The PR factor Hamlet controls heterotypic epithelial tissue assembly in *Drosophila* reproduction system

Huazhen Wang[1], Ludivine Bertonnier-Brouty[2,3], Isabella Artner[2,3], Jiayu Wen[4,*] and Qi Dai[1,*]

1 Department of Molecular Bioscience, the Wenner-Gren Institute, Stockholm University, Sweden

2 Lund Stem Cell Center, Lund University, Lund, Sweden

3 Lund University Diabetes Center, Lund University, Malmö, Sweden

4 Division of Genome Sciences and Cancer, The John Curtin School of Medical Research, The Australian National University, and Australian Research Council Centre of Excellence for the Mathematical Analysis of Cellular Systems.

[*] Co-corresponding authors

Abstract

Epithelial tissue fusion is essential for animal development and wound healing, yet its regulatory mechanisms remain elusive. The *Drosophila* reproductive system (RS) offers a model to study such fusion events, as it integrates independently formed gonadal- and genital disc-derived structures for efficient sperms and egg delivery. In this study, we unveil the pivotal role of Hamlet (Ham), a PR-domain containing transcription factor, in orchestrating sequential epithelial assembly steps in the *Drosophila* RS. Even mild reductions in *ham* expression lead to sterility and disruption of the connection between the testis (TE) and seminal vesicle (SV). To elucidate the regulatory network governing RS epithelial assembly, we systematically identified Ham-downstream effector genes, encompassing cytoskeletal, metabolic regulators, Wnt2, Toll signaling pathway genes, and numerous uncharacterized molecules. Notably, we discovered that *Drosophila* RS epithelial tube assembly represents a process of heterotypic epithelium fusion, with distinct gene expression profiles in TE terminal and SV epithelial cells. Furthermore, TE epithelial cells exhibit high sensitivity to E-Cadherin levels, as reduction of E-Cad induces epithelial-to-mesenchymal-transition-like behaviors. Our findings present a comprehensive gene network crucial for heterotypic epithelial tissue fusion. Given the mammalian Ham orthologs in epithelial tissues, our results suggest an evolutionarily conserved role for these factors across species.

Introduction

Complex organ formation requires the connection of cells of different types and origins. Epithelial tissue fusion is a crucial process for building epithelial tubular network and uniting separate ends together [1, 2].

Epithelial cells are apically-basally polarized, and connect with tight junctions in between. This means that before, during and after epithelial fusion, such apico-basolateral polarization needs to be established, maintained and re-established. It was shown that specialized cells at the leading edge migrate and interact using cytoskeleton protrusions [3-9] and small GTPases like Rac and Cdc42 [10]. Upon contact, the leading cells re-establish polarity and connections with the help of extracellular matrix molecules, apical proteins and E-Cadherins [11]. However, the exact regulators and mechanisms controlling epithelial fusion remain to be determined, especially in a context-specific manner.

Formation of the *Drosophila* RS requires a complex process of tissue assembly. In males, structures from the gonad and genital-disc merge to form a tubular system, while in females, the oviduct and ovary must also connect [12]. TE terminal epithelial cells from male-specific somatic gonadal precursors (msSGPs) form the joint site at the TE [13]. The SV and AG develop from mesenchymal cells in the genital disc (GD), transitioning into epithelial cells through mesenchymal-to-epithelial transition (MET) during the larval stage (**Fig. 1A**) [14, 15]. During the early pupal stage, the terminal epitheliums of TE and SV fuse (**Fig. 1B**), and pigment cells from the TE and myoblasts from the SV migrate and cover both ends [16]. This process serves as a model of epithelial tissue patterning, fusion and collective cell migration.

Ham belongs to the PR domain containing protein family whose members are evolutionarily conserved cell fate regulators [17, 18]. Dysfunction of the mammalian orthologs of Ham, PRDM16 (Mel1) and PRDM3 (Evi1, MECOM), has been linked to a variety of human abnormalities including congenital diseases, metabolic disorders and cancers [19-26]. Ham has mainly been studied in the central and periphery nervous system where it is required for neuronal dendrite development [27] and cell fate maintenance [28-30]. At the molecular level, Ham modulates chromatin states at gene loci regulated by Notch signaling [31] and repress gene expression via interacting with the co-repressor CtBP [28, 31]. Molecular functions of PRDM16 and PRDM3 are versatile, with their PR domains possessing intrinsic histone

methyltransferase activity [20, 32] and the ability to interact with various co-activators and co-repressors [17, 33].

In this study, we found that Ham PR-containing isoforms are dispensable for viability but essential for male and female fertility. Even mild reduction of Ham compromised epithelial cell integrity and tissue assembly. Using RNA-sequencing (RNA-seq) and Cleavage Under Targets and Tagmentation (CUT&TAG) and multiplexed in situ hybridization approaches, we defined Ham downstream effector genes during the fusion process. By elucidating the regulatory network controlled by Ham, our study provides novel insights into the molecular mechanisms underlying epithelial tissue fusion.

Results

Drosophila RS development requires the PR-containing isoforms of Ham

The Ham mammalian orthologs PRDM16 and MECOM are critical cell fate regulators in many developmental contexts. The genetic roles of Ham were mainly found in the central and periphery nervous systems [27-30]. To systematically study Ham function, we generated new frameshift alleles using CRISPR/Cas9. The sgRNA targets the third exon shared by the three long isoforms B, D and G (**Supplementary Fig. 1A**). The resultant new alleles specifically disrupted the three PR-domain containing isoforms, and we therefore named them $ham^{PR\Delta}$. The original allele ham^{1} and the null allele ham^{SK1} shifted open reading frame (ORF) for all predicted isoforms [27, 29].

We validated nucleotide deletions in the new alleles by sanger sequencing (**Supplementary Fig. 1B**) and protein deletion by western-blot using an antibody against Ham. The Ham antibody detected several bands in the blot (**Supplementary Fig. 1C**). However, there are two bands at and above the 150KD marker position in the two heterozygotic controls but not in the ham^{SK1} homozygous mutant, suggesting that these two bands represent the longest and the second longest isoforms respectively. The $ham^{PR\Delta}$ homozygous samples lack the largest band, verifying the depletion of the longest isoforms.

We then compared viability and sensory organ development of these alleles. In contrast to the original report showing that *ham* mutant animals died at the larval stage [27], we found that ham^{SK1}, $ham^{PR\Delta}$ and ham^{1} alleles all survived to hatch when combined in trans with the *ham* deficiency line (Df) which deletes the entire *ham* locus (**Supplementary Fig. 1D**). Although ham^{1}/Df and the ham^{SK1}/Df animals cannot develop further after hatching, about two thirds of the $ham^{PR\Delta}$/Df mutants are viable without obvious gross defects. Furthermore, we examined sensory organ formation in ham mutant alleles. It is known that Ham specifies internal cell types in the sensory organ lineage as the two internal cell types (the neuron and sheath) transform into the external cell types (the socket and shaft) in ham^{1} mutants [28]. Unexpectedly, $ham^{PR\Delta}$ homozygous animals or $ham^{PR\Delta}$ in trans with the *ham* deficiency displayed wild-type sensory organs, in contrast to defective sensory organs in the ham^{1} and

ham^{SK1} alleles (**Supplementary Fig. 1E**). Thus, the PR-containing isoforms of Ham are dispensable for sensory organ development and viability.

However, *ham*$^{PR\Delta}$ mutant animals showed impaired fertility and abnormal morphology of the RS in both males and females (**Fig. 1D-E, Supplementary 2A-B**). The tubular structure of the TE, AG and SV is all disrupted in males, and the mature eggs in females are stuck in the mutant ovaries, indicating a defect in egg delivery. To confirm that the defects are due to *ham* mutations, we performed a rescue experiment by utilizing a pBAC transgene that carries the genomic region of *ham* with the super-folding GFP (sfGFP*)* coding sequence in frame with the *ham* ORF before the stop codon. The pBAC transgene rescued the viability of all *ham* mutant conditions, and at least partly restored male fertility of these mutants. These results indicate that the *ham*$^{PR\Delta}$ allele represents a hypomorphic condition that maintains the shorter isoforms and partial activity of Ham. Although viability and sensory organ development do not require the PR-containing isoforms, formation of the RS depends on them.

Ham is expressed in the epithelial cells connecting the gonadal and supportive tissues

The expression and function of *ham* had not been studied in the RS. To examine whether Ham directly controls RS development, we checked Ham expression in the developing RS. A V5 antibody detects all Ham isoforms in the *ham_V5* allele which contains endogenously tagged V5 epitope at the C-terminus of Ham [34]. In the adult male, Ham expression was detected in the SV and testis terminal epithelium (**Fig. 1F**). At the late L3 stage Ham_V5 was detected in the terminal epithelial cells of the TE and btl-Gal4-expressing mesenchymal cells in the genital disc which will give rise to the SV and AG (**Fig. 1G-H**). At the pupal stage, a high level of Ham persists in the TE terminal epithelial cells, SV and AG. In contrast, Ham is barely detectable in the myoblasts of the SV labeled by the reporter of mef2-Gal4-driving UAS-GFP (**Fig. 1I-J**). Similarly, in the female, high levels of Ham_V5 were found at the joint sites of oviducts and ovaries during both pupal and adult stages (**Supplementary Fig 2C**). These results suggest that Ham is an intrinsic factor in cells at the joint sites between the gonadal- and genital-disc derived epithelial tissues in both males and females.

Ham is required for heterotypic epithelial tissue fusion in the developing male RS

The specific expression of Ham and mutant phenotypes in the RS prompted us to investigate its function in RS epithelial tissues. We decided to focus on the male, as the assembly of the TE and SV represents a remarkable process involving a series of tissue reorganization events (**Fig. 1A-C**). When the two ends of TE and SV get close enough to each other, the epithelial cells at the leading edges meet and form a common epithelial tube. In the meanwhile, myoblasts from the SV and pigment cells from the TE migrate toward opposite directions, eventually forming distinct cell layers covering both structures.

We found that already at 40-50h APF the TE and SV tubes in *ham$^{PR\Delta}$/Df and hamSK1/Df* mutant pupa appeared shorter and wider compared to that in the wild-type (**Fig. 2A**). The *ham* null allele *hamSK1/Df* showed more severe defects, as the AG and SV tubes did not separate and appeared as a pair of round-shaped epithelial pouches. Thus, the severity of RS abnormality correlates with extent of reduction in *ham* gene products. To examine epithelial integrity, we performed immunofluorescent staining using the antibodies against the epithelial cell markers, E-Cad and the septate junction (equivalent to mammalian tight junction) protein Coracle (Cor). We also incorporated an SV-specific transgenic reporter, btlGal4-driving UAS-EGFP or UAS-CD8GFP, to the control and *ham* mutant genotypes. E-Cad is enriched at the apical side (the lumen side, **Fig. 2B**) of the TE and SV epithelial tubes. Notably, the TE terminal epithelium accumulates a higher level of E-Cad compared to the SV epithelium, while Cor is mainly expressed at the SV but not much in the TE. These observations suggest that the epithelial ends of the TE and SV possess distinct properties despite both being epithelial cell types.

Control animals always showed connected TE and SV. However, *ham* mutant TE and SV tissues are often separated from each other. And even when some mutant TE and SV tubes were in close contact, they were not connected (**Fig. 2B**). Importantly, this phenotype can be fully rescued by the *pBAC_ham_sfGFP* transgene, confirming that the phenotype is caused by the *ham* mutation. The next question is how Ham regulates the fusion process.

Depletion of Ham abolishes TE terminal epithelial cell formation

A close inspection on the pupal TE revealed that most of TE epithelial cells were missing in *ham* mutant animals (**Fig. 2B**). TE terminal epithelial cells derive from male specific gonadal precursors (msSGPs) specified in the early embryo (**Supplementary Fig. 3B**) [14] and

expanded during larval and early pupal stages. We found that *ham* mutant lava either lacked TE epithelial cells or showed abnormal location of these cells, which can be again rescued by the *pBAC_ham_sfGFP* transgene (**Supplementary Fig. 3A**). This result indicates that Ham regulates specification or survival of these cells at an earlier developmental timepoint. Indeed, we detected Ham expression in msSGPs and other SGPs in the embryonic gonad. The number of msSGPs labeled by Abd-B or Sox100B-GFP is dramatically reduced in the *ham* mutant embryo already at S13 (**Supplementary Fig. 3C-G**). Thus, Ham is a master regulator for the specification and/ or survival of msSGPs, and the absence of msSGPs should account for loss of TE terminal epithelial cells in larval and pupal stages.

Patterning and differentiation of the SV epithelial tubes depends on Ham

On the other end, epithelial cells in the SV and AG originate from a group of mesenchymal cells recruited by the FGF pathway mediated by Btl and Bnl (**Fig. 1A** and **1H**) [14, 15]. After residing in the GD, these cells undergo MET, form an epithelial mass with inner cavity and eventually elongate into an epithelial tube. In comparison with the control, the signal intensity of both E-Cad and Crumbs are significantly reduced in *ham* mutant SV epithelial cells, which coincides with the shortening of septate junctions labeled by Cor (**Fig. 2C-E**). This result suggests that Ham regulates epithelial cell differentiation in the SV. To check whether Ham regulates epithelial cell differentiation in the TE, we generated mosaic somatic mutant clones using the FRT/FLP system (see Methods) to overcome the issue of epithelial cell loss in the full knockout mutant. In these animals, the mutant cells are GFP negative, while wild-type and *ham* heterozygous cells express GFP (**Fig. 2F**). E-Cad levels at the border between two mutant cells are lower than those between wild-type cells, indicating that *ham* depletion impaired differentiation of TE terminal epithelial cells.

In summary, these phenotypic analyses suggest that Ham controls formation of TE terminal epithelial cells in the early embryo and regulates epithelial tube elongation and cell differentiation at the pupal stage. Thus, failed epithelium fusion in *ham* mutant RS could result from: 1) absence of TE terminal epithelial cells, 2) defects in epithelial tube elongation and cell differentiation of the SV and 3) a combination of both.

Ham directly controls the epithelial fusion process between the TE and SV

Next, we sought to assess whether and how Ham directly regulates the fusion process. We first attempted to generate a mild mutant condition which can bypass defects of TE epithelial cell loss and SV-AG tube elongation. RNAi knockdown is a good tool because knockdown efficiency can be enhanced by overexpressing the microprocessor Dicer2 (Dcr2) and modulated by temperature when RNAi is controlled by the GAL4/UAS system.

However, there was no known GAL4 line that specifically acts in the TE terminal and SV epithelial cells. To identify a Gal4 driver expressed in these cells, we screened the available Gal4 lines in which GAL4 expression is driven by a series of genomic fragments spanning the *ham* locus [35] (**Supplementary Fig. 4A**), and found that one of them induced high expression of *UAS-GFP* in the TE and SV epithelial cells and neglectable expression in the myoblasts (**Supplementary Fig. 4B**).

We confirmed that this GAL4 line (thereafter named as *hamRSGAL4*), when driving a *ham* RNAi line, nearly abolished expression of Ham when the animals carried the *UAS-dcr2* transgene (**Fig. 2G-H,** the middle panels). These animals also had a reduced number of TE epithelial cells and displayed a ball-like structure of the GD, similar to the *ham* null allele. When UAS-dcr2 was omitted, *ham* RNAi knockdown reduced Ham expression but preserved TE epithelial cells and normal-appearing SV. Remarkably, the TE and SV epitheliums still failed to fuse (**Fig. 2H,** the right panel), suggesting Ham also controls the late step of fusion. Furthermore, E-Cad staining intensity was also significantly reduced in *ham* RNAi TE but not in *ham* RNAi SV (**Fig. 2I-J**), suggesting that expression of E-Cad in the TE is sensitive to even mild reduction of Ham protein levels.

Ham regulates expression of epithelial cell genes in the testis and SV.

Ham is a transcription factor, and its function in the developing RS should be mediated by its downstream target gene network. To dissect this network, we performed RNA-seq experiments from dissected TE and SV tissues during the stage of 18-25h APF. We used two pairs of genotypes including the wild-type (*wt*), and $ham^{PR\Delta}$ homozygous and the *wt/Df* and *hamSK1/Df* trans-heterozygous (**Fig. 3A**). The comparison between the first pair will determine dysregulated genes in the $ham^{PR\Delta}$ hypomorphic condition, while the comparison between the

second pair will identify genes that change expression between the null and *ham* heterozygous conditions.

Using an FDR < 10%, we found 2,833 and 2,546 up-regulated genes and 3,684 and 2,379 down-regulated genes in the GD for $ham^{PR\Delta}$ *vs wt* and *hamSK1/Df vs wt/Df* respectively. In mutant TEs, there were 3,330 and 2,274 up-regulated and 2,405 and 2,057 down-regulated genes for the same comparisons (**Fig. 3B**). The genes that only changed expression in one condition could result from unspecific genetic background, differential dependence on the PR-containing isoforms or different sensitivity to *ham* heterozygous condition. Nevertheless, these genes served as a starting point for us to dissect the regulatory network required for epithelial fusion in the developing RS.

Ham activates expression of epithelial genes

Expression of many of the dysregulated genes can be indirectly affected by Ham. We next intended to identify genes that are directly regulated by Ham at the transcriptional level. To this end, we determined Ham genomic binding sites using Cleavage Under Targets and Tagmentation assays (CUT&TAG) [36] on finely-dissected TE and grossly-dissected abdominal tissues in which Ham is mainly expressed in the TE and SV epithelium (**Fig. 3A**). We used a Ham antibody for wildtype tissues and a GFP antibody for tissues from the *pBAC_ham_SFGFP* transgene and IgG as a negative control in parallel for all the samples. Using the SEACR software, we identified peaks and intersections of peaks across the Ham and GFP antibody conditions. We then selected the top 6000 and top 1800 peaks for TG and TE samples, respectively (see Methods). For subsequent analyses, only the 3433 and 691 peaks shared between these conditions were included (**Fig. 3C-D**).

To identify Ham-regulated genes in the developing RS, we intersected the RNA-seq data from the hamSK1/Df versus wt/Df comparison with the Cut&Tag data. Among the 2,546 up-regulated genes in the GD, 441 (17.3%) contained Ham CUT&TAG peaks. Among the 2379 downregulated genes, 569 (23.9%) showed Ham binding (**Fig. 3E**). In the TE, a lower fraction of dysregulated genes contains Ham CUT&TAG peaks, including 4.7% (108/2274) upregulated genes and 7% (144/2057) downregulated genes (**Fig. 3F**), suggesting that dysregulated genes in the mutant TE are mainly indirectly controlled by Ham.

Ham was known for its repressive activity via interacting with the co-repressor CtBP and influencing histone modifications [31]. The observation that downregulated genes contain Ham binding peaks prompted us to examine whether Ham can also activate gene expression. To test this, we chose three downregulated target genes (**Fig. 4A**), *dwnt2*, *shot* and *shg*, for transcriptional reporter assays. *shg* encodes E-cad, while Shot is a member of the spectraplakin family of large cytoskeletal linker molecules that can simultaneously bind actin and microtubules. The function of *wnt2* was reported in the developing RS development [37, 38]. We cloned four genomic fragments from the *wnt2* locus, two each from *shg* and *shot* into a luciferase construct, and co-transfected each reporter with individual plasmids expressing a Ham long isoform and a PR truncation version of Ham, the NZF and the CZF fused with the activation domain of VP16 (NZF_VP16 and CZF_VP16) (**Fig. 4B**). while the two ZF domains of Ham are putative DNA binding domains, it remains undetermined whether they help to recruit Ham to DNA. This experimental setting allows to assess which of the ZF domains is able to target Ham-regulated enhancers. Interestingly, the full-length, the PR truncation and the NZF_VP16 fusion all activated expression of the reporters driven by 7 out of 8 Ham-bound elements (**Fig. 4C-F, Supplementary 5A-B**). By contrast, the CZF_VP16 had little activity on the reporters. This result demonstrates that Ham can directly activate target gene expression and that the NZF domain directly or indirectly recruits Ham to enhancers.

To further identify co-factors with which Ham interacts to regulate gene expression, we performed *de novo* motif discovery from Ham CUT&TAG peaks in mixed TE and GD tissues close to up-regulated and down-regulated genes in the *hamSK1* mutant using the MEME-ChIP software. The parameters used were: -ccut 100, -meme-p 5, -dna, -meme-mod anr, -minw 5, -maxw 15, and -filter-thresh 0.05. This resulted in nine and five significantly enriched motifs from up- and down-regulated gene loci respectively (**Fig. 4G**). Although five motifs were shared between these two sets, the up-regulated genes are uniquely enriched for motifs of Dpp/BMP signaling factor MAD, the zinc-finger protein ZNF263 and two unknown factors. The result implies that Ham may switch activation to repression activity when interacting with these co-factors.

Ham controls distinct molecular pathways to promote tissue assembly.

To systematically explore the function of Ham-regulated genes, we performed gene network analysis using STRING [39]. We focused on the genes bound by Ham and also dysregulated in the genital disc (GD) of the null condition *hamSK1*, considering that most of TE genes are only indirectly influenced in the ham^{SK1} mutant. The String MSL clustering analysis generated 114 functional clusters from down-regulated genes (**Fig. 4H, Supplementary Table 1**) and 62 clusters from up-regulated genes (**Supplementary Fig. 5C, Supplementary Table 1**). Interestingly, many of downregulated genes form into clusters enriched for cytoskeleton regulators, metabolic factors, cell adhesion and extracellular matrix proteins, suggesting that Ham activates genes to promote epithelium differentiation and remodeling. In contrast, upregulated genes are enriched for patterning factors, cell cycle genes and transcriptional and translational regulators, suggesting that Ham suppresses early developmental processes. Notably, although metabolic factors are enriched in both categories, these factors belong to distinct functional groups. Thus, epithelial fusion requires complex and balanced metabolic processes. Together, this network analysis confirms that Ham positions on top of distinct molecular pathways to facilitate TE and SV assembly.

Ham downstream effector genes required for male RS fertility.

We next intended to identify genes crucial for the fusion process. We reasoned that as the majority of TE terminal epithelial cells are lost in both *ham* mutant alleles, the common downregulated genes in $ham^{PR\Delta}$ and ham^{SK1} TE (**Fig. 3B**) may be expressed and functionally important in the *wt* TE epithelial cells. We thus focused on the top 241 downregulated genes that show a fold change greater than 1.5, and the three validated Ham target genes: *shot*, *wnt2* and *shg*. Using all available RNAi lines from the Bloomington stock center that cover 113 out of 244 candidate genes, we performed a UAS-RNAi screen with the *UAS-dcr2*; *hamRSGal4* driver for male fertility (see Methods).

The RNAi screen led to identification of 26 genes whose knockdown abolished or reduced male fertility (**Supplementary Table 2**). Knocking down four of these genes, *shg*, *AP2-alpha*, *pebble (pbl)* and *CG11406*, resulted in complete sterility, indicating their importance for male fertility. We define this category as 100% penetrance of sterility. Knockdown of other twenty-two genes reduced fertility to different extent (**Fig. 5A-B**).

Consistent with the gene network analysis (**Fig. 4H**), these genes include cytoskeleton regulators (*pbl*, *shg*, *shark*) [40-42], and metabolism factors (*CG11406*, *Ph4alphaMP*, *GstE11*, *CG17477*, *CG11069*, *AdSL*, *CG17283*, *CG15117*, *CG7484*). Other genes are known or predicted to be involved in intracellular membrane trafficking (*AP-2alpha*, *CG31935* and *CG7456*) [43] and signaling transduction (*wnt2*, *Toll*, *AkhR*, *Idgf4* and *Hr46*) [44-47].

We further confirmed that knocking down *shg*, *pbl*, *AP-2alpha* and *CG11406* led to failed fusion with abnormal TE and SV morphology at the pupal stage (**Fig. 5C-G)**. Remarkably, *shg* knockdown also led to aberrant mobilization of TE epithelial cells to the middle of the TE, while no such effect was found in the SV and AG (**Fig. 5D)**. Downregulation of E-Cad is a key event during epithelial to mesenchymal transition (EMT), and E-Cad promotes cohesive cell migration of mesenchymal cells but suppresses cell motility of epithelial cells [48, 49]. This result suggests that the SV/AG and the TE terminal epitheliums have distinct dependence on E-Cad and that TE epithelial cells are prone to becoming motile when E-Cad is reduced.

pbl encodes a Rho guanine nucleotide exchange factor whose function was implicated in multiple processes involving cytoskeleton reorganization, like in wound healing, cytokinesis and gastrulation [40, 50, 51]. Our result demonstrates that such function of *pbl* is also required for the epithelium fusion process in the RS (**Fig. 5E)**.

CG11406 encodes a predicted triglyceride lipase, predicted to cleave triacylglycerol (TAG) to produce diacylglycerol (DAG). In insects, TAGs are used for storage of energy and fatty acid precursors of semiochemicals, while DAGs are the main form of transported lipid. This result (**Fig. 5F)** reveals that CG11406 and its mediated lipid metabolism contributes to epithelial tissue fusion.

AP-2alpha, a subunit in the Adaptor complex 2, is involved in intracellular trafficking and secretion. It was shown that AP-2 regulates the apical-basal polarity of epithelial cells via controlling endocytosis of the apical protein Crumbs [52]. Our result shows its importance in epithelium assembly of the RS (**Fig. 5G)**.

Together, this screen identified a collection of known and novel factors involved in epithelial tissue fusion, confirming that the fusion process requires coordinated activity of several classes of molecules.

Spatial and temporal expression dynamics of Ham downstream effectors in the male RS

During epithelial fusion, cells at the leading edge become specialized and express genes required for tissue interaction, remodeling and assembly [1, 2]. This means some genes must express at a high level in the leading edge. Moreover, as we have observed, the TE terminal epithelium and SV epithelium have distinct properties, suggesting that there should be specific genes that exhibit spatial and temporal dynamics in these epithelial cells. We thus intended to identify such genes and assess spatial distribution of distinct gene classes. We conducted a high-resolution multiplexed RNA *in situ* hybridization assay known as SCRINSHOT (single cell RNA *in situ* hybridization on tissues) [53] to visualize the spatial distribution of candidate genes including *wnt2*, *shg*, *shot* and other positive genes from the RNAi screen, along with the receptor genes in Wnt2 signaling (*fz*, *fz2*, *fz3* and *fz4*) and ligand genes in Tl-signaling (*spz*, *spz4*, *spz5* and *spz6*) (**Supplementary Table 3**). Notably, some of Wnt2 and Tl signaling genes displayed dysregulation in at least one of the ham mutant alleles according RNA-seq data (**Supplementary Fig 6A-B**). Other receptor genes in Wnt signaling and ligand genes in Tl signaling were omitted due to low read numbers in the RNA-seq data, indicating negligible or minimal expression in the TE and GD.

Each SCRINSHOT reaction took place within a mini-chamber, accommodating pairs of control and *ham* mutant tissues, or GFP KD control and ham RNAi KD samples (refer to Methods for details on probe design and experimental procedures). This setup was designed to minimize technical variability between the genotypes. mRNAs are visualized as fluorescent dots in SCRINSHOT images, with each dot representing a single mRNA molecule. However, due to the densely packed nature of cells in the TE and SV tissues, quantifying mRNA molecules at a single-cell resolution is challenging. Therefore, we opted to quantify fluorescent signal intensity within defined areas (see Methods). In total, 37 genes were tested, including *ham* itself (**Supplementary Table 3**). Resembling Ham protein expression (**Fig. 1G**), *ham* mRNAs were detected in the TE terminal epithelium (**Fig. 6A**). Similarly, *wnt2*, *shg* and *shot* all exhibit high expression in the TE terminal epithelium (**Fig. 6A**).

Probes for 7 genes gave either no or low signal in the TE or GD, likely attributed to poor probe quality, given their RNA-seq read numbers not being low. Among the remaining 26 genes, 21 exhibited detectable SCRINSHOT signal in the wt SV and AG, while 25 displayed signals in

the wt TE (**Supplementary Table 3)**. Within the TE, 13 out of 25 were specifically localized in the TE terminal epithelial cells, with the remaining 13 detected only in other TE cells (**Fig. 6, Supplementary Fig. 7, Supplementary Table 3**). Thus, SCRINSHOT mapped 30 candidate genes in the developing GD and TE.

Subsequently, our analysis focused on genes expressed in the TE terminal and SV epitheliums. While 17 genes exhibited expression in both epithelial ends (**Fig. 6A-B**), 8 were exclusively expressed in SV epithelial cells (**Fig. 6C-F**, **Supplementary Fig. 7C-D**, **Supplementary Table 3**), underscoring the heterotypic nature of the TE and SV epitheliums. Notably, certain genes (CG11406, *AP-2alpha*, *GstE11*, *Tl*, *spz5* and *spz6*) demonstrated temporal dynamic expression in the TE epithelium: spz5 and CG11406 are present during the 20-25h APF timeframe, prior to fusion, but gone during the 28-32h APF period when the TE and SV are undergoing fusion; *AP-2alpha*, *GstE11*, *Tl* and *spz6* show the opposite trend, being expressed only later (**Supplementary Fig. 7B**).

In the GD, while the majority of genes exhibited ubiquitous expression, *wnt2* and *spz* were predominantly expressed in the leading cells within the SV, contrasting with their expression in the remaining GD (**Fig. 6B**). Additionally, the expression patterns of *fz2*, *fz3* and *Tl* displayed spatial specificity: fz2 exhibited higher expression levels in the SV compared to the AG, *fz3* demonstrated the opposite pattern with higher expression in the AG, and *Tl* was primarily expressed in the myoblasts (**Fig. 6B**). These findings imply that Wnt2 and Tl signaling may have cell-type-specific roles in the fusion process.

Wnt2 and Tl signaling pathways are key mediators of Ham function in RS tissue assembly

Next, we aimed to identify genes exhibiting altered expression in *ham* mutant SV and TE. A deficiency of TE terminal epithelial cells in the *hamSK1* mutant can indirectly influence gene expression in the TE. We thus quantified SCRINSHOT signal levels only in the developing SV between *wt* and the *hamSK1* mutant. In contrast, *ham* RNAi knockdown resulted in fusion defects but only mildly impacted SV morphology and TE terminal epithelial cell number (**Fig. 2G-H**). These samples allow to assess gene expression alterations in both the SV and TE epitheliums.

ham fluorescent intensity appeared comparable between *wt* and mutant samples (**Fig. 6E, Supplementary Fig. 7E**), suggesting that *ham* mutant RNAs, containing the CRISPR-generated indels, remain stable throughout development. In the *hamSK1* mutant SV, six genes became downregulated, while two exhibited upregulation. The downregulated genes were four Wnt2 pathway genes (*wnt2, fz, fz2, fz4*), one Tl pathway gene (*spz*), and *shg* (**Fig. 6B-**C). Notably, *wnt2, fz, spz* and *shg* also displayed downregulation in *ham* RNAi SV (**Fig. 6D-E**), suggesting that these genes and their associated signaling pathways may be directly involved in the fusion process downstream of Ham. The upregulated genes in the *hamSK1* mutant included *GstE11* and *CG11069* (**Supplementary Fig.7C-E**), which exhibited low expression in the leading cells of wt SV epithelium, but increased expression under the *hamSK1* condition. In *ham* RNAi knockdown animals, *CG11069* was again upregulated, but *GstE11* did not change expression. Instead, another ligand gene in Tl signaling (*spz5*) displayed increased expression (**Supplementary Fig.7D, F**). On the TE side, only four genes showed expression alteration in *ham* RNAi animals, including *ham* itself, *shg*, *fz* and *fz2* **Supplementary Fig.7G**), suggesting that *shg*, *fz* and *fz2* may function in both TE and SV epitheliums.

The observations that knocking down Tl resulted in decreased male fertility (**Fig. 5B**, **Supplementary Table 2**), coupled with the expression of multiple *spz* ligand genes in the developing TE and SV, suggest that the Tl signaling pathway might be required for assembling the TE and SV. To further investigate this possibility, we examined male flies with *spz* knockdown driven by *hamRSGal4*. These animals produced a comparable number of progenies to RNAi control animals. The lack of phenotype in this condition could be attributed to genetic compensation by another *spz* gene, like *spz5*, whose expression increased in *ham* RNAi animals where *spz* expression was reduced. In addition to Tl, two other Tl related genes (Toll-7 and Toll-9) exhibited moderate number of reads in the RNA-seq data. We then tested male fertility in Toll-7 and Toll-9 knockdown animals, and found a drastic reduction in progeny numbers compared to the RNAi control (**Supplementary Fig. 8**). Together, these results pinpoint an essential role of Tl signaling in male RS development.

To validate the role of Wnt2 signaling in RS development, we investigated four genetic conditions, aimed at modulating its level and activity: $wnt2^{O}$ (a *wnt2* null allele), *wnt2* overexpression driven by *hamRSGal4*, *fz* and *fz2* RNAi knockdown driven by *hamRSGal4*. Homozygous mutant pupae of $wnt2^{O}$ exhibited abnormality in both TE and GD (**Fig. 7A**), underscoring the crucial role of Wnt2 in male RS development. Overexpression of *wnt2* also

resulted in fusion defects (**Fig. 7B**), with some TE epithelial cells exhibiting mis-localization, a phenotype reminiscent of *shg* knockdown animals. Indeed, *wnt2*-overexpressing cells displayed reduced levels of *shg*, suggesting that Wnt2 signaling might suppress *shg* expression to facilitate cell motility. Consistently, *fz2* knockdown animals exhibited comparable phenotypes, including mobilized TE epithelial cells and failed tissue fusion (**Fig. 7C**). *fz2* and *fz* knockdown also decreased male fertility (**Fig. 7D**). Collectively, these results suggest the importance of optimal Wnt2 signaling levels in male RS development and epithelial tissue fusion.

Discussion

In this study we demonstrate that the transcription factor Ham plays a crucial role in controlling epithelial fusion during *Drosophila* RS development. Early in embryonic development, Ham determines the cell fate of msSGPs, which are progenitor cells for TE terminal epithelial cells. Later on, Ham directly regulates the differentiation of epithelial cells in both the TE and SV by orchestrating a gene expression program in epithelial cells. While our primary focus is on male RS, we found that Ham is also expressed in cells at the fusion site of the oviduct and ovary. Mutant females lacking functional Ham exhibit a non-functional oviduct epithelial tube, suggesting that Ham likely employs a similar mechanism to control epithelium fusion in females.

By examining *ham* mutant phenotypes and its downstream target genes, we identified several classes of novel factors involved in fusion process of SV and TE. These genes, directly or indirectly controlled by Ham, encompass a diverse array of molecular functions. This discovery provides a comprehensive gene network for future investigation of epithelial fusion processes. The cell-type-specific and tissue-specific spatial distribution of these genes suggest that they may regulate various aspects of fusion, including cell migration, intercellular communication between myoblasts and epithelial cells, specification of the leading cells, and other steps in the fusion process. For instance, Tl and its related receptors is known to mediate cell adhesions [54]. We can speculate that Tl signaling and the molecule itself promote interaction between myoblasts and myoblast and epithelial cells (**Fig. 7E**), based on high expression of Tl in myoblasts and presence of multiple Spz ligands in the TE and SV epithelial cells. Moreover, *Wnt2*, *fz* and *fz2* are all highly expressed in the leading cells, and also required for TE and SV connection, suggesting that Wnt2 signaling is an essential component mediating epithelial remodeling and fusion.

Despite being discovered more than twenty years ago [27], research on the function of Ham function has primarily focused on its roles in the nervous system. Our findings highlight a crucial function of Ham in epithelial tissues. Apart from the RS, we observed high expression of the Ham protein in other epithelial tissues including the Malpighian tubes (akin to the kidney in insects) and the intestine (data not shown). Furthermore, orthologs of Ham, PRDM16 and MECOM, exhibit high expression in epithelial tubular structures, such as the pancreas, kidney, urinary bladder and gastrointestinal tract and lung, according to observations in the Human

Protein Atlas. This suggests a potential role for these proteins in controlling epithelial cell integrity and fusion in mammalian organogenesis. Indeed, we observed that the pancreas duct volume in mouse embryos lacking *Prdm16* is significantly reduced (data not shown). Although the exact cellular changes and the underlying molecular mechanisms in this context await further investigation, this phenotype underscores a crucial function of *Prdm16* in pancreatic epithelial duct formation.

We show that compared to Ham function in the nervous system, the PR-containing isoforms are indispensable for RS epithelial fusion and differentiation, and that Ham can directly activate epithelial gene expression. It remains to be determined whether Ham has distinct molecular functions or it regulates a subset of common molecular pathways in these two contexts. Dendrite morphogenesis involves outgrowth and branching, somewhat analogous to epithelial tube growth and branching. For instance, molecular mechanisms regulating polarized protein trafficking are crucial for both dendrite and epithelial tube branching [55]. Additionally, several transcription factors like Sequoia [56, 57], Tramtrack [58, 59] and Cut [60, 61] play roles in both dendrite morphogenesis and tracheal tube fusion. These examples suggest that there might be more unrecognized common mechanisms between dendrite and epithelial morphogenesis. It is of interest to test whether the effectors we identified in the RS epithelium are involved in dendrite morphogenesis in the nervous system.

We observed a unique aspect of epithelial fusion in the *Drosophila* RS. Unlike many epithelial fusion processes where two homologous ends merge, such as those in mammalian neural tube closure and *Drosophila* tracheal formation *etc* [1, 8, 10], the two epithelial ends in the *Drosophila* RS are heterotypic. Specifically, the TE terminal epithelial cells express higher levels of E-Cad but lower levels of septate junction protein Cor compared to SV epithelial cells. In response to the reduction of E-Cad or ectopic expression of wnt2, TE terminal epithelial cells become motile, resembling a process akin to epithelial-to-mesenchymal transition. On the other hand, SV epithelial cells, despite originating from mesenchymal cells, exhibit typical epithelial cell characteristics, including regular Cor and E-Cad expression. Notably, many genes are expressed exclusively in the SV epithelium, distinguishing it from the TE epithelium. Heterotypic epithelial fusion occurs in many developmental processes, such as digestive system formation, vertebrate optic fissure closure, and mammalian urogenital development. Failure of epithelial fusion can lead to a wide variety of developmental diseases such as oesophageal atresia, spina bifida, ocular coloboma, congenital abnormalities of the kidney and

urinary tract and persistent cloaca. Studying the basic processes and key players of normal epithelial tissue fusion in an *in vivo* context will help to understand the pathogenesis of these diseases.

Acknowledgements

We thank the imaging facility (IFSU) of Stockholm University, the BEA sequencing facility at Karolinska Insitute, and the National Genomic Institute of Scilife Laboratories, Sweden, for providing service and support. The Bloomington Stock Center and the *Drosophila* community provided important reagents and fly stocks. We thank Christos Samakovlis's lab for helping to set up SCRINSHOT and Eric Lai's lab for the initial Ham reagents. Funding from Australian Research Council Centre of Excellence for the Mathematical Analysis of Cellular Systems (CE230100001) and ANU Future Scheme to J.W. The project was supported by project grant from Swedish Research Council (Vetenskapsrådet, 2020-03543) and Swedish Cancerfonden (CAN 22 2161 Pj) to Q.D.

Author contributions

Conceptualization and project design by QD, experiments conduction by HW, preliminary test in pancreas by IA and LB, computation analysis by JW. Data interpretation by HW, JW and QD. Manuscript writing by QD with input from HW and JW.

Materials and methods

Fly strain culturing and generation of transgenes

All fly stocks were maintained at 22°C. The *ham*$^{PR\Delta}$ mutants were created using CRISPR: transgenic flies carrying single guide RNA targeting the PR domain containing isoforms were crossed into the vas-Cas9 transgenic flies. Frame-shift mutations were identified by PCR and Sanger sequencing. Transgenic flies were created at BestGene, Inc.

The ham Gal4 lines and RNAi stocks were obtained from Bloomington Stock Center. Knocking down experiments were performed at 25 °C and with a regular 12-hr light–dark cycle.

Fertility test

For ham mutants, single male or female with indicated genotype was mated with two w^{1118} virgin females or two w^{1118} males on day 0 of the experiment. Crosses were flipped every 2 days for 3 consecutive times. Progenies from the second vial with eggs laid between day 2-4 were counted after 2 weeks.

For the RNAi screen, individual UAS-RNAi males were crossed with females carrying UAS-dcr2 and hamRSGal4. The resultant F1 male progenies (2-5 days old) with the genotype UAS-dcr2; hamRSGal4>UAS-RNAi were crossed to w^{1118} virgin females (1-2 days old) on day 0 of the experiment. In each test, four males were crossed with three females. Crosses were flipped on day2 and day3 for pre-laying. After flipping on day3, crosses were kept in the new vial for egg laying for a specific time period (GroupA: 48hours; GroupB: 16-18h; GroupC: 7-8h). If at least three males and two females were alive, then the flies were transferred to a new vial, the old vial was kept, and all progenies were counted.

For fertility test of the Wnt2 and Tl signaling pathways components, UAS-RNAi males were crossed to virgin females of the genotype UAS-dcr2; hamRSGal4. Two F1 male progenies (1-2 days old) with the genotype UAS-dcr2; hamRSGal4>UAS.RNAi were crossed to three w^{1118} virgin females (1-2 days old) on day 0 of the experiment. Crosses were flipped every day until day10. Progenies in vial 1 to 9 were counted separately.

Immunostaining

For embryo staining: embryos were collected, dechorionated with bleach and fixed with 4% formaldehyde in PEM (0.1M PIPES, 2mM EGTA, 1 mM MgSO4, pH 7) for 20 min. For immunostaining, embryos were rehydrated, washed with PBT (0.3% Triton X-100 in PBS) and blocked with PBSBT (0.3% Triton X-100, 1% BSA in PBS).

For tissue staining: tissues were dissected from stage defined larvae or pupae in 1xPBS within 30 min and fixed in 4% paraformaldehyde in PBS for 20min. For immunostaining, dissected tissues were washed with PBT (0.1% Triton X-100 in PBS) and blocked with PBSBT (0.1% Triton X-100, 1% BSA in PBS).

For both embryo and tissue staining, the primary antibodies were diluted with PBSBT (0.1% Triton X-100, 1% BSA in PBS) and incubated at 4°C over night. After extensive washing with PBSBT (15min x 4 times), the secondary antibodies were added in 1:500 dilution and incubated at room temperature for 2h. After washing with PBT (0.1% Triton X-100 in PBS), the samples were counter-stained with DAPI and mounted in Vectashield (Vector Laboratories).

The following primary antibodies were used: Rabbit anti-Ham (1:300) (Gift from Eric Lai's lab), Ms anti-V5 (1:500, R960-25, invitrogen), chk anti-GFP (1:2000, ab13970, Abcam), Rat anti-vasa (1:50, anti-vasa-c, Developmental Studies Hybridoma Bank, DSHB), Ms anti-ABD-B (1:50, 1A2E9-c), Ms anti-SXL (1:50, M18-c), rat anti-DE-Cad (1:50, DCAD2, DSHB), mouse anti-Crb (1:10, Cq4 DSHB), mouse anti-Coracle (1:100, C615.16 DSHB). Secondary antibodies conjugated to Cy3 or Cy5 or Alexa Fluor-488 (Jackson ImmunoResearch) were used.

Quantification of signal intensity from fluorescent immunostaining

For E-Cad and Crb in Figure 2, "raw intensity density" from manually defined regions (using free-hand drawing) along the apical side of the SV and TE terminal epitheliums was obtained in ImageJ. For each sample, 2 regions in the SV and 3 regions in the TE terminal epithelium were averaged. The regions used to determine relative E-Cad fluorescence in *hamSK1* TE somatic clones were at the apical side between cell junction. The length of Cor-stained cell junction was measured using ImageJ. A straight line was drawn over the Cor staining with the 'Freehand Line tool'. For each sample, the average length from 10 cell-cell junctions were used. In all experiments, images for the genotypes in comparison were always processed using exact settings. At least three biological replicates were included in statistical analysis.

Western Blot

Embryos from *w1118*, *ham$^{PR\Delta}$/CTG*, *ham^{SK1}/CTG* were collected and aged to embryonic stage S13 to S16. In each blot, 50-100st GFP+ or GFP- embryos were used. The selected embryos were homogenized directly in 1xSDS loading buffer (62.5 mM Tris-HCl pH 6.8, 2.5 % SDS, 0.002 % Bromophenol Blue, 0.7135 M (5%) β-mercaptoethanol, 10 % glycerol) and denatured at 98 °C heat blot for 5 min. After centrifuged, supernatant was taken for Western-Blotting with 8% of SDS-PAGE. Rabbit anti-Ham (1:2000) primary antibody was used in blotting. To minimize unspecific signal from IgG heavy chains, light chain specific secondary antibody of mouse anti-Rabbit IgG (1:40000 Jackson ImmunoResearch) was used, and the signals were developed with ECL Plus reagent (GE Healthcare).

Cell culture and Luciferase assays

To generate different variants of Ham, the open reading frames of HamFL, Ham-NZF-vp16, Ham-CZF-vp16, and *ham$^{PR\Delta}$* were PCR amplified and cloned into the pAC_V5 (nitrogen) vector. The enhancer fragments were cloned in to the p2xTetO-Firefly luciferase vector. All transfections were performed using *Drosophila* S2-R+ cells grown in Schneider *Drosophila* medium containing 10% fetal calf serum. Cells were co-transfected with the Ham expression construct, 2xTetO-Firefly luciferase and pAc-Renilla plasmids in 96-well plate using the Effectene Transfection kit (Qiagene). Luciferase assays were performed and measured using the Dual Luciferase Assay System (Promega). Expression was calculated as the ratio between the firefly and *Renilla* luciferase activities. Three rounds of transfections were performed on different days and different cell populations and considered as three biological replicates. Within each biological replicate, 4 wells of cells were transfected with the same DNA mixture and measured and thus considered as 4 technique replicates. In total, for each DNA combination, the values from the 12 wells of transfection were calculated, averaged and plotted in Fig. 5D-F. Cloning oligos are provided in **Supplementary table 4**.

SCRINSHOT

The Drosophila reproductive system was dissected at defined ages in 1xPBS and fixed in 4% paraformaldehyde in PBS for 20min. After washing with PBST 0.1% (Triton X-100) for 2 x 15 min, fine dissected samples were placed on Poly-lysine coverslip and freezed at -80°C after drying on a heater.

The SCRINSHOT experiments were carried out according to the published method [53]. In brief, 3 padlock probes and 3 corresponding detection probes were designed for each gene of interest. The slides with fine dissected samples were pretreated at 45˚°C to reduce moisture and fixed in 4% PFA in 1X PBS, followed by washing in PBS Tween-20 0.05% twice. Samples were then dehydrated with 70%, 85%, 100%ETOH and air dry. The SecureSeal hybridization chamber (GRACE BIO-LABS, 621501) was then mounted to cover samples on the slide. Permeabilization of tissues were first done by treating with 10µg/mL Proteinase K (Invitrogen) in PBSTween-20 0.05% for 7 minutes followed by washing with Glycine (2mg/mL in PBS Tween-20 0.05%) once and with PBS Tween-20 0.05% for 3 x 5 minutes. After this, samples were further Permeabilized in 0.1M HCl for 3min, followed by washing in PBS Tween-20 0.05% twice. Samples were then blocked in a probe-free hybridization reaction mixture of 1XAmplifase buffer (Lucigen, A1905B), 0.05M KCl, 20% Formamide deionized (Millipore S4117), 0.1uM Oligo-dT, 0.1ug/ul BSA (New England Biolabs, B9000S), 1U/ul RiboLock (Thermo, EO0384), and 0.2ug/ul tRNAs (Ambion, AM7119). Then hybridization of padlock probes was done by incubating samples with padlock probes (with the concentration of each one 0.01uM) mixed in blocking reagents used before (no oligo dT used in this step). The slide was then put into a PCR machine to denature at 55˚C for 15 minutes and hybridize at 45˚C for 120 minutes. Padlock probes were then ligated by using SplintR ligase (NEB M0375) at 25˚C for 16 hours followed by RCA (rolling cycle amplification) at 30˚C for 16 hours by using phi29 polymerase (Lucigen, 30221–2) and RCA primer1. Then a fixation step was applied to stabilize RCA product in 4% PFA for 15 minutes followed by washing in PBS Tween-20 0.05%. Then the hybridization of the first 3 genes was done by mixing all 3 3' fluorophore-conjugated detection probes of each gene in reaction reagent (2XSSC, 20% Formamide deionized, 0.1ug/ul BSA, and 0.5 ng/µl DAPI (Biolegend, 422801) followed by hybridization at 30˚C for 1 hour. The slides were washed in 20% formamide in 2X SSC and then in 6X SSC, followed by dehydration in 70%, 85%, 100% Ethanol until the chamber was removed. Then the samples were preserved in SlowFade Gold Antifade mountnant (Thermo, S36936) and kept in dark before imaging. After image acquisition, the first 3 detection probes were removed by using Uracil-DNA Glycosylase (Thermo, EN0362), and the slides were ready for the next round of detection probe hybridization. The procedure was repeated until all genes were hybridized and imaged. Images were acquired with an airyscan-equipped confocal microscope system (Zeiss LSM 800, Carl Zeiss), and ZEN 2.5 Blue edition software was used for image processing.

Quantification of SCRINSHOT signals

Similar to quantification of immunofluorescent signal, fluorescence intensity of SCRINSHOT for each gene was measured in defined areas within images of the SV, EjD, and TE in 2D projections. Squares were drawn to define areas at the leading side of the SV, the middle region of the Ejd and the TE terminal. The total fluorescence intensity was first normalized to area size. The resultant value for *ham* mutant or RNAi samples was then contrasted to the normalized intensity of control samples.

CUT&TAG

CUT&TAG was performed according to the published method [36]. In brief, TE and grossly tissue of the posterior abdomen were dissected from 18-25h APF *w1118* and *pBAC_ham_sfGFP* animals in PBS within 30 min. For grossly tissues, the dissected samples were homogenized in 2ml NE1 buffer followed by incubation on ice for 10 min. After filtering out the grossly cuticles, nuclei were collected by spinning at 1300xg and resuspended in 1ml PBS. For TE, nuclei extraction was done by simply pipetted samples in 200ul NE1 up and down for 20-30 times and incubated on ice for 10 min. Light fixation was performed with 0.1% formaldehyde at RT for 2 minutes and neutralized with 75mM Glycine. Then the nuclei were washed by PBS and wash buffer 1 time each and kept in 200ul wash buffer on ice until enough samples were collected (30-50st animals per/antibody). Resuspended nuclei in 1.2ml wash buffer and mixed with 20ul Concanavalin A-coated magnetic beads (Polyscience 86057) with soft vortexing. Aliquoted samples into two parts immediately and incubated for 10 min with rotation to allow the binding of the nuclei with the beads. The nuclei/beads mix was then blocked with 1mL cold antibody buffer for 10 min and resuspended in 150ul cold antibody buffer. For w1118 samples, 1uL Rabbit anti-Ham or 1ul Rabbit IgG (SIGMA-ALDRICH, I5006) was added for an overnight incubation at 4°C. For *pBAC_ham_sfGFP* samples, 1ul Rabbit anti-GFP (G10362, invitrogen) or Rabbit IgG was used. The beads/nuclei/antibody mix was then washed with Dig-wash buffer and incubated with secondary antibody (1:100 dilution) for 1 hour. After further washes with the Dig-wash buffer, a pA-Tn5 adaptor complex in Dig-300 buffer was added to the beads for 1 hour reaction. The beads were then washed with Dig-300 buffer before the incubation with 300ul Tagmentation buffer at 37°C for 1 hour. Then 10ul 0.5M EDTA, 3uL 10% SDS and 2.5uL 20 mg/mL proteinase K (Invitrogen) were used to stop the reaction. Then the resultant DNA was purified with DNA clean & Concentrator kit (ZYMO Research D4013) and eluted in 25uL Elution buffer. To generate libraries, 21uL fragment DNA was mixed with 2uL 10uM Universal i5 primer, 2uL 10uM uniquely barcoded i7 primer and

25uL PCR master mix (NEB Phusion® High-Fidelity PCR Master Mix with HF Buffer, M0531). The PCR condition was as follows: 72°C 5min, 98°C 30s, repeat 12 times (98°C 10s, 63°C 10s), 72°C 1 min and hold 4°C. The libraries were cleaned with standard Ampure XP beads as previously described. Libraries from four biological replicates were produced for each condition.

We mapped the CUT&TAG samples of *w1118* and *pBAC_ham_sfGFP* in TE and TG tissues to the dm6 genome assembly using Bowtie2 with the following parameters: bowtie2 --end-to-end --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700. The coverage for these samples was generated using bedtools genomeCoverageBed and normalized to library depth to yield reads per million per base (RPM). For peak calling, we employed SEACR [62], which is tailored for sparse chromatin profiling data like CUT&TAG. We first used a relaxed FDR cutoff of < 0.5 for each replicate (SEACR_1.3.sh normalized_coverage 0.5 norm stringent output_peaks), we then identified common peaks across replicates for each condition. To select confident peaks for *w1118* and *pBAC_ham_sfGFP*, we overlapped their peaks within TE and TG separately, choosing the top 1800 peaks from TE and the top 6000 from TG. Peak annotation was performed by HOMER's annotatePeaks.pl. Peak overlap analysis was performed using HOMER's mergePeaks function with the default -d given parameter, ensuring that only directly overlapping peaks were merged. This method preserved the integrity of the original peak boundaries and maintained the specified distances between peaks from the initial peak calls. *De novo* motif discovery for peaks in up-regulated or down-regulated genes, positioned 2000 bp upstream and 500 bp downstream of the TSS in TE and TG tissues, was performed using the MEME-ChIP software. The parameters used were: -ccut 100, -meme-p 5, -dna, -meme-mod anr, -minw 5, -maxw 15, and -filter-thresh 0.05.

RNA-seq

Total RNA was extracted from finely dissected genital discs or testis samples from 18-25h APF animals using Trizol reagent (Invitrogen). RNA quality was tested by Agilent Bioanalyzer.

RNA-seq libraries were made using the Illumina Truseq Total RNA library Prep Kit LT. Sequencing was performed on the Illumina Hiseq2500 platform. Following the trimming of adapter sequences with Trimmomatic, RNA-seq reads from replicated wild type (x3) and mutant samples (x3) were aligned to the *Drosophila melanogaster* (dm6) genome assembly utilizing HISAT2. Gene annotations were sourced from UCSC. The DESeq2 Bioconductor package was employed to identify differentially expressed mRNAs between *w1118* and the $ham^{PR\Delta}/ham^{PR\Delta}$ mutant, as well as between +/Df and ham^{SK1}/Df mutant, with a subsequent

estimation of the false discovery rate (FDR). Genes with an FDR < 10% and absolute Log2FC>0 were considered differentially expressed. Details of the calculated differential expression are provided in **Supplementary Data**.

STRING gene network analyses

Ham activated (genes downregulated in *hamSK1* and bound by Ham in TG) and repressed (genes upregulated in *hamSK1* and bound by Ham in TG) gene networks are constructed using String, MCL clustering (find natural clusters the stochastic flow, inflation parameter, 2. Disconnected genes are hidden. The position of the clusters is manually adjusted to fit related GO categories. The links to the gene networks in STRING:

https://string-db.org/cgi/network?taskId=bbF8IgMTqf1M&sessionId=bzDIAZvMgWsi

https://string-db.org/cgi/network?taskId=bhKCcMoLlbwh&sessionId=bzDIAZvMgWsi

References

1. Bernascone, I., M. Hachimi, and F. Martin-Belmonte, *Signaling Networks in Epithelial Tube Formation.* Cold Spring Harb Perspect Biol, 2017. **9**(12).
2. Chung, S. and D.J. Andrew, *The formation of epithelial tubes.* J Cell Sci, 2008. **121**(Pt 21): p. 3501-4.
3. Samakovlis, C., et al., *Development of the Drosophila tracheal system occurs by a series of morphologically distinct but genetically coupled branching events.* Development, 1996. **122**(5): p. 1395-407.
4. Samakovlis, C., et al., *Genetic control of epithelial tube fusion during Drosophila tracheal development.* Development, 1996. **122**(11): p. 3531-6.
5. Sutherland, D., C. Samakovlis, and M.A. Krasnow, *branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching.* Cell, 1996. **87**(6): p. 1091-101.
6. Tanaka-Matakatsu, M., et al., *Cadherin-mediated cell adhesion and cell motility in Drosophila trachea regulated by the transcription factor Escargot.* Development, 1996. **122**(12): p. 3697-705.
7. Gervais, L., G. Lebreton, and J. Casanova, *The making of a fusion branch in the Drosophila trachea.* Dev Biol, 2012. **362**(2): p. 187-93.
8. Jacinto, A., A. Martinez-Arias, and P. Martin, *Mechanisms of epithelial fusion and repair.* Nat Cell Biol, 2001. **3**(5): p. E117-23.
9. Vasioukhin, V., et al., *Directed actin polymerization is the driving force for epithelial cell-cell adhesion.* Cell, 2000. **100**(2): p. 209-19.
10. Pai, Y.J., et al., *Epithelial fusion during neural tube morphogenesis.* Birth Defects Res A Clin Mol Teratol, 2012. **94**(10): p. 817-23.
11. Caviglia, S. and S. Luschnig, *Tube fusion: making connections in branched tubular networks.* Semin Cell Dev Biol, 2014. **31**: p. 82-90.
12. Jemc, J.C., *Somatic gonadal cells: the supporting cast for the germline.* Genesis, 2011. **49**(10): p. 753-75.
13. DeFalco, T., S. Le Bras, and M. Van Doren, *Abdominal-B is essential for proper sexually dimorphic development of the Drosophila gonad.* Mech Dev, 2004. **121**(11): p. 1323-33.
14. Whitworth, C., E. Jimenez, and M. Van Doren, *Development of sexual dimorphism in the Drosophila testis.* Spermatogenesis, 2012. **2**(3): p. 129-136.
15. Ahmad, S.M. and B.S. Baker, *Sex-specific deployment of FGF signaling in Drosophila recruits mesodermal cells into the male genital imaginal disc.* Cell, 2002. **109**(5): p. 651-61.
16. Rothenbusch-Fender, S., et al., *Myotube migration to cover and shape the testis of Drosophila depends on Heartless, Cadherin/Catenin, and myosin II.* Biol Open, 2017. **6**(12): p. 1876-1888.
17. Fog, C.K., G.G. Galli, and A.H. Lund, *PRDM proteins: important players in differentiation and disease.* Bioessays, 2012. **34**(1): p. 50-60.
18. Chi, J. and P. Cohen, *The Multifaceted Roles of PRDM16: Adipose Biology and Beyond.* Trends Endocrinol Metab, 2016. **27**(1): p. 11-23.

945 19. Bard-Chapeau, E.A., et al., *Ecotopic viral integration site 1 (EVI1) regulates multiple*
946 *cellular processes important for cancer and is a synergistic partner for FOS protein in*
947 *invasive tumors.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 2168-73.
948 20. Zhou, B., et al., *PRDM16 Suppresses MLL1r Leukemia via Intrinsic Histone*
949 *Methyltransferase Activity.* Mol Cell, 2016.
950 21. Corrigan, D.J., et al., *PRDM16 isoforms differentially regulate normal and leukemic*
951 *hematopoiesis and inflammatory gene signature.* J Clin Invest, 2018. **128**(8): p. 3250-
952 3264.
953 22. Cibi, D.M., et al., *Prdm16 Deficiency Leads to Age-Dependent Cardiac Hypertrophy,*
954 *Adverse Remodeling, Mitochondrial Dysfunction, and Heart Failure.* Cell Rep, 2020.
955 **33**(3): p. 108288.
956 23. Kundu, A., et al., *PRDM16 suppresses HIF-targeted gene expression in kidney cancer.*
957 J Exp Med, 2020. **217**(6).
958 24. Horvath, C. and C. Scheele, *PRDM16 stability and metabolically healthy adipose*
959 *tissue.* Nat Metab, 2022. **4**(9): p. 1095-1096.
960 25. Wang, Q., et al., *Post-translational control of beige fat biogenesis by PRDM16*
961 *stabilization.* Nature, 2022. **609**(7925): p. 151-158.
962 26. Hurwitz, E., et al., *Antagonism between Prdm16 and Smad4 specifies the trajectory*
963 *and progression of pancreatic cancer.* J Cell Biol, 2023. **222**(4).
964 27. Moore, A.W., L.Y. Jan, and Y.N. Jan, *hamlet, a binary genetic switch between single-*
965 *and multiple- dendrite neuron morphology.* Science, 2002. **297**(5585): p. 1355-8.
966 28. Moore, A.W., et al., *Conversion of neurons and glia to external-cell fates in the*
967 *external sensory organs of Drosophila hamlet mutants by a cousin-cousin cell-type*
968 *respecification.* Genes Dev, 2004. **18**(6): p. 623-8.
969 29. Rives-Quinto, N., et al., *Sequential activation of transcriptional repressors promotes*
970 *progenitor commitment by silencing stem cell identity genes.* Elife, 2020. **9**.
971 30. Eroglu, E., et al., *SWI/SNF complex prevents lineage reversion and induces temporal*
972 *patterning in neural stem cells.* Cell, 2014. **156**(6): p. 1259-73.
973 31. Endo, K., et al., *Chromatin modification of Notch targets in olfactory receptor neuron*
974 *diversification.* Nat Neurosci, 2012. **15**(2): p. 224-33.
975 32. Pinheiro, I., et al., *Prdm3 and Prdm16 are H3K9me1 methyltransferases required for*
976 *mammalian heterochromatin integrity.* Cell, 2012. **150**(5): p. 948-60.
977 33. Hohenauer, T. and A.W. Moore, *The Prdm family: expanding roles in stem cells and*
978 *development.* Development, 2012. **139**(13): p. 2267-82.
979 34. Gokcezade, J., G. Sienski, and P. Duchek, *Efficient CRISPR/Cas9 plasmids for rapid and*
980 *versatile genome editing in Drosophila.* G3 (Bethesda), 2014. **4**(11): p. 2279-82.
981 35. Jenett, A., et al., *A GAL4-driver line resource for Drosophila neurobiology.* Cell Rep,
982 2012. **2**(4): p. 991-1001.
983 36. Kaya-Okur, H.S., et al., *CUT&Tag for efficient epigenomic profiling of small samples*
984 *and single cells.* Nat Commun, 2019. **10**(1): p. 1930.
985 37. Kozopas, K.M., C.H. Samos, and R. Nusse, *DWnt-2, a Drosophila Wnt gene required*
986 *for the development of the male reproductive tract, specifies a sexually dimorphic cell*
987 *fate.* Genes Dev, 1998. **12**(8): p. 1155-65.
988 38. Deshpande, G., A. Nouri, and P. Schedl, *Wnt Signaling in Sexual Dimorphism.*
989 Genetics, 2016. **202**(2): p. 661-73.

990 39. Szklarczyk, D., et al., *The STRING database in 2023: protein-protein association*
991 *networks and functional enrichment analyses for any sequenced genome of interest.*
992 Nucleic Acids Res, 2023. **51**(D1): p. D638-D646.
993 40. van Impel, A., et al., *Regulation of the Rac GTPase pathway by the multifunctional*
994 *Rho GEF Pebble is essential for mesoderm migration in the Drosophila gastrula.*
995 Development, 2009. **136**(5): p. 813-22.
996 41. Oda, H., et al., *A Drosophila homolog of cadherin associated with armadillo and*
997 *essential for embryonic cell-cell adhesion.* Dev Biol, 1994. **165**(2): p. 716-26.
998 42. Ferrante, A.W., Jr., R. Reinke, and E.R. Stanley, *Shark, a Src homology 2, ankyrin*
999 *repeat, tyrosine kinase, is expressed on the apical surfaces of ectodermal epithelia.*
1000 Proc Natl Acad Sci U S A, 1995. **92**(6): p. 1911-5.
1001 43. Gonzalez-Gaitan, M. and H. Jackle, *Role of Drosophila alpha-adaptin in presynaptic*
1002 *vesicle recycling.* Cell, 1997. **88**(6): p. 767-76.
1003 44. Weber, A.N., et al., *Binding of the Drosophila cytokine Spatzle to Toll is direct and*
1004 *establishes signaling.* Nat Immunol, 2003. **4**(8): p. 794-800.
1005 45. Staubli, F., et al., *Molecular identification of the insect adipokinetic hormone*
1006 *receptors.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3446-51.
1007 46. Kawamura, K., et al., *A new family of growth factors produced by the fat body and*
1008 *active on Drosophila imaginal disc cells.* Development, 1999. **126**(2): p. 211-9.
1009 47. Lam, G.T., C. Jiang, and C.S. Thummel, *Coordination of larval and prepupal gene*
1010 *expression by the DHR3 orphan receptor during Drosophila metamorphosis.*
1011 Development, 1997. **124**(9): p. 1757-69.
1012 48. Campbell, K. and J. Casanova, *A role for E-cadherin in ensuring cohesive migration of*
1013 *a heterogeneous population of non-epithelial cells.* Nat Commun, 2015. **6**: p. 7998.
1014 49. Chen, H., et al., *E-cadherin mediates adhesion and suppresses cell motility via distinct*
1015 *mechanisms.* J Cell Sci, 1997. **110 (Pt 3)**: p. 345-56.
1016 50. Prokopenko, S.N., et al., *A putative exchange factor for Rho1 GTPase is required for*
1017 *initiation of cytokinesis in Drosophila.* Genes Dev, 1999. **13**(17): p. 2301-14.
1018 51. Nakamura, M., J.M. Verboon, and S.M. Parkhurst, *Prepatterning by RhoGEFs governs*
1019 *Rho GTPase spatiotemporal dynamics during wound repair.* J Cell Biol, 2017. **216**(12):
1020 p. 3959-3969.
1021 52. Lin, Y.H., et al., *AP-2-complex-mediated endocytosis of Drosophila Crumbs regulates*
1022 *polarity by antagonizing Stardust.* J Cell Sci, 2015. **128**(24): p. 4538-49.
1023 53. Sountoulidis, A., et al., *SCRINSHOT enables spatial mapping of cell states in tissue*
1024 *sections with single-cell resolution.* PLoS Biol, 2020. **18**(11): p. e3000675.
1025 54. Umetsu, D., *Cell mechanics and cell-cell recognition controls by Toll-like receptors in*
1026 *tissue morphogenesis and homeostasis.* Fly (Austin), 2022. **16**(1): p. 233-247.
1027 55. Corty, M.M., B.J. Matthews, and W.B. Grueber, *Molecules and mechanisms of*
1028 *dendrite development in Drosophila.* Development, 2009. **136**(7): p. 1049-61.
1029 56. Araujo, S.J. and J. Casanova, *Sequoia establishes tip-cell number in Drosophila*
1030 *trachea by regulating FGF levels.* J Cell Sci, 2011. **124**(Pt 14): p. 2335-40.
1031 57. Brenman, J.E., et al., *Sequoia, a tramtrack-related zinc finger protein, functions as a*
1032 *pan-neural regulator for dendrite and axon morphogenesis in Drosophila.* Dev Cell,
1033 2001. **1**(5): p. 667-77.
1034 58. Araujo, S.J., C. Cela, and M. Llimargas, *Tramtrack regulates different morphogenetic*
1035 *events during Drosophila tracheal development.* Development, 2007. **134**(20): p.
1036 3665-76.

59. Li, S., et al., *Photoreceptor cell differentiation requires regulated proteolysis of the transcriptional repressor Tramtrack.* Cell, 1997. **90**(3): p. 469-78.
60. Pitsouli, C. and N. Perrimon, *The homeobox transcription factor cut coordinates patterning and growth during Drosophila airway remodeling.* Sci Signal, 2013. **6**(263): p. ra12.
61. Grueber, W.B., L.Y. Jan, and Y.N. Jan, *Different levels of the homeodomain protein cut regulate distinct dendrite branching patterns of Drosophila multidendritic neurons.* Cell, 2003. **112**(6): p. 805-18.
62. Meers, M.P., D. Tenenbaum, and S. Henikoff, *Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling.* Epigenetics Chromatin, 2019. **12**(1): p. 42.

Figure legends

Figure 1. *ham* is expressed and required for the development of Drosophila male RS

(**A**). Cartoons of the developing male RS. The testis (TE) terminal epithelial cells localize at the opposite side of the germline stem cell niche. The genital disc (GD) contains mesenchymal cells (MC) that will develop into the SV and accessory gland (ag). (**B**). Illustration of the TE and SV fusion process from 24-36h after pupal formation (APF), the two structures grow toward each other and form a continuous tube. (**C**). The adult male RS. ejd, ejaculatory duct. (**D**). White-field images of adult male RS, showing defective morphology in the *ham* mutants. (**E**). Quantification of male fertility. Ham_sfGFP with the *ham* genomic region was used to rescue the sterility. Statistical significance was calculated using unpaired t-test (***, $p<0.001$; **, $p< 0.01$; *, $p<0.05$; n.s., non-significant). (**F**). Images of the adult RS stained with Ham_V5 in red. DNA is stained with DAPI in blue. (**G**). Images of the larval TE (top) and GD (bottom) stained with Ham_V5 in red, the reporter in green (btlGal4 driving UAS-GFP) and DAPI in blue. AEL, after egg laying. (**H**). Images of the pupal GD stained with Ham_V5 in red, the reporter in green (btlGal4 driving UAS-GFP) and DAPI in blue. APF, after pupal formation. (**I**) Images of the pupal with Ham_V5 in red and DAPI in blue. (**J**). Higher magnification image from I, indicating the joint site between the SV and TE.

Figure 2. Ham is required for differentiation of the TE terminal and SV epithelial cells

(**A**). Images of the male RS at the pupal stage, showing morphological defects of the TE and SV in the ham mutants. Epithelial cells are labeled with Coracle (Cor). (**B**). High magnification images of the connection site of the TE and SV, showing that the ham mutant RS failed to form a continuous tube. Ham-sfGFP restored the normal morphology. Epithelial cells are marked by E-Cad in red, Cor in yellow, and the epithelium on the SV side also marked by btlGal4 driving UAS-GFP. DNA is labeled by DAPI in blue. (**C**). Images and quantification of the E-Cad signal in red in the *wt* and *hamPRdel* mutant SV. (**D**). Images and quantification of the Crumbs signal in red in the wt and *hamPRdel/Df* mutant SV. (**E**). Images and quantification of the length of Cor signal in red in the wt and *hamPRdel* mutant SV. (**F**) Images of *ham* mutant mosaic clones in the TE terminal epithelial cells. Note, the mutant cells are GFP negative and within the area surrounded by the white dashed line. The white arrows indicate junctions between mutant cells, while the yellow arrows highlight the junctions between wild-type cells. (**G**). Images of pupal male GD showing that *hamRNAi*-mediated knockdown reproduces *ham* mutant phenotypes.

(**H**). Images of pupal male RS, stained with E-Cad in red, Ham in yellow, GFP in green and DAPI in blue. UAS-dcr2 enhanced knockdown efficiency and led to more severe morphological defects. (**I**). Images and quantification of E-Cad signal in red and Cor length in blue in control and ham RNAi SV. (**J**) Images and quantification of E-Cad signal in red in control and ham RNAi TE terminal epithelium. Statistical significance was calculated using unpaired t-test except for TE mosaic clones which used paired t-test. (****, $p<0.0001$; ***, $p<0.001$; **, $p< 0.01$; *, $p<0.05$; ns, non-significant).

Figure 3. Ham-regulated genes in the developing TE and GD

(**A**). Illustration of the sample types and antibodies used in the RNA-seq and CUT&TAG experiments. (**B**) Venn diagram of dysregulated genes between the two mutant conditions in the GD (left) and TE (right) samples. (**C**) Heatmaps of Ham CUT&TAG peak density and IgG control, ranked from high to low and centered at peak summits. (**D**) Venn diagrams of CUT&TAG peaks from the two types of samples. (**E-F**). Volcano plots of dysregulated genes in GD (E) and TE (F) in the *hamSK1/Df* mutant, and the ones with at least one Ham CUT&TAG peak are highlighted.

Figure 4. Ham actives enhancers of *wnt2*, *shg* and *shot*

(**A**). A bar plot of the Log2 fold-change value from RNA-seq for *wnt2*, *shot* and *shg*, showing downregulation of these genes in the *hamSK1/Df* mutant. (B). Schemes of the Ham_FL and other truncation proteins, and illustration of the luciferase reporter assay. (**C-D**). A genome browser snapshot of the *shg* and *wnt2* loci, showing the CUT&TAG tracks from the TE and GD (TG) mixed samples and the enhancer fragments in the luciferase assays. (**E-F**). Normalized luciferase activity to the empty reporter and then to the control vector that expresses GFP and RFP proteins. Statistical significance was calculated using unpaired t-test (****, $p<0.0001$; ***, $p<0.001$; **, $p< 0.01$; *, $p<0.05$; ns, non-significant). (**G**). Significantly enriched motifs from Ham binding peaks in up-regulated genes (left) and downregulated genes (right). The dashed rectangles highlight MAD and ZNF263 motifs. (**H**) Ham gene network analysis from STRING. The position of MCL Clusters were manually adjusted to place genes with similar function in close proximity. The functional categories were determined by gene function tools in STRING.

Figure 5. A candidate gene screen identifies key regulators in RS epithelial tissue fusion.

(**A**). Scheme for the RNAi screen. (**B**) Summary of positive hits. The extent of sterility is determined by the number of progenies from the F1 male crossing to the wild-type female (see Methods). 100% means no progeny. >70% means the number of progenies is less than 30% of the number of progenies produced by the control cross (the same Gal4 driver x UAS-GFPRNAi). 40-70% means the number of progenies is between 30-60% of the number of progenies produced by the control. (**C**-**G**). Images of the control and RNAi knockdown RS. The left panels show the fusion site stained with E-Cad in red, GFP in green, Ham in yellow and DAPI in blue, the middle panels are GD, and the right ones are TE.

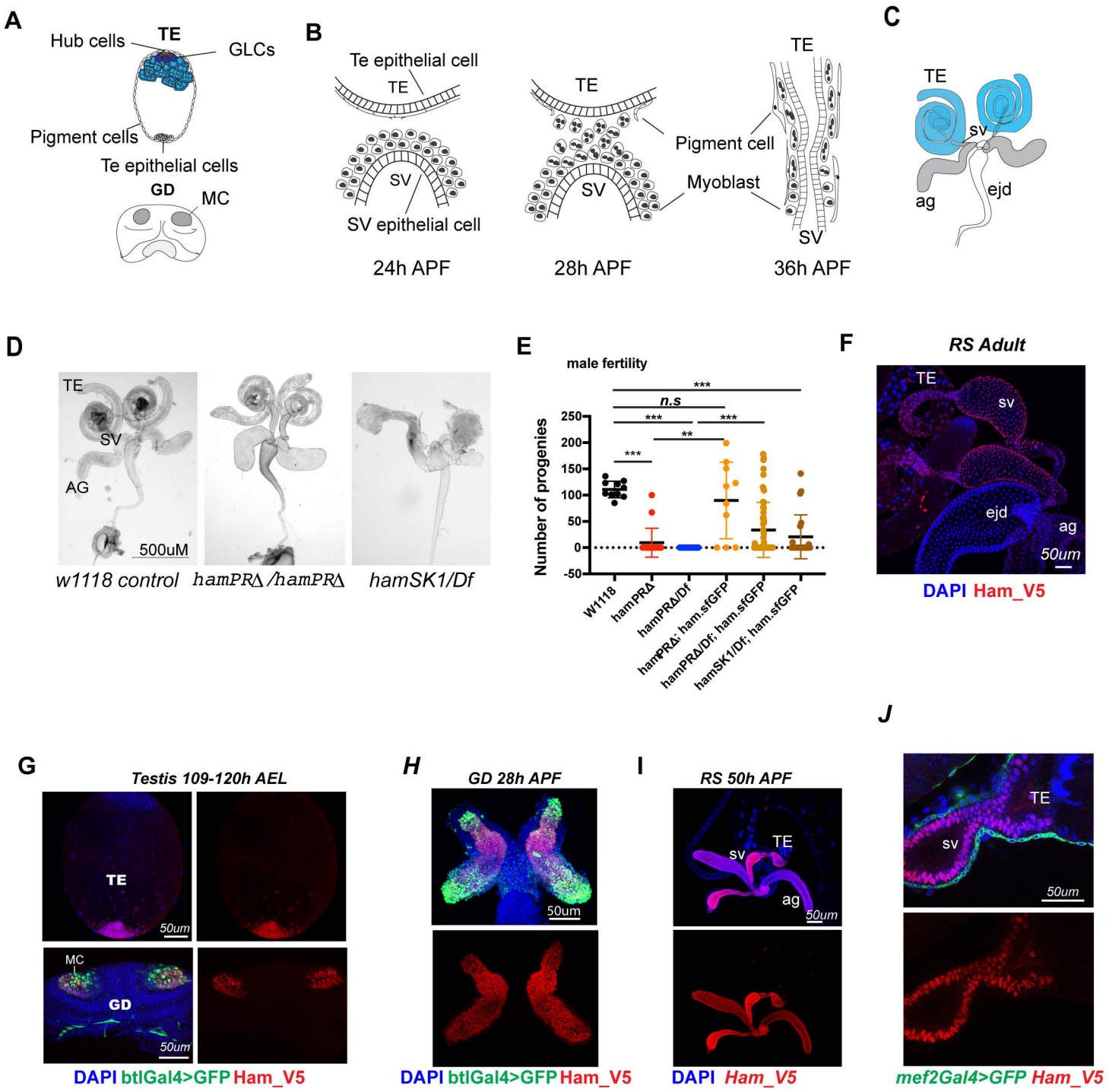
Figure 6. Spatial distribution of Ham-downstream target genes.

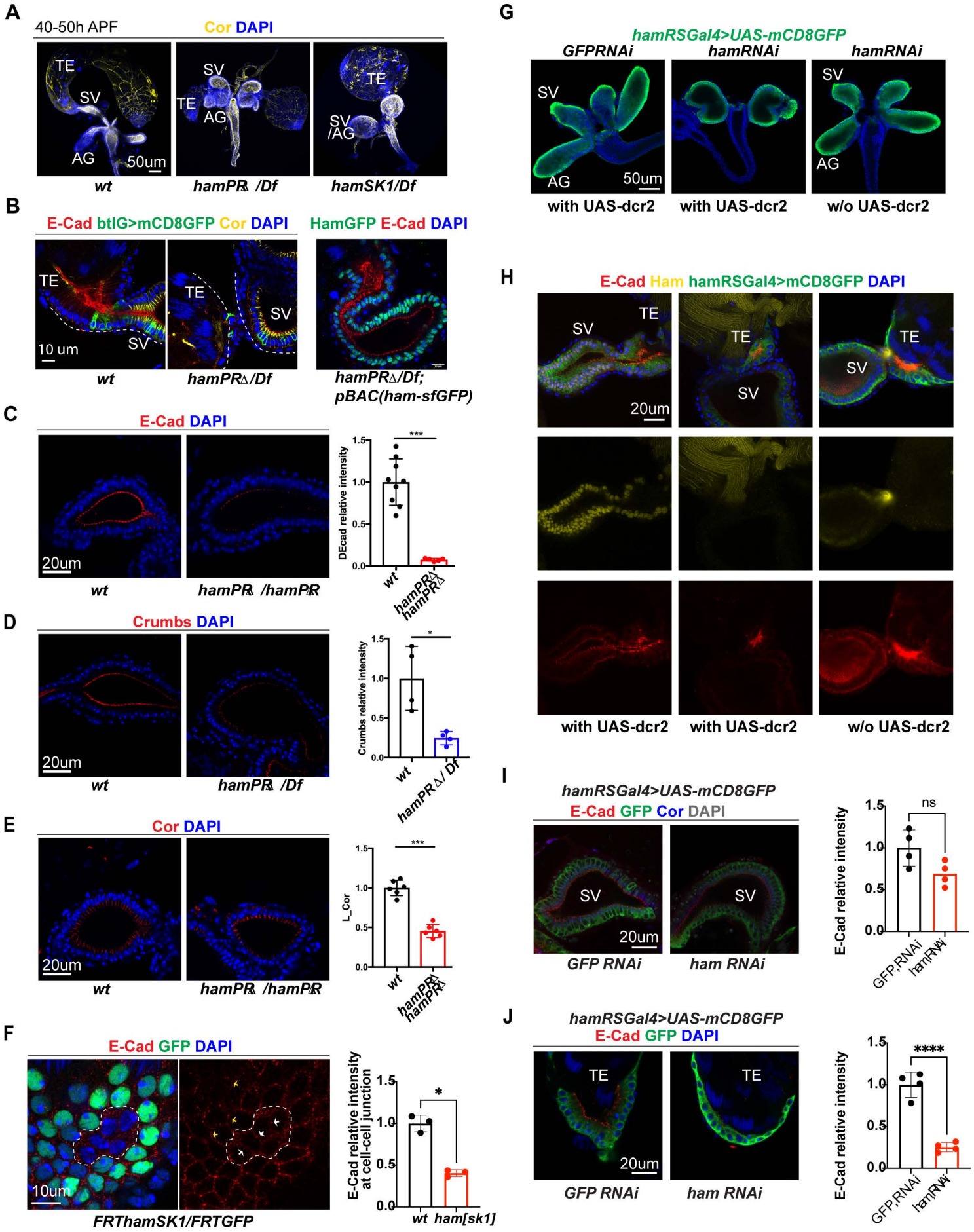
(**A**). Images from multiplexed in situ hybridization (the SCRINSHOT method), showing mRNA signal of the indicated genes in *wt* TE samples. DNA is marked by DAPI in grey. (**B**). Images from SCRINSHOT, showing mRNA signals of indicated genes in *wt/Df* and *hamSK1/Df* mutant GD. The last panel is an enlarged view from the sample with *Tl*, *ham* and *spz* probes. The yellow arrowheads indicate myoblast cells with Tl signal. (**C**). Quantification of signal intensity for genes shown in C in two areas, the SV and EjD. Statistical significance was calculated using unpaired t-test (****, $p<0.0001$; ***, $p<0.001$; **, $p< 0.01$; *, $p<0.05$; ns, non-significant). (**D**) Images from SCRINSHOT, showing mRNA signals of indicated genes in *GFP* and *ham RNAI* GD samples. (**E**). Quantification of signal intensity for genes shown in F in two areas, the SV and EjD. (**F**). A summary of genes expressed in the TE terminal and SV epithelium. Highlighted genes show temporal or spatial dynamic expression in the two tissues. Statistical significance was calculated using unpaired t-test (****, $p<0.0001$; ***, $p<0.001$; **, $p< 0.01$; *, $p<0.05$; ns, non-significant).

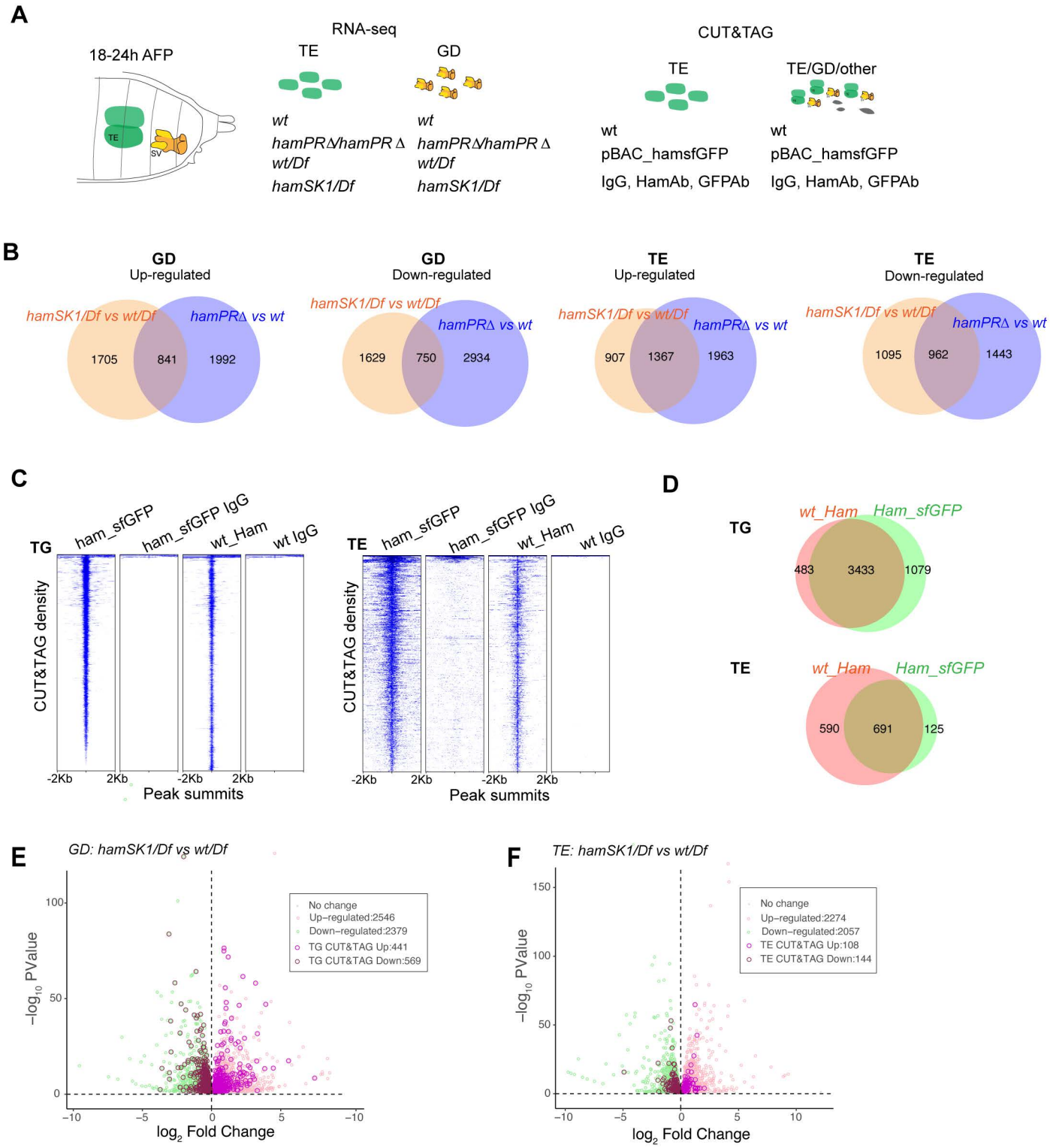
Figure 7. Wnt2 and its ligand genes are required for male fertility and RS development

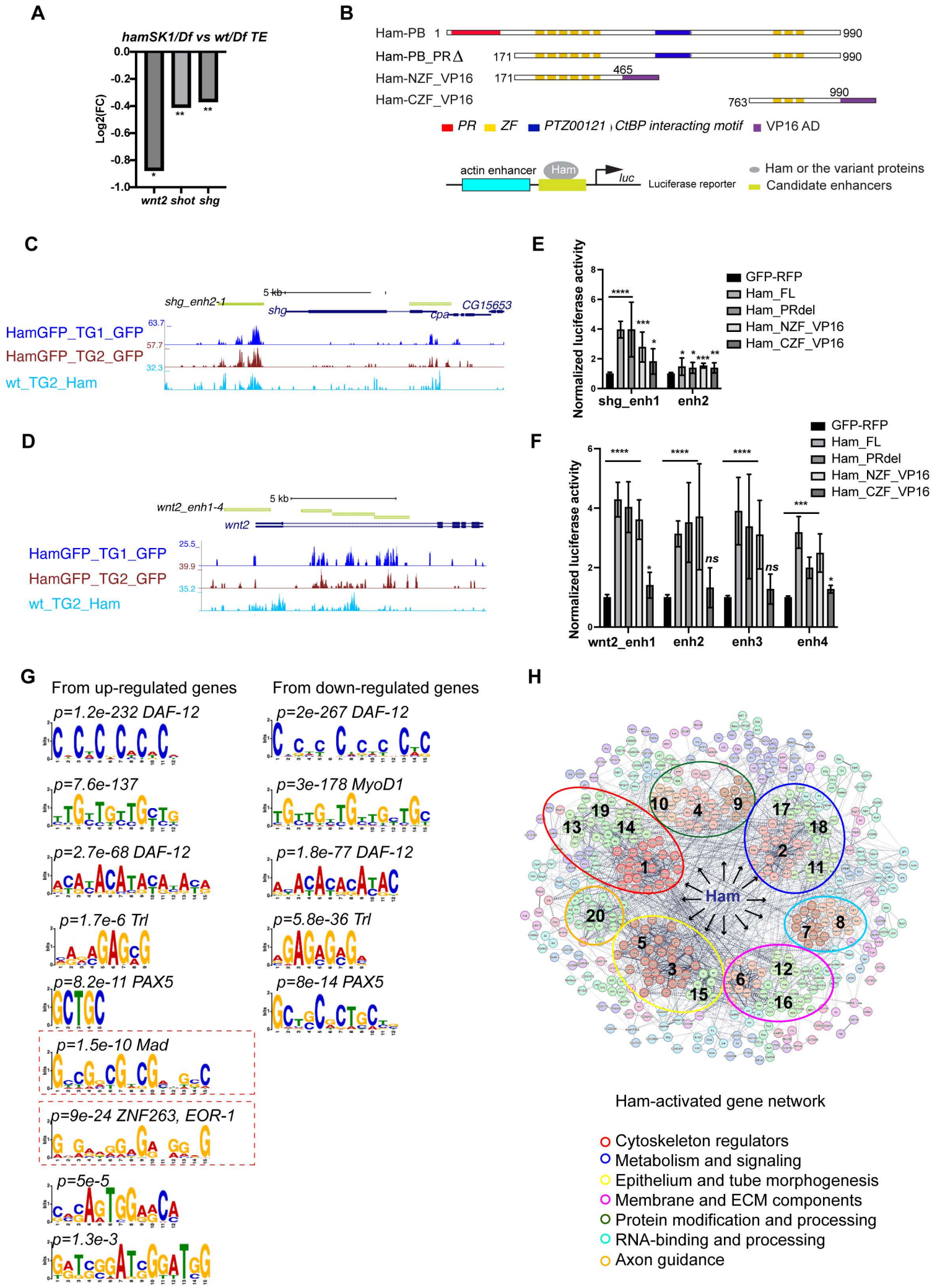
(**A**) Images of *wnt2o* mutant TE and GD, showing disrupted morphology in these tissues. (**B**). Images of the TE and SV fusion site from the control and *wnt2* overexpression animals. The epithelial cells of the TE and SV are labeled with hamRSGal4 driving UAS-mCD8GFP in green and E-Cad in red and Ham in yellow. The white arrows indicate mobilized GFP positive cells in the middle of the TE, and these cells have a lower level of E-Cad. (**C**). Images of the TE and SV fusion site from the control and *fz2* RNAi animals. The white arrows indicate

mobilized GFP positive cells in the middle of the TE, and these cells have a lower level of E-Cad. (**D**) Graphs of progeny number from male fertility tests for *fz* and *fz2* RNAi lines in comparison to GFP RNAi. Number of tested males in each genotype, 3. (**E**) Illustration of cell-cell interactions in the developing RS and speculated functional molecules. Wnt2 signaling may be involved in epithelial cell interaction between the two ends. Tl signaling and Tl may be involved in the interaction between epithelial cells and myoblasts.


A
Hub cells
TE
GLCs
Pigment cells
Te epithelial cells
GD
MC
B
Te epithelial cell
TE
SV
SV epithelial cell
24h APF
TE
SV
28h APF
TE
Pigment cell
Myoblast
SV
36h APF
C
TE
sv
ag
ejd
D
TE
SV
AG
500uM
w1118 control
hamPRΔ/hamPRΔ
hamSK1/Df
E
male fertility
Number of progenies
n.s
***
**
W1118
hamPRΔ
hamPRΔ/Df
hamPRΔ; ham.sfGFP
hamPRΔ/Df; ham.sfGFP
hamSK1/Df; ham.sfGFP
F
RS Adult
TE
sv
ejd
ag
50um
DAPI Ham_V5
G
Testis 109-120h AEL
TE
50um
MC
GD
50um
DAPI btlGal4>GFP Ham_V5
H
GD 28h APF
50um
DAPI btlGal4>GFP Ham_V5
I
RS 50h APF
TE
sv
ag
50um
DAPI Ham_V5
J
TE
sv
50um
mef2Gal4>GFP Ham_V5


A
40-50h APF
Cor DAPI
TE
SV
AG
50um
wt
hamPRΔ /Df
hamSK1/Df
B
E-Cad btlG>mCD8GFP Cor DAPI
HamGFP E-Cad DAPI
10 um
hamPRΔ/Df
hamPRΔ/Df;
pBAC(ham-sfGFP)
C
E-Cad DAPI
20um
hamPRΔ /hamPRΔ
DEcad relative intensity
***
D
Crumbs DAPI
Crumbs relative intensity
*
hamPRΔ/ Df
E
Cor DAPI
L_Cor
F
E-Cad GFP DAPI
10um
FRThamSK1/FRTGFP
E-Cad relative intensity at cell-cell junction
ham[sk1]
G
hamRSGal4>UAS-mCD8GFP
GFPRNAi
hamRNAi
with UAS-dcr2
w/o UAS-dcr2
H
E-Cad Ham hamRSGal4>mCD8GFP DAPI
I
hamRSGal4>UAS-mCD8GFP
E-Cad GFP Cor DAPI
GFP RNAi
ham RNAi
ns
E-Cad relative intensity
J
E-Cad GFP DAPI
****


A
18-24h AFP
TE
SV
RNA-seq
TE
GD
wt
hamPRΔ/hamPR Δ
wt/Df
hamSK1/Df
wt
hamPRΔ/hamPR Δ
wt/Df
hamSK1/Df
CUT&TAG
TE
TE/GD/other
wt
pBAC_hamsfGFP
IgG, HamAb, GFPAb
wt
pBAC_hamsfGFP
IgG, HamAb, GFPAb
B
GD
Up-regulated
hamSK1/Df vs wt/Df
hamPRΔ vs wt
1705
841
1992
GD
Down-regulated
hamSK1/Df vs wt/Df
hamPRΔ vs wt
1629
750
2934
TE
Up-regulated
hamSK1/Df vs wt/Df
hamPRΔ vs wt
907
1367
1963
TE
Down-regulated
hamSK1/Df vs wt/Df
hamPRΔ vs wt
1095
962
1443
C
TG
ham_sfGFP
ham_sfGFP IgG
wt_Ham
wt IgG
CUT&TAG density
-2Kb
2Kb
Peak summits
TE
ham_sfGFP
ham_sfGFP IgG
wt_Ham
wt IgG
CUT&TAG density
-2Kb
2Kb
Peak summits
D
TG
wt_Ham
Ham_sfGFP
483
3433
1079
TE
wt_Ham
Ham_sfGFP
590
691
125
E
GD: hamSK1/Df vs wt/Df
No change
Up−regulated:2546
Down−regulated:2379
TG CUT&TAG Up:441
TG CUT&TAG Down:569
−log10 PValue
log2 Fold Change
F
TE: hamSK1/Df vs wt/Df
No change
Up−regulated:2274
Down−regulated:2057
TE CUT&TAG Up:108
TE CUT&TAG Down:144
−log10 PValue
log2 Fold Change


A
hamSK1/Df vs wt/Df TE
Log2(FC)
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
*
**
**
wnt2 shot shg
B
Ham-PB 1 990
Ham-PB_PRΔ 171 990
Ham-NZF_VP16 171 465
Ham-CZF_VP16 763 990
PR ZF PTZ00121 CtBP interacting motif VP16 AD
actin enhancer Ham luc Luciferase reporter
Ham or the variant proteins
Candidate enhancers
C
5 kb
shg_enh2-1
shg
cpa
CG15653
HamGFP_TG1_GFP
HamGFP_TG2_GFP
wt_TG2_Ham
63.7
57.7
32.3
D
5 kb
wnt2_enh1-4
wnt2
HamGFP_TG1_GFP
HamGFP_TG2_GFP
wt_TG2_Ham
25.5
39.9
35.2
E
Normalized luciferase activity
GFP-RFP
Ham_FL
Ham_PRdel
Ham_NZF_VP16
Ham_CZF_VP16
****
***
*
*
*
***
**
shg_enh1
enh2
F
Normalized luciferase activity
GFP-RFP
Ham_FL
Ham_PRdel
Ham_NZF_VP16
Ham_CZF_VP16
****
****
****
***
*
ns
ns
*
wnt2_enh1
enh2
enh3
enh4
G
From up-regulated genes
From down-regulated genes
p=1.2e-232 DAF-12
p=2e-267 DAF-12
p=7.6e-137
p=3e-178 MyoD1
p=2.7e-68 DAF-12
p=1.8e-77 DAF-12
p=1.7e-6 Trl
p=5.8e-36 Trl
p=8.2e-11 PAX5
p=8e-14 PAX5
p=1.5e-10 Mad
p=9e-24 ZNF263, EOR-1
p=5e-5
p=1.3e-3
H
Ham
Ham-activated gene network
Cytoskeleton regulators
Metabolism and signaling
Epithelium and tube morphogenesis
Membrane and ECM components
Protein modification and processing
RNA-binding and processing
Axon guidance

A

hamRSGal4>UAS-dcr2; UAS-mCD8GFP

X

UAS-RNAi
(135 lines for 113 genes)

↓

Screen for F1 male sterility

↓

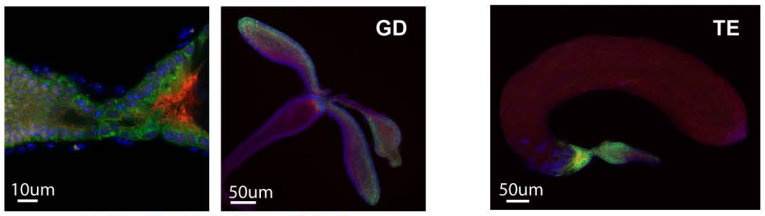
26 positive gene hits

B

Sterility	Genes
100%	*pbl, AP-2alpha, CG11406, shg*
>70%	*GstE11, CG7456, Gr64c, AkhR, Idgf4*
40%-70%	*CG17477, CG11069, CG17186, CG4872, AdSL, Tl, CG17283, decay,PH4alphaMP, Hr46, CG15117, gsb-n, shark, CG17018, cag, CG7484, CG31935*

C

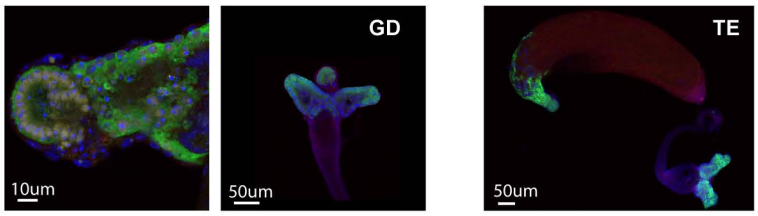
hamRSGal4>UASGFPRNAi;UASmCD8GFP



E-Cad GFP Ham DAPI

D

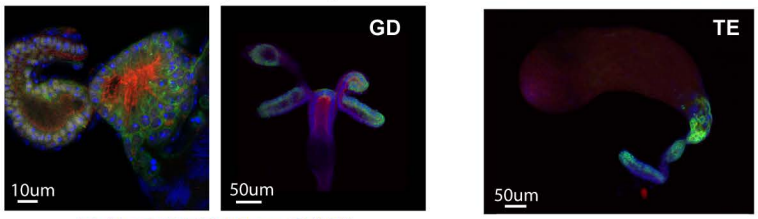
hamRSGal4>UASshgRNAi;UASmCD8GFP



E-Cad GFP Ham DAPI

E

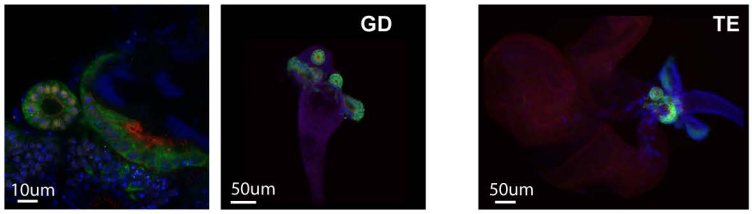
hamRSGal4>UASpblRNAi;UASmCD8GFP



E-Cad GFP Ham DAPI

F

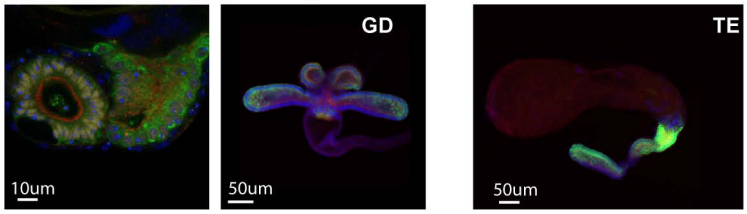
hamRSGal4>UASCG11406RNAi;UASmCD8GFP



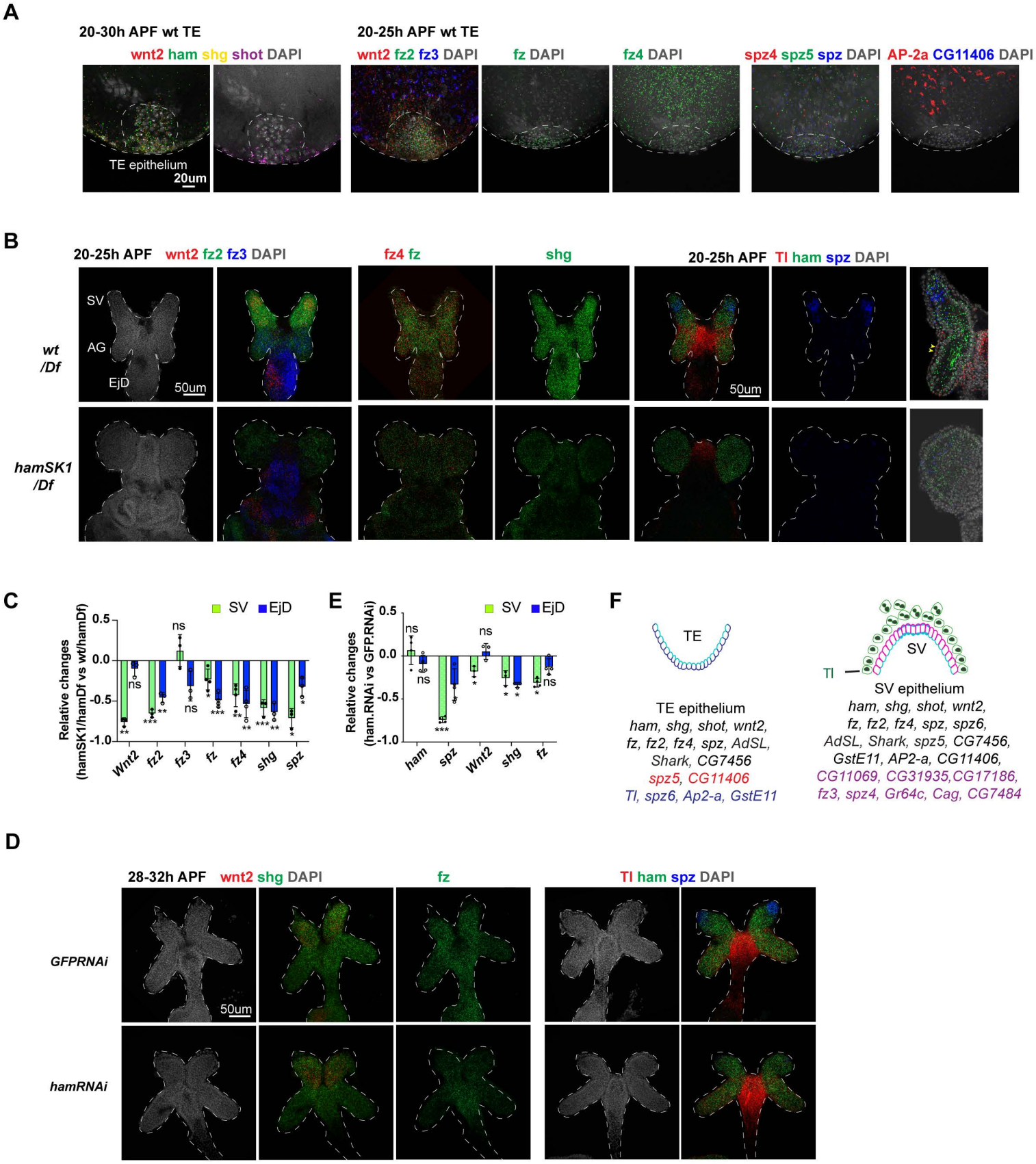
E-Cad GFP Ham DAPI

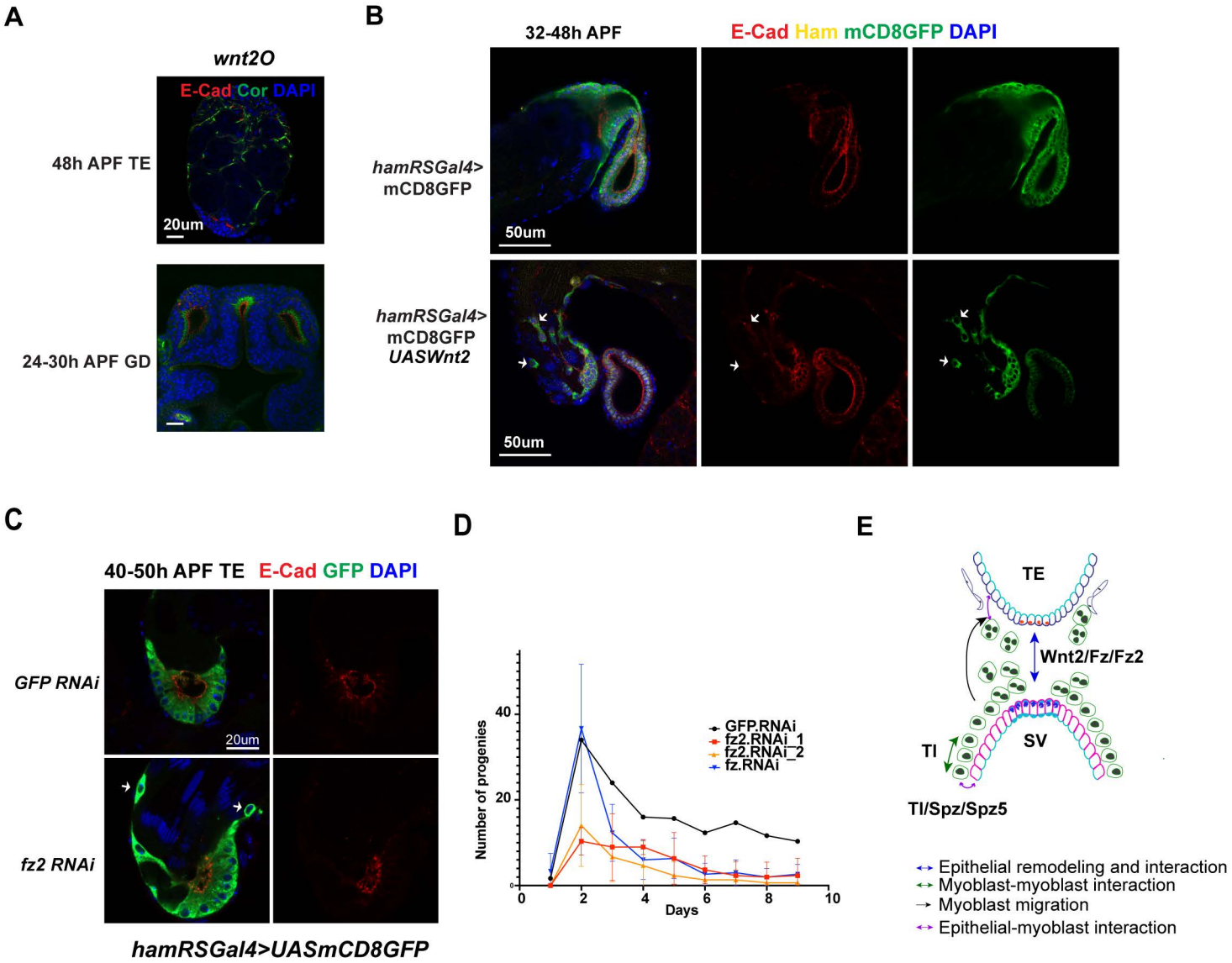
G

hamRSGal4>UASAP-2alphaRNAi;UASmCD8GFP



E-Cad GFP Ham DAPI


A
20-30h APF wt TE
wnt2 ham shg shot DAPI
TE epithelium
20um
20-25h APF wt TE
wnt2 fz2 fz3 DAPI
fz DAPI
fz4 DAPI
spz4 spz5 spz DAPI
AP-2a CG11406 DAPI
B
20-25h APF wnt2 fz2 fz3 DAPI
fz4 fz
shg
20-25h APF Tl ham spz DAPI
wt /Df
hamSK1 /Df
SV
AG
EjD
50um
C
SV EjD
Relative changes (hamSK1/hamDf vs wt/hamDf)
Wnt2 fz2 fz3 fz fz4 shg spz
E
Relative changes (ham.RNAi vs GFP.RNAi)
ham spz Wnt2 shg fz
F
TE
TE epithelium
ham, shg, shot, wnt2, fz, fz2, fz4, spz, AdSL, Shark, CG7456
spz5, CG11406
Tl, spz6, Ap2-a, GstE11
Tl
SV
SV epithelium
ham, shg, shot, wnt2, fz, fz2, fz4, spz, spz6, AdSL, Shark, spz5, CG7456, GstE11, AP2-a, CG11406, CG11069, CG31935,CG17186, fz3, spz4, Gr64c, Cag, CG7484
D
28-32h APF wnt2 shg DAPI
fz
Tl ham spz DAPI
GFPRNAi
hamRNAi
50um


A
wnt2O
E-Cad Cor DAPI
48h APF TE
20um
24-30h APF GD
B
32-48h APF
E-Cad Ham mCD8GFP DAPI
hamRSGal4>
mCD8GFP
50um
hamRSGal4>
mCD8GFP
UASWnt2
50um
C
40-50h APF TE E-Cad GFP DAPI
GFP RNAi
20um
fz2 RNAi
hamRSGal4>UASmCD8GFP
D
Number of progenies
Days
GFP.RNAi
fz2.RNAi_1
fz2.RNAi_2
fz.RNAi
E
TE
Wnt2/Fz/Fz2
SV
Tl
Tl/Spz/Spz5
Epithelial remodeling and interaction
Myoblast-myoblast interaction
Myoblast migration
Epithelial-myoblast interaction