A genetic analysis of tumour progression in *Drosophila* identifies the cohesin complex as a suppressor of individual and collective cell invasion

Brenda Canales Coutino†, Zoe E. Cornhill†, Africa Couto†, Natalie A. Mack†, Alexandra D. Rusu†, Usha Nagarajan†, Yuen Ngan Fan†, Marina R. Hadjicharalambous†, Marcos Castellanos Uribe†, Amy Burrows†, Anbarasu Lourdusamy†, Ruman Rahman†, Sean T. May†, Marios Georgiou*†

Affiliations:

1 School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK
2 School of Biosciences, University of Nottingham, Sutton Bonington, Leicestershire LE12 5RD, UK
3 School of Chemical and Biotechnology, SASTRA University, Thanjavur, 613401, India
4 Faculty of Biology, Medicine & Health, University of Manchester, Manchester M13 9PL, UK
5 Department of Pharmacy and Pharmacology, University of Bath, Bath, BA2 7AY, UK
6 School of Medicine, University of Nottingham, Nottingham NG7 2UH, UK

*Corresponding author and lead contact: marios.georgiou@nottingham.ac.uk
†These authors contributed equally to this work.
Abstract:
Metastasis is the leading cause of death for cancer patients. Consequently it is imperative that we improve our understanding of the molecular mechanisms that underlie progression of tumour growth towards malignancy. Advances in genome characterisation technologies have been very successful in identifying commonly mutated or misregulated genes in a variety of human cancers. However, the difficulty in evaluating whether these candidates drive tumour progression remains a major challenge. Using the genetic amenability of Drosophila melanogaster we generated tumours with specific genotypes in the living animal and carried out a detailed systematic loss-of-function analysis to identify conserved genes that enhance or suppress epithelial tumour progression. This enabled the discovery of functional cooperative regulators of invasion and the establishment of a network of conserved invasion suppressors. This includes constituents of the cohesin complex, whose loss-of-function either promotes individual or collective cell invasion, depending on the severity of effect on cohesin complex function.

Keywords: Cancer, tumour suppressor, invasion, collective invasion, metastasis, Drosophila melanogaster, cell polarity, cell-cell junctions, cohesin
Introduction

Metastasis is the major cause of mortality in human cancers, yet we know relatively little about the biology that underlies the important transition to invasive malignancy [1, 2] and currently few genes have been identified that suppress this process [3, 4]. Most human cancers are epithelial in origin; consequently cancer cell invasion, where individual cells or groups of cells break away from the primary tumour to invade the surrounding tissue, is a key hallmark of tumour progression.

Invasion is highly complex, involving concurrent dramatic changes in cytoskeletal organisation, cell polarity, cell-cell junctions and focal contacts, as cells within the developing tumour collectively destroy the normal architecture of the host epithelium and deregulate the local microenvironment [5]. Understanding and dissecting the molecular mechanisms that promote tumour progression and cancer cell invasion will be important for the development of new therapeutic strategies in our battle against this disease.

Drosophila melanogaster has become an increasingly important model system in the study of cancer biology. Conservation of major signalling pathways related to tumourigenesis and metastasis, coupled with the genetic amenability of this organism, has directly led to advances in our understanding of this disease [6, 7]. Its short lifespan and low running costs make the organism particularly amenable to large scale screens, and there is now a vast array of published literature using the fly to study cancer [6, 8, 9].
We have developed a novel in vivo system in *Drosophila* that allows us to study epithelial cell and tissue morphogenesis in real time [10-13]. This system allows the shape, dynamics and behaviour of labelled mutant epithelial cells to be followed in high resolution in the living animal. In this current study, we use this in vivo system to generate tumours with specific genotypes on the dorsal thorax epithelium of the fly and to observe tumour cell morphology and behaviour in high spatial and temporal resolution. Although several large-scale cancer screens have been carried out in the fly (for example [14-18]) our focus was to image and detail primary tumour behaviour and progression in the living animal. By combining sophisticated *Drosophila* genetic techniques with transgenic RNAi technology we present here a detailed systematic loss-of-function analysis that has identified novel genes that enhance or suppress tumour progression in this epithelium. We identify a number of conserved invasion suppressors that promote tumour cell invasion upon loss of expression. We further characterise components of the cohesin complex, which we find to be an important invasion suppressor and show that cohesin loss-of-function can promote either individual or collective cell invasion, depending on the subunit that is mutated and the degree of effect on cohesin function.

**Results**

We developed an in vivo genetic system in the fly that allows us to: (1) generate a patch of tissue on the dorsal thorax that is homozygous mutant for a tumour suppressor, surrounded by wild-type tissue; (2) specifically label the mutant tissue with GFP:Moe (the actin binding domain of moesin fused to GFP), thereby labelling the actin cytoskeleton of these cells; (3) overexpress an RNAi transgene to deplete...
expression of a gene of interest specifically within the mutant, labelled tissue.

Coupled with our ability to image this epithelium in the living animal in high
temporal and spatial resolution [13], this system allowed us to conduct a large-scale
genetic screen to identify genes that affect tumour behaviour and tumour
progression in a wide variety of ways.

**Design of an in vivo assay to identify modulators of epithelial tumour progression**

We combined the Flp/FRT system [19] the MARCM technique [20] and Pannier-Gal4
to generate positively marked homozygous mutant clones specifically within the
epithelium of the fly pupal notum (the dorsal thorax). When imaging GFP:Moe
labelled WT clones within the pupal notum (at 20-24h APF [after puparium
formation]) we observed columnar epithelial cells that formed an organised
monolayer on the back of the fly (Figure 1a-a’). Preparatory experiments identified
lethal (2) giant larvae\(^4\) homozygous mutant clones (lgl\(^4\)) as a suitable genetic
background for our screen, as tumours lacking lgl were large, partially multilayered,
and presented a low-level invasive phenotype, representing an ideal scenario for an
enhancer/suppressor screen (Figure 1b-d). Lgl is highly conserved, critical for the
correct maintenance of cell polarity, and has also been found to control tissue
growth and differentiation [21]. Lgl is a member of the scribble polarity complex
(lgl, scribble, dlg) which have been termed ‘neoplastic’ tumour suppressors due to
the fact that mutations in these genes can generate highly disorganised
multilayered tumours that are immortal, fail to differentiate, and show a high
metastatic potential upon transplantation [22, 23]. In addition, expression of
scribble complex genes has been shown to be lost or downregulated in numerous types of human cancer [24].

Although multilayered, amorphous and invasive overgrowth is observed in lgl, scribble or dlg mutant tissue, overgrowth is not observed when small mutant clones are generated, surrounded by WT tissue; here clones are restrained from overgrowth via a process known as ‘cell competition’. Mutant cells, despite undergoing excessive cell proliferation, are eliminated from the epithelium by Jun N-terminal kinase (JNK) pathway-mediated apoptosis [25, 26]. Both scribble and lgl mutants have previously been shown to cooperate with oncogenic Notch overexpression to overcome the effects of cell competition and cause neoplastic overgrowths within the proliferative epithelial primordia known as the imaginal discs [25, 27]. We wanted to see whether we could observe a similar cooperative effect within the pupal notum, which at the developmental stage of our analysis (20–24h APF), is largely post-mitotic. When generating GFP:Moe-labelled clones of cells expressing activated Notch (N\text{intra}) in the notum, we observed relatively normal clones, with no effect on cell shape nor tissue organisation, and with no invasive characteristics (Figure 1e and i-j). When overexpressing N\text{intra} specifically within lgl clones however, we observed a strong cooperative effect – these clones showed strong hyperproliferation, with increased levels of cell division, loss of normal epithelial architecture, and with increased invasion when compared to lgl alone (Figure 1f-j). We therefore had generated an in vivo system that would allow us to identify mutations that work cooperatively with lgl to promote tumour progression.
During an initial pilot screen, candidate genes previously implicated in cancer were studied. These genes were well characterised and therefore were very likely to present a phenotype. Also included were negative controls, i.e. RNAi lines to genes that are not normally expressed in this tissue. We used transgenic UAS-RNAi lines, which together with pannier-Gal4 and MARCM, allowed us to restrict gene knockdown to lgl\(^{4}\) mutant tissue on the notum of the fly (Figure 2a). We used RNAi lines from two near genome-wide RNAi libraries (VDRC, Austria and NIG, Japan) and where possible used two independent RNAi transgenes to knock down gene expression for each gene. In total, the pilot consisted of 67 RNAi lines targeting 46 well-known genes (see Table S1 for a list of pilot genes). These candidates included various oncogenes, tumour suppressor genes, MMPs, and regulators of cell morphogenesis, with a range of biological functions (Figure 2b).

We observed a wide range of phenotypes in the pilot screen including hyperproliferation, multilayering, invasion, and effects on subcellular structures (junctions, microvilli, basal protrusions; Figure 2c-k). Negative controls failed to generate significant phenotypes. We saw a range of expected phenotypes, for example: increased clonal coverage following RNAi of the known tumour suppressor, Tsc1 (a negative regulator of Tor signalling); reduced clonal coverage following RNAi of a known promoter of the cell cycle, tkv (promotes Dpp signalling); increased multilayering following RNAi of the polarity determinants scrib, expanded and dlg; smaller apices following RNAi of Cdc42, as has been observed previously [10] (Table S1).
Following the successful completion of the pilot screen, we went on to screen a total of 764 RNAi lines corresponding to 497 individual genes. Recent advances in genome characterisation technologies have uncovered a plethora of candidate genes across numerous tumour types that have been found to be commonly mutated or misregulated in human cancers [28-30]. However, other than being implicated by these new technologies, many are completely uncharacterised. By screening *Drosophila* orthologues of these previously implicated cancer genes we sought to determine which of these genes affect tumour behaviour and drive tumour progression in our system.
Systematic high-throughput scoring and quality control

We generated a database, whereby we could systematically score specific aspects of tumour behaviour, allowing us to record an extremely detailed analysis of how each gene knockdown affected tumour behaviour (see Table S1 for full database). This database consists of 33 phenotypic categories where each animal with lgl^4+ RNAi knockdown clones is scored relative to animals with lgl^4 clones alone. Each category describes an aspect of tumour behaviour. Categories include clone size and shape, number of dividing cells, number of invading cells, apex size, junction defects, cytoskeletal defects, multilayering etc. The scoring system we employed reflected the fact that gene knockdown could either positively or negatively affect specific aspects of tumour behaviour (Figure S1). A minimum of 5 animals were analysed per gene knockdown and each animal was scored blind by two researchers. An online searchable database with all results from the screen, including all high-resolution images for each RNAi line, is available at https://flycancerscreen.nottingham.ac.uk (*see footnote below*)

To verify that our high throughput qualitative scoring system gave meaningful results that represented real changes in tumour behaviour, we performed a careful quantitative analysis on a selection of genes chosen at random for categories that were amenable to a simple quantitative analysis. As shown in Figure S2a-d, a strong positive correlation was observed for all categories measured (0.91 – 0.97, Spearman correlation test).

** This site is not yet publicly available. To access the site go to: https://flycancerscreen.nottingham.ac.uk/wp-admin and use the following login details: Username: reviewer; Password: flycancerscr33n Click on the Fly Cancer Screen link in the top left menu to access the site.
To further evaluate the quality of our dataset, we asked whether two independently generated RNAi lines targeting the same gene produced similar phenotypes. We compared scores across categories for each pair of RNAi lines and found that, of the 256 genes that were targeted by two independent RNAi lines, 224 (87.5%) gave statistically similar phenotypes (Figure S2e-j; Table S2).

**Identification of genes that affect tumour behaviour**

We used an unbiased approach to identify candidate genes that increase or decrease specific aspects of tumour progression in our system. We calculated a mean score for each of the 764 RNAi lines across each of the 33 phenotypic categories (see [https://flycancerscreen.nottingham.ac.uk](https://flycancerscreen.nottingham.ac.uk)). Using these averages, we determined the distribution of scores for all 33 categories. Genes with a mean score above or below the interquartile range from the median were selected as genes of interest. For categories with a two-tailed distribution we were able to identify genes that when knocked down, either positively or negatively regulate a specific aspect of tumour behaviour. For example, using this methodology we identified 66 RNAi lines that promote, and 49 RNAi lines that inhibit cancer cell invasion (mean scores range from +0.73 to +1.5, and -0.55 to -1.2, respectively). See Table S3 for a full list of hits for all categories.

In order to identify genes that regulate similar or related cell behaviours, we clustered RNAi lines based on phenotypes presented across all categories. This resulted in the identification of ten phenotypic clusters (Figure 3a). Analysis of the
hierarchical clustering revealed, for example, that Cluster 8 shows decreased clonal
tissue and increased tissue multilayering and cell body rounding (Figure 3a). Gene
ontology (GO) term analysis shows enrichment in junction assembly, cell adhesion,
cell differentiation and fate specification factors (Table S4). A more general
categorisation of gene function reveals an increase in apicobasal polarity and cell-
adhesion factors (Figure S3). Therefore, Cluster 8 includes factors that are crucial to
the maintenance of an ordered, monolayered and polarised epithelium. Thus,
cluster analysis reveals groups of genes with similar overall phenotypes that may
share similar or related molecular functions. Within these groups lie several
uncharacterised genes that we can classify as novel tumour suppressors.

We additionally clustered categories based on phenotypes presented across all
RNAi lines and identified three distinct category clusters (Figure 3b). 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.
We were particularly interested in the identification of novel genes that promote
cancer cell invasion. Interaction networks have become a powerful tool to identify
novel disease-associated genes [31]. To generate a functionally validated interaction
map of invasive genes, we combined all hits in three categories that clustered
strongly together (Figure 3b): 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 4). The resulting network includes 321 interactions between 140 genes, 99 of which have not been previously implicated in cancer cell invasion or migration, including 9 genes that are completely uncharacterised.

Using MCODE (Molecular Complex Detection) software [32] we found seven clusters of highly interconnected nodes (Figure 4). 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; Complex 5 is enriched in axon guidance molecules, whilst other identified complexes are enriched in 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.

Characterisation of invading cancer cells

With the aim of characterising the behaviour of individual invading cells, we followed cells within mutant clones over time, prior to, during and post-invasion. We found, in all genotypes studied, that pre-invasive cells would round up and form a characteristic actin-rich spot at one side of the cell prior to invasion (Figure 5a, Movie S1). By calculating the coefficient of determination using Spearman’s rho ($r_s$) we observed a high to moderate positive correlation between a polarised actin accumulation and invasion in all genotypes studied, irrespective of whether the
mutant clones were rarely invasive or highly invasive (Figure 5b-d). The number of cells presenting this polarised phenotype within the epithelial sheet is therefore an indicator of invasive potential.

A major advantage of our in vivo model is that the directionality and speed of invading cells can be studied and quantified in real time (Figure 5a-i). It was notable that in many cases, invading cells, although viable, have no directionality to their migration and randomly move about over a number of hours (Figure 5a, Movie S1). However in some cases, as in the case of SA1KD, invading cells appear to be very motile (Figure 5e-i, Movie S2). Single cell tracking of lgl4 and SA1KD invading cells was performed to determine the X, Y and Z trajectories and to calculate their speed and directionality. An illustration of representative trajectories is shown in Figure 5f-f’. To determine directionality, the trajectory of each cell was measured over 30 min. The total number of micrometres travelled was documented (Length in Figure 5g-h) as well as the distance an invading cell would have travelled if following a straight line (Displacement in Figure 5g-h). Figure 5h shows a significant increase in length and displacement for SA1KD cells (41.55µm length, p<0.01; 26.55µm displacement, p<0.05) when compared to lgl4 cells (16.07µm length; 4.16µm displacement). There is no significant difference between length and displacement in SA1KD cells, indicating that their trajectories are directional. Additionally, the speed of migration for SA1KD cells was 2.7-fold higher (1.46µm/min, p<0.01) when compared to lgl4 invading cells (0.53µm/min; Figure 5i). It also became apparent that those cells that migrated in a fast, directional fashion did not possess a single actin-rich spot, but multiple dynamic actin-rich spots (Figure 5e) and quantification of migrating cells
showed that those cells with multiple spots migrated at a significantly faster rate.

We additionally found that a low proportion of $lg^{4}$ invading cells can possess multiple actin-rich spots, which also migrate in a directional fashion (Figure 5j-k) indicating that this change in cytoskeletal organisation and behaviour is important to promote directional migration, irrespective of mutant background.

When imaging pre-invasive and invading cells in the $xz$ plane, we found that cells that are still attached to, or within, the epithelial sheet show very limited lateral movement, and only migrate once they are fully detached from the sheet (Figure 5l-m). We additionally found that invading cells detach from the epithelial sheet more readily in SA1KD clones than in $lg^{4}$ clones, which corresponds with SA1KD clones being highly invasive, with invading cells that exhibit directional migration (Figure 5n).

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 [33, 34]. This is in stark contrast to the behaviour of invading cells within highly invasive tumours in our screen, where invasion is observed irrespective of the clone’s position within the epithelial sheet, and invading cells do not undergo immediate cell death (we have imaged invading cells for up to 2-hours without observing cell death; for example see Figure 5a and Movie S1). To specifically test for the viability of invading cells within highly invasive tumours, we used the genetically encoded apoptosis reporter iCasper [35]. We expressed iCasper within WT clones, $lg^{4}$ clones, and in clones for five strong hits for
invasion from our screen, namely: \textit{lgl}^4; \textit{CG12268KD}, \textit{lgl}^4; \text{RhoGAP19DKD}, \textit{lgl}^4; \text{Sema1aKD}, \textit{lgl}^4; \text{CG10931KD}, \textit{lgl}^4; \text{CacKD}. We observed that in four of the five invasive genotypes tested, a high proportion (~70%) of invading cells were iCasper negative. Only WT, \textit{lgl}^4 alone and \textit{lgl}^4; \text{CG12268KD} mutant clones showed a high proportion of invading cells that were positive for apoptosis (~64%; Figure 5-o-p).

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 \cite{36}. Genes included were \text{RhoGAP19D}, \text{Rim}, \text{S6kII}, \text{CG7379}, and \text{shot} (their closest human orthologues are \text{ARHGAP23}, \text{RIMS2}, \text{RPS6KA3}, \text{ING1}, \text{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 \text{RPS6KA3}, \text{ING1} and \text{DST}, and a significant increase in migration alone with gene KD of \text{RIMS2} (Figure S4).

These results provide strong evidence that our novel system can identify regulators of tumour progression and cancer cell invasion. Results show that in most cases invading cells are non-apoptotic, and that this model can provide additional insight on invading cell morphology and behaviour, which can indicate a tumour’s invasive
potential. Results also suggest that the invasion hits identified in our genetic screen are likely to have relevance to human disease.

The cohesin complex is an invasion suppressor

Cohesin is a multi-protein complex that forms a tripartite ring-like structure consisting of the proteins SMC1, SMC3 and RAD21 [37]. Additionally, RAD21 binds to a stromalin protein (SA1 or SA2, also known as STAG1 or 2 in humans) [38, 39] (Figure 6a). Therefore two cohesin complexes can form, with cohesin genomic distribution subject to a great degree on the SA/STAG protein that binds to the tripartite ring [40]. Cohesin is evolutionarily conserved, with functional cohesin complexes found ubiquitously in all Eukaryotic organisms, from yeast to humans [38, 41]. The cohesin complex is mainly known for its role in sister chromatid cohesion (SCC) [41] however current understanding of the possible and numerous roles cohesin may play in tumour initiation and cancer progression is limited [42].

Four subunits of the cohesin complex were studied in our genetic screen: SMC1, SMC3, RAD21 and SA1. Knockdown of these subunits induced significant cytoskeletal changes to lgl<sup>+</sup> tumours, including increased multilayering, cell body rounding and apex defects. Additionally, SA1KD significantly enhanced the lgl<sup>+</sup> invasive phenotype, with other cohesin subunits having no effect on invasion (Figure 6b-f). We next knocked down the expression of specific cohesin subunits in WT clones and found that SA1 and SA2KD strongly promoted invasion even in the absence of the lgl<sup>+</sup> mutation, whilst the other subunits did not; all subunits however promoted multilayering (Figure 6g-i). Using iCasper we also saw that a high
proportion of invading cells evaded apoptosis (Figure 6j-l) and as shown earlier,
showed fast directional migration (Figure 5e-k; Movie S2).

Our screen identified cohesin subunits as affecting epithelial architecture, cell
shape, and in the case of SA subunits, promoting frequent cell delamination. These
phenotypes therefore implicate effects on adhesion, polarity and actin regulation as
possible underlying influences on the observed cell behaviour. We investigated cell-
cell adhesion and polarity using antibodies to proteins that localise to the adherens
junction (AJ), septate junction (SJ) and the sub-apical region. We generated SA1 and
SA2KD clones and directly compared junction composition inside and outside the
clones within the same tissue. A significant reduction in the cortical localisation of E-
cadherin, α-catenin, β-catenin and FasIII was observed at the junctional level in both
SA1 and SA2KD clones, when compared to the surrounding wild type tissue, with
evidence of junctional breaks, ectopic structures (puncta, tubules) and
mislocalisation of junction components (Figure 6m-p), which are phenotypes that
are commonly observed when junctional integrity is compromised [10]. In contrast,
KD had no effect on the polarity proteins investigated (dlg and aPKC; Figure S5).
These results suggest that SA1 and SA2 act as invasion suppressors in part through
the correct localisation of junction determinants, thereby maintaining cell-cell
junction integrity.

To determine if the role of SA1 and SA2 as invasion suppressors is conserved, we
next studied the effect that the loss of their human orthologues, STAG1 and STAG2,
would have on MCF7 cell invasion and migration using an in vitro invasion assay.
Loss of function (LOF) mutations of STAG2 are significantly elevated in metastatic breast cancer tumours when compared to lower grades [43], suggesting that STAG2 has a role in preventing tumour transition to malignancy. STAG2 is also commonly mutated in several cancer types, including bladder cancer and Ewing’s sarcoma [44, 45]. When analysing each cohesin subunit in turn we found that only STAG1 and STAG2KD promoted invasion and migration, with the core components of the tripartite ring failing to affect cell behaviour (Figure S6a-i) thereby mirroring the effect we see in vivo in the fly (Figure 6g-h).

Cohesin is known to influence gene expression. It has been shown in yeast and flies that substantial reductions in cohesin dosage of more than 85% are required to disrupt cohesion and chromosome segregation, while small to moderate reductions can affect gene expression [46]. Therefore, the invasive effects that we see in SA/STAG mutants could be due to changes in the expression of genes that affect cell-cell junctions and/or the cytoskeleton. Since STAG2 is the most abundant and most mutated cohesin gene in human cancers we performed a microarray gene expression analysis, comparing gene expression in MCF-7 cells post STAG2KD with untreated cells (unt) and with cells treated with non-targeting siRNA (non-T). Out of 21448 genes analysed, the expression of 23 genes was significantly altered as a result of STAG2KD (p<0.01, FC≥1.5 or FC≤-1.5; Figure S6, Table S5). We additionally used RT-qPCR on a selection of genes (STAG2, PCDH1, EHD2 and AKR1B10) to verify the microarray results, with qPCR showing the same or stronger expression change in all cases (Figure S6n).
GO term analysis identified six biological processes that were significantly enriched within the 23 differentially expressed genes, including cell-cell adhesion, protein localisation and cell projection organisation (Figure S6o). Additionally, an interaction network was generated, using the Cytoscape plugin GeneMania, to display any genetic and physical interactions, verified by experimental data, between the differentially expressed genes and members of the AJ KEGG pathway (Figure S6p). 95 interactions between 20 differentially expressed genes and 20 AJ KEGG pathway genes indicate that the differentially expressed genes in STAG2KD cells extensively interact with members of the AJ pathway. Furthermore, EHD2 was significantly downregulated in STAG2KD cells. EHD2 has been linked to E-Cadherin localisation and expression, and lower EHD2 expression is associated with metastatic tumours [47, 48]. EHD2 links endocytosis to the actin cytoskeleton [49] and could therefore be influencing E-Cadherin’s ability to recycle at the junction.

An additional GO term analysis was performed on differentially expressed genes found in two studies that depleted STAG2 expression in cell lines of epithelial origin (MCF10A [40] and HCT116 [50]). Here we found statistically enriched terms including regulation of cell-cell adhesion, regulation of cellular protein localisation, regulation of cell-matrix adhesion [40] and positive regulation of cell migration [50].

Cohesin loss-of-function induces the formation of a supracellular actomyosin ring

Although SA1KD, SA2KD and SMC3KD promote multilayering (Figure 6i), at an apical level they present a phenotype very similar to WT, with cells presenting an
organised geometric shape (Figure 7a, b, d-e). By contrast, we see a very different phenotype for three cohesin loss of function genotypes: smc3A (an ethyl methane sulfonate induced truncating mutation within smc3, K575term [51, 52]); combined SA1 + SA2KD; and NipBKD (loss of NippedB prevents cohesin from interacting with DNA [53]). These mutants induced a highly distinctive phenotype with drastic cytoskeletal changes, including the formation of a supracellular actin ring (Figure 7c, f-h), eventually followed by clonal extrusion (Figure S7c). It therefore appears that a more severe disruption to cohesin function leads to a very different phenotype to that observed when a single SA subunit is KD. Here individual cell invasion is not observed, rather apical constriction and basal clonal extrusion occurs, which is likely to have relevance to the poorly understood process of collective cell invasion in cancer. We further characterised the phenotype using both GFP:Moe to label actin and mCherry:spaghetti squash (sqh; the fly orthologue of the regulatory light chain of non-muscle myosin II). We found that the supracellular ring is enriched with actomyosin, which induces the invagination of the mutant tissue, forming a ball of cells with a central lumen (Figure S7b-d). We also found significantly elevated levels of E-cadherin within smc3A clones (Figure S7d and f), which could also promote clonal invagination through differential adhesion properties between cell types [54].

Long time-lapse movies show that over a number of hours the actomyosin ring contracts, inducing a basal clonal extrusion from the epithelial sheet (Figure S7c). Using the caspase sensor, iCasper, we found no significant difference in the levels of apoptosis in smc3A clones, irrespective of whether the clone was still connected to the epithelial sheet or had already extruded (Figure S7g). Further, time-lapse
imaging was performed on extruded clones with little increase in iCasper signal observed over 1h post-extrusion (Figure S7h), indicating that the basal extrusion of $smc3^A$ clones does not trigger extensive cell death.

Known mechanisms that trigger apical constriction during development include the apical localisation of activated Rho1, which recruits and activates myosin II [55]. We found that Rho1 and Sqh are essential for the determination of $smc3^A$ cell morphology and actin ring formation, since dominant negative Rho (RhoN) and SqhKD both inhibit actin ring formation and clonal extrusion, whilst phosphomimetic Sqh (Sqh-EE) significantly increases the prevalence of this phenotype (Figure 7i-o).

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 localisation of myosin II and Rho1 to the apex of the cell was performed. Six candidate genes were KD and, where possible, overexpressed, both alone and in combination with the $smc3$ mutation, to determine if these genes enhance or rescue the actin ring and clonal extrusion phenotype. Although four genes promoted actin ring formation in WT clones when overexpressed, only Mad had any significant effect within $smc3^A$ clones. Mad overexpression within $smc3^A$ clones significantly increased the number of actin rings and delaminated clones (1.196, n=8, p<0.05) when compared to $smc3^A$ alone (0.393, n=8), whereas MadKD in $smc3^A$ tissue had the opposite effect (0.196, n=8, p<0.01; Figure 7p-q).
Mad is the main effector of the *Drosophila* Dpp signalling pathway. An increase in Dpp signalling has been directly implicated in apical constriction and actin ring formation [56]. Using a phospho-Mad antibody (pMad) we detected a significant increase in pMad levels in *smc3Δ* clones and SA1 + SA2KD clones, specifically when these clones contained actin rings (Figure 7t) suggesting that an increase in Mad activity is necessary to induce apical constriction in cohesin LOF clones. It therefore appears that an upregulation of Dpp signalling is a key determinant for the collective invasion observed in cohesin LOF clones.

Given the known pleiotropic effects of the cohesin complex (on SCC, homologous recombination, genome organisation and gene transcription, amongst others) and given our findings showing that cohesin subunits can regulate individual or collective cell invasion in an apparent dose-dependent manner, we studied the dynamics of chromosomal architecture in dividing cells in vivo. We generated WT, *smc3Δ*, SA1KD, and SA2KD clones, which were labelled with both GFP:Moe and Histone:RFP and carried out live imaging of dividing cells within these clones. We found the vast majority of *smc3Δ* mutant cell divisions were defective in chromosome alignment and/or chromosome separation during metaphase and anaphase respectively. In contrast, the vast majority of divisions in SA1 and SA2KD cells appeared normal (Figure S8; Movies S3-S6) adding to the growing body of evidence to suggest that only a major reduction of cohesin function leads to cohesion and segregation defects [42].
In summary, this work has: (1) identified numerous genes that affect tumour behaviour in a wide variety of ways; (2) generated a functionally validated network of invasion-suppressor genes; (3) identified the cohesin complex as an important invasion suppressor that can promote individual or collective invasion; (4) established the fly pupal notum as an excellent in vivo system to study tumour progression.

Discussion

By combining the genetic amenability of *Drosophila melanogaster* with the power of RNAi transgenics, we were able to generate tumours with specific genotypes and to monitor tumour behaviour in the living animal. The in vivo system we have developed offers a number of significant advantages, and is particularly suitable to the study of tumour progression and invasion. It enables us to: (1) monitor GFP:Moe labelled tumours in situ, surrounded by wild-type tissue and the native local microenvironment; (2) image tumours in high spatial and temporal resolution over a number of hours or even days post-tumour induction; (3) knockdown gene expression specifically within the developing tumour, allowing us to investigate the tumour promoting potential of numerous genes that would be developmentally lethal under classic mutation conditions.

Cancer genomes show extreme heterogeneity, with individual solid organ tumours possessing on average >50 non-silent mutations in the coding regions of different genes [57-60]. Breast and colorectal cancers have been found to be the most heterogeneous, with an average of 84 and 76 mutations/tumour respectively [61,
Further complexity is evident when considering epigenetic alterations that can contribute to tumourigenesis and tumour progression [63]. The challenge is to identify those genes, from the many that have been implicated in human cancer, which drive cancer progression. We used our in vivo system to investigate a set of almost 500 genes, whose human orthologues have previously been implicated in cancer, and have now identified numerous genes that either positively or negatively regulate specific aspects of tumour behaviour within an epithelium in a living animal.

One limitation of the screen, as is the case for any cancer screen, is the fact that the results presented here describe tumour behaviour within a specific tissue and anatomical location (the fly notum) and against a specific genetic background (the underlying mutation being \(lg^+/\)). In the fly, just as in humans, one would expect tumours with the same genotype to behave differently in different tissues, and additionally expect different combinations of mutations to result in different phenotypes. Despite this, work carried out in the human breast cancer cell line MCF7 shows that the majority of hits tested give the same phenotypes and thereby will have relevance to human disease. This is most clearly seen when testing cohesin subunits in the fly and in MCF7 cells: STAG1 and STAG2 both promote invasion when their expression is knocked down, whereas other cohesin subunits do not – recapitulating the effect seen within the fly screen.

To understand tumour transition to malignancy, and to develop new therapeutic strategies, it will be key to paint a detailed picture of the complex signalling
processes that occur during tumour progression. Our database incorporates 33 phenotypic categories and therefore offers a unique starting point to elucidate the molecular mechanisms of multiple aspects of tumour progression.

However, our primary focus was invasion, and our screen identified numerous genes that regulate epithelial cancer cell invasion. We generated a functionally validated network of invasive genes; GO term analysis of this network identified several terms that are significantly enriched, indicating processes that are likely to be important for invasion to take place. This includes adhesion, cytoskeletal remodelling, signalling and intriguingly many axon guidance molecules. The Slit, Robo and Semaphorin families have been previously implicated as both tumour and metastasis suppressors in breast cancer. SLIT/ROBO signalling has been postulated to prevent invasion by maintaining proper cell-cell adhesion, thereby inhibiting the detachment of tumour cells [64]. Many other axon guidance genes have been found to be invasion suppressors in our screen, as have uncharacterised genes that genetically interact with axon guidance genes, opening up an intriguing avenue of future research. It is clear that a loss of polarity and a disruption to normal adhesion are pivotal to promoting the process of invasion. Axon guidance proteins, being heavily involved in developmental processes that require cell movement, could be promoting invasive characteristics via these two fundamental processes.

Our in vivo system is furthermore particularly suited to imaging the invasive process. Our observation of characteristic cell shape changes (cell rounding and a polarised actin enrichment) that accompany invasion has been previously reported.
and associated with invasion [65, 66]. However, an important avenue of future research will be to investigate the morphological and molecular processes that underlie the differential behaviour between invading cells with and without directional migration. Cell body rounding would indicate an amoeboid type migration, but the characteristic blebbing of amoeboid migration is only clearly obvious in those cells undergoing directional migration. The use of a membrane (rather than actin-associated) marker together with high resolution microscopy would help to determine whether the extent of membrane blebbing is an important attribute for directionality in this system. An additional consideration is the genetic simplicity of these tumours. It is evident that, in the fly, where there is less redundancy in key regulatory genes, we are able to generate multilayered, invasive tumours, with just two key mutations, but for many invasion suppressors further cooperative mutations are likely to be required to promote directional migration. ECM composition and the presence/absence of a chemotactic gradient are also important considerations for directed migration, and will be influencing cell behaviour here [67].

Our work on the cohesin complex provides an example of how specific phenotypes observed in our screen can inform downstream characterisation analyses and provides further validation that our screen is picking up important regulators of tumour progression.

Cohesin was initially identified for its role in SCC in yeast [41, 68] and *Xenopus* [69], but has subsequently been found to be involved in homologous recombination-
mediated DNA repair, higher order-chromatin structure and transcriptional regulation [70-75]. How cohesin performs these multiple roles is not fully understood, but is thought to be largely due to cohesin’s ability to hold DNA strands in either trans (during cell division) or cis (generating chromatin loops) [42]. This wide variety of functions complicates our understanding of how cohesin mutations may contribute to cancer progression. Inactivating mutations in genes that encode either the core cohesin subunits, or regulatory proteins that impact on cohesin function (e.g. PDS5A/B, WAPL, CDCA5, NIPBL, MAU2, etc.) are common in numerous cancer types, including bladder, melanoma, colorectal, lung, Ewing sarcoma and myeloid malignancies. Importantly, there is no clear correlation between the presence of cohesin mutations and aneuploidy in many tumour types, with recent studies implicating effects on chromatin structure, transcription, DNA repair and stem cell/progenitor differentiation as important phenotypes that could promote cancer progression [42, 76]. Although cohesin is essential for cell viability, mutations are likely to reduce the amount of total functional cohesin within the cell, which will impact on these diverse cohesin-mediated tasks in different ways, depending on the subunit that is mutated, the nature of the mutation, and the cell type affected. Our work shows that, since each specific mutation impacts cohesin function in different ways, effects on tumour cell behaviour can range from defects in epithelial architecture, to the promotion of either individual or collective invasion; the phenotype observed will depend on whether the mutation leads to a modification or a disruption of cohesin function, and the degree of any such disruption.
We found loss of cohesin function to induce different phenotypes related to actin cytoskeleton rearrangement. KD of one subcellular localisation subunit, SA1 or SA2, increased invasion, multilayering and apex defects. Reduced expression of the core subunits, SMC1, RAD21 and SMC3, increased multilayering and apex defects, yet had no effect on invasion. A more severe loss of cohesin function (a LOF smc3 allele, SA1 + SA2 simultaneous KD or NipBKD) induced clonal extrusion and collective invasion. Differences in cohesin subunit function (SA1 and SA2 provide subcellular localisation; SMC1, SMC3 and RAD21 form the core of the ring) [37], isoform redundancy (SA1/SA2, SMC1A/SMC1B) [38, 77], in combination with the specific dose required for each subunit to efficiently perform its role in either gene expression regulation or SCC [78], 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 [79-81] and that the two SA subunits are not fully functionally interchangeable [40]. Therefore, loss of one specific SA subunit will have drastic effects on how cohesin interacts with chromatin and on gene expression. Our in vivo experiments in the fly and transcriptomics experiments in vitro suggest that loss of SA1 or SA2 induces single cell invasion by affecting cohesin mediated gene expression during interphase, with strong effects on junction stability. Our live cell imaging of SA1 and SA2KD cells provides further evidence to suggest that aneuploidy is unlikely to make a major contribution to this phenotype. By contrast, a severe loss of cohesin function due to a loss of functional SMC3 does lead to chromosomal instability, which ultimately leads to a misregulation of DPP signalling and increased E-cadherin levels, followed
by clonal extrusion. This phenotype could be due to a combination of chromosomal instability, aneuploidy and chromatin rearrangement defects.

**Acknowledgments**

We wish to thank the fly community for their generosity with reagents, especially the Bloomington, VDRC and NIG stock centres, without whom this project would not have been possible. We thank Anna Grabowska for the MCF-7 cell line, the School of Life Sciences Imaging (SLIM) for invaluable help with the confocal microscopes and Louise Cheng and Peter Shaw for critical reading of the manuscript. We additionally thank Keith Spriggs for help with preparing image files for the online resource. This work was supported by Cancer Research UK [grant numbers C36430, A12891]. B.C.C. was supported by a CONACYT award; A.D.R. was supported by a Nottingham Vice-Chancellor’s Scholarship for Research Excellence Award.

**Author Contributions**

Conceptualisation, M.G.; Methodology, M.G., B.C.C., A.C., N.A.M., and Z.E.C.; Investigation, B.C.C., Z.E.C, A.C., N.A.M., A.D.R, U.N., Y.N.F., M.H., M.C.U. and A.B.; Formal Analysis, A.L., M.C.U., B.C.C., and A.D.R.; Writing – Original Draft, M.G.; Writing – Review & Editing, M.G., B.C.C., A.C., A.D.R., A.L. Funding

Acquisition, M.G.; Resources, M.G., R.R., and S.T.M.; Supervision, M.G. and S.T.M.

**Declaration of Interests**

The authors declare no competing interests.
Data availability
The accession number for the microarray data reported in this paper is GEO: GSE137773.

Figure Legends
Figure 1: *lgl*4 mutant clones provide an ideal genetic background for an enhancer/suppressor screen for tumour progression

(a-b) 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 *lgl* (b-b'). (c-d) Quantification of average clonal area (c) (n=10 (WT); 18 (*lgl*4)) and the number of invading cells / the total number of labelled cells (d) (n= 30 (WT); 41 (*lgl*4)). Quantification shows *lgl*4 mutant clones to be similar to WT clones in size, with a significant increase in the number of invading cells. (e-h)

GFP:Moe labelled genetic clones in the dorsal thorax epithelium of living fly pupae.

Clones shown are overexpressing activated Notch (N\text{\textsuperscript{intra}}; e) or simultaneously homozygous mutant for *lgl*4 and overexpressing N\text{\textsuperscript{intra}} (f-h). Highlighted are effects on cell division (f), invasion (g) and multilayering (h). (i-j) Quantification of the number of dividing cells (i) and the number of invading cells (j) over the total number of labelled cells for clones with the genotypes shown (n= 30 (WT); 41(*lgl*4); 7 (N\text{\textsuperscript{intra}}); 13 (*lgl*4; N\text{\textsuperscript{intra}})). Error bars represent ± s.e.m. Student’s T test (e) and Kruskall-Wallis test (f, k-l) were performed to determine statistical significance. Red arrow: dividing cell; red arrowhead: cell doublet following cytokinesis; white arrows: invading cells. White scale bar: 50μm; red scale bar: 10μm.
Figure 2: Pilot screen identifies several modulators of tumour behaviour.

(a) Schematic illustrating how clones with distinct genotypes were generated on the back of the fly. The MARCM system was employed to generate mutant clones specifically within the fly dorsal thorax, through the use of Ubx-Flp. This generated GFP:Moe labelled $lgl^4$ homozygous mutant clones. RNAi transgene expression, and therefore gene knockdown, was restricted to the labelled $lgl^4$ mutant tissue. (b) Pie chart illustrating the range of biological functions from those genes included in the pilot screen. A: apicobasal polarity; B: cell adhesion; C: cytoskeleton; D: axon guidance; E: cell cycle; F: gene expression; G: signalling; H: mitochondria; I: others; J: unknown. (c-k) Examples of phenotypes observed within the pilot screen. In the pilot screen we observed effects on clonal size (d-e), tissue morphology (e-f), cell morphology (i and k), and cell behaviour (g-h and j). These are just a few examples of the many distinct phenotypes that we observed. Arrows: (g) invading cells; (h) dividing cells; (j) a blebbing dividing cell; (k) very long basal protrusions. Arrowheads: (h) cell doublet following cytokinesis; (k) long protrusions joining to form a fascicle. White scale bar: 50μm; red scale bar: 10μm; yellow scale bar: 10μm in xz plane.

Figure 3: Clustering analyses identify ten RNAi line clusters and three distinct phenotypic subgroups

(a) Heat-map representation of supervised clustering of 764 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 enriched representative GO categories are shown on the right-hand side of each cluster. (b) Consensus clustering of average scores of 29 phenotypic categories reveals three distinct subgroups. 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.

**Figure 4: An interaction network of invasion suppressors**

Interactions between genes for which knock down 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.

**Figure 5: Characterisation of selected invasion suppressors**
An example of a highly invasive mutant clone (genotype: \textit{lgl}^4; CG7379KD) labelled with GFP:Moe. Highlighted is a pre-invasive cell that rounds up and forms a characteristic actin-rich spot at one side of the cell prior to invasion (0 mins). The cell then detaches from the mutant clone and migrates away (arrow). (b-d) Correlation between the percentage of clonal cells with a polarised actin accumulation and the percentage of invading cells per animal (n=10 animals/genotype). The two parameters show a significant correlation, irrespective of whether the mutant clones were rarely invasive or highly invasive. (e) Stills from a time-lapse showing the basal surface of a GFP:Moe labelled SA1KD clone. Yellow star marks the initial location of an invading cell; magenta dot shows the location of the invading cell at the indicated time. The cell shown has moved 38\textmu m in 8 minutes. (f-f’) Representative single cell trajectories from \textit{lgl}^4 (orange) and SA1KD invading cells (blue) shown in \textit{xy} (f) and \textit{xz} (f’). Each cell was measured every 3 minutes for 30 minutes. (g) Illustration showing the two trajectories measured for each invading cell in order to determine directionality. Length (blue) follows the full trajectory of an invading cell. Displacement (red) measures a straight line from the initial to the final point. (h) Quantification of length and displacement from \textit{lgl}^4 and SA1KD cells (n=25 cells from 5 animals/genotype). Cells that have directionality have no significant difference between length and displacement. (i) Quantification of speed of migration, showing average \textmu m travelled per minute (n=25 cells from 5 animals/genotype). (j) Quantification of speed of migration (\textmu m/minute) for \textit{lgl}^4 and SA1KD cells that present either a single actin spot, or multiple actin spots (n=5 cells/group). Those with multiple spots travel faster irrespective of genotype. (k) SA1KD cells have a significantly higher proportion of invading cells with multiple
Orthogonal view of invading cells showing that cells only migrate once detached from the epithelial sheet (yellow dot). Red asterisk: pre-invasive cell within sheet; red dot: delaminated cell still attached to sheet. (n) Quantification of the percentage of pre-invasive cells that detach from the epithelial sheet and migrate, in WT, \( lgl^{4} \) and SA1KD clones (n=3 animals/genotype). (o-p) iCasper (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 p; n=10 animals/genotype).

**Figure 6: SA1 or SA2KD promotes invasion**

(a) Somatic cells simultaneously express two different Cohesin rings, differentiated by the presence of either SA1/STAG1 or SA2/STAG2. (b) Heat map illustrating qualitative scores given to cohesin subunits included in the genetic screen. A subset of categories is shown. Red: enhancement of a phenotype; yellow: no phenotype change; blue: inhibition of a phenotype. (c-f) GFP:moe positively marked \( lgl^{4} \) mutant clones with additional cohesin complex subunit KD, showing invading cells (arrows; c) and multilayering (e), quantified in (d) and (f); n=5 animals/genotype.

Red dashed line highlights edge of clone. Yellow line shows position of xz slice shown. (g) Basal confocal slice of GFP:moe positively marked WT, SA1 or SA2KD clones, highlighting invading cells (arrows). (h-i) Quantification of % invading cells...
(h) and % multilayering (i) following KD of each cohesin subunit, compared to WT.

(j-l) Confocal images of the basal surface of iCasper (red) and GFP:Moe (green) labelled WT clones (j) and SA2KD clones (k). Arrows highlight invading cells that are iCasper negative. Quantified in (l): Grey: % invading cells / total number of labelled cells; blue: % non-apoptotic invading cells / total number of labelled cells; n=50 cells from 10 animals/genotype. Young WT pupae were used as a control (j) as older WT animals have little to no invading cells. (m-p) SA1 or SA2KD clones, highlighted by magenta and cyan dashed lines, respectively, show disrupted E-cadherin (m), armadillo (n), α catenin (o), fasIII (p), localisation. Arrowheads highlight junctional breaks. Quantification shows fluorescence intensity at the level of the junction (n=100 junctions from 10 animals for each genotype). Scale bars: 10μm. Error bars = ± s.e.m. Student’s T test or One-way ANOVA with Dunnett’s post hoc test for multiple comparisons was performed to determine statistical significance.

Figure 7: A more severe cohesin LOF induces actin ring formation.

(a-g) GFP:moe positively marked clones (genotype indicated on the bottom left of panel). Actin rich rings (yellow arrows) were observed in smc3A, SA1 and SA2 simultaneous KD, and NipBKD clones. (h) Quantification of the number of actin rings per mm² of clonal tissue. Eight animals were analysed for each genotype. (i-o) GFP:moe positively marked clones (genotype indicated on the bottom left of panel). Dominant negative Rho (RhoN) and SqhKD inhibit actin ring formation in smc3A clones; phosphomimetic Sqh (SqhEE) increases the number of clones with actin rings. Quantified in (l) and (o) showing the number of actin rings or delaminated clones per mm² clonal tissue. Each dot represents one animal. smc3A + RhoV14
resulted in very small unicellular clones (j) or no clones at all and could not be quantified. (p-q) Genes involved in apical constriction were either knocked down or overexpressed in GFP:moe positively marked clones, either on their own (p) or within smc3A clones (q). Quantification shows the number of actin rings or delaminated clones per mm² clonal tissue. Each dot represents 1 animal. (r-s) GFP:moe labelled smc3A (r) and SA1 + SA2KD (s) clones stained for the active form of the Dpp signalling effector, phosphorylated Mad (pMad). (t) Quantification of mean fluorescence intensity from the nuclei of cells within 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. Scale bars: 10μm. Error bars = ± s.e.m. Statistical analysis: Student’s T test.

Supplementary Tables

Table S1: Full database
Table S2: Level of similarity between two RNAi lines targeting the same gene
Table S3: Hits for all categories
Table S4: Lists of genes within clusters and associated GO terms
Table S5: Genes showing a significant change in expression following STAG2KD in MCF7 cells

Other Supplementary material

Movie S1: Non-directional migration
Time-lapse movie of a highly invasive mutant clone (genotype: \textit{lg}^{l4}; CG7379KD) labelled with GFP:Moe, showing invading cells with non-directional migration. Time stamp: top left; scale bar: 10\,\mu m.

**Movie S2: Directional migration**

Time-lapse movie of an SA1KD clone labelled with GFP:Moe, showing invading cells with fast, directional migration. Time stamp: top left; scale bar: 10\,\mu m.

**Movies S3 – S6: In vivo imaging of cell division**

Time-lapse movies of WT (Movie S3), smc3^{4} (Movie S4), SA1KD (Movie S5), and SA2KD (Movie S6) clones, labelled with GFP:Moe and Histone:RFP. Time stamp: bottom right; scale bar: 5\,\mu m.

**Cytoscape network file for interaction map of invasive genes**

**Cytoscape network file for interaction map of genes misregulated by STAG2KD that affect cell-cell junctions**

**References**

1. Sporn MB. The war on cancer. Lancet. 1996;347:1377-81.
2. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57-70.
3. Steeg PS. Metastasis suppressors alter the signal transduction of cancer cells. Nat Rev Cancer. 2003;3:55-63.
4. Yan J, Yang Q, Huang Q. Metastasis suppressor genes. Histol Histopathol. 2013;28:285-92.
[5] Hurst DR, Welch DR. Metastasis suppressor genes at the interface between the environment and tumor cell growth. Int Rev Cell Mol Biol. 2011;286:107-80.

[6] Rudrapatna VA, Cagan RL, Das TK. Drosophila cancer models. Dev Dyn. 2012;241:107-18.

[7] Brumby AM, Richardson HE. Using Drosophila melanogaster to map human cancer pathways. Nat Rev Cancer. 2005;5:626-39.

[8] Gonzalez C. Drosophila melanogaster: a model and a tool to investigate malignancy and identify new therapeutics. Nat Rev Cancer. 2013;13:172-83.

[9] Mirzoyan Z, Sollazzo M, Allocca M, Valenza AM, Grifoni D, Bellotta P. Drosophila melanogaster: A Model Organism to Study Cancer. Front Genet. 2019;10:51.

[10] Georgiou M, Marinari E, Burden J, Baum B. Cdc42, Par6, and aPKC Regulate Arp2/3-Mediated Endocytosis to Control Local Adherens Junction Stability. Current Biology. 2008;18:1631-8.

[11] Georgiou M, Baum B. Polarity proteins and Rho GTPases cooperate to spatially organise epithelial actin-based protrusions. J Cell Sci. 2010;123:1089-98.

[12] Cohen M, Georgiou M, Stevenson NL, Miodownik M, Baum B. Dynamic filopodia transmit intermittent Delta-Notch signaling to drive pattern refinement during lateral inhibition. Dev Cell. 2010;19:78-89.

[13] Couto A, Mack NA, Favia L, Georgiou M. An apicobasal gradient of Rac activity determines protrusion form and position. Nat Commun. 2017;8:15385.

[14] Moberg KH, Bell DW, Wahrer DC, Haber DA, Hariharan IK. Archipelago regulates Cyclin E levels in Drosophila and is mutated in human cancer cell lines. Nature. 2001;413:311-6.

[15] Tapon N, Ito N, Dickson BJ, Treisman JE, Hariharan IK. The Drosophila tuberous sclerosis complex gene homologs restrict cell growth and cell proliferation. Cell. 2001;105:345-55.

[16] Woodhouse EC, Fisher A, Bandle RW, Bryant-Greenwood B, Charboneau L, Petricoin EF, 3rd, et al. Drosophila screening model for metastasis: Semaphorin 5c is required for l(2)gl cancer phenotype. Proc Natl Acad Sci U S A. 2003;100:11463-8.

[17] Pagliarini RA, Xu T. A genetic screen in Drosophila for metastatic behavior. Science. 2003;302:3227-31.

[18] Zoranovic T, Manent J, Willoughby L, Matos de Simoes R, LaMarca JE, Golenkina S, et al. A genome-wide Drosophila epithelial tumorigenesis screen identifies Tetraspanin 29Fb as an evolutionarily conserved suppressor of Ras-driven cancer. PLoS Genet. 2018;14:e1007688.

[19] Xu T, Rubin GM. Analysis of genetic mosaics in developing and adult Drosophila tissues. Development. 1993;117:1223-37.

[20] Lee T, Luo L. Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. Trends Neurosci. 2001;24:251-4.

[21] Stephens R, Lim K, Portela M, Kvansakul M, Humbert PO, Richardson HE. The Scribble Cell Polarity Module in the Regulation of Cell Signaling in Tissue Development and Tumorigenesis. J Mol Biol. 2018.

[22] Bilder D. Epithelial polarity and proliferation control: links from the Drosophila neoplastic tumor suppressors. Genes Dev. 2004;18:1909-25.

[23] Gateff E. Malignant neoplasms of genetic origin in Drosophila melanogaster. Science. 1978;200:1448-59.
[24] Lo P, Hawrot H, Georgiou M. Apicobasal polarity and its role in cancer progression. Biomol Concepts. 2012;3:505-21.

[25] Brumby AM, Richardson HE. scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila. Embo J. 2003;22:5769-79.

[26] Nagata R, Igaki T. Cell competition: Emerging mechanisms to eliminate neighbors. Dev Growth Differ. 2018;60:522-30.

[27] Khan SJ, Bajpai A, Alam MA, Gupta RP, Harsh S, Pandey RK, et al. Epithelial neoplasia in Drosophila entails switch to primitive cell states. Proc Natl Acad Sci U S A. 2013;110:E2163-72.

[28] Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010;463:899-905.

[29] International Cancer Genome C, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. Nature. 2010;464:993-8.

[30] Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014;505:495-501.

[31] Sevimoglu T, Arga KY. The role of protein interaction networks in systems biomedicine. Comput Struct Biotechnol J. 2014;11:22-7.

[32] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2.

[33] Marinari E, Mehonic A, Curran S, Gale J, Duke T, Baum B. Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding. Nature. 2012;484:542-5.

[34] Fujisawa Y, Kosakamoto H, Chihara T, Miura M. Non-apoptotic function of Drosophila caspase activation in epithelial thorax closure and wound healing. Development. 2019;146.

[35] To T-L, Piggott BJ, Makhijani K, Yu D, Jan YN, Shu X. Rationally designed fluorogenic protease reporter visualizes spatiotemporal dynamics of apoptosis in vivo. Proceedings of the National Academy of Sciences. 2015;112:3338-43.

[36] Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011;12:357.

[37] Gruber S, Haering CH, Nasmyth K. Chromosomal cohesin forms a ring. Cell. 2003;112:765-77.

[38] Losada A, Yokochi T, Kobayashi R, Hirano T. Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. J Cell Biol. 2000;150:405-16.

[39] Zhang N, Jiang Y, Mao Q, Demeler B, Tao YJ, Pati D. Characterization of the interaction between the cohesin subunits Rad21 and SA1/2. PLoS One. 2013;8:e69458.

[40] Kojic A, Cuadrado A, De Koninck M, Gimenez-Llorente D, Rodriguez-Corsino M, Gomez-Lopez G, et al. Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. Nat Struct Mol Biol. 2018;25:496-504.

[41] Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell. 1997;91:35-45.
[42] De Koninck M, Losada A. Cohesin Mutations in Cancer. Cold Spring Harb Perspect Med. 2016;6.

[43] Repo H, Loyttyanniemi E, Nykanen M, Lintunen M, Karra H, Pitkanen R, et al. The Expression of Cohesin Subunit SA2 Predicts Breast Cancer Survival. Appl Immunohistochem Mol Morphol. 2016;24:615-21.

[44] Aquila L, Ohm J, Woloszynska-Read A. The role of STAG2 in bladder cancer. PharmacoL Res. 2018;131:143-9.

[45] Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov. 2014;4:1342-53.

[46] Dorsett D, Merkenschlager M. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. Curr Opin Cell Biol. 2013;25:327-33.

[47] Shi Y, Liu X, Sun Y, Wu D, Qiu A, Cheng H, et al. Decreased expression and prognostic role of EHD2 in human breast carcinoma: correlation with E-cadherin. J Mol Histol. 2015;46:221-31.

[48] Yang X, Ren H, Yao L, Chen X, He A. Role of EHD2 in migration and invasion of human breast cancer cells. Tumour Biol. 2015;36:3717-26.

[49] Guilherme A, Soriano NA, Bose S, Holik J, Bose A, Pomerleau DP, et al. EHD2 and the novel EH domain binding protein EHBP1 couple endocytosis to the actin cytoskeleton. J Biol Chem. 2004;279:10593-605.

[50] Casa V, Moronta Gines M, Gade Gusmao E, Slotman JA, Zirkel A, Josipovic N, et al. Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control. Genome Res. 2020;30:515-27.

[51] Haelterman NA, Jiang L, Li Y, Bayat V, Sandoval H, Ugur B, et al. Large-scale identification of chemically induced mutations in Drosophila melanogaster. Genome Res. 2014;24:1707-18.

[52] Yamamoto S, Jaiswal M, Chang WL, Gambin T, Karaca E, Mirzaa G, et al. A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. Cell. 2014;159:200-14.

[53] Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, et al. Cohesin’s binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Mol Cell. 2000;5:243-54.

[54] Steinberg MS. Differential adhesion in morphogenesis: a modern view. Curr Opin Genet Dev. 2007;17:281-6.

[55] Padash Barmchi M, Rogers S, Hacker U. DRhoGEF2 regulates actin organization and contractility in the Drosophila blastoderm embryo. J Cell Biol. 2005;168:575-85.

[56] Jidigam VK, Srinivasan RC, Patthey C, Gunhaga L. Apical constriction and epithelial invagination are regulated by BMP activity. Biol Open. 2015;4:1782-91.

[57] Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007;446:153-8.

[58] Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008;321:1801-6.

[59] Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008;455:1069-75.
[60] Network TCGAR. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474:609-15.

[61] Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, et al. The consensus coding sequences of human breast and colorectal cancers. Science. 2006;314:268-74.

[62] Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. Science. 2007;318:1108-13.

[63] Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128:683-92.

[64] Yuasa-Kawada J, Kinoshita-Kawada M, Rao Y, Wu JY. Deubiquitinating enzyme USP33/VDU1 is required for Slit signaling in inhibiting breast cancer cell migration. Proc Natl Acad Sci U S A. 2009;106:14530-5.

[65] Sahai E, Marshall CJ. Differing modes of tumour cell invasion have distinct requirements for Rho/ROCK signalling and extracellular proteolysis. Nat Cell Biol. 2003;5:711-9.

[66] Yamazaki D, Kurisu S, Takenawa T. Regulation of cancer cell motility through actin reorganization. Cancer Sci. 2005;96:379-86.

[67] Talkenberger K, Cavalcanti-Adam EA, Voss-Bohme A, Deutsch A. Amoeboid-mesenchymal migration plasticity promotes invasion only in complex heterogeneous microenvironments. Sci Rep. 2017;7:9237.

[68] Guacci V, Hogan E, Koshland D. Centromere position in budding yeast: evidence for anaphase A. Mol Biol Cell. 1997;8:957-72.

[69] Losada A, Hirano M, Hirano T. Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. Genes Dev. 1998;12:1986-97.

[70] Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. Annu Rev Genet. 2009;43:525-58.

[71] Mizuguchi T, Fudenberg G, Mehta S, Belton JM, Taneja N, Folco HD, et al. Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. pombe. Nature. 2014;516:432-5.

[72] Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell. 2013;153:1281-95.

[73] Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, et al. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature. 2009;460:410-3.

[74] Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, et al. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res. 2013;23:2066-77.

[75] Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci U S A. 2014;111:996-1001.

[76] Hill VK, Kim JS, Waldman T. Cohesin mutations in human cancer. Biochim Biophys Acta. 2016;1866:1-11.

[77] Revenkova E, Eijpe M, Heyting C, Hodges CA, Hunt PA, Liebe B, et al. Cohesin SMC1 beta is required for meiotic chromosome dynamics, sister chromatid cohesion and DNA recombination. Nat Cell Biol. 2004;6:555-62.

[78] Laugsch M, Seebach J, Schnittler H, Jessberger R. Imbalance of SMC1 and SMC3 cohesins causes specific and distinct effects. PLoS One. 2013;8:e65149.

[79] Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, et al. Cohesin Loss Eliminates All Loop Domains. Cell. 2017;171:305-20 e24.
Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, et al. Two independent modes of chromatin organization revealed by cohesin removal. Nature. 2017;551:516.

Wutz G, Varnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J. 2017;36:3573-99.