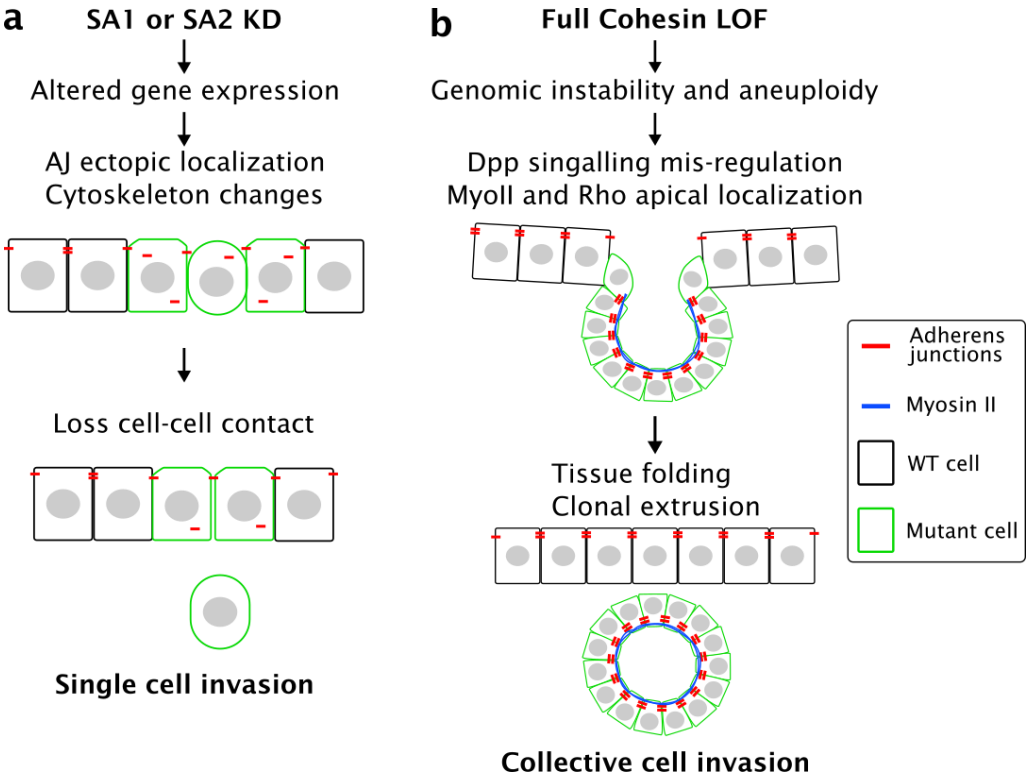


Figure 6.1 Proposed model of the role of cohesin in single and collective cell invasion

a) Individual KD of SA1 or SA2 induces chromatin re-arrangement and changes in gene expression. Adherens junction and cytoskeleton regulatory genes are affected. Mutant cells lose contact with neighbouring cells and single cell invasion occurs. **b)** Full cohesin loss of function, caused by a mutation in the cohesin ring core subunits: SMC1, SMC3 or RAD21; or by the simultaneous loss of SA1 and SA2; or by the loss of the cohesin loader NipB. This is likely to lead to defective sister chromatid cohesion and chromosome instability. If alterations that lead to activated Myosin II localization at the apex of the cells are induced, such as upregulated Dpp signalling, actin ring formation will be induced causing tissue folding and, for small clones, full clone extrusion and collective migration.

6.1.3 Apical constriction and its potential role in collective cell invasion.

An additional serendipitous finding from this research is the potential role of apical constriction genes in tumour progression by promoting mutant tissue folding that leads to collective cell invasion. Activation of the Dpp, Shrm and Cut signalling pathways induced apical actin ring formation and mutant clonal extrusion in the living epithelium of *Drosophila;* indicating that aberrant activation of these pathways in adult epithelial tissues could have a role in collective cell invasion.

The Dpp signalling pathway (BMP in humans) is found upregulated in various different cancers, including metastatic prostate, colon, lung and breast cancer (Darby *et al.*, 2008; Yokoyama *et al.*, 2017; Zabkiewicz *et al.*, 2017). Accordingly, the BMP signalling inhibitor, Noggin, has been proposed in different studies as a potential therapeutic drug for metastatic prostate and lung cancers (Haudenschild *et al.*, 2004; Feeley *et al.*, 2006). While the Hh pathway could not be overexpressed in this study, several research studies show that Dpp and Hh are redundant in their role regulating apical constriction (Fu *et al.*, 2003; Corrigall *et al.*, 2007). Therefore, Hh activation is expected to have a similar effect as Dpp. In accordance, Hh signalling is

found activated in metastatic pancreatic, esophageal, gastric and prostate cancer (Berman *et al.*, 2003; Thayer *et al.*, 2003; Stecca *et al.*, 2005).

CUX1 (the Cut human orthologue) has been identified as a TSG in acute myeloid leukaemia (AML), gastric and colorectal cancers (McNerney *et al.*, 2013; Wong *et al.*, 2013). In contrast, CUX1 acts as an oncogene in cancer of kidney, lung, ovary, endometrium and central nervous system (Sansregret *et al.*, 2008; Ramdzan *et al.*, 2014). Future studies of the role of CUT/CUX1 in tumour progression should take tissue specificity into consideration. Fog and SHRM are the least studied apical constriction pathways in cancer. FOG2 is found overexpressed in neuroblastoma (Hoene *et al.*, 2009). SHRMhas not been studied in cancer, yet extensive evidence shows how SHRM affects the actin cytoskeleton by the regulation of ROCKs (Hildebrand, 2005; Zalewski *et al.*, 2016).

Overall, the data presented here in combination with previous studies indicates that the genes involved in apical constriction are potential oncogenes that might be involved in collective cell migration. While some of them, such as the members of the Dpp signalling pathway have already been suggested as therapeutic targets, their additional role in collective cell migration should be explored to better understand all the different

phenotypes these pathways affect. Furthermore, the role of Fog and Shrm in tumour progression has not been studied and will provide novel information. As mentioned in Chapter 1, most of the cancer related processes are highly activated during development and wound healing, such as proliferation, angiogenesis, collective cell migration, among others. Therefore, apical constriction as an essential process for development and wound healing is a very promising line of research to study in relation to tumour progression.

6.2 Future perspectives

Long term future directions will aim for the translation of the results obtained from the genetic screen to a clinical context. Next steps should include a research model that resembles the human disease, since tissue specific mutations are key for each type of cancer, especially for genes involved in regulating gene expression. Further characterization of the effect that these genes have in cancer related phenotypes can be done in a tissue specific manner using cell culture of specific cancer types, mouse models, the emerging organoid models (Muthuswamy, 2018; Grassi *et al.*, 2019), among others. Bioinformatic tools will yield very informative data, since identification of the tissues where the "hit" genes from the screen are found commonly mutated will provide a starting point for future studies.

The role of SA1 and SA2 in single cell invasion can be further characterized in additional breast cancer cell lines. Moreover, the effect of SA1 or SA2 KD can be studied in bladder cancer, glioblastoma and AML where Cohesin is also frequently found to be mutated (Kandoth *et al.*, 2013; Leiserson *et al.*, 2015). The differential effects that Cohesin loss induces in gene expression is another exciting line of research to continue, where transcriptomic analysis and chromatin immunoprecipitation can provide valuable information.

Furthermore, the characterization of clonal extrusion and its potential link to collective cell migration is an exciting novel phenotype for further study. The design of an appropriate model, to determine if this phenotype is observed in other organisms, will be a challenge; since the cooperation between the mutant cells and the WT cells is very important for the extrusion of the mutant clones. Therefore, a model that is highly resembling of a human tumour is necessary, such as an adaptation of current invasion models, i.e the 3D spheroid models could be done (Liu *et al.*, 2018; Mi *et al.*, 2019); mouse models can be designed, especially since evidence built in *Drosophila* justifies the use of larger organisms to study this process. Previously mentioned organoid models can be adapted as well.

6.3 Limitations

While *Drosophila* provides many advantages, mentioned in section 1.4 of Chapter 1, certain phenotypes of tumour progression cannot be studied in this model. For example, the adaptive immune system in *Drosophila* has been described to be much simpler than in humans. Therefore, the effect that the immune system has in tumorigenesis cannot be asesed fully. Additionally, as *Drosophila* has an open circulatory system, the process of intravasation, that is a highly relevant step in the invasion-metastasis cascade, cannot be studied. Because of this, our study focuses on the poorly understood early stages of invasion, prior to intravasation. Yet, genes that did not show a phenotype of tumour progression in the genetic screen might be involved in processes that cannot be studied in *Drosophila*.

In addition to that, cancer is tissue specific and genes that did not induce any changes on the epithelial tissue of the notum of the fly, might sill act as TSGs in different tissues. In a similar manner, genes that induced tumour progressing changes in the notum of the fly might not act as TSGs in different tissues and/or in different organisms. Bioinformatic tools can be used before future validation experiments, by identifying the types of cancer where specific genes are commonly mutated, the effect that LOF of these genes can be assessed in those specific tissues.

Imaging techniques introduce additional limitations. Confocal imaging allows the capture of images with high spatiotemporal resolution, yet depth is very limited. Invading cells cannot be followed after a certain point, of around 20μM deep, where light has scattered and the microscope detectors cannot differentiate signal from noise. For studies that are interested in following the trajectory of the invading cells further, different imaging techniques, that have higher depth but less resolution can be used, such as multi photon microscopy and optoacustic methods.

Limitations of the cell culture *in vitro* studies include that only one cell line was used, this was due to the limited time and resources available. Due to variability of the transcriptomic analysis results, additional biological replicates should be analysed to increase sample size. Additionally, in vitro invasion assays should have been accompanied by cell shape analysis and prolifereation assays to better understand the mechanisms of cohesin mediated cell invasion. Finally, 2D migration assays, such as scrap assays can be performed to validate the migratory potential of STAG1 and STAG2 KD cells.

6.4 Conclusions

Cancer research faces several challenges, due to the complexity and diversity of the different types of cancer, specified by the cell type affected and the rich diversity of the acquired mutations and epigenetic changes. This research contributes to a better understanding of this disease and proposes several genes identified in the genetic screen as potential tumour suppressor genes. Additionally, a novel role for the Cohesin complex as a regulator of single and collective cell invasion is presented, encouraging complementary studies of the Cohesin subunits in a tissue specific manner for their consideration as biomarkers. Furthermore, apical constriction genes are suggested as potential oncogenes, with great potential to become therapeutic targets. Overall, the research presented here is an important starting point for the future translation of our current knowledge of the several potential tumour suppressor genes and oncogenes identified in this study to applications of clinical relevance that will have an impact in the improvement of public health.

Chapter 7 – Bibliography

Abel, E. L. *et al.* (2009) 'Multi-stage chemical carcinogenesis in mouse skin: Fundamentals and applications', *Nature Protocols*. Nature Publishing Group, 4, p. 1350.

Adams, M. D. *et al.* (2000) 'The Genome Sequence of Drosophila melanogaster', *Science*, 287(5461), p. 2185 LP-2195. doi: 10.1126/science.287.5461.2185.

Aguilar-Aragon, M. *et al.* (2018) 'Pak1 Kinase Maintains Apical Membrane Identity in Epithelia', *Cell reports*. Cell Press, 22(7), pp. 1639–1646. doi: 10.1016/j.celrep.2018.01.060.

Alberts, B. *et al.* (2002) *Molecular Biology of the Cell*. 4th Editio.

Alexander, S. *et al.* (2008) 'Dynamic imaging of cancer growth and invasion: a modified skin-fold chamber model', *Histochemistry and Cell Biology*, 130(6), pp. 1147–1154. doi: 10.1007/s00418-008-0529-1.

Alvarado, C. G. *et al.* (2011) 'Nuclear translocation of β-catenin synchronized with loss of E-cadherin in oral epithelial dysplasia with a characteristic two-phase appearance', *Histopathology*. John Wiley & Sons, Ltd (10.1111), 59(2), pp. 283–291. doi: 10.1111/j.1365-2559.2011.03929.x.

Amano, M. *et al.* (1996) 'Phosphorylation and Activation of Myosin by Rho-associated Kinase (Rho-kinase)', *Journal of Biological Chemistry* , 271(34), pp. 20246–20249. Available at: http://www.jbc.org/content/271/34/20246.abstract.

Amano, M. *et al.* (1997) 'Formation of Actin Stress Fibers and Focal Adhesions Enhanced by Rho-Kinase', *Science*, 275(5304), p. 1308 LP-1311. doi: 10.1126/science.275.5304.1308.

Anan, K. *et al.* (1996) 'Vascular endothelial growth factor and platelet-derived growth factor are potential angiogenic and metastatic factors in human breast cancer', *Surgery*, 119(3), pp. 333–339. doi: https://doi.org/10.1016/S0039-6060(96)80120-6.

Andrew, D. J. and Ewald, A. J. (2010) 'Morphogenesis of epithelial tubes: Insights into tube formation, elongation, and elaboration', *Developmental Biology*, 341(1), pp. 34–55. doi: https://doi.org/10.1016/j.ydbio.2009.09.024.

Antunes, M. *et al.* (2013) 'Coordinated waves of actomyosin flow and apical cell constriction immediately after wounding', *The Journal of Cell Biology*, 202(2), p. 365 LP-379. doi: 10.1083/jcb.201211039.

Arboleda, M. J. *et al.* (2003) 'Overexpression of AKT2/Protein Kinase Bβ Leads to Up-Regulation of β1 Integrins, Increased Invasion, and Metastasis of Human Breast and Ovarian Cancer Cells', *Cancer Research*, 63(1), p. 196 LP-206.

Ashburner, M. (1989) *Drosophila: A Laboratory Handbook and Manual*.

Baggetto, L. G. (1992) 'Deviant energetic metabolism of glycolytic cancer cells', *Biochimie*, 74(11), pp. 959–974. doi: https://doi.org/10.1016/0300-9084(92)90016-8.

Baluk, P., Hashizume, H. and McDonald, D. M. (2005) 'Cellular abnormalities of blood vessels as targets in cancer', *Current Opinion in Genetics & Development*, 15(1), pp. 102–111. doi: https://doi.org/10.1016/j.gde.2004.12.005.

Barber, T. D. *et al.* (2008) 'Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers', *Proceedings of the National Academy of Sciences of the United States of America*. 2008/02/25. National Academy of Sciences, 105(9), pp. 3443–3448. doi: 10.1073/pnas.0712384105.

Batlle, E. *et al.* (2000) 'The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells', *Nature Cell Biology*. Macmillan Magazines Ltd., 2, p. 84. Available at: https://doi.org/10.1038/35000034.

Baumann, P. and West, S. C. (1998) 'Role of the human RAD51 protein in homologous recombination and double-stranded-break repair', *Trends in Biochemical Sciences*, 23(7), pp. 247–251. doi: https://doi.org/10.1016/S0968-0004(98)01232-8.

Behrens, J. *et al.* (1996) 'Functional interaction of β-catenin with the transcription factor LEF-1', *Nature*, 382(6592), pp. 638–642. doi: 10.1038/382638a0.

Behrndt, M. *et al.* (2012) 'Forces Driving Epithelial Spreading in Zebrafish Gastrulation', *Science*, 338(6104), p. 257 LP-260. doi: 10.1126/science.1224143.

Bekes, E. M. *et al.* (2011) 'Tumor-Recruited Neutrophils and Neutrophil TIMP-Free MMP-9 Regulate Coordinately the Levels of Tumor Angiogenesis and Efficiency of Malignant Cell Intravasation', *The American Journal of Pathology*, 179(3), pp. 1455–1470. doi: https://doi.org/10.1016/j.ajpath.2011.05.031.

Benedetti, L. *et al.* (2017) 'Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1', *Oncotarget*, 8(23), pp. 37619–37632.

Bennett, M. *et al.* (1998) 'Cell Surface Trafficking of Fas: A Rapid Mechanism of p53-Mediated Apoptosis', *Science*, 282(5387), p. 290 LP-293. doi: 10.1126/science.282.5387.290.

Berman, D. M. *et al.* (2003) 'Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours', *Nature*. Macmillan Magazines Ltd., 425, p. 846. Available at: https://doi.org/10.1038/nature01972.

Berndt, J. D. *et al.* (2008) 'Rho-kinase and myosin II affect dynamic neural crest cell behaviors during epithelial to mesenchymal transition in vivo', *Developmental Biology*, 324(2), pp. 236–244. doi: https://doi.org/10.1016/j.ydbio.2008.09.013.

Beroukhim, R. *et al.* (2010) 'The landscape of somatic copy-number alteration across human cancers', *Nature*, 463(7283), pp. 899–905. doi: 10.1038/nature08822.

Bersini, S. *et al.* (2014) 'In vitro models of the metastatic cascade: from local invasion to extravasation', *Drug Discovery Today*, 19(6), pp. 735–742. doi: https://doi.org/10.1016/j.drudis.2013.12.006.

Bienz, M. and Tremml, G. (1988) 'Domain of Ultrabithorax expression in Drosophila visceral mesoderm from autoregulation and exclusion', *Nature*, 333(6173), pp. 576–578. doi: 10.1038/333576a0.

Birkenbihl, R. P. and Subramani, S. (1992) 'Cloning and characterization of rad21 an essential gene of Schizosaccharomyces pombe involved in DNA double-strand-break repair', *Nucleic Acids Research*, 20(24), pp. 6605–6611. doi: 10.1093/nar/20.24.6605.

Blankenship, J. T. *et al.* (2006) 'Multicellular Rosette Formation Links Planar Cell Polarity to Tissue Morphogenesis', *Developmental Cell*, 11(4), pp. 459–470. doi: https://doi.org/10.1016/j.devcel.2006.09.007.

Blaser, H. *et al.* (2006) 'Migration of Zebrafish Primordial Germ Cells: A Role for Myosin Contraction and Cytoplasmic Flow', *Developmental Cell*, 11(5), pp. 613–627. doi: https://doi.org/10.1016/j.devcel.2006.09.023.

Blochlinger, K. *et al.* (1990) 'Patterns of expression of cut, a protein required for external sensory organ development in wild-type and cut mutant Drosophila embryos.', *Genes and Development*, 4, pp. 1322–1331.

Blochlinger, K., Jan, L. Y. and Jan, Y. N. (1993) 'Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila', *Development*, 117(2), p. 441 LP-450.

Blum, R. *et al.* (2005) 'Ras Inhibition in Glioblastoma Down-regulates Hypoxia-Inducible Factor-1α, Causing Glycolysis Shutdown and Cell Death', *Cancer Research*, 65(3), p. 999 LP-1006.

Brabletz, T. *et al.* (2005) 'Invasion and Metastasis in Colorectal Cancer: Epithelial-Mesenchymal Transition, Mesenchymal-Epithelial Transition, Stem Cells and β-Catenin', *Cells Tissues Organs*, 179(1–2), pp. 56–65. doi: 10.1159/000084509.

Brand, A. H. and Perrimon, N. (1993) 'Targeted gene expression as a means of altering cell fates and generating dominant phenotypes', *Development*, 118(2), p. 401 LP-415.

Breier, G. *et al.* (2002) 'Transforming growth factor-β and Ras regulate the VEGF/VEGF-receptor system during tumor angiogenesis', *International Journal of Cancer*. John Wiley & Sons, Ltd, 97(2), pp. 142–148. doi: 10.1002/ijc.1599.

Brito, D. A. and Rieder, C. L. (2006) 'Mitotic Checkpoint Slippage in Humans Occurs via Cyclin B Destruction in the Presence of an Active Checkpoint', *Current Biology*, 16(12), pp. 1194–1200. doi: https://doi.org/10.1016/j.cub.2006.04.043.

Bruckheimer, E. M. *et al.* (2000) 'Bcl-2 accelerates multistep prostate carcinogenesis in vivo', *Oncogene*. Macmillan Publishers Limited, 19, p. 5251.

Bryan, T. M. *et al.* (1995) 'Telomere elongation in immortal human cells without detectable telomerase activity', *The EMBO journal*, 14(17), pp. 4240–4248.

Canales Coutino, B. *et al.* (2019) *An in vivo systematic genetic analysis of tumour progression in Drosophila*.

Cano, A. *et al.* (2000) 'The transcription factor Snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression', *Nature Cell Biology*. Macmillan Magazines Ltd., 2, p. 76.

Cantley, L. C. *et al.* (1991) 'Oncogenes and signal transduction', *Cell*. Elsevier, 64(2), pp. 281–302. doi: 10.1016/0092-8674(91)90639-G.

Cantley, L. C. and Neel, B. G. (1999) 'New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway', *Proceedings of the National Academy of Sciences*, 96(8), p. 4240 LP-4245. doi: 10.1073/pnas.96.8.4240.

Canudas, S. and Smith, S. (2009) 'Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells', *Journal of Cell Biology*, 187(2), pp. 165–173. doi: 10.1083/jcb.200903096.

Carmena, A. *et al.* (2002) 'Reciprocal Regulatory Interactions between the Notch and Ras Signaling Pathways in the Drosophila Embryonic Mesoderm', *Developmental Biology*, 244(2), pp. 226–242. doi: https://doi.org/10.1006/dbio.2002.0606.

Carragher, N. O. *et al.* (2006) 'Calpain 2 and Src dependence distinguishes

mesenchymal and amoeboid modes of tumour cell invasion: a link to integrin function', *Oncogene*. Nature Publishing Group, 25, p. 5726.

Carvalhal, S. *et al.* (2018) 'A quantitative analysis of cohesin decay in mitotic fidelity', *The Journal of Cell Biology*, 217(10), p. 3343 LP-3353. doi: 10.1083/jcb.201801111.

Casey, T. *et al.* (2009) 'Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer', *Breast Cancer Research and Treatment*, 114(1), pp. 47–62. doi: 10.1007/s10549-008-9982-8.

Castellana, B. *et al.* (2012) 'ASPN and GJB2 Are Implicated in the Mechanisms of Invasion of Ductal Breast Carcinomas', (Idc). doi: 10.7150/jca.4120.

Chambers, A. F., Groom, A. C. and MacDonald, I. C. (2002) 'Dissemination and growth of cancer cells in metastatic sites', *Nature Reviews Cancer*. Nature Publishing Group, 2, p. 563.

Chang, J. *et al.* (1999) 'Proliferative hemangiomas: analysis of cytokine gene expression and angiogenesis', *Plastic and reconstructive surgery*. Division of Plastic Surgery at Stanford University Medical Center, Calif. 94305, USA., 103(1), p. 1–9; discussion 10. doi: 10.1097/00006534-199901000-00001.

Chao, W. C. H. *et al.* (2012) 'Structure of the mitotic checkpoint complex', *Nature*. Nature Publishing Group, 484(7393), pp. 208–213. doi: 10.1038/nature10896.

Chen, J., Aronow, B. J. and Jegga, A. G. (2009) 'Disease candidate gene identification and prioritization using protein interaction networks', *BMC Bioinformatics*, 10(1), p. 73. doi: 10.1186/1471-2105-10-73.

Cheng, E. H.-Y. . *et al.* (2001) 'BCL-2, BCL-XL Sequester BH3 Domain-Only Molecules Preventing BAX- and BAK-Mediated Mitochondrial Apoptosis', *Molecular Cell*, 8(3), pp. 705–711. doi: https://doi.org/10.1016/S1097-2765(01)00320-3.

Cheung, K. J. *et al.* (2013) 'Collective Invasion in Breast Cancer Requires a Conserved Basal Epithelial Program', *Cell*, 155(7), pp. 1639–1651. doi: https://doi.org/10.1016/j.cell.2013.11.029.

Cicchini, L. *et al.* (2016) 'Suppression of Antitumor Immune Responses by Human Papillomavirus through Epigenetic Downregulation of CXCL14', *mBio*. Edited by M. J. Imperiale, 7(3), pp. e00270-16. doi: 10.1128/mBio.00270-16.

Coles, A. H. *et al.* (2007) 'Deletion of p37^{Ing1} in Mice Reveals a p53-Independent Role for Ing1 in the Suppression of Cell Proliferation, Apoptosis, and Tumorigenesis', *Cancer Research*, 67(5), p. 2054 LP-2061. doi: 10.1158/0008-5472.CAN-06-3558.

Collado, M. *et al.* (2005) 'Senescence in premalignant tumours', *Nature*. Nature Publishing Group, 436, p. 642.

Consortium, T. I. C. G. *et al.* (2010) 'International network of cancer genome projects', *Nature*. Macmillan Publishers Limited. All rights reserved, 464, p. 993.

Corrigall, D. *et al.* (2007) 'Hedgehog Signaling Is a Principal Inducer of Myosin-II-Driven Cell Ingression in Drosophila Epithelia', *Developmental Cell*, 13(5), pp. 730–742. doi: https://doi.org/10.1016/j.devcel.2007.09.015.

Couto, A. *et al.* (2017) 'An apicobasal gradient of Rac activity determines protrusion form and position', *Nature Communications*. The Author(s), 8, p. 15385.

Cucco, F. *et al.* (2014) 'Mutant cohesin drives chromosomal instability in early colorectal adenomas', *Human Molecular Genetics*, 23(25), pp. 6773–6778. doi: 10.1093/hmg/ddu394.

Curtiss, J. and Mlodzik, M. (2000) *Morphogenetic furrow initiation and progression during eye development in Drosophila: The roles of decapentaplegic, hedgehog and eyes absent*, *Development (Cambridge, England)*.

Dameron, K. M. *et al.* (1994) 'Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1', *Science*, 265(5178), p. 1582 LP-1584. doi: 10.1126/science.7521539.

Daniloski, Z. and Smith, S. (2017) 'Loss of tumor suppressor STAG2 promotes telomere recombination and extends the replicative lifespan of normal human cells', *Cancer Research*, 77(20), pp. 5530–5542. doi: 10.1158/0008-5472.CAN-17-1260.

Darby, S. *et al.* (2008) 'BMP-6 over-expression in prostate cancer is associated with increased Id-1 protein and a more invasive phenotype', *The Journal of Pathology*. John Wiley & Sons, Ltd, 214(3), pp. 394–404. doi: 10.1002/path.2292.

Dekanty, A. *et al.* (2012) 'Aneuploidy-induced delaminating cells drive tumorigenesis in Drosophila epithelia', *Proceedings of the National Academy of Sciences*, 109(50), p. 20549 LP-20554. doi: 10.1073/pnas.1206675109.

Deryugina, E. I. and Quigley, J. P. (2015) 'Tumor angiogenesis: MMP-mediated induction of intravasation- and metastasis-sustaining neovasculature', *Matrix Biology*, 44–46, pp. 94–112. doi: https://doi.org/10.1016/j.matbio.2015.04.004.

Desai, R. A. *et al.* (2009) 'Cell polarity triggered by cell-cell adhesion via E-cadherin', *Journal of Cell Science*, 122(7), p. 905 LP-911. doi: 10.1242/jcs.028183.

Dietzl, G. *et al.* (2007) 'A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila', *Nature*. Nature Publishing Group, 448, p. 151.

DiGiovanni, J. (1992) 'Multistage carcinogenesis in mouse skin', *Pharmacology & Therapeutics*, 54(1), pp. 63–128. doi: https://doi.org/10.1016/0163-7258(92)90051-Z.

Dillon, M., R Y Cahn, L. and Huey, R. (2007) *Life history consequences of temperature transients in Drosophila melanogaster*, *The Journal of experimental biology*. doi: 10.1242/jeb.007591.

Djos, A. *et al.* (2013) 'Aneuploidy in neuroblastoma tumors is not associated with inactivating point mutations in the STAG2 gene', *BMC Medical Genetics*, 14(1), p. 102. doi: 10.1186/1471-2350-14-102.

Dogu, Y. and Díaz, J. (2009) 'Mathematical model of a network of interaction between p53 and Bcl-2 during genotoxic-induced apoptosis', *Biophysical Chemistry*, 143(1), pp. 44–54. doi: https://doi.org/10.1016/j.bpc.2009.03.012.

Dolznig, H. *et al.* (2004) 'Evidence for a size-sensing mechanism in animal cells', *Nature Cell Biology*. Nature Publishing Group, 6, p. 899.

Dorsett, D. *et al.* (2005) 'Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila', *Development*, 132(21), p. 4743 LP-4753. doi: 10.1242/dev.02064.

Enomoto, M. and Igaki, T. (2011) 'Deciphering tumor-suppressor signaling in

flies: Genetic link between Scribble/Dlg/Lgl and the Hippo pathways', *Journal of Genetics and Genomics*, 38(10), pp. 461–470. doi: https://doi.org/10.1016/j.jgg.2011.09.005.

Escudero, L. M., Bischoff, M. and Freeman, M. (2007) 'Myosin II Regulates Complex Cellular Arrangement and Epithelial Architecture in Drosophila', (November), pp. 717–729. doi: 10.1016/j.devcel.2007.09.002.

Fang, D. *et al.* (2007) 'Phosphorylation of β-Catenin by AKT Promotes β-Catenin Transcriptional Activity', *Journal of Biological Chemistry* , 282(15), pp. 11221–11229.

Farber, E. (1984) 'The Multistep Nature of Cancer Development', *Cancer Research*, 44(10), p. 4217 LP-4223.

Faure, A. J. *et al.* (2012) 'Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules', *Genome Research* , 22(11), pp. 2163–2175.

Fearon, E. R. and Vogelstein, B. (1990) 'A genetic model for colorectal tumorigenesis', *Cell*, 61(5), pp. 759–767. doi: https://doi.org/10.1016/0092-8674(90)90186-I.

Feeley, B. T. *et al.* (2006) 'Mixed Metastatic Lung Cancer Lesions in Bone Are Inhibited by Noggin Overexpression and Rank:Fc Administration', *Journal of Bone and Mineral Research*. John Wiley & Sons, Ltd, 21(10), pp. 1571–1580. doi: 10.1359/jbmr.060706.

Feinberg, A. P., Ohlsson, R. and Henikoff, S. (2006) 'The epigenetic progenitor origin of human cancer', *Nature Reviews Genetics*. Nature Publishing Group, 7, p. 21.

Feitelson, M. A. *et al.* (2015) 'Sustained proliferation in cancer: Mechanisms and novel therapeutic targets', *Seminars in Cancer Biology*, 35, pp. S25–S54. doi: https://doi.org/10.1016/j.semcancer.2015.02.006.

Fernández, B. G., Arias, A. M. and Jacinto, A. (2007) 'Dpp signalling orchestrates dorsal closure by regulating cell shape changes both in the amnioserosa and in the epidermis', *Mechanisms of Development*, 124(11), pp. 884–897. doi: https://doi.org/10.1016/j.mod.2007.09.002.

Ferrara, N. *et al.* (1996) 'Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene', *Nature*, 380(6573), pp. 439–442. doi: 10.1038/380439a0.

Fingar, D. C. *et al.* (2002) 'Mammalian cell size is controlled by mTOR and its downstream targets S6K1 and 4EBP1/eIF4E', *Genes & Development* , 16(12), pp. 1472–1487.

Forrester, K. *et al.* (1987) 'Detection of high incidence of K-ras oncogenes during human colon tumorigenesis', *Nature*, 327(6120), pp. 298–303. doi: 10.1038/327298a0.

Friedl, P. and Gilmour, D. (2009) 'Collective cell migration in morphogenesis, regeneration and cancer', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 10, p. 445.

Fu, W. and Baker, N. E. (2003) 'Deciphering synergistic and redundant roles of Hedgehog, Decapentaplegic and Delta that drive the wave of differentiation in Drosophila eye development', *Development*, 130(21), p. 5229 LP-5239. doi: 10.1242/dev.00764.

Fujimoto, A. *et al.* (2016) 'Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer', *Nature Genetics*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 48, p. 500.

Fukata, M. and Kaibuchi, K. (2001) 'Rho-family GTPases in cadherin-mediated cell — cell adhesion', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 2, p. 887.

Gadea, G. *et al.* (2008) 'DOCK10-Mediated Cdc42 Activation Is Necessary for Amoeboid Invasion of Melanoma Cells', *Current Biology*. Elsevier, 18(19), pp. 1456–1465. doi: 10.1016/j.cub.2008.08.053.

Gaggioli, C. *et al.* (2007) 'Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells', *Nature Cell Biology*. Nature Publishing Group, 9, p. 1392.

Gallo, G. (2006) 'RhoA-kinase coordinates F-actin organization and myosin II activity during semaphorin-3A-induced axon retraction', *Journal of cell science*, 119(Pt 16), pp. 3413–3423. doi: 10.1242/jcs.03084.

Garrod, D. R., Merritt, A. J. and Nie, Z. (2002) 'Desmosomal cadherins', *Current Opinion in Cell Biology*, 14(5), pp. 537–545. doi: https://doi.org/10.1016/S0955-0674(02)00366-6.

Gatenby, R. A. and Gillies, R. J. (2004) 'Why do cancers have high aerobic glycolysis?', *Nature Reviews Cancer*. Nature Publishing Group, 4, p. 891.

Gebeshuber, C. A., Sladecek, S. and Grunert, S. (2007) 'β-Catenin/LEF-1 Signalling in Breast Cancer – Central Players Activated by a Plethora of Inputs', *Cells Tissues Organs*, 185(1–3), pp. 51–60. doi: 10.1159/000101303.

Georgiou, M. *et al.* (2008) 'Cdc42, Par6, and aPKC Regulate Arp2/3-Mediated Endocytosis to Control Local Adherens Junction Stability', *Current Biology*. Elsevier, 18(21), pp. 1631–1638. doi: 10.1016/j.cub.2008.09.029.

Georgiou, M. and Baum, B. (2010) 'Polarity proteins and Rho GTPases cooperate to spatially organise epithelial actin-based protrusions', *Journal of Cell Science*, 123(7), p. 1089 LP-1098. doi: 10.1242/jcs.060772.

Gerlich, D. *et al.* (2006) 'Live-Cell Imaging Reveals a Stable Cohesin-Chromatin Interaction after but Not before DNA Replication', *Current Biology*, 16(15), pp. 1571–1578. doi: 10.1016/j.cub.2006.06.068.

Di Giacomo, S. *et al.* (2017) 'MYC, Cell Competition, and Cell Death in Cancer: The Inseparable Triad', *Genes*. MDPI, 8(4), p. 120. doi: 10.3390/genes8040120.

Gibson, M. C. and Perrimon, N. (2005) 'Extrusion and Death of DPP / BMP-Compromised Epithelial Cells in the Developing Drosophila Wing', (March).

Gilles, C. *et al.* (1997) 'Implication of collagen type I-induced membrane-type 1-matrix metalloproteinase expression and matrix metalloproteinase-2 activation in the metastatic progression of breast carcinoma', *Laboratory investigation; a journal of technical methods and pathology*. Department of Cell Biology, Georgetown University Medical Center, Washington D.C. 20007, USA., 76(5), pp. 651–660.

Glynn, E. F. *et al.* (2004) 'Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae', *PLoS Biology*, 2(9). doi: 10.1371/journal.pbio.0020259.

Go, M. J., Eastman, D. S. and Artavanis-Tsakonas, S. (1998) 'Cell proliferation

control by Notch signaling in Drosophila development', *Development*, 125(11), p. 2031 LP-2040. Available at: http://dev.biologists.org/content/125/11/2031.abstract.

Gocheva, V. *et al.* (2010) 'IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion', *Genes & Development* , 24(3), pp. 241–255.

Godoy, A. *et al.* (2006) 'Differential subcellular distribution of glucose transporters GLUT1–6 and GLUT9 in human cancer: Ultrastructural localization of GLUT1 and GLUT5 in breast tumor tissues', *Journal of Cellular Physiology*. John Wiley & Sons, Ltd, 207(3), pp. 614–627. doi: 10.1002/jcp.20606.

Göhrig, A. *et al.* (2014) 'Axon Guidance Factor SLIT2 Inhibits Neural Invasion and Metastasis in Pancreatic Cancer', *Cancer Research*, 74(5), p. 1529 LP-1540. doi: 10.1158/0008-5472.CAN-13-1012.

Goldstein, A. S. *et al.* (2010) 'Identification of a Cell of Origin for Human Prostate Cancer', *Science*, 329(5991), p. 568 LP-571. doi: 10.1126/science.1189992.

van Golen, K. L. *et al.* (2000) 'RhoC GTPase, a Novel Transforming Oncogene for Human Mammary Epithelial Cells That Partially Recapitulates the Inflammatory Breast Cancer Phenotype', *Cancer Research*, 60(20), p. 5832 LP-5838.

Gómez-Cuadrado, L. *et al.* (2017) 'Mouse models of metastasis: progress and prospects', *Disease Models & Mechanisms*, 10(9), p. 1061 LP-1074. doi: 10.1242/dmm.030403.

Gottlieb, T. M. *et al.* (2002) 'Cross-talk between Akt, p53 and Mdm2: possible implications for the regulation of apoptosis', *Oncogene*. Macmillan Publishers Limited, 21, p. 1299.

Grassi, L. *et al.* (2019) 'Organoids as a new model for improving regenerative medicine and cancer personalized therapy in renal diseases', *Cell Death & Disease*, 10(3), p. 201. doi: 10.1038/s41419-019-1453-0.

Gratz, S. J. *et al.* (2013) 'Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease', *Genetics*. Genetics Society of America, 194(4), pp. 1029–1035. doi: 10.1534/genetics.113.152710.

Greaves, M. and Maley, C. C. (2012) 'Clonal evolution in cancer', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 481, p. 306.

Green, D. R. and Reed, J. C. (1998) 'Mitochondria and Apoptosis', *Science*, 281(5381), p. 1309 LP-1312. doi: 10.1126/science.281.5381.1309.

Grinstein, S., Rotin, D. and Mason, M. J. (1989) 'Na+/H+ exchange and growth factor-induced cytosolic pH changes. Role in cellular proliferation', *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 988(1), pp. 73–97. doi: https://doi.org/10.1016/0304-4157(89)90004-X.

Gruber, S., Haering, C. H. and Nasmyth, K. (2003) 'Chromosomal cohesin forms a ring', *Cell*, 112(6), pp. 765–777. doi: 10.1016/S0092-8674(03)00162-4.

Grumolato, L. *et al.* (2013) 'b -Catenin-Independent Activation of TCF1 / LEF1 in Human Hematopoietic Tumor Cells through Interaction with ATF2 Transcription Factors', 9(8). doi: 10.1371/journal.pgen.1003603.

Grunwald, G. B. (1993) 'The structural and functional analysis of cadherin calcium-dependent cell adhesion molecules', *Current Opinion in Cell Biology*,

5(5), pp. 797–805. doi: https://doi.org/10.1016/0955-0674(93)90028-O.

Grzeschik, N. A. *et al.* (2010) 'Lgl, aPKC, and Crumbs Regulate the Salvador/Warts/Hippo Pathway through Two Distinct Mechanisms', *Current Biology*, 20(7), pp. 573–581. doi: https://doi.org/10.1016/j.cub.2010.01.055.

Guilford, P. *et al.* (1998) 'E-cadherin germline mutations in familial gastric cancer', *Nature*. Macmillan Magazines Ltd., 392, p. 402.

Gunasinghe, N. P. A. D. *et al.* (2012) 'Mesenchymal–epithelial transition (MET) as a mechanism for metastatic colonisation in breast cancer', *Cancer and Metastasis Reviews*, 31(3), pp. 469–478. doi: 10.1007/s10555-012-9377-5.

Gurniak, C. B., Perlas, E. and Witke, W. (2005) 'The actin depolymerizing factor n-cofilin is essential for neural tube morphogenesis and neural crest cell migration', *Developmental Biology*, 278(1), pp. 231–241. doi: https://doi.org/10.1016/j.ydbio.2004.11.010.

H. Watson, P. *et al.* (2001) 'Genetic model of multi-step breast carcinogenesis involving the epithelium and stroma: clues to tumour–microenvironment interactions', *Human Molecular Genetics*, 10(18), pp. 1907–1913. doi: 10.1093/hmg/10.18.1907.

Haelterman, N. A. *et al.* (2014) 'Large-scale identification of chemically induced mutations in Drosophila melanogaster', *Genome research*. Cold Spring Harbor Laboratory Press, 24(10), pp. 1707–1718. doi: 10.1101/gr.174615.114.

Hagiwara, A. *et al.* (2012) 'Hepatic mTORC2 Activates Glycolysis and Lipogenesis through Akt, Glucokinase, and SREBP1c', *Cell Metabolism*, 15(5), pp. 725–738. doi: https://doi.org/10.1016/j.cmet.2012.03.015.

Haigo, S. L. *et al.* (2003) 'Shroom Induces Apical Constriction and Is Required for Hingepoint Formation during Neural Tube Closure', *Current Biology*, 13, pp. 2125–2137. doi: 10.1016/j.cub.2003.11.054.

Hajra, K. M., Chen, D. Y.-S. and Fearon, E. R. (2002) 'The SLUG Zinc-Finger Protein Represses E-Cadherin in Breast Cancer', *Cancer Research*, 62(6), p. 1613 LP-1618.

Hall, A. (1998) 'Rho GTPases and the Actin Cytoskeleton', *Science*, 279(5350), p. 509 LP-514. doi: 10.1126/science.279.5350.509.

Hanahan, D. and Weinberg, R. A. (2000) 'The Hallmarks of Cancer', *Cell*. Elsevier, 100(1), pp. 57–70. doi: 10.1016/S0092-8674(00)81683-9.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell*. Elsevier, 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Hartwell, L. H. and Weinert, T. A. (1989) 'Checkpoints: controls that ensure the order of cell cycle events', *Science*, 246(4930), p. 629 LP-634. doi: 10.1126/science.2683079.

Haudenschild, D. R. *et al.* (2004) 'Bone Morphogenetic Protein (BMP)-6 Signaling and BMP Antagonist Noggin in Prostate Cancer', *Cancer Research*, 64(22), p. 8276 LP-8284. doi: 10.1158/0008-5472.CAN-04-2251.

Hauf, S. (2001) 'Cohesin Cleavage by Separase Required for Anaphase and Cytokinesis in Human Cells', *Science*, 293(5533), pp. 1320–1323. doi: 10.1126/science.1061376.

Hay, B. A., Wolff, T. and Rubin, G. M. (1994) 'Expression of baculovirus P35

prevents cell death in Drosophila', 2129, pp. 2121–2129.

Hayflick, L. and Moorhead, P. S. (1961) 'The serial cultivation of human diploid cell strains', *Experimental Cell Research*, 25(3), pp. 585–621. doi: https://doi.org/10.1016/0014-4827(61)90192-6.

Heberlein, U., Wolff, T. and Rubin, G. M. (1993) 'The TGFβ homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the Drosophila retina', *Cell*, 75(5), pp. 913–926. doi: https://doi.org/10.1016/0092-8674(93)90535-X.

Hegerfeldt, Y. *et al.* (2002) 'Collective Cell Movement in Primary Melanoma Explants', *Cancer Research*, 62(7), p. 2125 LP-2130.

Heitzler, P. *et al.* (1996) 'A genetic analysis of pannier, a gene necessary for viability of dorsal tissues and bristle positioning in Drosophila', *Genetics*, 143(3), pp. 1271–1286.

Hildebrand, J. D. (2005) 'Shroom regulates epithelial cell shape via the apical positioning of an actomyosin network', *Journal of Cell Science*, 118(22), p. 5191 LP-5203. doi: 10.1242/jcs.02626.

Hinds, P. W. *et al.* (1994) 'Function of a human cyclin gene as an oncogene.', *Proceedings of the National Academy of Sciences*, 91(2), p. 709 LP-713. doi: 10.1073/pnas.91.2.709.

Hirano, S. *et al.* (1987) 'Calcium-dependent cell-cell adhesion molecules (cadherins): subclass specificities and possible involvement of actin bundles.', *The Journal of Cell Biology*, 105(6), p. 2501 LP-2510. doi: 10.1083/jcb.105.6.2501.

Ho, A. and Dowdy, S. F. (2002) 'Regulation of G1 cell-cycle progression by oncogenes and tumor suppressor genes', *Current Opinion in Genetics & Development*, 12(1), pp. 47–52. doi: https://doi.org/10.1016/S0959-437X(01)00263-5.

Hodges, C. A. *et al.* (2005) 'SMC1B-deficient female mice provide evidence that cohesins are a missing link in age-related nondisjunction', *Nature Genetics*, 37(12), pp. 1351–1355. doi: 10.1038/ng1672.

Hoene, V. *et al.* (2009) 'GATA factors in human neuroblastoma: distinctive expression patterns in clinical subtypes', *British Journal Of Cancer*. The Author(s), 101, p. 1481. Available at: https://doi.org/10.1038/sj.bjc.6605276.

Holz, M. K. *et al.* (2005) 'mTOR and S6K1 Mediate Assembly of the Translation Preinitiation Complex through Dynamic Protein Interchange and Ordered Phosphorylation Events', *Cell*, 123(4), pp. 569–580. doi: https://doi.org/10.1016/j.cell.2005.10.024.

Hood, J. D. and Cheresh, D. A. (2002) 'Role of integrins in cell invasion and migration', *Nature Reviews Cancer*. Nature Publishing Group, 2, p. 91.

Hoque, M. T. and Ishikawa, F. (2002) 'Cohesin defects lead to premature sister chromatid separation, kinetochore dysfunction, and spindle-assembly checkpoint activation', *Journal of Biological Chemistry*, 277(44), pp. 42306–42314. doi: 10.1074/jbc.M206836200.

Horn, C. and Handler, A. M. (2005) 'Site-specific genomic targeting in Drosophila', *Proceedings of the National Academy of Sciences of the United States of America*, 102(35), p. 12483 LP-12488. doi: 10.1073/pnas.0504305102.

Hoskin, P. J. *et al.* (2003) 'GLUT1 and CAIX as intrinsic markers of hypoxia in bladder cancer: relationship with vascularity and proliferation as predictors of outcome of ARCON', *British Journal Of Cancer*. The Author(s), 89, p. 1290.

Hou, F. and Zou, H. (2005) 'Two Human Orthologues of Eco1/Ctf7 Acetyltransferases Are Both Required for Proper Sister-Chromatid Cohesion', *Molecular Biology of the Cell*, 16, pp. 3908–3918.

Hsu, H. *et al.* (1996) 'TRADD–TRAF2 and TRADD–FADD Interactions Define Two Distinct TNF Receptor 1 Signal Transduction Pathways', *Cell*, 84(2), pp. 299–308. doi: https://doi.org/10.1016/S0092-8674(00)80984-8.

Huang, J. *et al.* (2005) 'The Hippo Signaling Pathway Coordinately Regulates Cell Proliferation and Apoptosis by Inactivating Yorkie, the Drosophila Homolog of YAP', *Cell*, 122(3), pp. 421–434. doi: https://doi.org/10.1016/j.cell.2005.06.007.

Huber, M. A., Kraut, N. and Beug, H. (2005) 'Molecular requirements for epithelial–mesenchymal transition during tumor progression', *Current Opinion in Cell Biology*, 17(5), pp. 548–558. doi: https://doi.org/10.1016/j.ceb.2005.08.001.

Humbert, P., Russell, S. and Richardson, H. (2003) 'Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer', *BioEssays*. John Wiley & Sons, Ltd, 25(6), pp. 542–553. doi: 10.1002/bies.10286.

Hurst, R. *et al.* (2004) 'Actin-Related Protein 2/3 Complex Is Required for Actin Ring Formation', *JOURNAL OF BONE AND MINERAL RESEARCH*, 19(3). doi: 10.1359/JBMR.0301238.

Ishiguro, K. I. *et al.* (2011) 'A new meiosis-specific cohesin complex implicated in the cohesin code for homologous pairing', *EMBO Reports*. Nature Publishing Group, 12(3), pp. 267–275. doi: 10.1038/embor.2011.2.

Jaglarz, M. K. and Howard, K. R. (1995) 'The active migration of Drosophila primordial germ cells', *Development*, 121(11), p. 3495 LP-3503.

Jauffred, B. and Bellaiche, Y. (2012) 'Analyzing Frizzled Signaling Using Fixed and Live Imaging of the Asymmetric Cell Division of the Drosophila Sensory Organ Precursor Cell', *Planar Cell Polarity: Methods and Protocols*, 839.

Jeanes, A., Gottardi, C. J. and Yap, A. S. (2008) 'Cadherins and cancer: how does cadherin dysfunction promote tumor progression?', *Oncogene*, 27(55), pp. 6920–6929. doi: 10.1038/onc.2008.343.

Jidigam, V. K. *et al.* (2015) 'Apical constriction and epithelial invagination are regulated by BMP activity', *Biology Open*, 4(12), p. 1782 LP-1791. doi: 10.1242/bio.015263.

Johnson, L. *et al.* (2001) 'Somatic activation of the K-ras oncogene causes early onset lung cancer in mice', *Nature*. Macmillan Magazines Ltd., 410, p. 1111.

Kalkavan, H. and R Green, D. (2017) *MOMP, cell suicide as a BCL-2 family business*, *Cell Death & Differentiation*. doi: 10.1038/cdd.2017.179.

Kamai, T. *et al.* (2004) 'Overexpression of RhoA, Rac1, and Cdc42 GTPases Is Associated with Progression in Testicular Cancer', *Clinical Cancer Research*, 10(14), p. 4799 LP-4805. doi: 10.1158/1078-0432.CCR-0436-03.

Kandoth, C. *et al.* (2013) 'Mutational landscape and significance across 12 major cancer types', *Nature*. Nature Publishing Group, 502(7471), pp. 333–339. doi: 10.1038/nature12634.

Kaplan, R. N. *et al.* (2005) 'VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche', *Nature*, 438(7069), pp. 820–827. doi: 10.1038/nature04186.

Kaplan, R. N., Psaila, B. and Lyden, D. (2006) 'Bone marrow cells in the "pre-metastatic niche": within bone and beyond', *Cancer and Metastasis Reviews*, 25(4), pp. 521–529. doi: 10.1007/s10555-006-9036-9.

Karim, F. D. and Rubin, G. M. (1998) 'Ectopic expression of activated Ras1 induces hyperplastic growth and increased cell death in Drosophila imaginal tissues', *Development*, 125(1), p. 1 LP-9. Available at: http://dev.biologists.org/content/125/1/1.abstract.

Khalil, A. A. and Friedl, P. (2010) 'Determinants of leader cells in collective cell migration', *Integrative Biology*, 2(11–12), pp. 568–574. doi: 10.1039/c0ib00052c.

Kiehart, D. P. *et al.* (2000) 'Multiple Forces Contribute to Cell Sheet Morphogenesis for Dorsal Closure in Drosophila', *The Journal of Cell Biology*, 149(2), p. 471 LP-490. doi: 10.1083/jcb.149.2.471.

Killander, D. and Zetterberg, A. (1965) 'A quantitative cytochemical investigation of the relationship between cell mass and initiation of DNA synthesis in mouse fibroblasts in vitro', *Experimental Cell Research*, 40(1), pp. 12–20. doi: https://doi.org/10.1016/0014-4827(65)90285-5.

Kim, H. Y., Varner, V. D. and Nelson, C. M. (2013) 'Apical constriction initiates new bud formation during monopodial branching of the embryonic chicken lung', *Development*, 140(15), p. 3146 LP-3155. doi: 10.1242/dev.093682.

Kim, J. *et al.* (1998) 'Requirement for Specific Proteases in Cancer Cell Intravasation as Revealed by a Novel Semiquantitative PCR-Based Assay', *Cell*, 94(3), pp. 353–362. doi: https://doi.org/10.1016/S0092-8674(00)81478-6.

Kim, J. *et al.* (2011) 'AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1', *Nature Cell Biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 13, p. 132.

Kim, J. S. *et al.* (2016) 'Intact Cohesion, Anaphase, and Chromosome Segregation in Human Cells Harboring Tumor-Derived Mutations in STAG2', *PLoS Genetics*, 12(2), pp. 1–18. doi: 10.1371/journal.pgen.1005865.

Kim, M. S. *et al.* (2012) 'Mutational and expressional analyses of STAG2 gene in solid cancers', *Neoplasma*, 59. doi: 10.4149/neo_2012_067.

Kim, N. W. *et al.* (1994) 'Specific association of human telomerase activity with immortal cells and cancer', *Science*, 266(5193), p. 2011 LP-2015. doi: 10.1126/science.7605428.

Kim, W. Y. *et al.* (2009) 'HIF2α cooperates with RAS to promote lung tumorigenesis in mice', *The Journal of Clinical Investigation*. The American Society for Clinical Investigation, 119(8), pp. 2160–2170. doi: 10.1172/JCI38443.

Kim, Y. H. *et al.* (2017) 'Senescent tumor cells lead the collective invasion in thyroid cancer', *Nature Communications*. The Author(s), 8, p. 15208.

Kinch, M. S. *et al.* (1995) 'Tyrosine phosphorylation regulates the adhesions of ras-transformed breast epithelia.', *The Journal of Cell Biology*, 130(2), p. 461 LP-471. doi: 10.1083/jcb.130.2.461.

King, M.-C., Marks, J. H. and Mandell, J. B. (2003) 'Breast and Ovarian Cancer

Risks Due to Inherited Mutations in BRCA1 and BRCA2', *Science*, 302(5645), p. 643 LP-646. doi: 10.1126/science.1088759.

Kischkel, F. C. *et al.* (1995) 'Cytotoxicity-dependent APO-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor', *The EMBO journal*, 14(22), pp. 5579–5588.

Kitajima, T. S. *et al.* (2006) 'Shugoshin collaborates with protein phosphatase 2A to protect cohesin', *Nature*, 441(1), pp. 46–52. doi: 10.1038/nature04663.

Klein, F. *et al.* (2005) 'Cohesin's Binding to Chromosomes Depends on a Separate Complex Consisting of Scc2 and Scc4 Proteins', *Mutat Res Fund Mol Mech Mutagen*, 15(March), pp. 399–410. doi: 10.1016/j.ceb.2012.10.013.

Knudson, A. G. (1971) 'Mutation and Cancer: Statistical Study of Retinoblastoma', *Proceedings of the National Academy of Sciences*, 68(4), p. 820 LP-823. doi: 10.1073/pnas.68.4.820.

Knudson, A. G. (1973) 'Mutation and Human Cancer', in Kleln, G., Weinhouse, S., and Haddow, A. B. T.-A. in C. R. (eds). Academic Press, pp. 317–352. doi: https://doi.org/10.1016/S0065-230X(08)60534-5.

Knust, E. and Bossinger, O. (2002) 'Composition and Formation of Intercellular Junctions in Epithelial Cells', *Science*, 298(5600), pp. 1955–1959.

Koch, A. E. (1998) 'Angiogenesis: Implications for rheumatoid arthritis', *Arthritis & Rheumatism*. John Wiley & Sons, Ltd, 41(6), pp. 951–962. doi: 10.1002/1529-0131(199806)41:6<951::AID-ART2>3.0.CO;2-D.

Koff, A. *et al.* (1992) 'Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle', *Science*, 257(5077), p. 1689 LP-1694. doi: 10.1126/science.1388288.

Kojic, A. *et al.* (2018) 'Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization', *Nature Structural and Molecular Biology*, 25(6), pp. 496–504. doi: 10.1038/s41594-018-0070-4.

Kong, X. *et al.* (2014) 'Distinct Functions of Human Cohesin-SA1 and Cohesin-SA2 in Double-Strand Break Repair', *Molecular and Cellular Biology*, 34(4), pp. 685–698. doi: 10.1128/MCB.01503-13.

Korsmeyer, S. J. *et al.* (2000) 'Pro-apoptotic cascade activates BID, which oligomerizes BAK or BAX into pores that result in the release of cytochrome c', *Cell Death And Differentiation*. Macmillan Publishers Limited, 7, p. 1166.

Kramer, R. H. and Nicolson, G. L. (1979) 'Interactions of tumor cells with vascular endothelial cell monolayers: a model for metastatic invasion', *Proceedings of the National Academy of Sciences*, 76(11), p. 5704 LP-5708. doi: 10.1073/pnas.76.11.5704.

Krantz, I. D. *et al.* (2004) 'Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B', *Nature Genetics*, 36(6), pp. 631–635. doi: 10.1038/ng1364.

Krek, W., Xu, G. and Livingston, D. M. (1995) 'Cyclin A-kinase regulation of E2F-1 DNA binding function underlies suppression of an S phase checkpoint', *Cell*, 83(7), pp. 1149–1158. doi: https://doi.org/10.1016/0092-8674(95)90141-8.

Krugmann, S. *et al.* (2001) 'Cdc42 induces filopodia by promoting the formation of an IRSp53:Mena complex', *Current Biology*, 11(21), pp. 1645–1655. doi: https://doi.org/10.1016/S0960-9822(01)00506-1.

Kueng, S. *et al.* (2006) 'Wapl Controls the Dynamic Association of Cohesin with Chromatin', *Cell*, 127(5), pp. 955–967. doi: 10.1016/j.cell.2006.09.040.

Kumar, N. M. and Gilula, N. B. (1996) 'The Gap Junction Communication Channel', *Cell*. Elsevier, 84(3), pp. 381–388. doi: 10.1016/S0092-8674(00)81282-9.

Kumar, R. and Yarmand-Bagheri, R. (2001) 'The role of HER2 in angiogenesis', *Seminars in Oncology*, 28, pp. 27–32. doi: https://doi.org/10.1016/S0093-7754(01)90279-9.

Kunda, P. *et al.* (2008) 'Moesin Controls Cortical Rigidity, Cell Rounding, and Spindle Morphogenesis during Mitosis', *Current Biology*, 18(2), pp. 91–101. doi: https://doi.org/10.1016/j.cub.2007.12.051.

Kuphal, S. *et al.* (2006) 'Expression of Hugl-1 is strongly reduced in malignant melanoma', *Oncogene*, 25(1), pp. 103–110. doi: 10.1038/sj.onc.1209008.

de la Cova, C. *et al.* (2004) 'Drosophila Myc Regulates Organ Size by Inducing Cell Competition', *Cell*. Elsevier, 117(1), pp. 107–116. doi: 10.1016/S0092-8674(04)00214-4.

Lai, S. *et al.* (2008) 'Diaphanous-Related Formin 2 and Profilin I Are Required for Gastrulation Cell Movements', 3(10), pp. 1–16. doi: 10.1371/journal.pone.0003439.

Lambacher, N. J. *et al.* (2015) 'TMEM107 recruits ciliopathy proteins to subdomains of the ciliary transition zone and causes Joubert syndrome', *Nature Cell Biology*. Nature Publishing Group, 18, p. 122.

Landén, N. X., Li, D. and Ståhle, M. (2016) 'Transition from inflammation to proliferation: a critical step during wound healing', *Cellular and molecular life sciences : CMLS*. 2016/05/14. Springer International Publishing, 73(20), pp. 3861–3885. doi: 10.1007/s00018-016-2268-0.

Laprise, P. *et al.* (2009) 'Yurt, Coracle, Neurexin IV and the Na+,K+-ATPase form a novel group of epithelial polarity proteins', *Nature*. Macmillan Publishers Limited. All rights reserved, 459, p. 1141.

Larsen, C. W. *et al.* (2003) 'Segment boundary formation in Drosophila embryos', pp. 5625–5635. doi: 10.1242/dev.00867.

Laugsch, M. *et al.* (2013) 'Imbalance of SMC1 and SMC3 Cohesins Causes Specific and Distinct Effects', *PLoS ONE*.

Lawrence, M. S. *et al.* (2014) 'Discovery and saturation analysis of cancer genes across 21 tumour types', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 505, p. 495.

Lawson, C. D. and Ridley, A. J. (2018) 'Rho GTPase signaling complexes in cell migration and invasion', *The Journal of Cell Biology*, 217(2), p. 447 LP-457. doi: 10.1083/jcb.201612069.

Lee, C.-H. *et al.* (2010) 'Epidermal growth factor receptor regulates beta-catenin location, stability, and transcriptional activity in oral cancer', *Molecular cancer*. BioMed Central, 9, p. 64. doi: 10.1186/1476-4598-9-64.

Lee, T. and Luo, L. (1999) 'Mosaic Analysis with a Repressible Cell Marker for Studies of Gene Function in Neuronal Morphogenesis', *Neuron*. Elsevier, 22(3), pp. 451–461. doi: 10.1016/S0896-6273(00)80701-1.

Lee, T. and Luo, L. (2001) 'Mosaic analysis with a repressible cell marker

(MARCM) for Drosophila neural development', *Trends in Neurosciences*, 24(5), pp. 251–254. doi: https://doi.org/10.1016/S0166-2236(00)01791-4.

Leiserson, M. D. M. *et al.* (2015) 'Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes', *Nature Genetics*, 47(2), pp. 106–114. doi: 10.1038/ng.3168.

Lengronne, A. *et al.* (2004) 'Cohesin relocation from sites of chromosomal loading to places of convergent transcription', *Nature*. Macmillan Magazines Ltd., 430, p. 573.

Leskow, F. C. *et al.* (2006) 'The zebrafish homologue of mammalian chimerin Rac-GAPs is implicated in epiboly progression during development', 103(14), pp. 5373–5378.

Levayer, R., Carole, D. and Eduardo, M. (2016) 'Tissue Crowding Induces Caspase-Dependent Competition for Space', *Current Biology*, 26(5), pp. 670–677.

Levine, B. *et al.* (2010) 'Opposing interactions between Drosophila Cut and the C/EBP encoded by Slow Border Cells direct apical constriction and epithelial invagination', *Developmental Biology*, 344(1), pp. 196–209. doi: https://doi.org/10.1016/j.ydbio.2010.04.030.

Levy, M. Z. *et al.* (1992) 'Telomere end-replication problem and cell aging', *Journal of Molecular Biology*, 225(4), pp. 951–960. doi: https://doi.org/10.1016/0022-2836(92)90096-3.

Li, J. *et al.* (2017) 'AKR1B10 promotes breast cancer cell migration and invasion via activation of ERK signaling', 8(20), pp. 33694–33703.

Li, M. *et al.* (2010) 'The ATM–p53 pathway suppresses aneuploidy-induced tumorigenesis', *Proceedings of the National Academy of Sciences*, 107(32), p. 14188 LP-14193. doi: 10.1073/pnas.1005960107.

Li, X. *et al.* (2015) 'Loss of STAG2 causes aneuploidy in normal human bladder cells', 2(1), pp. 2638–2646.

Lin, C. *et al.* (2016) 'FNDC3B promotes cell migration and tumor metastasis in hepatocellular carcinoma', 7(31).

Litjens, G. *et al.* (2016) 'Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis', *Scientific Reports*. The Author(s), 6, p. 26286.

Little, C. D. *et al.* (1983) 'Amplification and expression of the c-myc oncogene in human lung cancer cell lines', *Nature*, 306(5939), pp. 194–196. doi: 10.1038/306194a0.

Liu, C. *et al.* (2018) 'Hybrid collagen alginate hydrogel as a platform for 3D tumor spheroid invasion', *Acta Biomaterialia*, 75, pp. 213–225. doi: https://doi.org/10.1016/j.actbio.2018.06.003.

Liu, J. *et al.* (2009) 'Analysis of Drosophila Segmentation Network Identifies a JNK Pathway Factor Overexpressed in Kidney Cancer', *Science*, 323(5918), p. 1218 LP-1222. doi: 10.1126/science.1157669.

LIU, J. and LIN, A. (2005) 'Role of JNK activation in apoptosis: A double-edged sword', *Cell Research*. Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 15, p. 36.

Liu, M. and Pile, L. A. (2017) 'The Transcriptional Corepressor SIN3 Directly

Regulates Genes Involved in Methionine Catabolism and Affects Histone Methylation, Linking Epigenetics and Metabolism', *Journal of Biological Chemistry* , 292(5), pp. 1970–1976. Available at: http://www.jbc.org/content/292/5/1970.abstract.

Liu, Y. *et al.* (2017) 'RAD51 Mediates Resistance of Cancer Stem Cells to PARP Inhibition in Triple-Negative Breast Cancer', *Clinical Cancer Research*, 23(2), p. 514 LP-522. doi: 10.1158/1078-0432.CCR-15-1348.

Liu, Z. *et al.* (1996) 'Dissection of TNF Receptor 1 Effector Functions: JNK Activation Is Not Linked to Apoptosis While NF-κB Activation Prevents Cell Death', *Cell*, 87(3), pp. 565–576. doi: https://doi.org/10.1016/S0092-8674(00)81375-6.

Lodish, H. F. (2013) *Molecular cell biology.* New York: W.H. Freeman and Co.

Losada, A. *et al.* (2000) 'Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes', *Journal of Cell Biology*, 150(3), pp. 405–416. doi: 10.1083/jcb.150.3.405.

Losada, A. (2005) 'Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts', *Journal of Cell Science*, 118(10), pp. 2133–2141. doi: 10.1242/jcs.02355.

Losada, A. (2014) 'Cohesin in cancer: chromosome segregation and beyond', *Nature Reviews Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 14, p. 389. Available at: https://doi.org/10.1038/nrc3743.

Lou, X.-L. *et al.* (2015) 'Interaction between circulating cancer cells and platelets: clinical implication', *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu*. AME Publishing Company, 27(5), pp. 450–460. doi: 10.3978/j.issn.1000-9604.2015.04.10.

Lu, X. *et al.* (2009) 'Aberrant Splicing of Hugl-1 Is Associated with Hepatocellular Carcinoma Progression', *Clinical Cancer Research*, 15(10), p. 3287 LP-3296. doi: 10.1158/1078-0432.CCR-08-2078.

Lunt, S. Y. and Vander Heiden, M. G. (2011) 'Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation', *Annual Review of Cell and Developmental Biology*. Annual Reviews, 27(1), pp. 441–464. doi: 10.1146/annurev-cellbio-092910-154237.

Luo, X. *et al.* (2017) 'Emerging roles of lipid metabolism in cancer metastasis'. Molecular Cancer, pp. 1–10. doi: 10.1186/s12943-017-0646-3.

Luzzi, K. J. *et al.* (1998) 'Multistep Nature of Metastatic Inefficiency: Dormancy of Solitary Cells after Successful Extravasation and Limited Survival of Early Micrometastases', *The American Journal of Pathology*, 153(3), pp. 865–873. doi: https://doi.org/10.1016/S0002-9440(10)65628-3.

Ly, D. L. and Lumsden, C. J. (2009) '3D amoeboid migration of a eukaryotic cell in a fiber matrix', *Artificial Life and Robotics*, 14(1), pp. 1–6. doi: 10.1007/s10015-008-0708-0.

Ma, J. *et al.* (2012) 'AKR1B10 overexpression in breast cancer : Association with tumor size , lymph node metastasis and patient survival and its', 871, pp. 862–871. doi: 10.1002/ijc.27618.

Maacke, H. *et al.* (2000) 'Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer', *International Journal of Cancer*. John Wiley & Sons, Ltd, 88(6), pp. 907–913. doi: 10.1002/1097-

0215(20001215)88:6<907::AID-IJC11>3.0.CO;2-4.

Malumbres, M. and Barbacid, M. (2003) 'RAS oncogenes: the first 30 years', *Nature Reviews Cancer*. Nature Publishing Group, 3, p. 459.

Mani, S. A. *et al.* (2008) 'The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells', *Cell*, 133(4), pp. 704–715. doi: https://doi.org/10.1016/j.cell.2008.03.027.

Margetts, P. J. *et al.* (2002) 'Inflammatory Cytokines, Angiogenesis, and Fibrosis in the Rat Peritoneum', *The American Journal of Pathology*, 160(6), pp. 2285–2294. doi: https://doi.org/10.1016/S0002-9440(10)61176-5.

Margolis, B. and Borg, J.-P. (2005) 'Apicobasal polarity complexes', *Journal of Cell Science*, 118(22), p. 5157 LP-5159. doi: 10.1242/jcs.02597.

Marinari, E. *et al.* (2012) 'Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding', *Nature*. Nature Publishing Group, 484(7395), pp. 542–545. doi: 10.1038/nature10984.

Martin-Belmonte, F. and Perez-Moreno, M. (2011) 'Epithelial cell polarity, stem cells and cancer', *Nature Reviews Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 12, p. 23. Available at: https://doi.org/10.1038/nrc3169.

Martín-Castellanos, C. and Edgar, B. A. (2002) 'A characterization of the effects of Dpp signaling on cell growth and proliferation in the Drosophila wing', *Development*, 129(4), p. 1003 LP-1013.

Martin, A. C. and Goldstein, B. (2014) 'Apical constriction: themes and variations on a cellular mechanism driving morphogenesis', *Development (Cambridge, England)*. The Company of Biologists, 141(10), pp. 1987–1998. doi: 10.1242/dev.102228.

Martin, A. C., Kaschube, M. and Wieschaus, E. F. (2008) 'Pulsed contractions of an actin–myosin network drive apical constriction', *Nature*. Macmillan Publishers Limited. All rights reserved, 457, p. 495. Available at: https://doi.org/10.1038/nature07522.

Martínez-Morentin, L. *et al.* (2015) 'Cardiac deficiency of single cytochrome oxidase assembly factor scox induces p53-dependent apoptosis in a Drosophila cardiomyopathy model', *Human Molecular Genetics*, 24(13), pp. 3608–3622. doi: 10.1093/hmg/ddv106.

Massimi, P. *et al.* (2008) 'Regulation of the hDlg/hScrib/Hugl-1 tumour suppressor complex', *Experimental Cell Research*, 314(18), pp. 3306–3317. doi: https://doi.org/10.1016/j.yexcr.2008.08.016.

Matter, K. and Mellman, I. (1994) 'Mechamisms of cell polarity: sorting and transport in epithelial cells', *Current Opinion in Cell Biology*, 6(4), pp. 545–554. doi: https://doi.org/10.1016/0955-0674(94)90075-2.

McNerney, M. E. *et al.* (2013) 'CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia', *Blood*, 121(6), p. 975 LP-983. doi: 10.1182/blood-2012-04-426965.

Mechler, B. M., McGinnis, W. and Gehring, W. J. (1985) 'Molecular cloning of lethal(2)giant larvae, a recessive oncogene of Drosophila melanogaster.', *The EMBO Journal*. John Wiley & Sons, Ltd, 4(6), pp. 1551–1557. doi: 10.1002/j.1460-2075.1985.tb03816.x.

Medici, D., Hay, E. D. and Goodenough, D. A. (2006) 'Cooperation between Snail and LEF-1 Transcription Factors Is Essential for TGF-β1-induced Epithelial-Mesenchymal Transition', *Molecular Biology of the Cell*. American Society for Cell Biology (mboc), 17(4), pp. 1871–1879. doi: 10.1091/mbc.e05-08-0767.

Medici, D. and Nawshad, A. (2010) 'Type I collagen promotes epithelial–mesenchymal transition through ILK-dependent activation of NF-κB and LEF-1', *Matrix Biology*, 29(3), pp. 161–165. doi: https://doi.org/10.1016/j.matbio.2009.12.003.

Médina, E. *et al.* (2002) 'Role of the Crumbs complex in the regulation of junction formation in Drosophila and mammalian epithelial cells', *Biology of the Cell*. John Wiley & Sons, Ltd (10.1111), 94(6), pp. 305–313. doi: 10.1016/S0248-4900(02)00004-7.

Mège, R.-M., Gavard, J. and Lambert, M. (2006) 'Regulation of cell–cell junctions by the cytoskeleton', *Current Opinion in Cell Biology*, 18(5), pp. 541–548. doi: https://doi.org/10.1016/j.ceb.2006.08.004.

Menéndez, J. *et al.* (2010) 'A tumor-suppressing mechanism in Drosophila involving cell competition and the Hippo pathway', *Proceedings of the National Academy of Sciences*, 107(33), p. 14651 LP-14656. doi: 10.1073/pnas.1009376107.

Mi, S. *et al.* (2019) 'Three-dimensional microfluidic tumor–macrophage system for breast cancer cell invasion', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 0(0). doi: 10.1002/bit.26961.

Michaelis, C., Ciosk, R. and Nasmyth, K. (1997) 'Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids', *Cell*, 91(1), pp. 35–45. doi: 10.1016/S0092-8674(01)80007-6.

Mittal, S. *et al.* (2009) 'Cooperation of Notch and Ras/MAPK signaling pathways in human breast carcinogenesis', *Molecular Cancer*, 8(1), p. 128. doi: 10.1186/1476-4598-8-128.

Miyake, N. *et al.* (2008) 'Human CHN1 Mutations Hyperactivate α2-Chimaerin and Cause Duane's Retraction Syndrome', *Science*, 321(5890), p. 839 LP-843. doi: 10.1126/science.1156121.

Moohr, O. and Bridges, C. (1919) 'The Inheritance of the Mutant Character "Vortex"', *Genetics*, (4:252.).

Moon, R. T. *et al.* (2004) 'WNT and β-catenin signalling: diseases and therapies', *Nature Reviews Genetics*. Nature Publishing Group, 5, p. 691.

Moreno, E., Basler, K. and Morata, G. (2002) 'Cells compete for Decapentaplegic survival factor to prevent apoptosis in Drosophila wing development', *Nature*. Nature Publishing Group, 416, p. 755.

Mouri, K., Horiuchi, S. and Uemura, T. (2012) 'Cohesin controls planar cell polarity by regulating the level of the seven-pass transmembrane cadherin Flamingo', *Genes to Cells*. John Wiley & Sons, Ltd (10.1111), 17(6), pp. 509–524. doi: 10.1111/j.1365-2443.2012.01604.x.

Mulcahy, L. S., Smith, M. R. and Stacey, D. W. (1985) 'Requirement for ras proto-oncogene function during serum-stimulated growth of NIH 3T3 cells', *Nature*, 313(5999), pp. 241–243. doi: 10.1038/313241a0.

Müller, M. *et al.* (1998) 'p53 Activates the CD95 (APO-1/Fas) Gene in Response to DNA Damage by Anticancer Drugs', *The Journal of Experimental Medicine*, 188(11), p. 2033 LP-2045. doi: 10.1084/jem.188.11.2033.

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

Munjal, A. *et al.* (2015) 'A self-organized biomechanical network drives shape changes during tissue morphogenesis'. doi: 10.1038/nature14603.

Murai, J. *et al.* (2012) 'Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors', *Cancer Research*, 72(21), p. 5588 LP-5599. doi: 10.1158/0008-5472.CAN-12-2753.

Muthuswamy, S. K. (2018) 'Organoid Models of Cancer Explode with Possibilities', *Cell Stem Cell*, 22(3), pp. 290–291. doi: https://doi.org/10.1016/j.stem.2018.02.010.

Naber, H. P. H. *et al.* (2013) 'Snail and Slug, key regulators of TGF-β-induced EMT, are sufficient for the induction of single-cell invasion', *Biochemical and Biophysical Research Communications*, 435(1), pp. 58–63. doi: https://doi.org/10.1016/j.bbrc.2013.04.037.

Nabeshima, K. *et al.* (1999) 'Cohort migration of carcinoma cells: differentiated colorectal carcinoma cells move as coherent cell clusters or sheets.', *Histology and histopathology*. Spain, 14(4), pp. 1183–1197. doi: 10.14670/HH-14.1183.

Neisch, A. L. *et al.* (2010) 'Rho1 regulates apoptosis via activation of the JNK signaling pathway at the plasma membrane', *THE JOURNAL OF CELL BIOLOGY*, 189(2), pp. 311–323. doi: 10.1083/jcb.200912010.

Nelson, W. J. (2009) 'Remodeling Epithelial Cell Organization: Transitions Between Front–Rear and Apical–Basal Polarity', *Cold Spring Harbor Perspectives in Biology* , 1(1). Available at: http://cshperspectives.cshlp.org/content/1/1/a000513.abstract.

Niessen, C. M. (2007) 'Tight Junctions/Adherens Junctions: Basic Structure and Function', *Journal of Investigative Dermatology*, 127(11), pp. 2525–2532. doi: https://doi.org/10.1038/sj.jid.5700865.

Nishimura, T. and Takeichi, M. (2008) 'Shroom3-mediated recruitment of Rho kinases to the apical cell junctions regulates epithelial and neuroepithelial planar remodeling', 1502, pp. 1493–1502. doi: 10.1242/dev.019646.

Nör, J. E. *et al.* (1999) 'Vascular Endothelial Growth Factor (VEGF)-Mediated Angiogenesis Is Associated with Enhanced Endothelial Cell Survival and Induction of Bcl-2 Expression', *The American Journal of Pathology*, 154(2), pp. 375–384. doi: https://doi.org/10.1016/S0002-9440(10)65284-4.

O'Brien, C. A. *et al.* (2006) 'A human colon cancer cell capable of initiating tumour growth in immunodeficient mice', *Nature*. Nature Publishing Group, 445, p. 106.

Oda, H. and Tsukita, S. (2003) 'Real-time imaging of cell-cell adherens junctions reveals that Drosophila mesoderm invagination begins with two phases of apical constriction of cells', pp. 493–501.

Okamura, K. *et al.* (2008) 'The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs', *Nature*. Nature Publishing Group, 453, p. 803.

Okarmoto, T. *et al.* (1989) 'Multi-step Carcinogenesis Model for Adult T-cell Leukemia', *Japanese Journal of Cancer Research*. John Wiley & Sons, Ltd (10.1111), 80(3), pp. 191–195. doi: 10.1111/j.1349-7006.1989.tb02289.x.

Organista, M. F. and De Celis, J. F. (2013) 'The Spalt transcription factors regulate cell proliferation, survival and epithelial integrity downstream of the Decapentaplegic signalling pathway', *Biology Open*, 2(1), p. 37 LP-48. doi: 10.1242/bio.20123038.

Ozdamar, B. *et al.* (2005) 'Regulation of the Polarity Protein Par6 by TGFß Receptors Controls Epithelial Cell Plasticity', *Science*, 307(5715), p. 1603 LP-1609. doi: 10.1126/science.1105718.

Padash Barmchi, M., Rogers, S. and Häcker, U. (2005) 'DRhoGEF2 regulates actin organization and contractility in the Drosophila blastoderm embryo', *The Journal of Cell Biology*, 168(4), p. 575 LP-585. doi: 10.1083/jcb.200407124.

Palucka, K. and Banchereau, J. (2013) 'Dendritic-Cell-Based Therapeutic Cancer Vaccines', *Immunity*, 39(1), pp. 38–48. doi: https://doi.org/10.1016/j.immuni.2013.07.004.

Parelho, V. *et al.* (2008) 'Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms', *Cell*, 132(3), pp. 422–433. doi: 10.1016/j.cell.2008.01.011.

Pasierbek, P. *et al.* (2001) 'A Caenorhabditis elegans cohesion protein with functions in meiotic chromosome pairing and disjunction', *Genes & Development*, pp. 1349–1360. doi: 10.1101/gad.192701.linking.

Patel, M. S., Shah, H. S. and Shrivastava, N. (2017) 'c-Myc-Dependent Cell Competition in Human Cancer Cells', *Journal of Cellular Biochemistry*. John Wiley & Sons, Ltd, 118(7), pp. 1782–1791. doi: 10.1002/jcb.25846.

Pećina-Slaus, N. (2003) 'Tumor suppressor gene E-cadherin and its role in normal and malignant cells', *Cancer cell international*. BioMed Central, 3(1), p. 17. doi: 10.1186/1475-2867-3-17.

Peglion, F. and Etienne-Manneville, S. (2013) 'p120 catenin alteration in cancer and its role in tumour invasion', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 368(1629), p. 20130015. doi: 10.1098/rstb.2013.0015.

Peifer, M. *et al.* (1992) 'The vertebrate adhesive junction proteins beta-catenin and plakoglobin and the Drosophila segment polarity gene armadillo form a multigene family with similar properties', *The Journal of cell biology*. The Rockefeller University Press, 118(3), pp. 681–691.

Peinado, H. *et al.* (2017) 'Pre-metastatic niches: organ-specific homes for metastases', *Nature Reviews Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 17, p. 302.

Peinado, H., Olmeda, D. and Cano, A. (2007) 'Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?', *Nature Reviews Cancer*. Nature Publishing Group, 7, p. 415.

Peter, M. and Herskowitz, I. (1994) 'Joining the complex: Cyclin-dependent kinase inhibitory proteins and the cell cycle', *Cell*. Elsevier, 79(2), pp. 181–184. doi: 10.1016/0092-8674(94)90186-4.

Plageman, T. F. *et al.* (2010) 'Pax6-dependent Shroom3 expression regulates apical constriction during lens placode invagination', 415, pp. 405–415. doi: 10.1242/dev.045369.

Polager, S. and Ginsberg, D. (2009) 'p53 and E2f: partners in life and death', *Nature Reviews Cancer*. Nature Publishing Group, 9, p. 738.

Pollard, J. W. (2004) 'Tumour-educated macrophages promote tumour progression and metastasis', *Nature Reviews Cancer*. Nature Publishing Group, 4, p. 71.

Poste, G. and Fidler, I. J. (1980) 'The pathogenesis of cancer metastasis', *Nature*, 283(5743), pp. 139–146. doi: 10.1038/283139a0.

Potter, C. J., Huang, H. and Xu, T. (2001) 'Drosophila Tsc1 Functions with Tsc2 to Antagonize Insulin Signaling in Regulating Cell Growth, Cell Proliferation, and Organ Size', *Cell*, 105(3), pp. 357–368. doi: https://doi.org/10.1016/S0092-8674(01)00333-6.

Potter, C. J., Turenchalk, G. S. and Xu, T. (2000) 'Drosophila in cancer research: an expanding role', *Trends in Genetics*. Elsevier, 16(1), pp. 33–39. doi: 10.1016/S0168-9525(99)01878-8.

Powell, S. N. and Kachnic, L. A. (2003) 'Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation', *Oncogene*. Nature Publishing Group, 22, p. 5784. Available at: https://doi.org/10.1038/sj.onc.1206678.

Pratap Singh, V. and Gerton, J. (2015) *Cohesin and human disease: Lessons from mouse models*, *Current opinion in cell biology*. doi: 10.1016/j.ceb.2015.08.003.

Prescott, D. M. (1956) 'Relation between cell growth and cell division III. Changes in nuclear volume and growth rate and prevention of cell division in Amoeba proteus resulting from cytoplasmic amputations', *Experimental Cell Research*, 11(1), pp. 94–98. doi: https://doi.org/10.1016/0014-4827(56)90193-8.

Prieto, I. *et al.* (2001) 'Mammalian STAG3 is a cohesin specific to sister chromatid arms in meiosis I', *Nature Cell Biology*, 3, pp. 761–766.

Prin, F. *et al.* (2014) 'Hox proteins drive cell segregation and non-autonomous apical remodelling during hindbrain segmentation', 20, pp. 1492–1502. doi: 10.1242/dev.098954.

Radisky, E. S. and Radisky, D. C. (2010) 'Matrix Metalloproteinase-Induced Epithelial-Mesenchymal Transition in Breast Cancer', *Journal of Mammary Gland Biology and Neoplasia*, 15(2), pp. 201–212. doi: 10.1007/s10911-010-9177-x.

Rajagopal, R. *et al.* (2009) 'The type I BMP receptors, Bmpr1a and Acvr1, activate multiple signaling pathways to regulate lens formation', *Dev Biol.*, 335(2), pp. 305–316. doi: 10.1016/j.ydbio.2009.08.027.The.

Rak, J. *et al.* (1995) 'Mutant ras Oncogenes Upregulate VEGF/VPF Expression: Implications for Induction and Inhibition of Tumor Angiogenesis', *Cancer Research*, 55(20), p. 4575 LP-4580.

Ramdzan, Z. M. and Nepveu, A. (2014) 'CUX1, a haploinsufficient tumour suppressor gene overexpressed in advanced cancers', *Nature Reviews Cancer*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 14, p. 673. Available at: https://doi.org/10.1038/nrc3805.

Rao, S. S. P. *et al.* (2017) 'Cohesin Loss Eliminates All Loop Domains', *Cell*, 171(2), p. 305–320.e24. doi: 10.1016/j.cell.2017.09.026.

Rauzi, M. *et al.* (2008) 'Nature and anisotropy of cortical forces orienting Drosophila tissue morphogenesis', *Nature Cell Biology*. Nature Publishing Group, 10, p. 1401.

Ravi, R. *et al.* (2000) 'Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1α', *Genes & Development* , 14(1), pp. 34–44.

Rees, R. (2014) *Tumor Immunology and Immunotherapy*.

Repo, H. *et al.* (2016) 'The Expression of Cohesin Subunit SA2 Predicts Breast Cancer Survival', *Applied Immunohistochemistry & Molecular Morphology*, 24(9), pp. 615–621.

Revenkova, E. *et al.* (2004) 'Cohesin SMC1β is required for meiotic chromosome dynamics, sister chromatid cohesion and DNA recombination', *Nature Cell Biology*, 6(6), pp. 555–562. doi: 10.1038/ncb1135.

Ridley, A. J. *et al.* (1992) 'The small GTP-binding protein rac regulates growth factor-induced membrane ruffling', *Cell*, 70(3), pp. 401–410. doi: https://doi.org/10.1016/0092-8674(92)90164-8.

Robertson, H. M. *et al.* (1988) 'A stable genomic source of P element transposase in Drosophila melanogaster', *Genetics*, 118(3), pp. 461–470.

Rocquain, J. *et al.* (2010) 'Alteration of cohesin genes in myeloid diseases', *American Journal of Hematology*. John Wiley & Sons, Ltd, 85(9), pp. 717–719. doi: 10.1002/ajh.21798.

Rodrigues, M., Griffith, L. G. and Wells, A. (2010) 'Growth factor regulation of proliferation and survival of multipotential stromal cells', *Stem Cell Research & Therapy*, 1(4), p. 32. doi: 10.1186/scrt32.

Rodriguez-Boulan, E. and Nelson, W. J. (1989) 'Morphogenesis of the polarized epithelial cell phenotype', *Science*, 245(4919), p. 718 LP-725. doi: 10.1126/science.2672330.

Rollins, R. A. *et al.* (2004) 'Drosophila Nipped-B Protein Supports Sister Chromatid Cohesion and Opposes the Stromalin / Scc3 Cohesion Factor To Facilitate Long-Range Activation of the cut Gene', 24(8), pp. 3100–3111. doi: 10.1128/MCB.24.8.3100.

Romero-Pérez, L. *et al.* (2013) 'ZEB1 overexpression associated with E-cadherin and microRNA-200 downregulation is characteristic of undifferentiated endometrial carcinoma', *Modern Pathology*. The United States & Canadian Academy of Pathology, 26, p. 1514. Available at: https://doi.org/10.1038/modpathol.2013.93.

Rubin, G. M. and Spradling, A. C. (1983) 'Vectors for P element-mediated gene transfer in Drosophila', *Nucleic acids research*, 11(18), pp. 6341–6351.

Rubio, E. D. *et al.* (2008) 'CTCF physically links cohesin to chromatin', *Proceedings of the National Academy of Sciences*, 105(24), pp. 8309–8314. doi: 10.1073/pnas.0801273105.

Rundhaug, J. E. (2005) 'Matrix metalloproteinases and angiogenesis', *Journal of Cellular and Molecular Medicine*. John Wiley & Sons, Ltd (10.1111), 9(2), pp. 267–285. doi: 10.1111/j.1582-4934.2005.tb00355.x.

Sabeh, F., Shimizu-Hirota, R. and Weiss, S. J. (2009) 'Protease-dependent versus -independent cancer cell invasion programs: three-dimensional amoeboid movement revisited', *The Journal of Cell Biology*, 185(1), p. 11 LP-19. doi: 10.1083/jcb.200807195.

Sansregret, L. and Nepveu, A. (2008) 'The multiple roles of CUX1: Insights from mouse models and cell-based assays', *Gene*, 412(1), pp. 84–94. doi:

https://doi.org/10.1016/j.gene.2008.01.017.

Santiago, L. *et al.* (2017) 'Wnt signaling pathway protein LEF1 in cancer, as a biomarker for prognosis and a target for treatment', *American journal of cancer research*. e-Century Publishing Corporation, 7(6), pp. 1389–1406.

Sarov, M. *et al.* (2016) 'A genome-wide resource for the analysis of protein localisation in Drosophila', *ELife*.

Sauter, E. R. *et al.* (2002) 'Cyclin D1 Is a Candidate Oncogene in Cutaneous Melanoma', *Cancer Research*, 62(11), p. 3200 LP-3206.

Schalbetter, S. A. *et al.* (2017) 'SMC complexes differentially compact mitotic chromosomes according to genomic context', *Nature Cell Biology*, 19(9), pp. 1071–1080. doi: 10.1038/ncb3594.

Schimanski, C. C. *et al.* (2005) 'Reduced expression of Hugl-1, the human homologue of Drosophila tumour suppressor gene lgl, contributes to progression of colorectal cancer', *Oncogene*, 24(19), pp. 3100–3109. doi: 10.1038/sj.onc.1208520.

Schleiffer, A. *et al.* (2003) 'Kleisins: A Superfamily of Bacterial and Eukaryotic SMC Protein Partners', *Molecular Cell*, 11(3), pp. 571–575. doi: https://doi.org/10.1016/S1097-2765(03)00108-4.

Schlichting, K. and Dahmann, C. (2008) 'Hedgehog and Dpp signaling induce cadherin Cad86C expression in the morphogenetic furrow during Drosophila eye development', *Mechanisms of Development*, 125(8), pp. 712–728. doi: https://doi.org/10.1016/j.mod.2008.04.005.

Schneider, S. *et al.* (1996) 'β-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos', *Mechanisms of Development*, 57(2), pp. 191–198. doi: https://doi.org/10.1016/0925-4773(96)00546-1.

Schwarzer, W. *et al.* (2017) 'Two independent modes of chromatin organization revealed by cohesin removal', *Nature*. Macmillan Publishers Limited, part of Springer Nature. All rights reserved., 551, p. 51.

Serrano, I. *et al.* (2012) 'Role of the integrin-linked kinase (ILK)/Rictor complex in TGFβ-1-induced epithelial–mesenchymal transition (EMT)', *Oncogene*. Macmillan Publishers Limited, 32, p. 50.

Seyfried, T. N. and Huysentruyt, L. C. (2013) 'On the origin of cancer metastasis', *Critical reviews in oncogenesis*, 18(1–2), pp. 43–73.

Shen, C.-H. *et al.* (2016) 'Loss of cohesin complex components STAG2 or STAG3 confers resistance to BRAF inhibition in melanoma', *Nat Med*, 22(9), pp. 1056–1061.

Shi, Y. *et al.* (2015) 'Decreased expression and prognostic role of EHD2 in human breast carcinoma: correlation with E-cadherin', *Journal of Molecular Histology*, 46(2), pp. 221–231.

Shindo, A. and Wallingford, J. B. (2014) 'PCP and Septins Compartmentalize Cortical Actomyosin to Direct Collective Cell Movement', *Science*, 343(6171), p. 649 LP-652. doi: 10.1126/science.1243126.

Silicheva, M. *et al.* (2010) 'Drosophila mini-white model system: new insights into positive position effects and the role of transcriptional terminators and gypsy insulator in transgene shielding', *Nucleic acids research*. 2009/10/23. Oxford University Press, 38(1), pp. 39–47. doi: 10.1093/nar/gkp877.

Smyth, M. J. *et al.* (2000) 'Perforin-Mediated Cytotoxicity Is Critical for Surveillance of Spontaneous Lymphoma', *The Journal of Experimental Medicine*, 192(5), p. 755 LP-760. doi: 10.1084/jem.192.5.755.

Snow, P. M., Bieber, A. J. and Goodman, C. S. (1989) 'Fasciclin III: A novel homophilic adhesion molecule in Drosophila', *Cell*, 59(2), pp. 313–323.

Soardi, F. C. *et al.* (2017) 'Familial STAG2 germline mutation defines a new human cohesinopathy', *npj Genomic Medicine*, 2(1), p. 7. doi: 10.1038/s41525-017-0009-4.

Solomon, D. A. *et al.* (2011) 'Mutational Inactivation of STAG2 Causes Aneuploidy in Human Cancer', *Science*, 333(6045), p. 1039 LP-1043. doi: 10.1126/science.1203619.

Solomon, D. A. *et al.* (2013) 'Frequent truncating mutations of STAG2 in bladder cancer', *Nature Genetics*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 45, p. 1428. Available at: https://doi.org/10.1038/ng.2800.

Song, J. *et al.* (2013) 'Hugl-1 induces apoptosis in esophageal carcinoma cells both in vitro and in vivo', *World journal of gastroenterology*. 2013/07/14. Baishideng Publishing Group Co., Limited, 19(26), pp. 4127–4136. doi: 10.3748/wjg.v19.i26.4127.

Sonoda, E. *et al.* (2001) 'Scc1/Rad21/Mcd1 Is Required for Sister Chromatid Cohesion and Kinetochore Function in Vertebrate Cells', *Developmental Cell*, 1(6), pp. 759–770. doi: 10.1016/S1534-5807(01)00088-0.

Stampfel, G. *et al.* (2015) 'Transcriptional regulators form diverse groups with context-dependent regulatory functions', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 528, p. 147. Available at: https://doi.org/10.1038/nature15545.

Stecca, B., Mas, C. and Altaba, A. R. i (2005) 'Interference with HH–GLI signaling inhibits prostate cancer', *Trends in Molecular Medicine*, 11(5), pp. 199–203. doi: https://doi.org/10.1016/j.molmed.2005.03.004.

Stewart, N. *et al.* (1995) 'Evidence for a second cell cycle block at G2/M by p53.', *Oncogene*, 10, pp. 109–115.

Street, S. E. A. *et al.* (2002) 'Suppression of Lymphoma and Epithelial Malignancies Effected by Interferon γ', *The Journal of Experimental Medicine*, 196(1), p. 129 LP-134. doi: 10.1084/jem.20020063.

Sweeton, D. *et al.* (1991) 'Gastrulation in Drosophila : the formation of the ventral furrow and posterior midgut invaginations', 789, pp. 775–789.

Talkenberger, K. *et al.* (2017) 'Amoeboid-mesenchymal migration plasticity promotes invasion only in complex heterogeneous microenvironments', *Scientific Reports*, 7(1), p. 9237. doi: 10.1038/s41598-017-09300-3.

Tamura, M. *et al.* (1999) 'Tumor Suppressor PTEN Inhibition of Cell Invasion, Migration, and Growth: Differential Involvement of Focal Adhesion Kinase and p130^{Cas}', *Cancer Research*, 59(2), p. 442 LP-449.

Tellez, G. F. *et al.* (2016) 'Protocadherin-1 Localization and Cell- Adhesion Function in Airway Epithelial Cells in Asthma', pp. 1–21. doi: 10.1371/journal.pone.0163967.

Teng, Y. *et al.* (2016) 'WASF3 provides the conduit to facilitate invasion and metastasis in breast cancer cells through HER2/HER3 signaling', *Oncogene*.

2016/01/25, 35(35), pp. 4633–4640. doi: 10.1038/onc.2015.527.

Tepass, U. *et al.* (1996) 'shotgun encodes Drosophila E-cadherin and is preferentially required during cell rearrangement in the neurectoderm and other morphogenetically active epithelia.', *Genes & Development* , 10(6), pp. 672–685.

Thayer, S. P. *et al.* (2003) 'Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis', *Nature*. Macmillan Magazines Ltd., 425, p. 851. Available at: https://doi.org/10.1038/nature02009.

Theodosiou, N. A. and Xu, T. (1998) 'Use of FLP/FRT System to StudyDrosophilaDevelopment', *Methods*, 14(4), pp. 355–365. doi: https://doi.org/10.1006/meth.1998.0591.

Theveneau, E. and Linker, C. (2017) 'Leaders in collective migration: are front cells really endowed with a particular set of skills?', *F1000Research*. F1000Research, 6, p. 1899. doi: 10.12688/f1000research.11889.1.

Thompson, S. L. and Compton, D. A. (2010) 'Proliferation of aneuploid human cells is limited by a p53-dependent mechanism', *The Journal of Cell Biology*, 188(3), p. 369 LP-381. doi: 10.1083/jcb.200905057.

Tie, J. *et al.* (2010) 'MiR-218 Inhibits Invasion and Metastasis of Gastric Cancer by Targeting the Robo1 Receptor', *PLOS Genetics*. Public Library of Science, 6(3), p. e1000879.

To, T.-L. *et al.* (2015) 'Rationally designed fluorogenic protease reporter visualizes spatiotemporal dynamics of apoptosis in vivo', *Proceedings of the National Academy of Sciences*, 112(11), pp. 3338–3343. doi: 10.1073/pnas.1502857112.

Tolde, O. *et al.* (2018) 'Quantitative phase imaging unravels new insight into dynamics of mesenchymal and amoeboid cancer cell invasion', *Scientific Reports*, 8(1), p. 12020. doi: 10.1038/s41598-018-30408-7.

Trumpp, A. *et al.* (2001) 'c-Myc regulates mammalian body size by controlling cell number but not cell size', *Nature*. Macmillan Magazines Ltd., 414, p. 768.

Twu, Y.-C., Gold, M. R. and Teh, H.-S. (2011) 'TNFR1 delivers pro-survival signals that are required for limiting TNFR2-dependent activation-induced cell death (AICD) in CD8+ T cells', *European Journal of Immunology*. John Wiley & Sons, Ltd, 41(2), pp. 335–344. doi: 10.1002/eji.201040639.

Udan, R. S. *et al.* (2003) 'Hippo promotes proliferation arrest and apoptosis in the Salvador/Warts pathway', *Nature Cell Biology*, 5(10), pp. 914–920. doi: 10.1038/ncb1050.

Umbas, R. *et al.* (1994) 'Decreased E-Cadherin Expression Is Associated with Poor Prognosis in Patients with Prostate Cancer', *Cancer Research*, 54(14), p. 3929 LP-3933.

Unal, E. *et al.* (2004) 'DNA Damage Response Pathway Uses Histone Modification to Assemble a Double-Strand Break-Specific Cohesin Domain CHK1/Chk1p and CHK2/Rad53p (reviewed in Jackson [2002] and Shiloh [2003]). Substrates of these kinases include DNA repair-related proteins such', *Molecular Cell*, 16, pp. 991–1002. doi: 10.1016/j.molcel.2004.11.027.

Vass, S. *et al.* (2003) 'Depletion of Drad21/Scc1 in Drosophila cells leads to instability of the cohesin complex and disruption of mitotic progression', *Current Biology*, 13(3), pp. 208–218. doi: 10.1016/S0960-9822(03)00047-2.

Vega, H. *et al.* (2005) 'Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion', *Nature Genetics*, 37(5), pp. 468–470. doi: 10.1038/ng1548.

Venugopala Reddy, G. *et al.* (1999) 'Irregular chiasm-C-roughest, a member of the immunoglobulin superfamily, affects sense organ spacing on the Drosophila antenna by influencing the positioning of founder cells on the disc ectoderm', *Development Genes and Evolution*, 209(10), pp. 581–591. doi: 10.1007/s004270050292.

Verdun, R. E. and Karlseder, J. (2007) 'Replication and protection of telomeres', *Nature*. Nature Publishing Group, 447, p. 924.

Vesuna, F. *et al.* (2008) 'Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer', *Biochemical and biophysical research communications*. 2007/12/04, 367(2), pp. 235–241. doi: 10.1016/j.bbrc.2007.11.151.

Vigil, D. *et al.* (2013) 'ROCK1 and ROCK2 are Required for Non-Small Cell Lung Cancer Anchorage-Independent Growth and Invasion', *Cancer Res*, 72(20), pp. 1–20. doi: 10.1158/0008-5472.CAN-11-2373.ROCK1.

Vincent, T. L. and Gatenby, R. A. (2008) 'An evolutionary model for initiation, promotion, and progression in carcinogenesis', *Journal of Oncology*, pp. 729–737.

de Visser, K. E., Eichten, A. and Coussens, L. M. (2006) 'Paradoxical roles of the immune system during cancer development', *Nature Reviews Cancer*. Nature Publishing Group, 6, p. 24.

Vogelstein, B. *et al.* (2013) 'Cancer genome landscapes', *Science (New York, N.Y.)*, 339(6127), pp. 1546–1558. doi: 10.1126/science.1235122.

Waizenegger, I. C. *et al.* (2000) 'Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase', *Cell*, 103(3), pp. 399–410. doi: 10.1016/S0092-8674(00)00132-X.

Walter, M. J. *et al.* (2012) 'Clonal Architecture of Secondary Acute Myeloid Leukemia', *New England Journal of Medicine*, 366(12), pp. 1090–1098. doi: 10.1056/NEJMoa1106968.

Wang, C. and Youle, R. J. (2009) 'The role of mitochondria in apoptosis*', *Annual review of genetics*, 43, pp. 95–118. doi: 10.1146/annurev-genet-102108-134850.

Wang, H. *et al.* (2017) 'Stromal antigen 2 functions as a tumor suppressor in bladder cancer cells', *Oncology Reports*, 38(2), pp. 917–925. doi: 10.3892/or.2017.5711.

Wang, Q. *et al.* (2003) 'Tumor suppressor p53 dependent recruitment of nucleotide excision repair factors XPC and TFIIH to DNA damage', *DNA Repair*, 2(5), pp. 483–499. doi: https://doi.org/10.1016/S1568-7864(03)00002-8.

Wang, Y.-L. *et al.* (2015) 'Snail promotes epithelial-mesenchymal transition and invasiveness in human ovarian cancer cells', *International journal of clinical and experimental medicine*. e-Century Publishing Corporation, 8(5), pp. 7388–7393. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26221280.

Wang, Y. *et al.* (2014) 'Clonal evolution in breast cancer revealed by single nucleus genome sequencing', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 512, p. 155.

Weijzen, S. *et al.* (2002) 'Activation of Notch-1 signaling maintains the neoplastic phenotype in human Ras-transformed cells', *Nature Medicine*, 8(9), pp. 979–986. doi: 10.1038/nm754.

Weinberg, R. A. (1991) 'Tumor suppressor genes', *Science*, 254(5035), p. 1138 LP-1146. doi: 10.1126/science.1659741.

Weinberg, R. A. (1995) 'The retinoblastoma protein and cell cycle control', *Cell*. Elsevier, 81(3), pp. 323–330. doi: 10.1016/0092-8674(95)90385-2.

Weinberg, R. A. (2014) *The Biology of Cancer.* 2nd edn.

Weiner, L. M., Surana, R. and Wang, S. (2010) 'Monoclonal antibodies: versatile platforms for cancer immunotherapy', *Nature Reviews Immunology*. Nature Publishing Group, 10, p. 317.

Wendt, K. S. *et al.* (2008) 'Cohesin mediates transcriptional insulation by CCCTC-binding factor', *Nature*. Nature Publishing Group, 451, p. 796.

Wiersdorff, V. *et al.* (1996) 'Mad acts downstream of Dpp receptors, revealing a differential requirement for dpp signaling in initiation and propagation of morphogenesis in the Drosophila eye', *Development*, 122(7), p. 2153 LP-2162.

WIKMAN, H., VESSELLA, R. and PANTEL, K. (2008) 'Cancer micrometastasis and tumour dormancy ', *APMIS*. John Wiley & Sons, Ltd (10.1111), 116(7-8), pp. 754–770. doi: 10.1111/j.1600-0463.2008.01033.x.

Williams, G. C. (1957) 'PLEIOTROPY, NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE', *Evolution*. John Wiley & Sons, Ltd (10.1111), 11(4), pp. 398–411. doi: 10.1111/j.1558-5646.1957.tb02911.x.

Willott, E. *et al.* (1993) 'The tight junction protein ZO-1 is homologous to the Drosophila discs-large tumor suppressor protein of septate junctions', *Proceedings of the National Academy of Sciences*, 90(16), p. 7834 LP-7838. doi: 10.1073/pnas.90.16.7834.

Witta, S. E. *et al.* (2006) 'Restoring E-Cadherin Expression Increases Sensitivity to Epidermal Growth Factor Receptor Inhibitors in Lung Cancer Cell Lines', *Cancer Research*, 66(2), p. 944 LP-950. doi: 10.1158/0008-5472.CAN-05-1988.

Wong, A. D. and Searson, P. C. (2014) 'Live-Cell Imaging of Invasion and Intravasation in an Artificial Microvessel Platform', *Cancer Research*, 74(17), p. 4937 LP-4945. doi: 10.1158/0008-5472.CAN-14-1042.

Wong, C. C. *et al.* (2013) 'Inactivating CUX1 mutations promote tumorigenesis', *Nature Genetics*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 46, p. 33. Available at: https://doi.org/10.1038/ng.2846.

Wong, K.-K., Engelman, J. A. and Cantley, L. C. (2010) 'Targeting the PI3K signaling pathway in cancer', *Current Opinion in Genetics & Development*, 20(1), pp. 87–90. doi: https://doi.org/10.1016/j.gde.2009.11.002.

Wong, R. W. and Blobel, G. (2008) 'Cohesin subunit SMC1 associates with mitotic microtubules at the spindle pole', *Proceedings of the National Academy of Sciences*, 105(40), pp. 15441–15445. doi: 10.1073/pnas.0807660105.

Wright, C. *et al.* (1989) 'Expression of c-erbB-2 Oncoprotein: A Prognostic Indicator in Human Breast Cancer', *Cancer Research*, 49(8), p. 2087 LP-2090.

Wu, B. *et al.* (2007) 'p53 independent induction of PUMA mediates intestinal

apoptosis in response to ischaemia-reperfusion', *Gut*. 2006/11/23. BMJ Group, 56(5), pp. 645–654. doi: 10.1136/gut.2006.101683.

Wu, J. S. and Luo, L. (2006) 'A protocol for mosaic analysis with a repressible cell marker (MARCM) in Drosophila', *Nature Protocols*, 1(6), pp. 2583–2589. doi: 10.1038/nprot.2006.320.

Wu, M. *et al.* (2010) 'Characterization of the roles of RHOC and RHOA GTPases in invasion, motility, and matrix adhesion in inflammatory and aggressive breast cancers', *Cancer*. John Wiley & Sons, Ltd, 116(S11), pp. 2768–2782. doi: 10.1002/cncr.25181.

Wu, M., Pastor-Pareja, J. C. and Xu, T. (2010) 'Interaction between RasV12 and scribbled clones induces tumour growth and invasion', *Nature*. Macmillan Publishers Limited. All rights reserved, 463, p. 545. Available at: https://doi.org/10.1038/nature08702.

Wu, X. and Levine, A. J. (1994) 'p53 and E2F-1 cooperate to mediate apoptosis', *Proceedings of the National Academy of Sciences of the United States of America*, 91(9), pp. 3602–3606.

Wutz, G. *et al.* (2017) 'Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins', *The EMBO journal*. 2017/12/07. John Wiley and Sons Inc., 36(24), pp. 3573–3599. doi: 10.15252/embj.201798004.

Xiong, Y. *et al.* (1993) 'p21 is a universal inhibitor of cyclin kinases', *Nature*, 366(6456), pp. 701–704. doi: 10.1038/366701a0.

Xu, T. *et al.* (1995) 'Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase', *Development*, 121(4), p. 1053 LP-1063. Available at: http://dev.biologists.org/content/121/4/1053.abstract.

Yamaguchi, H. and Condeelis, J. (2007) 'Regulation of the actin cytoskeleton in cancer cell migration and invasion', 1773, pp. 642–652. doi: 10.1016/j.bbamcr.2006.07.001.

Yamanaka, T. *et al.* (2006) 'Lgl mediates apical domain disassembly by suppressing the PAR-3-aPKC-PAR-6 complex to orient apical membrane polarity', *Journal of Cell Science*, 119(10), p. 2107 LP-2118. doi: 10.1242/jcs.02938.

Yamazaki, D., Kurisu, S. and Takenawa, T. (2009) 'Involvement of Rac and Rho signaling in cancer cell motility in 3D substrates', *Oncogene*. Macmillan Publishers Limited, 28, p. 1570.

Yan, C. *et al.* (2010) 'Epithelial to Mesenchymal Transition in Human Skin Wound Healing Is Induced by Tumor Necrosis Factor-α through Bone Morphogenic Protein-2', *The American Journal of Pathology*, 176(5), pp. 2247–2258. doi: https://doi.org/10.2353/ajpath.2010.090048.

Yan, L. and Bi, D. (2019) 'Multicellular Rosettes Drive Fluid-solid Transition in Epithelial Tissues', *Physical Review X*. American Physical Society, 9(1), p. 11029. doi: 10.1103/PhysRevX.9.011029.

Yang, J. and Weinberg, R. A. (2008) 'Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis', *Developmental Cell*, 14(6), pp. 818–829. doi: https://doi.org/10.1016/j.devcel.2008.05.009.

Yang, X. *et al.* (2015) 'Role of EHD2 in migration and invasion of human breast cancer cells', pp. 3717–3726. doi: 10.1007/s13277-014-3011-9.

Yokoyama, Y. *et al.* (2017) 'Autocrine BMP-4 Signaling Is a Therapeutic Target in Colorectal Cancer', *Cancer Research*, 77(15), p. 4026 LP-4038. doi: 10.1158/0008-5472.CAN-17-0112.

Yoshizaki, H. *et al.* (2003) 'Activity of Rho-family GTPases during cell division as visualized with FRET-based probes', *THE JOURNAL OF CELL BIOLOGY*, pp. 223–232. doi: 10.1083/jcb.200212049.

Young, K. G. and Kothary, R. (2007) 'Dystonin/Bpag1—A link to what?', *Cell Motility*. John Wiley & Sons, Ltd, 64(12), pp. 897–905. doi: 10.1002/cm.20235.

Zabkiewicz, C. *et al.* (2017) 'Bone morphogenetic proteins, breast cancer, and bone metastases: striking the right balance', *Endocrine-related cancer*. Bioscientifica Ltd, 24(10), pp. R349–R366. doi: 10.1530/ERC-17-0139.

Zalewski, J. K. *et al.* (2016) 'Structure of the Shroom-Rho Kinase Complex Reveals a Binding Interface with Monomeric Shroom That Regulates Cell Morphology and Stimulates Kinase Activity', *Journal of Biological Chemistry* , 291(49), pp. 25364–25374. Available at: http://www.jbc.org/content/291/49/25364.abstract.

Zallen, J. A. and Blankenship, J. T. (2008) 'Multicellular dynamics during epithelial elongation', *Seminars in Cell & Developmental Biology*, 19(3), pp. 263–270. doi: https://doi.org/10.1016/j.semcdb.2008.01.005.

Zaman, N. *et al.* (2013) 'Signaling Network Assessment of Mutations and Copy Number Variations Predict Breast Cancer Subtype-Specific Drug Targets', *Cell Reports*, 5(1), pp. 216–223. doi: https://doi.org/10.1016/j.celrep.2013.08.028.

Zervantonakis, I. K. *et al.* (2012) 'Three-dimensional microfluidic model for tumor cell intravasation and endothelial barrier function', *Proceedings of the National Academy of Sciences*, 109(34), p. 13515 LP-13520. doi: 10.1073/pnas.1210182109.

Zhan, L. *et al.* (2008) 'Deregulation of Scribble Promotes Mammary Tumorigenesis and Reveals a Role for Cell Polarity in Carcinoma', *Cell*, 135(5), pp. 865–878. doi: https://doi.org/10.1016/j.cell.2008.09.045.

Zhang, H. *et al.* (2003) 'Loss of Tsc1/Tsc2 activates mTOR and disrupts PI3K-Akt signaling through downregulation of PDGFR', *The Journal of Clinical Investigation*. The American Society for Clinical Investigation, 112(8), pp. 1223–1233. doi: 10.1172/JCI17222.

Zhang, N. *et al.* (2013) 'Characterization of the Interaction between the Cohesin Subunits Rad21 and SA1/2', *PLoS ONE*, 8(7). doi: 10.1371/journal.pone.0069458.

Zhang, N. and Pati, D. (2012) 'Sororin is a master regulator of sister chromatid cohesion and separation', *Cell Cycle*, 11(11), pp. 2073–2083. doi: 10.4161/cc.20241.

Zhang, X. *et al.* (2013) 'N-Cadherin Expression Is Associated with Acquisition of EMT Phenotype and with Enhanced Invasion in Erlotinib-Resistant Lung Cancer Cell Lines', *PLOS ONE*. Public Library of Science, 8(3), p. e57692.

Zhou, B.-B. S. and Elledge, S. J. (2000) 'The DNA damage response: putting checkpoints in perspective', *Nature*. Macmillan Magazines Ltd., 408, p. 433.

van Zijl, F., Krupitza, G. and Mikulits, W. (2011) 'Initial steps of metastasis: cell invasion and endothelial transmigration', *Mutation research*. Elsevier, 728(1–2), pp. 23–34. doi: 10.1016/j.mrrev.2011.05.002.

Zou, H. *et al.* (1999) 'An APAF-1·Cytochrome c Multimeric Complex Is a Functional Apoptosome That Activates Procaspase-9 ', *Journal of Biological Chemistry* , 274(17), pp. 11549–11556.

Zuin, J. *et al.* (2017) 'Regulation of the cohesin-loading factor NIPBL: Role of the lncRNA NIPBL-AS1 and identification of a distal enhancer element', *PLoS genetics*. Public Library of Science, 13(12), pp. e1007137–e1007137. doi: 10.1371/journal.pgen.1007137.

Appendixes

.

Appendix 1 – Screen Phenotypes and scoring system

Category	Data Type	Scoring system	Description
No Clones	Yes/No	No (0), Yes (1)	No labelled cells are observed
Not Developed	Yes/No	No (0), Yes (1)	Animal did not reach the pupal stage (lethal)
Malformed	Yes/No	Normal (0), malformed (1)	Animal was too malformed to mount
Clonal tissue	Number	0-5% (-2), 5-15% (-1), 15-35% (0), 35-50% (1), 50% + (2)	Area covered by the labelled cells
Individual clone size	Number	mostly small (-1), mixture/mostly medium (0), mostly large (1), very large (2)	Size of the individual clones
Clonal shape	Number	mostly round (-1), mixture (0), mostly long (1)	Clones were rounded or long in shape
Dividing cells	Number	none (-2), little (-1), medium (0), lots (1), extensive (2)	Amount cells in mitosis, identified by the mitotic cell rounding
Invading Cells	Number	none (-2), very few (-1), medium (0), lots (1), extensive (2)	Amount of cells that detached from the epithelium basally
Type of Invasion	Number	mostly non-polarised (-1), mixture (0), mostly polarised (1)	Describes the actin localization within the invading cell. Polarized describes the accumulation of actin on one side of the cell.
Invading cells with protrusions	Yes/No	No (0), Yes (1)	Presence of protrusions in invading cells.
Cuticle	Yes/No	normal (0), deep folding (1)	Folding in the supracellular cuticle
closure defects	Yes/No	No (0), Yes (1)	Failure in the dorsal closure process.
Apex size	Number	very small (-2), small (-1), normal (0), large (1)	Size of the apical area of the cells
Defective Apex	Number	none (0), up to 50% (1), >50% (2)	Amount of apex that lost their geometrical shape

Category	Data Type	Scoring system	Description
No Apex	Number	none (-1), 0-15% (0), 15-50% (1), >50% (2)	Amount of cells that lost their polarized shape and therefore lost their apex
Junction defects	Yes/No	No (0), Yes (1)	Junctional breaks
Cell body (rounding)	Number	none (-1), 0-15% (0), 15-50% (1), 50% (2)	Amount of rounded cells that lost their columnar shape
Polarised Cytoskeleton	Number	none (-1), very few (0), medium (1), lots (2)	Non-symmetrical actin accumulation
Basal bundles	Yes/No	No (0), Yes (1)	Formation of protrusions bundles
Long apical/intermediate protrusions	Yes/No	No (0), Yes (1)	Length of apical/intermediate protrusions
Very Long protrusions	Yes/No	No (0), Yes (1)	Presence of very long protrusions
Blebbing mitotic cells	Yes/No	No (0), Yes (1)	Presence of blebs in mitotic cells
Thick protrusions	Yes/No	No (0), Yes (1)	Protrusion width
Bright protrusions	Yes/No	No (0), Yes (1)	Protrusion brightness
Branched protrusions	Yes/No	No (0), Yes (1)	Presence of branches in the protrusions
Cyst-like clones	Yes/No	No (0), Yes (1)	Presence of rounded cyst-like clones
Apoptotic	Number	none(0), little (1), medium (2), lots (3)	Apoptotic estimation by identifying ruptured, fragmented or highly abnormal cells.
Basal Protrusions	Number	absent (-3), small (-2), normal (-1), long (0), very long (1)	Length of basal protrusions
Basal Actin Rich Spot	Number	No (0), Yes (1)	Actin rich spots present basally
Multi-layering	Number	none (-1), 0-15% (0), 15-50% (1), 50+% (2)	Loss of contact inhibition, presence of cells growing on top of each other
Cell Length	Number	short <7um (-2), normal 8-12um (-1), long (0) 13-19um, very long (1) 20um+	Cell length in microns (from apex to basal)
Single Cell Clones	Yes/No	No (0), Yes (1)	Presence of single cell clones

Appendix 2- Screen genes

Gene	Stock No	Pilot	Source
18w (1)	8896R3 (III)	No	NIG
18w (2)	30498 (III)	No	BDSC
Abi	36142 (III)	Yes	VDRC
Ac78C	55312 (III)	No	BDSC
Actn	4376R1 (III)	No	NIG
Actn (2)	7760 (III)	No	VDRC
Adar	7763 (III)	No	VDRC
Adar (2)	28311 (III)	No	BDSC
AdoR	1385 (III)	No	VDRC
AdoR (2)	27536 (III)	No	BDSC
ago	15010R-3 (III)	No	NIG
Akt1 (1)	2902 (III)	No	VDRC
Akt1 (2)	4006R-3 (III)	No	NIG
Alk	11446 (X)	No	VDRC
alpha-catenin	20123-	Yes	VDRC
alpha-spec	31209 (III)	No	BDSC
Ank	25945 (III)	No	VDRC
Ank2 (1)	26121 (III)	No	VDRC
Ank2 (2)	5115 (III)	No	VDRC
APC (1)	51469-	Yes	VDRC
APC (2)	1451R-1	Yes	NIG
Apc2	22290 (X)	Yes	VDRC
Apoltp	51937 (III)	No	BDSC
Archipelago	15010R-3 (III)	Yes	NIG
arm (1)	11579R-1 (III)	Yes	NIG
Arm (2)	35004	Yes	VDRC
asp	2910 (III)	No	VDRC
asp (2)	28741 (III)	No	BDSC
AstC-R2	50000 (III)	No	VDRC
AstC-R2 (2)	25940 (III)	No	BDSC
Asx	31192 (III)	No	BDSC
ATP7	8315 (III)	No	VDRC
ATPsyn-beta (1)	11154R-1 (III)	No	NIG
ATPsyn-beta (2)	37812 (X)	No	VDRC
Atx-1	21870 (III)	No	VDRC

Gene	Stock No	Pilot	Source
aub	30124 (III)	No	VDRC
babo (1)	3825 (III)	No	VDRC
babo (2)	853 (III)	No	VDRC
Bap170 (1)	3274R-2 (III)	No	NIG
Bap170 (2)	34582 (III)	No	VDRC
Baz	5055R-1 (III)	Yes	NIG
bbg	15974 (III)	No	VDRC
Bcl2 (debcl) (1)	47515 (III)	Yes	VDRC
Bcl2 (debcl) (2)	12397R-3 (III)	Yes	NIG
bdl	31974 (III)	No	BDSC
Beclin-1 (Atg6) (1)	22122 (III)	Yes	VDRC
Beclin-1 (Atg6) (2)	5429R-1 (III)	Yes	NIG
bel	6299 (III)	No	VDRC
bent	46252 (III)	No	VDRC
Blimp-1 (1)	5249R-2 (III)	No	NIG
Blimp-1 (2)	34978 (III)	No	VDRC
bmm	37880 (III)	No	VDRC
bnk	1480R-3 (III)	No	NIG
brm (1)	37720 (III)	No	VDRC
brm (2)	37721 (III)	No	VDRC
Bro	13679 (X)	No	VDRC
brv1	31496 (III)	No	BDSC
brv2 (1)	31670 (III)	No	BDSC
brv2 (2)	31671 (III)	No	BDSC
btl (1)	32134R-1 (III)	No	NIG
btl (2)	27108 (X)	No	VDRC
bun	19680 (X)	No	VDRC
bun (2)	28322 (III)	No	BDSC
bx	29454 (III)	No	BDSC
Ca-alp1T (1)	31961 (III)	No	VDRC
Ca-alp1T (2)	26251 (III)	No	BDSC
ca-alpha1D (1)	51491 (III)	No	VDRC
ca-alpha1D (2)	52644 (III)	No	VDRC
cac	48093 (X)	No	VDRC
Cad99C (1)	3733 (III)	No	VDRC
Cad99C (2)	27212 (III)	No	VDRC

Gene	Stock No	Pilot	Source
CadN	7100R3 (III)	No	NIG
CadN (2)	41982 (III)	No	BDSC
Calx	28306 (III)	No	BDSC
calypso (1)	47743 (III)	No	VDRC
calypso (2)	28904 (X)	No	VDRC
Cap	39207 (III)	No	VDRC
capt (1)	5061R-4 (III)	No	NIG
capt (2)	21995 (III)	No	VDRC
capu	32922 (III)	No	BDSC
CcapR	14767 (III)	No	VDRC
CcapR (2)	31490 (III)	No	BDSC
cdc42	12530R-2 (III)	Yes	NIG
cdep	2008R3 (III)	No	NIG
cdep (2)	31168 (III)	No	BDSC
CDK5	8203R-3	Yes	NIG
CG10186 (1)	30767 (III)	No	VDRC
CG10186 (2)	10186R-1 (III)	No	NIG
CG10211	12352 (III)	No	VDRC
CG10274 (1)	10274R-2 (X)	No	NIG
CG10274 (2)	52407 (III)	No	VDRC
CG10348	10348R1 (III)	No	NIG
CG10348 (2)	44091 (III)	No	BDSC
CG10585 (1)	31272 (III)	No	VDRC
CG10585 (2)	31273 (III)	No	VDRC
CG10600 (1)	10600R-2 (III)	No	NIG
CG10600 (2)	31276 (III)	No	VDRC
CG10621	31291 (III)	No	VDRC
CG10623 (1)	10623R-3 (III)	No	NIG
CG10623 (2)	10623R-4 (X)	No	NIG
CG10669 (1)	10669R-1 (III)	No	NIG
CG10669 (2)	10669R-4 (III)	No	NIG
CG10702	3691 (III)	No	VDRC
CG10702 (2)	27052 (III)	No	VDRC
CG10777	10777R-3 (III)	No	NIG
CG10916	31379 (III)	No	VDRC
CG10931	52107 (X)	No	VDRC

Gene	Stock No	Pilot	Source
CG11168	18171 (X)	No	VDRC
CG11180	11180R2 (III)	No	NIG
CG11180 (2)	28629 (III)	No	BDSC
CG11247	52116 (III)	No	VDRC
CG11318	11318R1 (III)	No	NIG
CG11318 (2)	51792 (III)	No	BDSC
CG11444 (1)	11444R-1 (III)	No	NIG
CG11444 (2)	31501 (III)	No	VDRC
CG11593	11593R2 (III)	No	NIG
CG11593 (2)	46728 (III)	No	VDRC
CG11836	43308 (III)	No	BDSC
CG12268	12268R1 (III)	No	NIG
CG12268 (2)	1166 (III)	No	VDRC
CG12484 (1)	12484R1 (III)	No	NIG
CG12484 (2)	12484R2 (III)	No	NIG
CG12869	4875 (III)	No	VDRC
CG12950	12950Ra1 (III)	No	NIG
CG12950 (2)	10011 (III)	No	VDRC
CG1311 (2)	12242 (III)	No	VDRC
CG13185	18782 (III)	No	VDRC
CG13579	13579R1 (III)	No	NIG
CG13579 (2)	9366 (III)	No	VDRC
CG13597 (1)	13597R-4 (X)	No	NIG
CG13597 (2)	32225 (III)	No	VDRC
CG14372	29806 (III)	No	VDRC
CG14446	14446R2 (III)	No	NIG
CG14446 (2)	32487 (III)	No	BDSC
CG14838	14838R3 (III)	No	NIG
CG15117 (2)	47530 (III)	No	VDRC
CG15269	51506 (III)	No	BDSC
CG15395 (1)	15395R-3 (III)	No	NIG
CG15395 (2)	39929 (III)	No	VDRC
CG15556	1791 (III)	No	VDRC
CG15744	28516 (III)	No	BDSC
CG16721	25201 (III)	No	VDRC
CG16868	16868R1 (III)	No	NIG

Gene	Stock No	Pilot	Source
CG16868 (2)	29617 (III)	No	BDSC
CG17002	17002R-2 (III)	No	NIG
CG1724 (1)	30315 (III)	No	VDRC
CG1724 (2)	30317 (III)	No	VDRC
CG17739	36542 (III)	No	VDRC
CG18335 (1)	18335R-1 (III)	No	NIG
CG18335 (2)	52471 (III)	No	VDRC
CG18476 (1)	6527 (III)	No	VDRC
CG18476 (2)	17822R-1 (III)	No	NIG
CG2135	16625 (III)	No	VDRC
Cg25C	28369 (III)	No	VDRC
CG2807	28965 (III)	No	BDSC
CG30268	42705 (III)	No	VDRC
CG30280 (1)	34743 (III)	No	VDRC
CG30280 (2)	34744 (III)	No	VDRC
CG30417 (1)	25619 (III)	No	VDRC
CG30417 (2)	48182 (X)	No	VDRC
CG31626	1975R-2 (III)	No	NIG
CG31731 (1)	11082 (X)	No	VDRC
CG31731 (2)	31731R-2 (III)	No	NIG
CG31999	31999R1 (III)	No	NIG
CG31999 (2)	31587 (III)	No	BDSC
CG32066	44825 (III)	No	VDRC
CG32238 (1)	32238R-1 (III)	No	NIG
CG32238 (2)	29424 (III)	No	VDRC
CG32333	26118 (III)	No	VDRC
CG32373	32373R-1 (III)	No	NIG
CG32669	47916 (III)	No	VDRC
CG33288 (1)	51189 (X)	No	VDRC
CG33288 (2)	51190 (III)	No	VDRC
CG3339 (1)	41917 (III)	No	VDRC
CG3339 (2)	41918 (III)	No	VDRC
CG34113	10989R2 (III)	No	NIG
CG34113 (2)	22742 (III)	No	VDRC
CG34120 (1)	11673 (III)	No	VDRC
CG34120 (2)	1819R-1 (III)	No	NIG

Gene	Stock No	Pilot	Source
CG34371	44997 (III)	No	VDRC
CG34380	31640R1 (III)	No	NIG
CG34380 (2)	39446 (III)	No	VDRC
CG34422 (1)	7274R-3 (III)	No	NIG
CG34422 (2)	40440 (III)	No	VDRC
CG4096	44522 (III)	No	BDSC
CG42255	32092R-3 (III)	No	NIG
CG42342	28648 (III)	No	BDSC
CG42672	31737 (III)	No	BDSC
CG42817	3849 (III)	No	VDRC
CG42817 (2)	25882 (III)	No	BDSC
CG4360	4360R2 (III)	No	NIG
CG4360 (2)	51813 (III)	No	BDSC
CG4393 (1)	4393R1 (III)	No	NIG
CG4393 (2)	4393R4 (III)	No	NIG
CG4438	50033 (III)	No	VDRC
CG45263	18706 (III)	No	VDRC
CG4587 (1)	4587R2 (III)	No	NIG
CG4587 (2)	3839 (III)	No	VDRC
CG4887 (1)	4887R-3 (III)	No	NIG
CG4887 (2)	21970 (III)	No	VDRC
CG4896 (1)	26652 (III)	No	VDRC
CG4896 (2)	48197 (III)	No	VDRC
CG5037	42787 (III)	No	VDRC
CG5037 (2)	31169 (III)	No	BDSC
CG5155	31568 (III)	No	BDSC
CG5245	27382 (III)	No	VDRC
CG5245 (2)	27381 (X)	No	VDRC
CG5389 (1)	5389R-3 (X)	No	NIG
CG5389 (2)	22111 (III)	No	VDRC
CG5550 (1)	31000 (III)	No	VDRC
CG5550 (2)	31001 (III)	No	VDRC
CG5645	9289 (X)	No	VDRC
CG5984	52489 (X)	No	VDRC
CG6153	22257 (III)	No	VDRC
CG6272 (1)	34156 (III)	No	VDRC

Gene	Stock No	Pilot	Source
CG6272 (2)	34158 (III)	No	VDRC
CG6480	23447 (III)	No	VDRC
CG6867	37416 (III)	No	VDRC
CG7229 (1)	7229R-1 (III)	No	NIG
CG7229 (2)	37671 (III)	No	VDRC
CG7379 (1)	27988 (III)	No	VDRC
CG7379 (2)	27989 (III)	No	VDRC
CG7394	9210 (III)	No	VDRC
CG7420 (1)	46316 (III)	No	VDRC
CG7420 (2)	7420R-1 (III)	No	NIG
CG7607	7607R1 (III)	No	NIG
CG7607 (2)	9208 (III)	No	VDRC
CG7708	28613 (III)	No	BDSC
CG7911	23075 (III)	No	VDRC
CG8027	8027R-3 (III)	No	NIG
CG8184	8184R1 (X)	No	NIG
CG8197	42126 (III)	No	VDRC
CG8290 (1)	8290R-2 (III)	No	NIG
CG8290 (2)	12738 (III)	No	VDRC
CG89095	23159 (III)	No	VDRC
CG8918	8918R-3 (III)	No	NIG
CG9279	45052 (III)	No	VDRC
CG9492	51725 (III)	No	BDSC
CG9572 (1)	44802 (III)	No	VDRC
CG9572 (2)	44804 (III)	No	VDRC
CG9701 (1)	9701R-2 (III)	No	NIG
CG9701 (2)	3359 (III)	No	VDRC
CG9784 (1)	9784R-1 (III)	No	NIG
CG9784 (2)	30098 (III)	No	VDRC
Chd3	13636 (III)	No	VDRC
cher	3937R2 (III)	No	NIG
CIC-c	6466 (III)	No	VDRC
CIC-c (2)	27034 (III)	No	BDSC
Cirl	8639R1 (III)	No	NIG
Cirl (2)	34821 (III)	No	BDSC
cnc (1)	17894R-3 (X)	No	NIG

Gene	Stock No	Pilot	Source
cnc (2)	37674 (III)	No	VDRC
cnn	4832R1 (III)	No	NIG
Cont	40613 (III)	No	VDRC
Cont (2)	28923 (III)	No	BDSC
cora	9788 (III)	No	VDRC
crb	6383R2 (III)	No	NIG
crb (2)	40869 (III)	No	BDSC
CSN8 (1)	13383R-3 (III)	No	NIG
CSN8 (2)	50566 (X)	No	VDRC
csw	21757 (III)	No	VDRC
ct (1)	4138 (III)	No	VDRC
ct (2)	5687 (III)	No	VDRC
CTCF	30713 (III)	No	VDRC
Cubn (1)	14613 (III)	No	VDRC
Cubn (2)	14614 (III)	No	VDRC
Cul-4	8711R-1 (III)	No	NIG
cv-c	14847R2 (X)	No	NIG
cv-c (2)	32329 (III)	No	VDRC
CycD	29024 (III)	No	VDRC
cype	13404 (III)	No	VDRC
cype (2)	33878 (III)	No	BDSC
Cyt-c-d	17128 (III)	Yes	VDRC
Ddr	29720 (III)	No	VDRC
Ddr (2)	55906 (III)	No	BDSC
dE2f2 (1)	45743 (X)	Yes	VDRC
dE2f2 (2)	1071R1 (III)	Yes	VDRC
Death caspase 1	5370R-2 (III)	Yes	NIG
Dgk	29459 (III)	No	BDSC
Dgk (2)	36745 (III)	No	BDSC
Dhc36C (1)	27451 (III)	No	VDRC
Dhc36C (2)	5526R-1 (III)	No	NIG
Dhc62B	15804R-1 (III)	No	NIG
Dhc93AB (1)	41947 (III)	No	VDRC
Dhc93AB (2)	3723R-2 (III)	No	NIG
dia (1)	1768R-1 (III)	No	NIG
dia (2)	20518 (III)	No	VDRC

Gene	Stock No	Pilot	Source
Dis3 (1)	6413R-2 (III)	No	NIG
Dis3 (2)	35090 (III)	No	VDRC
Dlg (1)	41136 (III)	Yes	VDRC
Dlg (2)	1725R-1 (III)	Yes	NIG
dlt	41876 (III)	No	VDRC
DNApol-epsilon	6768R-2 (III)	No	NIG
dob	14814 (III)	No	VDRC
dom	7787 (III)	No	VDRC
dom (2)	41674 (III)	No	BDSC
Dp (1)	7836 (III)	No	VDRC
Dp (2)	15637R-2 (III)	No	NIG
ds	4313 (III)	No	VDRC
Dscam1 (1)	25622 (III)	No	VDRC
Dscam1 (2)	36233 (III)	No	VDRC
Dscam2 (1)	1003 (III)	No	VDRC
Dscam2 (2)	41890 (X)	No	VDRC
Dscam3	31190R1 (III)	No	NIG
Dscam3 (2)	6685 (III)	No	VDRC
Dscam4 (1)	25366 (III)	No	VDRC
Dscam4 (2)	42882 (III)	No	VDRC
Dsor1 (1)	40025 (III)	No	VDRC
Dsor1 (2)	40026 (III)	No	VDRC
Dys	19230 (III)	No	VDRC
E(z) (1)	6502R-4 (III)	No	NIG
E(z) (2)	27646 (III)	No	VDRC
E2F1 (1)	15886 (III)	Yes	VDRC
E2F1 (2)	15887 (III)	Yes	VDRC
Eaat1	3747R-2 (III)	No	NIG
ebi (1)	40862 (III)	No	VDRC
ebi (2)	4063R-2 (III)	No	NIG
E-cadherin (shg)	27082-	Yes	VDRC
Egfr (1)	43267 (III)	No	VDRC
Egfr (2)	10079R-2 (III)	No	NIG
egg (1)	12196R-1 (III)	No	NIG
egg (2)	21172 (III)	No	VDRC
Ehbp1	31143 (III)	No	BDSC

Gene	Stock No	Pilot	Source
eIF-2beta (1)	4153R-2 (III)	No	NIG
eIF-2beta (2)	9416 (III)	No	VDRC
Elongin-C	15302 (III)	No	VDRC
emb	31353 (III)	No	BDSC
emb (2)	34021 (III)	No	BDSC
Ent1	11907R2 (III)	No	NIG
Ent1 (2)	49328 (X)	No	VDRC
Eph (1)	1511R-1 (III)	No	NIG
Eph (2)	4771 (III)	No	VDRC
Ephrin	1862R2 (III)	No	NIG
Ephrin (2)	27039 (III)	No	BDSC
Ex	22994 (III)	Yes	VDRC
Exn	41960 (III)	No	VDRC
Exn (2)	33373 (III)	No	BDSC
fas	42237 (III)	Yes	VDRC
Fbp2	33173 (III)	No	VDRC
fd102C	11152R-2 (III)	No	NIG
fkh (1)	37062 (III)	No	VDRC
fkh (2)	49960 (III)	No	VDRC
fra (1)	6557 (III)	No	VDRC
fra (2)	29909 (III)	No	VDRC
frac (1)	31577 (III)	No	BDSC
frac (2)	31578 (III)	No	BDSC
fred	46180 (III)	No	VDRC
fred (2)	42621 (III)	No	BDSC
ft	29566 (III)	No	BDSC
ft (2)	34970 (III)	No	BDSC
fw	39575 (III)	No	VDRC
Gbeta13F (1)	10545R-2 (III)	No	NIG
Gbeta13F (2)	31257 (III)	No	VDRC
GI	27721 (III)	No	BDSC
grk	17610R2 (III)	No	NIG
grk (2)	4331 (III)	No	VDRC
grn	33746 (III)	No	BDSC
grn (2)	34578 (III)	No	BDSC
Gtp-bp	14877 (III)	No	VDRC

Gene	Stock No	Pilot	Source
Gug	13687 (III)	No	VDRC
ham	15907R1 (III)	No	NIG
ham (2)	32470 (III)	No	BDSC
hbs	7449R1 (III)	No	NIG
hbs (2)	27065 (III)	No	VDRC
hig (1)	13266 (X)	No	VDRC
hig (2)	2040R-1 (III)	No	NIG
hiw	26998 (III)	No	VDRC
hiw (2)	28031 (III)	No	BDSC
Hml (1)	37005 (III)	No	VDRC
Hml (2)	37006 (III)	No	VDRC
how	10293R-1 (III)	No	NIG
hppy	7097R-1 (III)	No	NIG
Hsp83	7716 (III)	No	VDRC
IA-2	11344R3 (III)	No	NIG
IA-2 (2)	7560 (III)	No	VDRC
Idh (1)	42915 (III)	No	VDRC
Idh (2)	42916 (III)	No	VDRC
if (aPS2) Integrin (1)	44885 (III)	Yes	VDRC
if (aPS2) Integrin (2)	5600 (X)	Yes	VDRC
ih	35836 (III)	No	VDRC
ih (2)	29574 (III)	No	BDSC
InR	991 (III)	No	VDRC
In-R	31037 (III)	No	BDSC
IP3K2	12724R-2 (III)	No	NIG
Ir25a	15627R2 (III)	No	NIG
Ir8a	29917 (III)	No	VDRC
Ir8a	25813 (III)	No	BDSC
Irp-1B (2)	30153 (III)	No	VDRC
Irp-1B (3)	55224	No	BDSC
itgbn	40895 (III)	No	VDRC
jbug	11605R2 (III)	No	NIG
jbug (2)	39070 (III)	No	BDSC
jim (1)	11352R2 (III)	No	NIG
jim (2)	11352R3 (III)	No	NIG
jub	30806 (III)	No	NIG

Gene	Stock No	Pilot	Source
Kirre (1)	6695 (III)	Yes	VDRC
Kirre (2)	6696 (III)	Yes	VDRC
kis (1)	18326R-1 (III)	No	NIG
kis (2)	3696R-1 (III)	No	NIG
kl-2	19181 (III)	No	VDRC
kl-3	53317 (III)	No	BDSC
klg	39515 (III)	No	VDRC
klg (2)	28746 (III)	No	BDSC
klu (1)	12296R-2 (III)	No	NIG
klu (2)	12296R-3 (III)	No	NIG
kon	10275R-1 (III)	No	NIG
Krn	8056R4 (III)	No	NIG
kst (2)	33933 (III)	No	BDSC
kto (1)	8491R-1 (III)	No	NIG
kto (2)	23142 (III)	No	VDRC
kug (1)	5098 (III)	No	VDRC
kug (2)	7749R-1 (III)	No	NIG
LanB1 (1)	23119 (III)	No	VDRC
LanB1 (2)	23121 (III)	No	VDRC
Lap1	10255R1 (III)	No	NIG
Lap1 (2)	27036 (III)	No	BDSC
Lar	36270 (X)	No	VDRC
lea	5574R2 (III)	No	NIG
lea (2)	34589 (III)	No	BDSC
lectin-22C (1)	15378R1 (III)	No	NIG
lectin-22C (2)	15378R6 (III)	No	NIG
lgl-IR (1)	51247 (III)	Yes	VDRC
lgl-IR (2)	51249 (III)	Yes	VDRC
lid (1)	42203 (III)	No	VDRC
lid (2)	42204 (III)	No	VDRC
lilli (1)	13081 (III)	No	VDRC
lilli (2)	8817R-4 (III)	No	NIG
Lpt	22171 (III)	No	VDRC
LRP1	8397 (III)	No	VDRC
LRP1 (2)	44579 (III)	No	BDSC
ltgbn	1762R1 (III)	No	NIG

Gene	Stock No	Pilot	Source
lz	27985 (III)	No	BDSC
Mad (1)	12399R-1	Yes	NIG
Mad (2)	12399R-2	Yes	NIG
MAGE	44514 (III)	No	BDSC
Marcal1	34701 (III)	No	VDRC
mbc	16044 (III)	No	VDRC
mbc (2)	51446 (III)	No	BDSC
Mcr	2785 (III)	No	VDRC
MED23 (1)	28361 (III)	No	VDRC
mei-38	51805 (III)	No	VDRC
Merlin	7161-	Yes	VDRC
Mes-4	10837 (III)	No	VDRC
mesh (1)	40940 (X)	No	VDRC
mesh (2)	12074R-1 (III)	No	NIG
mew (aPS1) Integrin	1771R-1 (III)	Yes	NIG
mfas	37889 (III)	No	VDRC
mgl	29324 (III)	No	BDSC
mgl (2)	33940 (III)	No	BDSC
mGluRA	11144R3 (III)	No	NIG
mGluRA (2)	1793 (III)	No	VDRC
Mhc	17927R4 (III)	No	NIG
Mhc (2)	7164 (III)	No	VDRC
Mkk4 (1)	26929 (III)	No	VDRC
Mkk4 (2)	9738R-1 (III)	No	NIG
Mnt (1)	10971 (III)	Yes	VDRC
Mnt (2)	2856R-2 (X)	Yes	NIG
Moesin (1)	10701R-1 (III)	Yes	NIG
Moesin (2)	10701R-3 (X)	Yes	NIG
Mrtf	34503 (III)	No	VDRC
msk	38963 (III)	No	VDRC
Msp-300 (1)	25906 (III)	No	VDRC
Msp-300 (2)	18252R-2 (III)	No	NIG
mtt	32376 (III)	No	BDSC
mub	34870 (III)	No	BDSC
Muc14A	44527 (III)	No	BDSC
mus205	24472 (III)	No	VDRC

Gene	Stock No	Pilot	Source
Myb (1)	37710 (III)	No	VDRC
Myb (2)	9045R-1 (III)	No	NIG
Myd88	25399 (III)	No	VDRC
Myosin II (sqh)	7917 (III)	Yes	VDRC
mys	1560R-1 (III)	Yes	NIG
na	3306 (III)	No	VDRC
nAcRbeta-64B (1)	39421 (X)	No	VDRC
nAcRbeta-64B (2)	31883 (III)	No	BDSC
Ndae1 (1)	4675R-1 (III)	No	NIG
Ndae1 (2)	3664 (III)	No	VDRC
nej	46535 (III)	No	VDRC
Nep3	9565R-2 (III)	No	NIG
NetA	18657R-2 (III)	No	NIG
Neto	17302 (III)	No	VDRC
Neto (2)	19594 (X)	No	VDRC
NF1 (1)	35877-	Yes	VDRC
NF1 (2)	8318R-1	Yes	NIG
Nipped-A	2905R3 (III)	No	NIG
Nipped-A (2)	40789 (III)	No	VDRC
Nlg1	42616 (III)	No	VDRC
Nlg2	44236 (III)	No	VDRC
Nlg3	1062R1 (III)	No	NIG
Nmdar2	14794R1 (III)	No	NIG
Nmdar2 (2)	3196 (III)	No	VDRC
nompC	31512 (III)	No	BDSC
nopo	22013 (III)	No	VDRC
Nos	6713R3 (III)	No	NIG
Nos (2)	50675 (III)	No	BDSC
Nrg (1)	27201 (III)	No	VDRC
Nrg (2)	27202 (III)	No	VDRC
nrm	979 (III)	No	VDRC
Nrx-1	36326 (III)	No	VDRC
Nrx-IV	28715 (III)	No	BDSC
Nrx-IV (2)	32424 (III)	No	BDSC
nub (1)	6217 (III)	No	VDRC
Nup62 (1)	44806 (III)	No	VDRC

Gene	Stock No	Pilot	Source
Nup62 (2)	6251R-1 (III)	No	NIG
Or33a (1)	1461 (III)	No	VDRC
Or33a (2)	1462 (III)	No	VDRC
Or59b (1)	43808 (III)	No	VDRC
Or59b (2)	43809 (III)	No	VDRC
osa (1)	7467R-1 (III)	No	NIG
osa (2)	7810 (III)	No	VDRC
otk	8967R4 (III)	No	NIG
otk (2)	30833 (III)	No	VDRC
otu (1)	47431 (III)	Yes	VDRC
otu (2)	12743R-3 (III)	Yes	NIG
p120ctn	51729 (III)	No	BDSC
p53 (1)	10692-	Yes	VDRC
p53 (2)	14873R-1	Yes	NIG
Pak	12553 (X)	No	VDRC
PAK1	12553 (X)	Yes	VDRC
PAK3 (1)	39843 (III)	Yes	VDRC
PAK3 (2)	14895R-1 (III)	Yes	NIG
pan	3014 (III)	No	VDRC
para (1)	6131 (III)	No	VDRC
para (2)	6132 (III)	No	VDRC
patj (1)	31620 (III)	No	VDRC
patj (2)	12021R-4 (III)	No	NIG
pdm2 (1)	30708 (III)	No	VDRC
pdm2 (2)	52272 (III)	No	VDRC
Pex13	39544 (III)	No	VDRC
Pex13 (2)	50697 (III)	No	BDSC
phI	31038 (III)	No	BDSC
Pi3K21B	33556 (III)	No	VDRC
Pi3K92E (1)	38985 (III)	No	VDRC
Pi3K92E (2)	4141R-1 (III)	No	NIG
piwi	6122R-1 (III)	No	NIG
Pkd2 (1)	31296 (III)	No	BDSC
Pkd2 (2)	51502 (III)	No	BDSC
Plc21C	4574R2 (III)	No	NIG
Plc21C (2)	33719 (III)	No	BDSC

Gene	Stock No	Pilot	Source
plexA	11081R2 (III)	No	NIG
plexA (2)	4740 (III)	No	VDRC
plexB	17245R2 (III)	No	NIG
plexB (2)	6873 (III)	No	VDRC
PMCA	31572 (III)	No	BDSC
Pp2A-29B (1)	23886 (III)	No	VDRC
Pp2A-29B (2)	49671 (III)	No	VDRC
Pp2C1	40827 (III)	No	BDSC
PpV	31690 (III)	No	VDRC
Prestin (1)	5485R-1 (III)	No	NIG
Prestin (2)	5485R-3 (III)	No	NIG
pTEN (1)	35731-	Yes	VDRC
pTEN (2)	5671R-1	Yes	NIG
Ptp10D	1817R1 (III)	No	NIG
Ptp10D (2)	1101 (III)	No	VDRC
Ptp36E (1)	34369 (III)	No	VDRC
Ptp36E (2)	34368 (III)	No	VDRC
Ptp4E	1012 (II)	No	VDRC
Ptp4E (2)	38369 (III)	No	BDSC
ptp52F	3116 (III)	No	VDRC
Puc	3019 (X)	Yes	VDRC
put (1)	7904R-2 (III)	No	NIG
put (2)	37279 (III)	No	VDRC
Pvr (1)	8222R-2 (X)	No	NIG
Pvr (2)	13502 (III)	No	VDRC
Pxn	15276 (III)	No	VDRC
Pxn (2)	15277 (X)	No	VDRC
pyx	31297 (III)	No	BDSC
pyx (2)	51836 (III)	No	BDSC
R	33437 (III)	No	VDRC
Rab40 (1)	1900R-1 (III)	No	NIG
Rab40 (2)	29259 (III)	No	VDRC
Rac1 (1)	2248R-2 (X)	No	NIG
Rac1 (2)	49247 (X)	No	VDRC
Rac2 (1)	8556R-3 (III)	No	NIG
Rac2 (2)	50349 (III)	No	VDRC

Gene	Stock No	Pilot	Source
Ral	2849R-1 (III)	Yes	NIG
Ras85D	28129 (III)	No	VDRC
Rbf1	10696-	Yes	VDRC
Rbf2	36697 (III)	No	BDSC
rdx	9924R-2 (X)	No	NIG
ref(2)P	10360R-1 (III)	No	NIG
Rfabg	6879 (III)	No	VDRC
Rheb	1081R-4 (X)	No	NIG
RhoGAP19D	43955 (X)	No	VDRC
RhoGAP19D (2)	32361 (III)	No	BDSC
RhoGAPp190	28877 (III)	No	VDRC
RhoGEF3	31716 (III)	No	VDRC
RhoGEF3 (2)	31581 (III)	No	BDSC
Ric (1)	8418R-3 (III)	No	NIG
Ric (2)	35930 (X)	No	VDRC
Rim	7305R2 (III)	No	NIG
Rim (2)	39385 (III)	No	VDRC
Rm62 (1)	10279R-2 (X)	No	NIG
Rm62 (2)	46908 (III)	No	VDRC
robo	35768 (III)	No	BDSC
robo3	29398 (III)	No	BDSC
RpL5 (1)	17489R-4 (III)	No	NIG
RpL5 (2)	38929 (III)	No	VDRC
RpS15	35416 (III)	No	VDRC
rst	951 (III)	No	VDRC
rst (2)	28672 (III)	No	BDSC
run	28673 (III)	No	BDSC
run (2)	34707 (III)	No	BDSC
RunxA	33353 (III)	No	BDSC
RyR	28919 (III)	No	BDSC
S6k (1)	10539R-2 (III)	No	NIG
S6k (2)	18126 (X)	No	VDRC
S6kII	17596R-1 (III)	No	NIG
SA (1)	3423R-2 (III)	No	NIG
SA (2)	3423R-3 (III)	No	NIG
salm	33714 (III)	No	BDSC

Gene	Stock No	Pilot	Source
salr	4881R3 (III)	No	NIG
salr (2)	28386 (III)	No	VDRC
salvador	13831R-2 (III)	Yes	NIG
sax (1)	1891R-1 (III)	No	NIG
sax (2)	42457 (III)	No	VDRC
Scar	4636R-1	Yes	VDRC
Sce	5595R-1 (III)	No	NIG
Scox	8885R2 (III)	No	NIG
Scox (2)	7861 (III)	No	VDRC
Scrib (1)	45556 (III)	Yes	VDRC
Scrib (2)	5462R-2 (III)	Yes	NIG
sdk	33412 (III)	No	BDSC
Sdr	3837R1 (III)	No	NIG
Sdr (2)	44575 (III)	No	VDRC
sei	3606 (III)	No	VDRC
sema-1a	18405R1 (III)	No	NIG
Sema-1b	6446R3 (III)	No	NIG
Sema-1b (2)	28588 (III)	No	BDSC
Sema-2a	15810 (III)	No	VDRC
Sema-2b	48057 (III)	No	VDRC
Sema-2b (2)	28932 (III)	No	BDSC
Sema-5c	5661R1 (III)	No	NIG
Sema-5c (2)	1052 (X)	No	VDRC
sens-2	21386 (III)	No	VDRC
Sep2	26413 (III)	No	VDRC
Ser	27174 (X)	No	VDRC
shn	34689 (III)	No	BDSC
shot	28336 (III)	No	BDSC
sick	53315 (III)	No	BDSC
side	31062R2 (III)	No	NIG
side (2)	1284 (III)	No	VDRC
sif	5406R3 (III)	No	NIG
sif (2)	27406 (III)	No	VDRC
Sirt4	40295 (X)	No	VDRC
SK	24597 (III)	No	VDRC
SK (2)	27238 (III)	No	BDSC

Gene	Stock No	Pilot	Source
sl (1)	4200R-3 (III)	No	NIG
sl (2)	7173 (III)	No	VDRC
slbo	27043 (III)	No	BDSC
slbo (2)	53309 (III)	No	BDSC
sli (1)	31468 (III)	No	BDSC
sli (2)	31467 (III)	No	BDSC
SLIK (1)	4527R-2	Yes	NIG
SLIK (2)	4527R-3 (III)	Yes	NIG
SMC1 (2)	6532 (III)	No	VDRC
Smox	2262R-1 (III)	No	NIG
Smr	27068 (III)	No	BDSC
Smr (2)	34087 (III)	No	BDSC
Snr1	12645 (III)	No	VDRC
Sox15	45482 (III)	No	VDRC
spen (1)	18497R-1 (III)	No	NIG
spen (2)	48846 (III)	No	VDRC
Spn38F	9334R-3 (III)	No	NIG
spp	7247 (III)	No	VDRC
ss	10715 (III)	No	VDRC
ss (2)	33415 (III)	No	BDSC
stj	25807 (III)	No	BDSC
su(Hw)	8573R2 (III)	No	NIG
su(Hw) (2)	34006 (III)	No	BDSC
Su(var)2-HP2	25972 (III)	No	BDSC
su(var)2-HP2 (2)	31346 (III)	No	BDSC
Syx1A (1)	5448R-4 (III)	No	NIG
Syx1A (2)	33112 (III)	No	VDRC
Taf1	41099 (III)	No	VDRC
Taf1 (2)	32421 (III)	No	BDSC
tefu (1)	6535R-1 (III)	No	NIG
tefu (2)	22502 (III)	No	VDRC
Tehao (1)	17903 (III)	No	VDRC
Tehao (2)	7121R-3 (III)	No	NIG
Ten-a	11270R1 (III)	No	NIG
Ten-a (2)	32482 (III)	No	VDRC
Ten-m	29390 (III)	No	BDSC

Gene	Stock No	Pilot	Source
Tep1	30873 (III)	No	VDRC
Tep1 (2)	32856 (III)	No	BDSC
Tep2	18589R1 (III)	No	NIG
Tep4	13466 (III)	No	VDRC
Tep5	31866 (III)	No	VDRC
Tg	7356R1 (III)	No	NIG
Tg (2)	26100 (III)	No	VDRC
Tim17a2	11701 (III)	No	VDRC
Tim17B1	9455 (III)	No	VDRC
Tim17B2	3838 (III)	No	VDRC
TIMP1	6281R-2	Yes	NIG
tkv (1)	14026R-1	Yes	NIG
tkv (2)	14026R-3	Yes	NIG
tld	51507 (III)	No	BDSC
tol	2656 (III)	No	VDRC
Toll-7	30488 (III)	No	BDSC
Tollo (1)	9430 (III)	No	VDRC
Tollo (2)	27098 (III)	No	VDRC
TOR	5092R-1 (III)	Yes	NIG
Traf4 (1)	3048R-3 (III)	No	NIG
Traf4 (2)	55226	No	BDSC
Traf-like	34836 (III)	No	VDRC
Trax	49383 (III)	No	VDRC
Trf4-1	41096 (III)	No	VDRC
Trf4-2 (1)	19799 (III)	No	VDRC
Trf4-2 (2)	19801 (III)	No	VDRC
Trim9	21405 (III)	No	VDRC
trio	18214R1 (III)	No	NIG
trio (2)	40137 (III)	No	VDRC
trol	22642 (X)	No	VDRC
trol (2)	29440 (III)	No	BDSC
Trp1 (1)	8894 (III)	No	VDRC
Trp1 (2)	8895 (III)	No	VDRC
trr	10749 (X)	No	VDRC
trx	31092 (III)	No	BDSC
Trx-2	36298 (III)	No	VDRC

Gene	Stock No	Pilot	Source
TSC1 (1)	22252-	Yes	VDRC
TSC1 (2)	6147R-1	Yes	NIG
TTLL3B	31456 (X)	No	VDRC
tweek	19306 (III)	No	VDRC
tweek (2)	26645 (III)	No	VDRC
U2af38 (1)	3582R-2 (III)	No	NIG
U2af38 (2)	3582R-3 (III)	No	NIG
uif	38365 (III)	No	BDSC
unc-13	2999R3 (III)	No	NIG
unc-13 (2)	33606 (III)	No	VDRC
unc-5	8166R1 (III)	No	NIG
unc-5 (2)	8137 (III)	No	VDRC
unc79	51471 (III)	No	BDSC
Unc-89 (1)	29412 (III)	No	VDRC
Unc-89 (2)	29413 (III)	No	VDRC
upd-IR3282 (os)	5993R-2 (III)	Yes	NIG
usp (1)	4380R-3 (III)	No	NIG
usp (2)	16893 (III)	No	VDRC
Utx	5640R-2 (III)	No	NIG
vap	44638 (III)	No	VDRC
vav	6241 (III)	No	VDRC
vav (2)	39059 (III)	No	BDSC
Vhl	50727 (III)	No	BDSC
vkg (1)	16858R-3 (III)	No	NIG
vkg (2)	16986 (III)	No	VDRC
vn	10491R2 (III)	No	NIG
vn (2)	50358 (III)	No	VDRC
WASp (1)	13757 (III)	Yes	VDRC
WASp (2)	1520R-2 (III)	Yes	NIG
wb (1)	3141 (III)	No	VDRC
wb (2)	15288R-3 (III)	No	NIG
wit (1)	10776R-2 (III)	No	NIG
wit (2)	42244 (III)	No	VDRC
Wnt2	38077 (III)	No	VDRC
wrapper	10382R3 (III)	No	NIG
wrapper (2)	29561 (III)	No	BDSC

Gene	Stock No	Pilot	Source
wry	8021 (III)	No	VDRC
X11L-beta	1852R-3 (III)	No	NIG
Xpd	41021 (X)	No	VDRC
zfh2	13305 (III)	No	VDRC
Zir	40673 (X)	No	VDRC
Zir (2)	28005 (III)	No	BDSC
zormin (1)	11953R-2 (III)	No	NIG
zormin (2)	17563 (III)	No	VDRC

Appendix 3 - Screen hit genes

Top suppressor hits

Gene	Apex_size
CG2807	-2
su(Hw)	-1.666666667
Su(var)2-HP2	-1.666666667
Rbf1	-1.517241379
S6kII	-1.5
jbug (2)	-1.428571429
trol (2)	-1.222222222
Cap	-1.2
CTCF	-1.2
Gug	-1.2
Sema-1b (2)	-1.2
wrapper (2)	-1.2
lea	-1.142857143
Nrg (2)	-1.142857143
Sdr (2)	-1.142857143
Myosin II (sqh)	-1.111111111
asp (2)	-1.1
Blimp-1 (2)	-1.1
capt (1)	-1.1
CG15395 (2)	-1.1
CG34422 (2)	-1.1
CG9701 (2)	-1.1
Nrg (1)	-1.1
Ser	-1.1
SLIK (2)	-1.1
Akt1 (2)	-1
Blimp-1 (1)	-1
CG11180	-1
CG11180 (2)	-1
CG31731 (2)	-1
Dp (2)	-1
Ex-IR109281	-1
klg (2)	-1
Myospheroid (bPS)	-1
mys	-1
Or59b (2)	-1
pyx	-1

Gene	Apex_size (cont)
Sirt4	-1
sli (1)	-1
sli (2)	-1
SLIK (1)	-1
Toll-7	-1
trx	-1
patj (2)	-0.933333333
brv2 (1)	-0.909090909
bx (2)	-0.9
Ptp36E (1)	-0.9
run	-0.9
RunxA	-0.9
salm	-0.9
Sema-5c (2)	-0.9
shot	-0.9
ss (2)	-0.9
TSC1 (2)	-0.9
usp (1)	-0.9
arm (1)	-0.894736842
CG31999 (2)	-0.888888889
CG4393 (2)	-0.888888889
Scar	-0.888888889
tkv (2)	-0.882352941
CG34120 (1)	-0.875
CG5550 (1)	-0.875
RhoGAP19D	-0.875
Taf1 (2)	-0.875
tkv (1)	-0.85
CG42672	-0.818181818
E(z) (2)	-0.818181818

Gene	Basal Protrusions
otk	-2
lea	-1.625
CG4393 (1)	-1.6
CG7607	-1.6
su(Hw)	-1.6
CG7379 (1)	-1.545454545
Nos	-1.545454545
CG89095	-1.5
CG9279	-1.5
cv-c	-1.5
Gug	-1.5
put (2)	-1.5
Chd3	-1.4
CG4896 (2)	-1.363636364
Cg25C	-1.3
CG4896 (1)	-1.3
Cirl	-1.3
Eph (2)	-1.3
Idh (1)	-1.3
jim (2)	-1.3
CG4887 (2)	-1.272727273
cac	-1.2
msk	-1.2
CG7229 (2)	-1.181818182
Sema-2a	-1.142857143
Sdr (2)	-1.125
CG2135	-1.1
CG34113	-1.1
Dscam3	-1.1
jbug (2)	-1.1
Rac1 (2)	-1.1
cdep	-1.090909091
E(z) (2)	-1.090909091
18w	-1
CG15395 (2)	-1

Gene	Basal Protrusions
Ex-IR109281	-1
lea (2)	-1
RpL5 (2)	-1
vn	-1
CG1724 (2)	-0.9
CG4393 (2)	-0.9
CG5037	-0.9
Ephrin	-0.9
Ric (2)	-0.9
sema-1a	-0.9
CG10702 (2)	-0.8
CG16721	-0.7
CG7420 (2)	-0.7
egg (2)	-0.7
Su(var)2-HP2	-0.7
Tg	-0.7
unc-13	-0.7
Unc-89 (2)	-0.7

Gene	Cellbody_rounding
sei	-0.3
para (1)	-0.2
robo	-0.2
S6k (1)	-0.2
CG18335 (2)	-0.1
Mnt (2)	-0.1
vkg (2)	-0.1
atonal	0
CG6480	0
Elongin-C	0
Tollo (1)	0
Trf4-2 (1)	0
vap	0.090909091
CG10585 (2)	0.1
CG10621	0.1
CG32238 (2)	0.1
CG7911	0.1
Mkk4 (1)	0.1
Puc	0.1
tefu (2)	0.1
Trx-2	0.1
NF1 (2)	0.125

Gene	Multi_layering
para (1)	-0.5
sei	-0.4
robo	-0.3
Tollo (1)	-0.3
ATPsyn-beta (1)	-0.2
Dscam2 (1)	-0.2
Trf4-2 (1)	-0.2
vkg (2)	-0.2
CG11444 (1)	-0.142857143
CG10211	-0.1
CG34422 (1)	-0.1
CG4887 (1)	-0.1
E2F-IR15886 (1)	-0.1
Elongin-C	-0.1
emb (2)	-0.1
Mnt (2)	-0.1
otu (2)	-0.1
Pi3K92E (1)	-0.1
ref(2)P	-0.1
Smox	-0.1
wb (1)	-0.1

Gene	Cell_Length
brm (2)	-0.75
CG2135	-0.6
Rac2 (2)	-0.6
RpL5 (2)	-0.6
Sox15	-0.6
CG7229 (2)	-0.545454545
babo (1)	-0.5
Bap170 (2)	-0.5
calypso (2)	-0.5
CG10211	-0.5
nub	-0.5
Pak	-0.5
Snr1	-0.5

Gene	Clonal_shape
Cap	-0.7
Rbf1	-0.612903226
Blimp-1 (2)	-0.6
lea	-0.6
Sirt4	-0.6
Snr1	-0.6
Su(var)2-HP2	-0.6
tkv (1)	-0.6
TSC1 (2)	-0.6
tkv (2)	-0.555555556
calypso (1)	-0.5
calypso (2)	-0.5
CG2807	-0.5
Sdr (2)	-0.5

Gene	Clonal_tissue
tkv (2)	-1.444444444
cv-c	-1.4
Myospheroid (bPS)	-1.4
Sirt4	-1.4
Taf1	-1.4
tkv (1)	-1.35
brm (1)	-1.3
Nrg (1)	-1.3
Pxn (2)	-1.3
Snr1	-1.3
su(Hw)	-1.3
alpha-catenin (2)	-1.2
arm (1)	-1.2
CG11180	-1.2
CG13579	-1.2
put (2)	-1.2
CG5645	-1.166666667
Cap	-1.1
Chd3	-1.1
Eaat1	-1.1
GI	-1.1
lea	-1.1
Marcal1	-1.1
Myosin II (sqh)	-1.1
Ric (2)	-1.1
RpL5 (2)	-1.1
Sdr (2)	-1.1
sli (1)	-1.1
vn (2)	-1.1
Wnt2	-1.1
Rbf1	-1.03125
CG2807	-1
dia (2)	-1
Gtp-bp	-1
hbs	-1
Mcr	-1
Nrg (2)	-1
Scox (2)	-1

Gene	Individual clone_size
vn (2)	-1
Cap	-0.9
Snr1	-0.9
Taf1	-0.9
alpha-catenin (2)	-0.8
dom (2)	-0.8
SA (1)	-0.8
CG5645	-0.75
AdoR	-0.7
CG10702	-0.7
GI	-0.7
lea	-0.7
Marcal1	-0.7
Nrg (1)	-0.7
Sirt4	-0.7
Ca-alp1T	-0.6
CG11180	-0.6
CG4393 (2)	-0.6
Mcr	-0.6
Nrg (2)	-0.6
put (2)	-0.6
R	-0.6
Su(var)2-HP2	-0.6
arm (1)	-0.55
CG8197	-0.545454545
APC (2)	-0.5
bx (2)	-0.5
CG10777	-0.5
CG11180 (2)	-0.5
CG2807	-0.5
CG6153	-0.5
cv-c	-0.5
LanB1 (1)	-0.5
Mnt (1)	-0.5
Nipped-A (2)	-0.5
RpL5 (2)	-0.5
su(Hw)	-0.5
tkv (2)	-0.5

Gene	Dividing_cells	Gene	Dividing_cells (Cont)
			-0.833333333
Nrg (1)	-1.3	CG5645	
CG6153	-1.2	alpha-spec	-0.8
Myospheroid (bPS)	-1.2	CG13185	-0.8
pTEN (1)	-1.19047619	CG30417 (1)	-0.8
Myb (1)	-1.125	Chd3	-0.8
	-1.111111111		
Blimp-1 (2)		kug (1)	-0.8
ebi (2)	-1.1	mbc (2)	-0.8
Tep4	-1.1	mesh (1)	-0.8
trx	-1.1	Myosin II (sqh)	-0.8
	-1.1		-0.777777778
TTLL3B		Atx-1	
	-1		-0.777777778
18w (2)		CG5037	
atonal	-1	Nrg (2)	-0.75
ATPsyn-beta (1)	-1	pyx	-0.7
Beclin-1 (Atg6) (1)	-1	Rac1 (2)	-0.7
capt (1)	-1	Ras85D	-0.7
dia (1)	-1	RhoGEF3 (2)	-0.7
kug (2)	-1	Ric (2)	-0.7
Nlg2	-1	Snr1	-0.7
SK (2)	-1	Tg (2)	-0.7
	-1		-0.692307692
wrapper (2)		APC (2)	
	-0.933333333		-0.666666667
hig (2)		Xpd	
	-0.909090909		
otu (1)			
alpha-catenin (2)	-0.9		
brm (1)	-0.9		
Cap	-0.9		
Dhc93AB (1)	-0.9		
Ephrin (2)	-0.9		
fw	-0.9		
piwi	-0.9		
Pkd2	-0.9		
Rac2 (2)	-0.9		
Trax	-0.9		
trol	-0.9		
	-0.833333333		
Bap170 (1)			

Gene	Invading_Cells
tkv (1)	-1.05
tkv (2)	-0.941176471
trx	-0.9
CG8027	-0.818181818
Cad99C (1)	-0.8
capt (1)	-0.8
CG10777	-0.8
Mcr	-0.8
Myb (1)	-0.75
Cap	-0.7
CG10186 (1)	-0.7
CG9784 (1)	-0.7
Eph (2)	-0.7
Hml (2)	-0.7
nompC	-0.7
para (1)	-0.7
S6k (2)	-0.7
Dmel\Sep2	-0.7
Bcl2 (debcl) (1)	-0.636363636
CG7229 (2)	-0.636363636
18w (2)	-0.6
btl (1)	-0.6
CG12268	-0.6
CG12950 (2)	-0.6
CG7607 (2)	-0.6
CG8290 (1)	-0.6
E(z) (1)	-0.6
kto (1)	-0.6
kug (1)	-0.6
Myd88	-0.6
Myospheroid (bPS)	-0.6
NF1 (1)	-0.6
Pi3K21B	-0.6
put (1)	-0.6
Rab40 (2)	-0.6
Smox	-0.6

Gene	No_Apex
Ras85D	-0.9
CG13579 (2)	-0.7
Adar	-0.6
CG12268 (2)	-0.6
ln-R	-0.6
put (2)	-0.6
sei	-0.6
Trx-2	-0.6
CG5245 (2)	-0.571428571
asp	-0.5
atonal	-0.5
ATP7	-0.5
babo (1)	-0.5
CcapR	-0.5
CG10211	-0.5
CG15556	-0.5
CG34113 (2)	-0.5
fred	-0.5
mei-38	-0.5
msk	-0.5
Neto (2)	-0.5
para (1)	-0.5
PpV	-0.5
Ptp10D (2)	-0.5
RhoGAPp190	-0.5
Ric (2)	-0.5
robo	-0.5
hppy	-0.461538462

Gene	Polarised Cytoskeleton
SA (1)	-0.4
CG11444 (1)	-0.214285714
alpha-catenin (2)	-0.2
CG10777	-0.2
egg (1)	-0.2
Eph (1)	-0.2
mys	-0.2
TSC1 (2)	-0.2
usp (1)	-0.2
wit (1)	-0.2
Bap170 (1)	-0.166666667
arm (1)	-0.15
CG10274 (1)	-0.133333333
CG34113 (2)	-0.111111111
ATPsyn-beta (1)	-0.1
CG13579	-0.1
Dhc62B	-0.1
Dscam1 (1)	-0.1
Ex-IR109281	-0.1
Myosin II (sqh)	-0.1
Nrg (1)	-0.1
otu (1)	-0.090909091
DNApol-epsilon	-0.083333333
fd102C	-0.083333333
NF1 (2)	-0.066666667
tkv (1)	-0.05

Gene	Type_of_Invasion
CG12950 (2)	-0.9
bent	-0.8
CcapR	-0.7
CG10702	-0.7
tkv (1)	-0.684210526
CG11318 (2)	-0.6
E2F-IR15886 (2)	-0.6
slbo (2)	-0.6
CG32669	-0.5
ham	-0.5
Myb (1)	-0.5
para (1)	-0.5
TSC1 (2)	-0.5
CG12869	-0.444444444
jbug	-0.444444444
tkv (2)	-0.4375
18w	-0.4
CadN (2)	-0.4
CG10211	-0.4
CG10348 (2)	-0.4
CG11593 (2)	-0.4
CG14446	-0.4
CG14446 (2)	-0.4
CG7607 (2)	-0.4
Myosin II (sqh)	-0.4
otk (2)	-0.4
Rim (2)	-0.4
sif	-0.4
upd-IR3282 (os)	-0.4
Wnt2	-0.4

Top enhancer hits

Gene	Apex_size
mub	0.3
robo	0.3
CG9572 (1)	0.333333333
mgl (2)	0.375
Dgk (2)	0.4
dia (1)	0.4
para (1)	0.4
atonal	0.5
emb (2)	0.5
Myb (1)	0.75
nej	0.8

Gene	Basal Protrusions
CG10669 (1)	0.533333333
lectin-22C (2)	0.545454545
CG11593	0.6
cype	0.6
Nmdar2 (2)	0.6
p120ctn	0.6
para (1)	0.6
Rac2 (1)	0.6
Trf4-2	0.6
Myb (1)	0.625
lilli (2)	0.7
Pvr (1)	0.7
side (2)	0.7
TSC1 (1)	0.7
kug (2)	0.8
atonal	1

Gene	Cell_Length
para (2)	0.8
ptp52F	0.8
su(Hw) (2)	0.8
Trim9	0.8
Wnt2	0.8
Calx	0.818181818
pTEN (2)	0.888888889
18w	0.9
cac	0.9
CG10600 (2)	0.9
CG33288 (2)	0.9
CG5550 (2)	0.9
fra (1)	0.9
kug (1)	0.9
patj (1)	0.9
SLIK (1)	0.9
Tollo (1)	0.9
lectin-22C (1)	0.909090909
CG7420 (1)	1
Cubn (1)	1
fkh (1)	1
Myb (2)	1
pTEN (1)	1

Gene	Cellbody_rounding	Gene	Cellbody_rounding (Cont)
CG42672	1.636363636	Rim	1.8
Nrg (2)	1.666666667	run	1.8
Cad99C	1.7	S6kII	1.8
CadN (2)	1.7	sdk	1.8
cdep (2)	1.7	shot	1.8
CG10348 (2)	1.7	Tep1 (2)	1.8
CG11180 (2)	1.7	frac (2)	1.9
CG15744	1.7	ft	1.9
CG34380 (2)	1.7	Iz	1.9
CG4393 (1)	1.7	Or33a (2)	1.9
CG5037 (2)	1.7	pyx	1.9
CG5550 (2)	1.7	RhoGAP19D	1.9
CG7708	1.7	Sdr (2)	1.9
Ehbp1	1.7	Sema-2b (2)	1.9
Gug	1.7	Taf1 (2)	1.9
Mcr	1.7	Calx	1.909090909
Nmdar2 (2)	1.7	Nos	1.909090909
nompC	1.7	18w (2)	2
otk (2)	1.7	asp (2)	2
Pkd2	1.7	CTCF	2
PMCA	1.7	hiw (2)	2
Ptp4E (2)	1.7	jim (2)	2
Pvr (1)	1.7	mgl	2
Rbf1	1.7	Ser	2
Scar	1.7		
sli (1)	1.7		
Ddr	1.727272727		
klu (2)	1.727272727		
CG10600 (1)	1.75		
CG5245 (2)	1.785714286		
CG15556	1.8		
CG31731 (1)	1.8		
CG4393 (2)	1.8		
CG5984	1.8		
CG7420 (2)	1.8		
jbug (2)	1.8		

Gene	Clonal_shape
CG15269	0.7
CG30268	0.7
CG33288 (2)	0.7
CG42255	0.7
CG7420 (1)	0.7
Dhc93AB (1)	0.7
dia (1)	0.7
Dp (1)	0.7
ds	0.7
Dscam2 (2)	0.7
jub	0.7
robo	0.7
Tollo (1)	0.7
Ac78C	0.727272727
CG10186 (2)	0.733333333
kon	0.733333333
Adar	0.8
CG11836	0.8
CG13579	0.8
CG32333	0.8
CG9492	0.8
Dscam2 (1)	0.8
E2F-IR15886 (1)	0.8
egg (1)	0.8
ln-R	0.8
mesh (1)	0.8
Nlg2	0.8
para (1)	0.8
para (2)	0.8
Unc-89 (1)	0.8
wb (1)	0.8
Cubn (2)	0.8125
dlt	0.9
E2F-IR15886 (2)	0.9
Nos (2)	0.9
atonal	1
CG11318 (2)	1

Gene	Clonal_tissue
CG9784 (2)	0.333333333
SA (2)	0.333333333
CG1311 (2)	0.4
CG31999	0.4
grk	0.4
kto (2)	0.4
Nos (2)	0.4
RhoGAPp190	0.4
Trf4-1	0.4
wit (2)	0.4
wry	0.4
bbg	0.454545455
pyx (2)	0.5
zormin (1)	0.5
msk	0.555555556
CG7420 (1)	0.6
nub	0.6
sl (1)	0.6
Exn	0.7
fd102C	0.75
Mkk4 (2)	0.8
Tg	0.8
TSC1 (1)	0.8
Bcl2 (debcl) (1)	0.818181818

Gene	Cuticle
Tim17a2	0.9
tld	0.9
unc-13	0.9
zormin (1)	0.9
CG42672	0.909090909
Ddr	0.909090909
pTEN (1)	0.916666667
btl (1)	1
CG11318 (2)	1
CG15556	1
CG2807	1
CG32669	1
CG9572 (1)	1
Dgk (2)	1
Dscam1 (2)	1
Dscam3	1
Ex-IR109281	1
lea (2)	1
mgl (2)	1
mys	1
nAcRbeta-64B	1
nompC	1
Nos	1
Pex13 (2)	1
Plc21C	1
Pvr (1)	1
RhoGEF3 (2)	1
run	1
Sce	1
sdk	1
sema-1a	1
Sema-2b (2)	1
Sirt4	1
SLIK (2)	1
Taf1 (2)	1
trio (2)	1
Unc-89 (1)	1

Gene	Individual clone_size
Exn	0.4
grn (2)	0.4
if (aPS2) Integrin (1)	0.4
kto (1)	0.4
MED23	0.4
Mkk4 (2)	0.4
mub	0.4
otu (2)	0.4
Ptp10D	0.4
sick	0.4
tefu (1)	0.4
Tep1	0.4
Trf4-2	0.4
tweek	0.4
Mad (2)	0.466666667
CG11247	0.5
CG16868	0.5
CG4887 (1)	0.5
CG6867	0.5
CG8290 (1)	0.5
CG8290 (2)	0.5
lid (1)	0.5
msk	0.5
Pp2C1	0.5
RhoGAPp190	0.5
NF1 (2)	0.533333333
CG4896 (2)	0.545454545
Cirl	0.6
Nos (2)	0.6
salvador	0.6
Smox	0.6
upd-IR3282 (os)	0.6
CG4887 (2)	0.636363636
TSC1 (2)	0.7
Bcl2 (debcl) (1)	0.727272727
atonal	1
Tg	1
TSC1 (1)	1.2

Gene	Defective apex	Gene	Defective apex (Cont)
lectin-22C (1)	0.909090909	CG5550 (1)	1.1
pTEN (1)	0.916666667	CSN8 (2)	1.1
sax (1)	0.916666667	CTCF	1.1
arm (1)	0.95	Ex-IR109281	1.1
Bap170 (1)	1	ft	1.1
Blimp-1 (2)	1	Idh (1)	1.1
Ca-alpha1T	1	mbc	1.1
CG10600 (1)	1	SA (1)	1.2
CG2807	1	SA (2)	1.25
CG30280 (1)	1	CadN (2)	1.3
CG34380 (2)	1	cnc (1)	1.3
CG9279	1	cv-c	1.3
CG9701 (2)	1	emb	1.3
Eaat1	1	Scar	1.3
GI	1	Tep1 (2)	1.3
jim (2)	1	Cap	1.4
Mcr	1	fw	1.4
Mes-4	1	Rbf1	1.5
Nrg (2)	1	S6kII	1.5
nub	1	jbug (2)	1.6
Or33a (2)	1	mys	1.6
Or59b (1)	1		
Pak	1		
Ptp4E (2)	1		
Rab40 (1)	1		
RhoGAP19D (2)	1		
Sdr (2)	1		
Ser	1		
shot	1		
usp (2)	1		
vn	1		
zfh2	1		
CG18476 (2)	1.066666667		
pTEN (2)	1.076923077		
rdx	1.083333333		
Nos	1.090909091		
Blimp-1 (1)	1.1		
cac	1.1		

Gene	Dividing_cells
lr25a	0.5
nAcrbeta-64	0.5
Pxn	0.5
Ric (1)	0.5
S6kII	0.5
SA (2)	0.583333333
CG13579 (2)	0.6
CG4587 (2)	0.6
CG9492	0.6
Kirre (1)	0.6
ln-R	0.6
wb (1)	0.6
CG11318	0.666666667
CG12268	0.7
ham	0.7
MED23	0.7
mGluRA (2)	0.7
Or33a (1)	0.7
Sema-5c (2)	0.7
su(Hw) (2)	0.7
Traf4 (1)	0.7
CG11593	0.8
CG12484 (1)	0.8
CG5245	0.8
Mhc	0.8
Neto (2)	0.8
side (2)	0.8
CG34113	0.9
Dscam3 (2)	0.9
IA-2 (2)	0.9
plexB (2)	0.9
lectin-22C (2)	0.909090909
grk	1
Nmdar2 (2)	1
CG14446	1.1
CG5245 (2)	1.428571429

Gene	No apex
CG4393 (2)	0.555556
cher	0.555556
CG13579	0.6
CG15395 (2)	0.6
Dp (1)	0.6
fw	0.6
Gug	0.6
hbs	0.6
mys	0.6
sax (2)	0.6
Sema-5c	0.6
Ser	0.6
su(Hw)	0.6
CG31999	0.7
jbug	0.7
Scar	0.7
sema-1a	0.7
Unc-89 (2)	0.7
plexA	0.727273
CG30280 (1)	0.8
cv-c	0.8
Ex-IR109281	0.8
Gtp-bp	0.8
Myosin II (sqh)	0.8
otk	0.8
Rbf1	0.8
X11L-beta	0.8
pTEN (2)	0.846154
lea	0.888889
Cirl	0.9
patj (1)	0.9
18w	1
CG7420 (2)	1.1
vn	1.1
lectin-22C (1)	1.181818
Nos	1.181818
R	1.266667
jbug (2)	1.3
jim (2)	1.3

Gene	Invading_Cells	Gene	Invading_Cells (Cont)
bbg	0.727272727	mbc	1
cac	0.8	Ptp4E (2)	1
CG10931	0.8	S6kII	1
CG12268 (2)	0.8	Ser	1
CG6867	0.8	trio	1
cnn	0.8	Ptp36E (1)	1.1
Dsor2	0.8	su(Hw)	1.1
Ir8a	0.8	ct (2)	1.2
itgbn	0.8	Dscam3	1.2
Lap1 (2)	0.8	Rbf2	1.2
lea (2)	0.8	RhoGAP19D	1.2
p120ctn	0.8	Fbp2	1.3
Pex13	0.8	ih	1.3
plexA (2)	0.8	Pp2C1	1.3
shot	0.8	Sdr (2)	1.3
SK	0.8	tld	1.3
asp	0.9	CG7379 (1)	1.363636364
CG14372	0.9	CG4393 (1)	1.4
CG9572 (1)	0.9	Su(var)2-HP2	1.4
cv-c	0.9	Trax	1.5
Exn (2)	0.9		
Gtp-bp	0.9		
IP3K2	0.9		
klg	0.9		
lr25a	0.9		
LRP1 (2)	0.9		
mbc (2)	0.9		
Or59b (2)	0.9		
Pkd2 (2)	0.9		
Rac2 (1)	0.9		
Rim	0.9		
sema-1a	0.9		
Sema-2a	0.9		
Tep2	0.9		
Tep4	0.9		
Ac78C	1		
CG11836	1		
CG4096	1		
CG9572 (2)	1		
Ddr	1		

Gene	Multilayering
Ddr	1.545454545
CG10348 (2)	1.6
frac (2)	1.6
Gug	1.6
mbc	1.6
mgl	1.6
Or33a (2)	1.6
RhoGAP19D	1.6
run	1.6
Scar	1.6
Sdr (2)	1.6
sema-1a	1.6
Sirt4	1.6
sli (1)	1.6
SLIK (1)	1.6
Taf1 (2)	1.6
TSC1 (2)	1.6
CG5245 (2)	1.642857143
CG7420 (2)	1.7
Calx	1.727272727
Nos	1.727272727
CG10600 (1)	1.75
18w (2)	1.8
asp (2)	1.8
Cap	1.8
CTCF	1.8
Iz	1.8
jim (2)	1.8
mys	1.8
S6kII	1.8
sdk	1.8
Sema-2b (2)	1.8
shot	1.8
Tep1 (2)	1.8
18w	1.9
ft	1.9
Ser	1.9
pTEN (2)	1.923076923
hiw (2)	2
jbug (2)	2

Gene	Polarised Cytoskeleton
E(z) (2)	0.909090909
Asx	1
CG11836	1
cype (2)	1
Dsor2	1
Idh (1)	1
Ir8a	1
lid (2)	1
mbc (2)	1
mgl	1
Myd88	1
PpV	1
RhoGEF3 (2)	1
RyR	1
Trax	1
cnc (2)	1.090909091
AdoR (2)	1.1
CG31626	1.1
CG9492	1.1
Dys	1.1
LanB1 (2)	1.1
mei-38	1.1
Ras85D	1.1
Ric (2)	1.1
CG10274 (2)	1.2
CG4393 (1)	1.2
egg (2)	1.2
Eph (2)	1.2
nub	1.2
kto (2)	1.3
msk	1.3
Irp-1B (2)	1.4
lid (1)	1.5

Gene	Type_of_Invasion
CG10702 (2)	0.7
CG10931	0.7
CG4393 (1)	0.7
CG4393 (2)	0.7
CG4887 (1)	0.7
CG5389 (1)	0.7
CG7420 (1)	0.7
dE2f2 (1)	0.7
Dsor1	0.7
IP3K2	0.7
Lar	0.7
Mkk4 (1)	0.7
Mkk4 (2)	0.7
Ptp4E (2)	0.7
Rab40 (1)	0.7
Scar	0.7
DNApol-epsilon	0.75
CG10623 (2)	0.8
Dscam4 (2)	0.8
fas	0.8
mei-38	0.8
Myb (2)	0.8
Ras85D	0.8
Traf4 (1)	0.8
vn	0.8
Prestin (2)	0.888888889
Gtp-bp	0.9
Tehao (2)	0.9
CG18476 (2)	1
CG31626	1
Dhc62B	1

Gene	Very_Long protrusions
CG17002 (2)	0.8125
CG4887 (2)	0.818181818
cnc (2)	0.818181818
CG10669 (2)	0.857142857
CG5245 (2)	0.857142857
CG32373	0.866666667
Bro	0.9
CG10186 (1)	0.9
CG10931	0.9
CG30417 (1)	0.9
cype	0.9
fred (2)	0.9
if (aPS2) Integrin (1)	0.9
kto (2)	0.9
lilli (2)	0.9
Mkk4 (1)	0.9
Msp-300 (1)	0.9
Nmdar2 (2)	0.9
PpV	0.9
Pvr (1)	0.9
sax (2)	0.9
sei	0.9
Tollo (2)	0.9
Trf4-1	0.9
Trim9	0.9
usp (2)	0.9
zfh2	0.909090909
fkh (1)	0.916666667
Dscam2 (1)	0.933333333
Egfr (2)	0.933333333
atonal	1
CG10600 (2)	1
fw	1

Gene	Invading_cells with_protrusions
Nos	0.545454545
SMC1	0.545454545
brm (2)	0.555555556
pTEN (2)	0.555555556
pyx (2)	0.555555556
18w	0.6
APC (2)	0.6
babo (1)	0.6
brm (1)	0.6
calypso (2)	0.6
CG11836	0.6
CG32238 (2)	0.6
hbs	0.6
IA-2	0.6
kl-3	0.6
Lap1	0.6
lilli (2)	0.6
Nlg3	0.6
PAK3 (2)	0.6
sax (2)	0.6
sei	0.6
Snr1	0.6
tefu (2)	0.6
Trax	0.6
wb (2)	0.6
wit (2)	0.6
Blimp-1 (2)	0.7
S6k (2)	0.7
sl (2)	0.7
Tollo (2)	0.7
Trx-2	0.7
CG10348	0.8
CG11444 (2)	0.8
PpV	0.8
Tim17B2	0.8
Trf4-2	0.8
Tim17B1	0.9
Trim9	0.9
TSC1 (1)	0.9

Gene	closure_defects
CG15395 (1)	0.7
CG32066	0.7
CG4587	0.7
cnn	0.7
Idh (1)	0.7
IP3K2	0.7
lilli (2)	0.7
Or33a (2)	0.7
Rheb	0.7
tefu (1)	0.7
unc-13	0.7
CG15269	0.714285714
Ank2 (2)	0.75
Bro	0.75
Dhc93AB (2)	0.75
Nos (2)	0.75
Or59b (2)	0.75
Tehao (1)	0.75
Mad (2)	0.785714286
CG9701 (2)	0.8
PMCA	0.8
Tep1 (2)	0.8
Tollo (2)	0.8
Dys	0.833333333
Ank2 (1)	0.857142857
CG11836	0.875
CG9492	0.875
sif	0.888888889
CG30280 (2)	1
crb	1
mus205	1
pyx (2)	1

Gene	Junction_defects
lectin-22C (2)	0.818181818
jbug (2)	0.833333333
Sdr (2)	0.833333333
trio	0.875
Cap	0.888888889
lr25a	0.888888889
R	0.888888889
Rheb	0.888888889
CG12268	0.9
CG4438	0.9
Ephrin	0.9
Gtp-bp	0.9
Mhc	0.9
Mkk4 (1)	0.9
mtt	0.9
pan	0.9
patj (1)	0.9
slbo (2)	0.9
Tim17B2	0.9
Trf4-1	0.9
Trf4-2	0.9
plexA	0.909090909
CG11593	1
grn	1
Myd88	1
para (1)	1
pdm2 (1)	1
SLIK (1)	1
Trx-2	1

Gene	Basal_bundles
cnc (2)	0.727272727
CG15744	0.777777778
CG10931	0.8
CG1724 (1)	0.8
CG1724 (2)	0.8
CG30417 (2)	0.8
crb	0.8
fw	0.8
grk	0.8
mesh (1)	0.8
p120ctn	0.8
RyR	0.8
Sema-1b	0.8
Trf4-2	0.8
wit (2)	0.8
Bcl2 (debcl) (1)	0.818181818
CG10600 (2)	0.9
CG15117 (2)	0.9
Nos (2)	0.9
pdm2 (2)	0.9
Rab40 (2)	0.9
sl (2)	0.9
slbo (2)	0.9
Tim17B2	0.9
Trim9	0.9
Trx-2	0.9
usp (2)	0.9
zormin (2)	0.9
Ac78C	0.909090909
atonal	1
CG11593	1
sax (2)	1

Gene	Long_apical intermediate protrusions
CG4887 (2)	0.545454545
CG7379 (1)	0.545454545
CG11444 (2)	0.6
CG1724 (1)	0.6
Idh (1)	0.6
Irp-1B (2)	0.6
Lar	0.6
Myd88	0.6
pan	0.6
Rab40 (2)	0.6
RhoGAPp190	0.6
Tim17B1	0.6
SMC1	0.636363636
Dscam2 (1)	0.666666667
CG10600 (2)	0.7
CG4438	0.7
mei-38	0.7
sax (2)	0.7
Dmel\Sep2	0.7
Smox	0.7

Gene	Thick_protrusions
cnc (2)	0.636363636
SMC1	0.636363636
fkh (1)	0.666666667
CG4438	0.7
CG7420 (1)	0.7
sl (2)	0.7
Trf4-1	0.7
PpV	0.8

Gene	Blebbing mitotic_cells
Nrg (2)	0.714285714
Nos	0.727272727
CG10186 (2)	0.733333333
fkh (1)	0.75
Dscam2 (1)	0.769230769
emb (2)	0.777777778
CG14446	0.8
CG7420 (1)	0.8
Dhc36C (1)	0.8
Dhc36C (2)	0.8
ds	0.8
Dscam2 (2)	0.8
Egfr (2)	0.8
hig (2)	0.8
Hml (2)	0.8
Idh (1)	0.8
kon	0.8
mesh (2)	0.866666667
Cad99C	0.888888889
cv-c	0.9
lilli (2)	0.9
PpV	0.9
CG34120 (2)	0.923076923
CG32373	0.933333333
CG11593	1
CG33288 (2)	1

Gene	Bright_protrusions
CG32238 (2)	0.6
egg (2)	0.6
Eph (2)	0.6
lid (2)	0.6
Lpt	0.6
Myd88	0.6
nub	0.6
SK (2)	0.6
Smox	0.6
Tim17B2	0.6
CG4887 (2)	0.636363636
lid (1)	0.7
mei-38	0.7
SMC1	0.727272727
CG4438	0.8

Gene	Branched_protrusions
Idh (1)	0.7
Irp-1B (2)	0.7
mei-38	0.7
Myd88	0.7
nopo	0.7
pan	0.7
PpV	0.7
Ras85D	0.7
Tim17B2	0.7
SMC1	0.727272727
cype	0.8
Dmel\Sep2	0.8
Tim17B1	0.8
vkg (2)	0.9

Gene	Cyst_like_clones
CG10669 (2)	0.214285714
tkv (2)	0.266666667
Adar (2)	0.3
Bap170 (2)	0.3
brm (1)	0.3
CG13597 (2)	0.3
IA-2	0.3
Pak	0.3
Rab40 (2)	0.3
RpL5 (2)	0.3
Sirt4	0.3
tkv (1)	0.3
sax (1)	0.333333333
CG4887 (2)	0.363636364
calypso (2)	0.4
CG34380 (2)	0.4
CG4896 (1)	0.4
Chd3	0.4
mei-38	0.4
put (1)	0.4
Blimp-1 (2)	0.5
brm (2)	0.5
Cap	0.5
Ric (2)	0.5
Cg25C	0.6
ft	0.6
lea	0.6

Gene	Apoptotic
CG11836	1
CG31626	1
Dhc36C (2)	1
Ex-IR109281	1
lea	1
rdx	1
usp (2)	1
18w	1.1
CG31999	1.1
CSN8 (1)	1.1
cv-c	1.1
hbs	1.1
mys	1.1
Or33a (2)	1.1
Tehao (1)	1.1
Nrg (2)	1.125
Blimp-1 (2)	1.2
CG4393 (1)	1.2
cv-c (2)	1.2
S6kII	1.2
Ser	1.2
cac	1.3
CTCF	1.3
mbc	1.3
RhoGAP19D	1.3
sema-1a	1.4
CG4587	1.5
Nos	1.636363636

Gene	Basal_Actin_Rich_Spot
klg (2)	0.555555556
CG5645	0.583333333
babo (2)	0.6
Ca-alpha1T	0.6
CG30417 (2)	0.6
emb (2)	0.6
ln-R	0.6
Or59b (2)	0.6
Rim (2)	0.6
rst (2)	0.6
sema-1a	0.6
capt (1)	0.7
CG11247	0.7
CG15117 (2)	0.7
CG32238 (2)	0.7
Dgk	0.7
dia (1)	0.7
Ephrin (2)	0.7
Ptp36E (1)	0.7
robo3	0.7
Sox15	0.7
Su(var)2-HP2	0.7
Trax	0.7
AdoR (2)	0.8
babo (1)	0.8
CG1311 (2)	0.8
CG18335 (2)	0.9

Gene	Single_Cell_Clones
cnc (2)	0.727272727
Dp (2)	0.733333333
CG9784 (2)	0.75
CG34422 (2)	0.777777778
calypso (2)	0.8
Cap	0.8
CG7420 (1)	0.8
CG9701 (2)	0.8
Gtp-bp	0.8
LanB1 (1)	0.8
nub	0.8
pdm2 (2)	0.8
Rac2 (2)	0.8
SA (1)	0.8
Scar	0.8
Ser	0.8
Snr1	0.8
wit (2)	0.8
klu (2)	0.818181818
SMC1	0.818181818
vap	0.818181818
LanB1 (2)	0.9
mei-38	0.9
mys	0.9
S6kII	0.9

Lethal genes

Gene
bel
ca-alpha1D
CG6272 (1)
ClC-c
Cont
cora
Dis3
Dlg
dom
E-cadherin (shg)
eIF-2beta
how
Hsp83
MAGE
Mrtf
na (2)
Nrx-IV (2)
osa (1)
Pp2A-29B (2)
Rm62 (2)
salr (2)
Scrib
Smr
spen (1)
Syx1A (1)
U2af38 (1)
unc-5
WASp (1)

Appendix 4 – Hierarchical clustering GO terms (BP)

Cluster 1

GO.ID	Term	p-value	genes
GO:0022008	neurogenesis	0.00017	c("Pp2A-29B", "SA", "salr", "babo", "cora", "shg(1)", "Hsp83", Nrx-IV, "eIF-2beta", "brm", "Cont", "Snr1", "Rm62", "Taf1", "put", "osa", "Dis3", "WASp", "spen", "mys", "capt", "dom")
GO:0006357	regulation of transcription from RNA pol...	0.00051	c("Marcal1", "tkv", "salr", "babo", "Hsp83", "Blimp-1", "brm", "Snr1", "Taf1", "osa", "cic")
GO:0061343	cell adhesion involved in heart morphoge...	0.00116	c("cora", "Nrx-IV", "Cont")
GO:0048190	wing disc dorsal/ventral pattern formati...	0.00116	c("osa", "capt", "cic")
GO:0021782	glial cell development	0.00116	c("cora", "Nrx-IV", "Cont")
GO:0008586	imaginal disc-derived wing vein morphoge...	0.00184	c("salr", "brm", "Snr1", "osa", "spen")
GO:0008587	imaginal disc-derived wing margin morpho...	0.0021	c("brm", "Snr1", "osa", "mys")
GO:0035194	posttranscriptional gene silencing by RN...	0.00223	c("Rm62", "dom", "bel")

Cluster 2

GO.ID	Term	p-value	genes
GO:0016192	vesicle-mediated transport	0.0026	c("cac", "Rab40", "Cdc42", "Ex", "SCAR", "Dscam1", "Cirl", "Rim", "Cad99C", "unc-13")
GO:0007156	homophilic cell adhesion via plasma memb...	0.003	c("fw", "ft", "Dscam3", "Cad99C", "hbs")
GO:0045792	negative regulation of cell size	0.0063	c("Myb", "Sema-1a", "Pten")
GO:0001736	establishment of planar polarity	0.0094	c("fw", "ft", "Pten", "jbug", "Patj", "hbs")

Cluster 3

GO.ID	Term	p-value	genes
GO:0007350	blastoderm segmentation	0.00025	c("Moe", "upd1", "Mer", "run", "kis", "lilli", "Mad", "grk", "csw", "cic")
GO:0045935	positive regulation of nucleobase-contai...	0.00026	c("p53", "trr", "otu", "run", "Mad", "Scox", "shn", "egg", "E2F(2)", AdoR, "PlexA")
GO:0001745	compound eye morphogenesis	0.0006	c("trr", "rst", "Rala", "Moe", "Krn", "upd1", "Vav", "Mer", "lilli", Mad, "fred", "bdl", "Cyt-c-d", "CadN", "shn", "put", "Pten", "Apc", "csw", "CSN8")
GO:0044712	single-organism catabolic process	0.00088	c("dob", "ago", "bmm", "Atg6", "Pten")
GO:0000122	negative regulation of transcription fro...	0.0014	c("Mnt", "kis", "E2f2", "shn", "Blimp-1", "E2F(2)", "cic")
GO:0050829	defense response to Gram-negative bacter...	0.00181	c("Rala", "upd1", "Traf4", "kug", "Atg6", "Spn38F")
GO:0043065	positive regulation of apoptotic process	0.00208	c("p53", "Traf4", "Cyt-c-d", "Dp", "E2F(2)")
GO:0061564	axon development	0.00226	c("Moe", "if", "Vav", "run", "kis", "CadN", "Dscam1", "Wnt2", Ptp52F, "wrapper", "ago", "eIF-2beta", "put", "Akt1", "PlexA", "sif", "Pten", "Apc", "CG5092")

Cluster 4

GO.ID	Term	p-value	genes
GO:0044711	single-organism biosynthetic process	0.0024	c("usp", "IP3K2", "CG5037", "Dp", "ATPsynbetaL", "CG7708")
GO:0010628	positive regulation of gene expression	0.0058	c("Rheb", "usp", "Bap170", "Dp", "Taf1", "trx", "Pi3K92E", "cnc")
GO:0007423	sensory organ development	0.0072	c("sdk", "Moe", "ebi", "ft", "salm", "bun", "CadN", "ham", "sax", Asx, "Ehbp1", "klu", "rdx", "ss", "klg", "Sce", "CSN8")
GO:0016567	protein ubiquitination	0.0093	c("hiw", "CG10916", "rdx", "Sce")
GO:0045596	negative regulation of cell differentiat...	0.0093	c("sdk", "usp", "sl", "salm", "bun", "CadN")

Cluster 5

GO.ID	Term	p-value	genes
GO:0044711	G-protein coupled receptor signaling pat...	0.0011	c("CG15744", "Dgk", "mtt", "CG13579", "CG15556")
GO:0010628	locomotion	0.0027	c("Ten-a", "Neto", "hiw", "CG6867", "run", "Pvr", "otk", "unc-5", "Sema5c", "jbug", "Alk")
GO:0007423	TOR signaling	0.0028	c("Pvr", "hppy", "Tsc1")
GO:1903046	meiotic cell cycle process	0.0045	c("SA", "tefu", "asp")
GO:0006811	ion transport	0.0092	c("Nmdar2", "CG15627", "AChR(2)", "RyR")

Cluster 6

GO.ID	Term	p-value	genes
GO:0007417	central nervous system development	0.0018	c("robo3", "bun", "pdm2", "cnn", "Abi", "mbc", "Ephrin", trio, "robo2", "ct", "Apc")
GO:0035011	melanotic encapsulation of foreign targe...	0.0018	c("Zir", "dia", "Itgbn", "Rac2", "Exn", "Ephrin")
GO:0006909	phagocytosis	0.0028	c("Bx", "Zir", "Tep2", "Itgbn", "Nipped-A", "Rac2", "Nlg3", "Abi")
GO:0055114	oxidation-reduction process	0.0053	c("Pp2C1", "Fbp2", "CG30268", "klu", "Idh")
GO:0022412	cellular process involved in reproductio...	0.0081	c("Dsor1", "TTLL3B", "piwi", "bun", "dia", "Pex13", "cnn", "slbo", kst, "Rac2", "bbg", "Abi", "mbc", "crb", "tld", "nej", "ct", "Apc")

Cluster 7

GO.ID	Term	p-value	genes
GO:0008045	motor neuron axon guidance	0.00014	c("Ptp4E", "Ptp10D", "NetA", "Lar", "fra", "Cdk5", "grn", "side", "PlexB")
GO:0014706	striated muscle tissue development	0.00562	c("kirre", "vkg", "Msp300", "hbs")
GO:0007156	homophilic cell adhesion via plasma memb...	0.00586	c("kirre", "da", "Lar", "Dscam2", "Dscam3", "hbs")
GO:0071702	organic substance transport	0.00618	c("Ent1", "CG7420", "lilli", "Apoltp", "Lap1", "Prestin", "CG18402", "crb", "WASp")
GO:0008344	adult locomotory behavior	0.00639	c("Actn", "Plc21C", "Cirl", "Cdk5", "Nf1")
GO:0048812	neuron projection morphogenesis	0.00999	c("sav", "Ptp4E", "Ptp10D", "NetA", "Vav", "kis", "Lar", "sick", "fra", "Cdk5", "Dscam2", "Ank2", "grn", "Mkk4", "CG18402", "E2F(2)", "Apc2", "crb", "side", "WASp", "PlexB", "sif", "jim", "capt")

Cluster 8

GO.ID	Term	p-value	genes
GO:0000381	regulation of alternative mRNA splicing,...	0.00023	c("U2af38", "Rm62", "how", "dom")
GO:0019991	septate junction assembly	0.00122	c("Nrg", "dlg1", "Cont", "scrib")
GO:0008366	axon ensheathment	0.00137	c("Nrg", "Cont", "how")
GO:0048863	stem cell differentiation	0.00234	c("arm", "vn", "osa", "scrib", "robo2")
GO:0007475	apposition of dorsal and ventral imagina...	0.00263	c("shot", "how", "mys")
GO:0022008	neurogenesis	0.00334	c("arm", "Nrg", "dlg1", "S6kII", "Ca-alpha1D", "hig", "shot", Su(var)2-HP2, "unc-5", "CG6272", "Cont", "Rm62", "MAGE", "osa", "how", "scrib", "robo2", "mys", "na", "dom", "Gtp-bp")
GO:0022407	regulation of cell-cell adhesion	0.00443	c("arm", "dlg1", "robo2")
GO:0016203	muscle attachment	0.00498	c("CAP", "shot", "vn", "mys")
GO:0001708	cell fate specification	0.00653	c("arm", "S6kII", "vn", "robo2")
GO:0007379	segment specification	0.00683	c("CTCF", "osa", "Gug")

Cluster 9

GO.ID	Term	p-value	genes
GO:0061647	histone H3-K9 modification	0.00033	c("lid", "Lpt", "egg", "E(z)")
GO:0035220	wing disc development	0.00063	c("nub", "Bap170", "sax", "babo", "Sox15", "Lpt", "egg", "Rac1")
GO:0016319	mushroom body development	0.00138	c("usp", "vap", "RhoGAPp190", "babo", "SMC1", "Eph")
GO:0035099	hemocyte migration	0.00371	c("Pvr", "Rac1", "Rac2", "Ras85D")
GO:0048813	dendrite morphogenesis	0.00427	c("usp", "nub", "babo", "Rac1", "E(z)", "brm", "cnc", "SMC1", "wit")
GO:0008586	imaginal disc-derived wing vein morphoge...	0.00553	c("Bap170", "Lpt", "brm", "Ras85D", "spen", "msk")
GO:0045610	regulation of hemocyte differentiation	0.00622	c("Pvr", "Rac1", "Ras85D", "pan")
GO:0045448	mitotic cell cycle, embryonic	0.00689	c("nopo", "Xpd", "E(z)")
GO:0006626	protein targeting to mitochondrion	0.00689	c("CG1724", "Tim17b1", "Tim17b2")

Cluster 10

GO.ID	Term	p-value	genes
GO:0006979	response to oxidative stress	0.00042	c("ebi", "Trx-2", "CG10211", "S6k", "puc", "Pi3K92E")
GO:0010906	regulation of glucose metabolic process	0.00083	c("Adar", "Mnt", "Myb", "Egfr", "slbo", "Pi3K92E", "ATPsynbeta", Tollo)
GO:0031400	negative regulation of protein modificat...	0.00166	c("l(2)gl", "ebi", "FRG1", "puc")
GO:0046622	positive regulation of organ growth	0.0031	c("jub", "ebi", "S6k", "Pi3K92E")
GO:0070925	organelle assembly	0.0045	c("Myb", "Unc-89", "emb", "Mhc", "kon", "dia", "Nph")
GO:0007298	border follicle cell migration	0.00548	c("Raf", "sl", "baz", "Mhc", "Egfr", "slbo", "puc")
GO:0045936	negative regulation of phosphate metabol...	0.00599	c("l(2)gl", "ebi", "puc")
GO:2000112	regulation of cellular macromolecule bio...	0.007	c("Adar", "Raf", "Mnt", "Smox", "otu", "Myb", "ebi", "slbo", "FRG1")
GO:0001666	response to hypoxia	0.00774	c("Adar", "para", "emb")
GO:0051489	regulation of filopodium assembly	0.00776	c("baz", "kon", "dia")

Appendix 5 – Differentially expressed genes in STAG2 KD

STAG2 vs Untreated MCF7

Gene Symbol	p-value(Attribute)	Fold-Change(SA2 vs. Unt)
ACACB	0.00113836	-2.16129
AKR1B10	0.00983365	1.69266
AKR1B15	0.00013521	1.79308
AKR1C1	0.00245998	1.85242
AKR1C2	0.000920622	2.14892
C14orf132	0.00456957	-1.52615
CDR1	0.0011283	2.04201
CDR1	0.00305863	2.64551
CHN1	0.00386543	-1.59707
CSRNP2	0.00123759	-1.65863
EHD2	0.000324219	-1.97
GLO1	0.00912099	-1.62795
GLYR1	8.39E-05	-1.81307
HCAR1	0.0030783	-1.5732
LEF1	0.00322183	-1.7041
LXN	0.00289594	-2.26954
MINPP1	0.00308907	-1.66218
NPNT	0.000445752	1.55047
PCDH1	0.0039482	-1.51467
PCDH19	0.00158893	1.6463
PSMA7	0.00233964	-1.60218
SFT2D1	0.00779402	-2.01992
SHCBP1	0.00386107	-1.5
SLC37A2	0.00175792	-1.79236
SNX3	0.00401712	-1.57342
SOAT1	0.0033291	-1.70103
SPECC1L-ADORA2A	0.00286362	-1.81634
STAG2	0.00287305	-5.92288
TMEM107	0.00107138	-1.53369
TNRC6A	0.00625051	1.68854
UBL5	0.00162786	-1.64067
UST	0.00525705	-1.6965
VTCN1	0.000118633	-2.97795

STAG2 vs Non-targeting siRNA MCF7

Gene Symbol	p-value(Attribute)	Fold-Change(SA2 vs. Non-T)
ACACB	0.00213836	-1.82882
ADCY9	0.00137272	-1.52667
AKR1B10	0.00983365	1.6678
AKR1B15	0.00013521	1.73804
C14orf132	0.00456957	-1.6732
CARNMT1	0.00135822	1.83674
CHN1	0.00386543	-1.51582
CSRNP2	0.00123759	-1.55182
DCAF7	0.0027407	-1.56759
EHD2	0.000324219	-1.72594
FGD1	0.00280366	-1.6167
GHR	0.00141735	1.57566
GLYR1	8.39E-05	-1.82847
HIF1A	0.00338725	1.56291
HSPG2	0.00133605	-1.82191
LEF1	0.00322183	-1.56792
MINPP1	0.00308907	-1.73809
NPNT	0.000445752	1.59168
PCDH1	0.0039482	-1.96819
PCP4	0.00169416	1.9346
PSMA7	0.00233964	-1.57631
RDH10	0.00196145	1.6571
SFT2D1	0.00779402	-1.94641
SHCBP1	0.00386107	-1.56969
SNX3	0.00401712	-1.57723
SPECC1L-ADORA2A	0.00286362	-1.99992
STAG2	0.00287305	-3.73787
TESK2	0.00902922	-1.52983
TMEM107	0.00107138	-1.54311
UBL5	0.00162786	-1.63933
UST	0.00525705	-1.91902
VTCN1	0.000118633	-2.10456

Non-targeting siRNA MCF7 vs Untreated MCF7

Gene Symbol	p-value(Attribute)	Fold-Change(Non-T vs. Unt)
EFR3B	0.000434199	-1.50088
PUDP	0.00539124	1.71077
STRADB	0.00611152	1.59278
HCN2	0.00903802	1.52124