Single cell transcriptomic landscapes of pattern formation, proliferation and growth in *Drosophila* wing imaginal discs

Mingxi Deng[1,2], Ying Wang[1,2], Lina Zhang[1,2], Yang Yang[1,2], Shengshuo Huang[1,2,3], Jiguang Wang[1,2,3], Hao Ge[4], Toyotaka Ishibashi[1], Yan Yan[1,2]

[1]Division of Life Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
[2]Center of Systems Biology and Human Health, School of Science and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
[3]Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China
[4]Beijing International Center for Mathematical Research and Biomedical Pioneering Innovation Center, Peking University

To whom correspondences should be addressed:
Yan Yan, Division of Life Science, Center of Systems Biology and Human Health, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. Tel: (852)23585929; Email: yany@ust.hk

Abstract

Organ formation relies on the orchestration of pattern formation, proliferation and growth during development. How these processes are integrated at individual cell level remains unclear. Studies using *Drosophila* wing imaginal discs as a model system have provided valuable insights into pattern formation, growth control and regeneration in the past decades. Here we provided single cell transcriptomic landscapes of pattern formation, proliferation and growth of wing imaginal discs. We found that patterning information is robustly maintained in the single cell transcriptomic data and can provide reference matrices to computationally map single cells into discrete spatial domains. Assignment of wing disc single cells to spatial sub-regions facilitates examination of patterning refinement processes. We also clustered single cells into different proliferation and growth states and evaluated the correlation between cell proliferation/growth states and spatial patterning. Furthermore, the single cell transcriptomic analysis allowed us to quantitatively examine the disturbance of differentiation, proliferation and growth in a well-established tumor model. We made a database to explore these datasets at: http://drosophilayanlab-virtual-wingdisc.ust.hk:3838/v2/

Introduction

The *Drosophila* wing imaginal discs have been an important developmental model to study pattern formation and organ size control (Beira and Paro, 2016). The wing imaginal discs are a pair of sac-like epithelial organs in *Drosophila* larvae which later become adult wings. A wing disc is composed of a squamous layer of epithelial cells called peripodial membrane and a columnar epithelial layer called disc proper, with myoblast cells, tracheal cells and a few neurons attached to the disc proper (Beira and Paro, 2016). Over the 96-hour larval period wing disc cells rapidly proliferate from 30-50 cells to 30,000-50,000 cells (Hariharan, 2015; Milan et al., 1996). The wing disc is patterned along the anterior-posterior (AP), dorsal-ventral (DV), and proximal-distal (PD) axes during which cells acquire positional information through functions of morphogens like Hedgehog (Hh)(Basler and Struhl, 1994; Strigini and Cohen, 1997; Tabata and Kornberg, 1994), Wingless (Wg, a Wnt family member)(Neumann and Cohen, 1997; Zecca et al., 1996) and Decapentaplegic (Dpp, a TGFβ family member)(Lecuit et al., 1996; Nellen et al., 1996) (Tabata and Takei, 2004). Wing disc size is tightly controlled through biochemical signals, such as the Dpp pathway (Burke and

Basler, 1996; Rogulja and Irvine, 2005; Wartlick et al., 2011) and the insulin/PTEN pathway (Bohni et al., 1999; Brogiolo et al., 2001; Goberdhan et al., 1999; Leevers et al., 1996; Montagne et al., 1999; Weinkove et al., 1999), as well as structural and mechanical signals which likely function through tuning the Hippo signaling activity (Hariharan, 2015; Irvine and Shraiman, 2017).

Although the pattern formation process and the growth control process are generally thought as two independent processes, several lines of evidence suggest that wing disc pattern formation and size control processes have underlying links. For example, Dpp and wg signaling activity is required for both pattern formation and wing disc growth (Capdevila and Guerrero, 1994; Giraldez and Cohen, 2003; Harmansa et al., 2015; Lecuit et al., 1996; Martin-Castellanos and Edgar, 2002; Teleman and Cohen, 2000; Zecca et al., 1995). Another line of evidence comes from studies of wing disc tumor models. Wing discs mutant for a group of neoplastic tumor suppressor gene (nTSG) are known to fail to differentiate and lose growth control at the same time (Bilder, 2004; Gateff, 1978). The representative neoplastic tumor suppressor gene (nTSG) mutants, including *lethal(2) giant larvae* (*l(2)gl*), *discs-large* (*dlg*) and *scribble* (*scrib*), encode highly conserved scaffold proteins that form a basolaterally-localized cell polarity complex in epithelial cells (Bilder et al., 2000; Bilder and Perrimon, 2000).The tissue architecture of the *nTSG* mutant wing discs, which are composed of unpolarized and undifferentiated cells, is severely disrupted and the mutant wing discs are presumably devoid of patterns and differentiation (Bilder et al., 2000; Bilder and Perrimon, 2000; Gateff, 1978). Interestingly, it was shown that the *nTSG* mutant clones grow preferably in the wing hinge region while the same mutant clones are eliminated in the pouch region (Tamori et al., 2016). A third line of evidence comes from the study of wing disc regeneration. It has been shown that different wing disc regions have dramatically different regenerative capacities (Martin et al., 2017; Verghese and Su, 2016). These lines of evidence suggest that the patterning process and cell differentiation states might affect cell competency for proliferation and growth.

Here we provided a systematic analysis of single cell transcriptomic datasets for wild type and the *scrib* mutant wing imaginal discs. Wing imaginal discs, as a classical model to study pattern formation, provide an ideal setting for us to map single cells to distinct spatial domains with high accuracy. We then used gene co-expression correlation analysis to identify two gene sets related to DNA replication and protein translation to cluster cells into

distinctive proliferation and growth states. We found that cells of different proliferation and growth states are distributed uniformly in wild type imaginal discs. We also examined patterning, proliferation and growth in the *scrib* mutant wing imaginal discs, a well-established wing imaginal disc tumor model. Surprisingly, although the homozygous *scrib* mutant wing imaginal discs are entirely composed of cells lacking apicobasal polarity and the overall organ architecture is severely disrupted, the mutant imaginal discs retain partial patterning with notum identity cells becoming more distinguishable over time. In the *scrib* mutant imaginal discs, the distribution of cell proliferation and growth states is severely disrupted. Together, these data provided landscapes of pattern formation, proliferation and growth at single-cell level in wing imaginal discs and demonstrated broad potential applications of single cell transcriptomic data beyond cell type identification.

Results

Overview of 10x single cell RNA data from wing imaginal discs

We generated single-cell RNA (scRNA) data of dissociated wing imaginal disc cells using the 10x genomics platform from five samples (Supplemental Table 1). While the total cell number captured varies with different samples, the median number of genes detected per cell is consistent at around 3000 genes per cell and the median UMI (unique molecular identifier) counts per cell is stable at around 20,000 per cell (Supplemental table 1). For each sample, we plotted the distribution of UMI number vs gene number detected per cell (Figure S1). In addition to removing low-quality cells with few genes detected, we also removed cells with over 4000 genes detected per cell to avoid potential doublets or even cell clusters (Figure S1).

We first visualized single cells from the 96 hour after egg laying (AEL) wild type imaginal discs using the UMAP (Uniform manifold approximation and projection) dimension reduction method (Becht et al., 2018). We noticed that the myoblasts and tracheal cells were well-clustered expressing unambiguous markers such as myoblast markers *twist (twi)* and *Mef2* (Gunage et al., 2014; Sudarsan et al., 2001) and tracheal cell marker *breathless (btl)* (Klambt et al., 1992; Roy et al., 2014) (Figure 1A-C). The tracheal cells captured in our datasets are low (n=19) and therefore we did not perform any further analysis. We examined the myoblast cluster (n=689) and found that the myoblasts split into two clusters with differential Notch signaling activity signatures (Figure S2A-C and Supplemental table 2),

consistent with previous findings of heterogeneous myoblast populations in wing imaginal discs (Gunage et al., 2014).

The majority of single cells (n=4498), which are from the wing disc epithelia, form a continuum without clear-cut clustering boundaries to distinguish wing disc sub-regions (Figure 1B). We inspected known landmark genes for wing disc sub-regions and noticed that cells expressing genes that mark the wing disc regional identity along the proximal-distal (PD) axis (e.g. *tsh* and *nub*) were well-clustered (Figure 1C), while cells expressing anterior-posterior (AP) identity genes (e.g. *ci* and *en*) were less well-clustered in the UMAP plot (Figure S3). Similarly, the *dpp*- and *wg*- expressing cells, which form signaling centers in wing discs, are also scattered as pockets of cells in the UMAP plot (Figure S3).

Assignment of wing disc epithelial cells to wing disc sub-regions

Single cells from the stage 6 *Drosophila* embryos have previously been successfully mapped to their spatial origins at single-cell resolution (Karaiskos et al., 2017). A 96-hour AEL wing imaginal disc contains 10,000-20,000 cells, in contrast with 3000 cells in a stage 6 embryo. Moreover, wing imaginal discs do not have digitalized *in situ* patterns of marker genes as a spatial reference map available like stage 6 embryos do (Fowlkes et al., 2008), posing more challenges for spatial mapping at single cell resolution. However, decades of work using wing discs as a model for pattern formation has established a number of marker genes to distinguish wing sub-regions (Bryant, 1975; Held, 2002). Using these marker genes as seeds, we performed gene expression correlation analysis to identify additional landmark genes with region-restrictive expression patterns (Figure 1D). For additional landmark genes identified through gene expression correlation analysis, we further confirmed their expression patterns in published *in situ* datasets (Butler et al., 2003; Ibrahim et al., 2013) and the Janelia FlyLight database (Jory et al., 2012). We then used the combinatorial expression of binarized 40 landmark genes as a spatial reference map (Supplemental table 3) to assign the wing disc epithelial cells (n=4498) with high confidence into three regions including the pouch/hinge region, the notum region and the peripodial membrane (PM) region (Figure 1D-F). Specifically, we manually binarized the expression data for each landmark gene and adopted the DistMap strategy previously developed (Karaiskos et al., 2017) to compute the most likely region for each cell (Figure 1E-F).

After single cells were assigned to specific regions, wing disc patterns become readily visible on the UMAP plots. For the pouch/hinge region (n=2906), it is polarized along the PD axis with marker genes expressed in partially overlapping domains including *defective proventriculus (dve)*(Kolzer et al., 2003; Nakagoshi et al., 2002), *nab* (Terriente Felix et al., 2007), *rotund (rn)*(St Pierre et al., 2002), *nubbin* (*nub*) (Ng et al., 1995), *zinc finger homeodomain 2 (zfh2)* (Terriente et al., 2008), *Sox box protein 15 (Sox15)(Cremazy et al., 2001)* and *teashirt (tsh)* (Wu and Cohen, 2002). In particular, the pouch marker genes including *dve*, *nab*, *rn* and *nub* express in concentric circular domains that extend partially to the hinge region at different lengths, while the hinge marker genes including *zfh2*, *Sox15* and *tsh* express only in part of the hinge region (Ayala-Camargo et al., 2013; Azpiazu and Morata, 2000; Terriente et al., 2008; Wu and Cohen, 2002). The expression patterns for these PD axis marker genes are well maintained on the UMAP plot (Figure 2A-B). The *nub+rn+dve-nab-* region is enriched for *wg+* inner ring cells (Figure 2C). The *Sox15+zfh2+* hinge region (n=1448) exhibits active JAK/STAT signaling signatures, consistent with prior studies (Figure S4) (Ayala-Camargo et al., 2013). In the pouch region, the cells from the Dpp and Wg signaling stripes, as well as distinct regions responding to the spatially graded Dpp and Wg signals respectively, are also clearly visible on the UMAP plot (Figure 2D-E).

For the notum region (n=1231), the *pnr+* cells, the *wg+* cells and the *Iro-C+*(*ara+mirr+caup+*) cells are well clustered, defining the medial and lateral domains of the notum region (Figure 3A-F) (Calleja et al., 2000; Diez del Corral et al., 1999; Letizia et al., 2007; Sato and Saigo, 2000). Similarly, *en+hh+* cells, *ci+* cells and *eyg+* cells are clustered along the AP axis (Figure 3G-K)(Aldaz et al., 2003). Note that the *en+hh+* domain is small in the notum region, consistent with previous reports (Figure 3G-K)(Aldaz et al., 2003; Zecca and Struhl, 2002).

For the PM region (n=303), the *Ubx+* cells are enriched in the *en+ or hh+* posterior domain and the *Antp+* cells are enriched in the *ci+* anterior domain, consistent with previous reports (Pallavi and Shashidhara, 2005; Tripura et al., 2011) (Figure S5). Note that a significant percentage of PM region cells are *en-hh-ci-* (n=112) possibly due to dropout events, which are an inherent challenge of single cell RNA sequencing studies now due to the low amounts of mRNA captured and the stochastic occurrence of failure to detect an expressed gene.

Together, these data suggest that the wing disc patterning information is robustly maintained in the single cell transcriptomic data. In particular, the proximal-distal patterning process is a dominant source for 96-hour AEL wing disc single cell clustering. Assignment of wing disc cells into spatial sub-regions and secondary clustering within each sub-region facilitates further analysis of patterning refinement processes.

Partial differentiation occurs in the *scrib* mutant imaginal discs despite disruption of organ architecture.

The *scrib* mutant imaginal discs grow into neoplastic tumors, which were thought to largely lack differentiation capacity (Bilder, 2004; Gateff, 1978). These mutant discs are composed of cells lacking apicobasal polarity and the overall organ architecture is largely disrupted in comparison with wild type imaginal discs (Bilder and Perrimon, 2000). It has been difficult to examine to what degree the patterning and differentiation processes occur in the *scrib* mutant imaginal discs.

We generated single cells from the 4-day, 5-day, 8-day, and 14-day AEL *scrib* mutant imaginal discs, because the *scrib* mutants from these time points were shown to have distinctive global transcriptomic signatures from bulk transcriptome analysis in another study. These datasets provided us with an opportunity to systematically examine the patterning and differentiation processes in the *scrib* mutants in comparison with wild type imaginal discs. Single cells from the *scrib* mutant wing discs and wild type wing discs segregate on the UMAP plot (Figure 4A). The *scrib* mutant imaginal discs uniquely express genes like *pvf1*, *Ilp8* and *upd1*, consistent with previous findings (Figure S6) (Bunker et al., 2015).

To examine the patterning and differentiation processes in the *scrib* mutant imaginal discs, we examined the expression of landmark genes previously used for assigning wild-type single cells to wing disc sub-regions. These genes can be grouped into four clusters based on gene expression correlation in wild type imaginal disc single cells (Figure 1D). In the *scrib* mutant imaginal discs, the expression correlation within each gene cluster is severely disrupted (Figure 4B). Interestingly, the clustering of landmark genes improves over time, indicative of continuous PD patterning process as the *scrib* mutant imaginal discs develop over time (Figure 4B). Similar to wild type imaginal discs, we used patterning landmark

genes as a spatial reference map to assign the *scrib* mutant cells to wing disc sub-regions (Figure 4C). In comparison with the wild type wing disc cells (Figure 1F), the assignment of the *scrib* mutant cells is more uncertain as the MCC scores of three different regions become less differentiating for the *scrib* mutant cells, which can be quantified by an overall increase in entropy (Figure 4D). Note that as the *scrib* mutant imaginal discs develop, the entropy decreases, supporting that the PD patterns become more distinguishable over time. For example, the UMAP projection of the *scrib* mutant cells no longer reveals distinct expression domains along the PD axis marked by *Iro-C* genes (*ara*, *caup*, *mirr)* and *wg*, but the co-expression pattern of *ara, caup and mirr* become evident in the 14-day *scrib* mutant tumors (Figure S7).

For each of the patterning landmark genes used to define wing sub-regions and expressed in the *scrib* mutant imaginal discs at all stages, we also examined the percentage of cells expressing that specific marker at all stages (Figure 4E). We ranked all the expressed genes (7986 in total) based on the change of the percentage of cells expressing that specific gene between 4-day and 14-day AEL *scrib* mutant imaginal discs and found that the notum identity marker genes are among the top 5% genes that show significant percentage increase (Figure 4E). We stained the *scrib* mutant imaginal discs for a notum identity marker Tailup (Tup) (Del Signore et al., 2012) and found that Tup+ cells indeed increase significantly from the 4-day to 14-day *scrib* mutant imaginal discs (Figure 4F). Together, these data demonstrate that the notum identity cells increase in the *scrib* mutant imaginal discs over time.

Analysis of proliferation and growth states of wing disc single cells.

The cell proliferation and growth states can be potentially deduced from the single cell transcriptomic data (Buettner et al., 2015; Liu et al., 2017; McDavid et al., 2016). In *Drosophila* imaginal discs, Yki activity is a key proliferation regulator (Huang et al., 2005). Several well-established Yki targets, including *diap1*, *ex,* and *CycE*, show weak co-expression correlation among single cells (data not shown), consistent with published results and likely due to cross talks between Yki and other growth regulatory pathways (Zhang et al., 2017). Moreover, a few marker genes are not sufficient to determine cell proliferation and growth states due to the intrinsic stochasticity in single cell transcriptomic data (Liu et al., 2017). To find gene expression patterns as proxies to indicate cell proliferation and growth

states, we performed gene expression correlation analysis for all expressed genes in wing disc cells and identified two gene sets related to proliferation and growth which exhibit high co-expression correlation within the gene set itself (Figure 5A). The first gene set is composed of 47 genes which includes *cdc25/string*, *CycE* and other genes functionally important for DNA replication, which are likely co-regulated by E2F1and Polycomb group genes (Ji et al., 2012; Neufeld et al., 1998; Thacker et al., 2003) (Supplemental table 4). The second gene set is composed of 91 genes related to protein translation including ribosomal genes, translation initiation and elongation factors (Supplemental table 4). We then clustered wing disc cells into four proliferation/growth states based on a combinatorial expression pattern of these two gene sets (Figure 5B) and further found that cells of different proliferation and growth states distribute to all wing disc sub-regions without obvious bias (Figure 5C and 5F). These results are consistent with previous studies showing that proliferation and growth are largely uniform in wing imaginal discs (Rogulja and Irvine, 2005; Schwank et al., 2011).

We adopted the same strategy to analyze single cells from the *scrib* mutant imaginal discs. The co-expression correlation within the protein translation and DNA replication gene set is still observed among the *scrib* mutant cells (Figure 5D). A moderate anti-correlation relationship between the DNA replication gene set and the protein translation gene set observed in the wild type cells (Figure 5A) is no longer detectable in the *scrib* mutant cells (Figure 5B). We clustered the *scrib* mutant cells into four classes corresponding to those in the wild type cells based on k-nearest neighbor algorithm (Figure 5B and 5E) and the distribution of these four classes of cells changed significantly in the *scrib* mutant cells (Figure 5E). These results are consistent with previous studies demonstrating coordination of cell division and growth in wing imaginal discs (Neufeld et al., 1998) and a disruption of proliferation and growth coordination in cancerous cells. We then examined the distribution of the four proliferation/growth states for cell populations marked by specific patterning genes and did not detect any significant deviation from overall distribution (Figure 5F). These results suggest that proliferation and growth states are not biased in specific cell populations marked by single patterning genes in both wildtype and the *scrib* mutant wing imaginal discs.

Discussion

Here we have demonstrated single cell transcriptomes from *Drosophila* wing imaginal disc cells as potential resources for systematically analyzing processes such as pattern formation, proliferation and growth. We have made an accompanying database for exploring these datasets (http://drosophilayanlab-virtual-wingdisc.ust.hk:3838/v2/)

These datasets provided comprehensive transcriptomic signatures for different cell populations in the wing imaginal discs and enable the identification of additional marker genes for each cell population. For example, the myoblasts are clustered into two populations with distinctive Notch signaling signatures, consistent with previous findings (Gunage et al., 2014). Additional markers distinguishing these two myoblast sub-populations can provide candidate markers and genes for future study. We have provided two functions to explore gene expression patterns in the accompanying database. First, for any given gene detected in our datasets, its normalized expression pattern can be examined on UMAP plots of single cells from the wild-type wing imaginal discs, wing imaginal disc sub-regions and the *scrib* mutant imaginal discs of different stages. Second, we have built a virtual imaginal disc as a rough reference to explore gene expression pattern. For this we have divided the imaginal discs into 56 regions based on commonly used regional markers including *dpp, wg, en, ci, ap, salm, bi, brk, Dll, Sox15, Zfh2, pnr, ara* and *eyg*. Note that likely due to the coverage limitation of single cell RNA sequencing on 10x platform, even though wing discs in our datasets perform relatively well with a median of around 3000 genes detected per cell, we cannot find genes highly correlated with certain markers like *ap* and *bi* to correct for dropout events and therefore the virtual imaginal disc gene expression pattern is less accurate than the UMAP plot function.

We found that gene co-expression analysis in *single* cells is effective at identifying functionally-related gene sets or modules important for patterning. For any given gene detected in our datasets, its correlation coefficients with other genes in wild type imaginal discs can be calculated and shown in a table ranked by either descending or ascending orders in the accompanying database. In our experience, a correlation coefficient above 0.4 is meaningful in our datasets. For example, the correlation coefficients of dpp/ptc and

en/hh/ci/inv are around 0.4. We detected correlation coefficients of around 0.2 for *dpp* with a set of glycolytic genes including *Eno, Pglym78, Tpi, Gapdh2, Ald, Gapdh1, Pgk,* and *Pfk.* Moreover, when we performed differential gene expression analysis for *dpp+* cells vs *dpp-* cells in the pouch region, the glycolytic genes are identified as significantly differentially-expressed genes. However, unlike *dpp* and *ptc*, the expression of glycolytic genes is widely distributed and we could not detect quantitative differences of glycolytic genes between *dpp+* and *dpp-* regions through fluorescent *in situ* hybridization (data not shown). Although it is possible that the *in situ* hybridization method is not sensitive enough to detect subtle differences, correlation coefficients of around 0.2 between two genes should be approached with caution even when they are ranked among the highest for all genes.

We have identified two gene sets through gene co-expression analysis as proxies to define proliferation and growth states in single cells. The first gene set includes *cdc25/string*, *CycE*, *PCNA*, and other genes functionally important for DNA replication. The second gene set includes ribosomal genes, translation initiation and elongation factors. The DNA replication gene and protein translation gene sets are the only two proliferation/growth-related gene groups that we can identify as co-regulated and suitable for cell clustering and classification. However, the exact cell state that each class represents is unclear at the moment. The DNA replication gene set is likely co-regulated by E2F1and Polycomb group genes (Ji et al., 2012; Neufeld et al., 1998; Thacker et al., 2003). The high co-expression correlation of the 91 translation-related genes suggests an underlying co-regulatory mechanism that remains to be uncovered. Interestingly, in wing imaginal discs, the pS6 staining pattern and TORC activation is patchy and stochastic (Romero-Pozuelo et al., 2017). Likewise, the staining patterns of common proliferation markers like PH3+ and BrdU+ are known to be sporadic and uniformly distributed. It seems that a balanced distribution of cells in different proliferation and growth states exist in wild type imaginal discs. How this balance is achieved in wild type imaginal discs and becomes biased in the *scrib* mutant tumors remains to be explored in the future.

Materials and methods

Fly stocks

The fly strains used in this study were: FRT82B (BL2035) and *scrib*^{1}FRT82B/TM6B (a kind gift from Doe lab).

Immunohistochemistry

Imaginal discs were dissected, fixed and stained according to standard protocols. The primary antibody used was mouse anti-tup (1:1000, 40.3A4, DSHB). The secondary antibody conjugated with Alexa Fluor-568 (ThermoFisher) was used at 1:500. Phalloidin conjugated with Alexa Fluor-647 (1:1000, ThermoFisher) and Hoechst (1:10000, ThermoFisher) were used to stain F-actin and DNA, respectively. All images were acquired on a Leica TCS SP8 confocal microscope.

Wing imaginal discs staging, clone induction, dissection, single cell dissociation and sequencing

Fly embryos were collected on apple juice plates with yeast pastes within 3 hours before transferring into fly food vials. Embryo numbers are controlled around 50 to avoid crowding. Staged wild type and *scrib*^{1} wing imaginal discs were dissected and transferred to DPBS (ThermoFisher). Wild type larvae were dissected at 96-hour AEL. The *scrib*^{1} mutant larvae were dissected at 96-, 120-,192-,336-hour respectively. The wing imaginal discs were dissociated in 0.25% Trypsin-EDTA solution at 37 ℃ for 10 min. Cells were then washed in DPBS and passed through 35μm filter before library preparation. Construction of 10x single cell libraries and sequencing on Illumina Hiseq platform were performed by Novogene. The NCBI GEO accession numbers are GSE133204 for 96-hour AEL wild type imaginal discs and GSE130566 for the *scrib* mutant imaginal discs of different stages.

10x Genomics raw-data processing, quality filtering and standard analysis in the Seurat pipeline

The sequencing raw-data files were processed by the Cell Ranger pipeline (v 2.2.0) with default settings. Read alignment and UMI counts were based on a BDGP6 genome reference fasta file and annotated by a BDGP6.91 gtf file developed by Ensembl. Cell-UMI count tables were loaded into *Seurat* projects. Cells expressing 1000~4000 genes and genes detected in more than 10 cells were used as filtering gates to select cells for downstream analysis. The expression data normalization is performed using the default LogNormalize method in Seurat (v2.3.4). Variable genes were selected by the Seurat FindVariableGenes function and used for subsequent analysis. PCA analysis, dimension reduction, cell clustering, and cluster marker identification were performed with the standard Seurat pipeline (R script available at: https://github.com/drosophilayanlab1/disc_script/tree/master).

Wing disc sub-region spatial mapping

Well-established wing disc region-specific genes including *nub*, *rn*, *dve*, *nab*, *Sox15*, *zfh2*, *pnr*, *eyg*, *ara*, *caup*, *Ubx* and *odd* were used as seed genes to identify other landmark genes which show high correlation coefficients with seed genes. The expression levels of landmark genes were binarized into on/off states with the binarization threshold for each gene specified in Supplemental Table 3. The thresholds for a few landmark genes are manually adjusted using expression regions of seed genes as references. For example, the gene *fz* is expressed in all the disc proper cells and shows a high correlation coefficient with Sox15 and can serve as a good landmark gene using a high-cutoff threshold. A cell-bin score for each cell to three positional bins is calculated based on the Mattews correlation coefficient (MCC), a strategy adopted from the DistMap algorithm (Karaiskos et al., 2017). The assignment of an individual cell is determined by the highest score among the three scores. A small number of cells with identical cell-bin scores were designated as uncertain cells and excluded from sub-clustering analysis.

Calculation of gene expression correlation coefficients

The Pearson correlation efficient between gene *i* and gene *j* was calculated from normalized gene expression levels (calculated using the default LogNormalize method in Seurat) of gene *i* and gene *j* in single cells using the cor function in R. Gene correlation coefficients larger than that *en* and *hh* are retained for visualization in a correlation matrix constructed using hclust function in R.

Calculation of Shannon entropy

The Shannon entropy of a cell is used to quantify the uncertainty regarding to its spatial assignment and is calculated as: For each single cell, denote S_i as the MCC score of this single cell to be assigned into the i-th subregion. Let $S = \sum_i S_i$, then the entropy of the MCC scores for this single cell is $\text{Ent} = \sum_i \left(\frac{S_i}{S}\right) log \left(\frac{S_i}{S}\right)$.

Wing disc cell proliferation and growth state analysis

The DNA replication gene set (47 genes, Supplemental table 3) and the protein translation gene set (91 genes, Supplemental table 3) were identified as discrete gene clusters in the correlation matrix. These 138 genes are used as variable genes to cluster wild type single cells into four classes of distinctive proliferation/growth states in Figure 5B. The *scrib* mutant cells were allocated to the four classes defined in wild type cells through the KNN algorithm (k=11).

Author contributions:

Y.Y., Y.W. and L.Z. designed the experiments. Y.W. , L.Z. and Y.Yang. performed all the experiments. M.X.D. and Y.Y. analyzed all the data and prepared all the figures. S.S.H., J.G.W and H.G. provided bioinformatic analysis tools. Y.Y. wrote the manuscript with input from H.G..

Acknowledgements

We thank Dr. Chris Doe for fly stocks. We thank Dr. Ting Xie for providing the 10x single cell RNA-seq protocol. We thank Dr. Trudi Schupbach for helpful comments on the manuscript. This work was supported by grants to Yan Yan from the Research Grants Council of the Hong Kong Special Administrative Region (grants GRF16103314, 16103815, 16150016, AoE/M-09/12) and to Jiguang Wang from from a NSFC/RGC Grant No. N_HKUST601/17 and a CRF Grant No. C6002-17G.

Reference:

Aldaz, S., Morata, G., and Azpiazu, N. (2003). The Pax-homeobox gene eyegone is involved in the subdivision of the thorax of Drosophila. Development *130*, 4473-4482.
Ayala-Camargo, A., Anderson, A.M., Amoyel, M., Rodrigues, A.B., Flaherty, M.S., and Bach, E.A. (2013). JAK/STAT signaling is required for hinge growth and patterning in the Drosophila wing disc. Dev Biol *382*, 413-426.
Azpiazu, N., and Morata, G. (2000). Function and regulation of homothorax in the wing imaginal disc of Drosophila. Development *127*, 2685-2693.
Basler, K., and Struhl, G. (1994). Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. Nature *368*, 208-214.
Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.
Beira, J.V., and Paro, R. (2016). The legacy of Drosophila imaginal discs. Chromosoma *125*, 573-592.
Bilder, D. (2004). Epithelial polarity and proliferation control: links from the Drosophila neoplastic tumor suppressors. Genes Dev *18*, 1909-1925.
Bilder, D., Li, M., and Perrimon, N. (2000). Cooperative regulation of cell polarity and growth by Drosophila tumor suppressors. Science *289*, 113-116.
Bilder, D., and Perrimon, N. (2000). Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. Nature *403*, 676-680.

Bohni, R., Riesgo-Escovar, J., Oldham, S., Brogiolo, W., Stocker, H., Andruss, B.F., Beckingham, K., and Hafen, E. (1999). Autonomous control of cell and organ size by CHICO, a Drosophila homolog of vertebrate IRS1-4. Cell *97*, 865-875.
Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., and Hafen, E. (2001). An evolutionarily conserved function of the Drosophila insulin receptor and insulin-like peptides in growth control. Curr Biol *11*, 213-221.
Bryant, P.J. (1975). Pattern formation in the imaginal wing disc of Drosophila melanogaster: fate map, regeneration and duplication. J Exp Zool *193*, 49-77.
Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat Biotechnol *33*, 155-160.
Bunker, B.D., Nellimoottil, T.T., Boileau, R.M., Classen, A.K., and Bilder, D. (2015). The transcriptional response to tumorigenic polarity loss in Drosophila. Elife *4*.
Burke, R., and Basler, K. (1996). Dpp receptors are autonomously required for cell proliferation in the entire developing Drosophila wing. Development *122*, 2261-2269.
Butler, M.J., Jacobsen, T.L., Cain, D.M., Jarman, M.G., Hubank, M., Whittle, J.R., Phillips, R., and Simcox, A. (2003). Discovery of genes with highly restricted expression patterns in the Drosophila wing disc using DNA oligonucleotide microarrays. Development *130*, 659-670.
Calleja, M., Herranz, H., Estella, C., Casal, J., Lawrence, P., Simpson, P., and Morata, G. (2000). Generation of medial and lateral dorsal body domains by the pannier gene of Drosophila. Development *127*, 3971-3980.
Capdevila, J., and Guerrero, I. (1994). Targeted expression of the signaling molecule decapentaplegic induces pattern duplications and growth alterations in Drosophila wings. EMBO J *13*, 4459-4468.
Cremazy, F., Berta, P., and Girard, F. (2001). Genome-wide analysis of Sox genes in Drosophila melanogaster. Mech Dev *109*, 371-375.
Del Signore, S.J., Hayashi, T., and Hatini, V. (2012). odd-skipped genes and lines organize the notum anterior-posterior axis using autonomous and non-autonomous mechanisms. Mech Dev *129*, 147-161.
Diez del Corral, R., Aroca, P., JL, G.m.-S., Cavodeassi, F., and Modolell, J. (1999). The Iroquois homeodomain proteins are required to specify body wall identity in Drosophila. Genes Dev *13*, 1754-1761.
Fowlkes, C.C., Hendriks, C.L., Keranen, S.V., Weber, G.H., Rubel, O., Huang, M.Y., Chatoor, S., DePace, A.H., Simirenko, L., Henriquez, C.*, et al.* (2008). A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm. Cell *133*, 364-374.
Gateff, E. (1978). Malignant neoplasms of genetic origin in Drosophila melanogaster. Science *200*, 1448-1459.
Giraldez, A.J., and Cohen, S.M. (2003). Wingless and Notch signaling provide cell survival cues and control cell proliferation during wing development. Development *130*, 6533-6543.
Goberdhan, D.C., Paricio, N., Goodman, E.C., Mlodzik, M., and Wilson, C. (1999). Drosophila tumor suppressor PTEN controls cell size and number by antagonizing the Chico/PI3-kinase signaling pathway. Genes Dev *13*, 3244-3258.
Gunage, R.D., Reichert, H., and VijayRaghavan, K. (2014). Identification of a new stem cell population that generates Drosophila flight muscles. Elife *3*.
Hariharan, I.K. (2015). Organ Size Control: Lessons from Drosophila. Dev Cell *34*, 255-265.

Harmansa, S., Hamaratoglu, F., Affolter, M., and Caussinus, E. (2015). Dpp spreading is required for medial but not for lateral wing disc growth. Nature *527*, 317-322.
Held, L.I. (2002). Imaginal discs : the genetic and cellular logic of pattern formation (Cambridge ; New York: Cambridge University Press).
Huang, J., Wu, S., Barrera, J., Matthews, K., and Pan, D. (2005). The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila Homolog of YAP. Cell *122*, 421-434.
Ibrahim, D.M., Biehs, B., Kornberg, T.B., and Klebes, A. (2013). Microarray comparison of anterior and posterior Drosophila wing imaginal disc cells identifies novel wing genes. G3 (Bethesda) *3*, 1353-1362.
Irvine, K.D., and Shraiman, B.I. (2017). Mechanical control of growth: ideas, facts and challenges. Development *144*, 4238-4248.
Ji, J.Y., Miles, W.O., Korenjak, M., Zheng, Y., and Dyson, N.J. (2012). In vivo regulation of E2F1 by Polycomb group genes in Drosophila. G3 (Bethesda) *2*, 1651-1660.
Jory, A., Estella, C., Giorgianni, M.W., Slattery, M., Laverty, T.R., Rubin, G.M., and Mann, R.S. (2012). A survey of 6,300 genomic fragments for cis-regulatory activity in the imaginal discs of Drosophila melanogaster. Cell Rep *2*, 1014-1024.
Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., Kocks, C., Rajewsky, N., and Zinzen, R.P. (2017). The Drosophila embryo at single-cell transcriptome resolution. Science *358*, 194-199.
Klambt, C., Glazer, L., and Shilo, B.Z. (1992). breathless, a Drosophila FGF receptor homolog, is essential for migration of tracheal and specific midline glial cells. Genes Dev *6*, 1668-1678.
Kolzer, S., Fuss, B., Hoch, M., and Klein, T. (2003). Defective proventriculus is required for pattern formation along the proximodistal axis, cell proliferation and formation of veins in the Drosophila wing. Development *130*, 4135-4147.
Lecuit, T., Brook, W.J., Ng, M., Calleja, M., Sun, H., and Cohen, S.M. (1996). Two distinct mechanisms for long-range patterning by Decapentaplegic in the Drosophila wing. Nature *381*, 387-393.
Leevers, S.J., Weinkove, D., MacDougall, L.K., Hafen, E., and Waterfield, M.D. (1996). The Drosophila phosphoinositide 3-kinase Dp110 promotes cell growth. EMBO J *15*, 6584-6594.
Letizia, A., Barrio, R., and Campuzano, S. (2007). Antagonistic and cooperative actions of the EGFR and Dpp pathways on the iroquois genes regulate Drosophila mesothorax specification and patterning. Development *134*, 1337-1346.
Liu, Z., Lou, H., Xie, K., Wang, H., Chen, N., Aparicio, O.M., Zhang, M.Q., Jiang, R., and Chen, T. (2017). Reconstructing cell cycle pseudo time-series via single-cell transcriptome data. Nat Commun *8*, 22.
Martin, R., Pinal, N., and Morata, G. (2017). Distinct regenerative potential of trunk and appendages of Drosophila mediated by JNK signalling. Development *144*, 3946-3956.
Martin-Castellanos, C., and Edgar, B.A. (2002). A characterization of the effects of Dpp signaling on cell growth and proliferation in the Drosophila wing. Development *129*, 1003-1013.
McDavid, A., Finak, G., and Gottardo, R. (2016). The contribution of cell cycle to heterogeneity in single-cell RNA-seq data. Nat Biotechnol *34*, 591-593.
Milan, M., Campuzano, S., and Garcia-Bellido, A. (1996). Cell cycling and patterned cell proliferation in the wing primordium of Drosophila. Proc Natl Acad Sci U S A *93*, 640-645.
Montagne, J., Stewart, M.J., Stocker, H., Hafen, E., Kozma, S.C., and Thomas, G. (1999). Drosophila S6 kinase: a regulator of cell size. Science *285*, 2126-2129.

Nakagoshi, H., Shirai, T., Nabeshima, Y., and Matsuzaki, F. (2002). Refinement of wingless expression by a wingless- and notch-responsive homeodomain protein, defective proventriculus. Dev Biol *249*, 44-56.
Nellen, D., Burke, R., Struhl, G., and Basler, K. (1996). Direct and long-range action of a DPP morphogen gradient. Cell *85*, 357-368.
Neufeld, T.P., de la Cruz, A.F., Johnston, L.A., and Edgar, B.A. (1998). Coordination of growth and cell division in the Drosophila wing. Cell *93*, 1183-1193.
Neumann, C.J., and Cohen, S.M. (1997). Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing. Development *124*, 871-880.
Ng, M., Diaz-Benjumea, F.J., and Cohen, S.M. (1995). Nubbin encodes a POU-domain protein required for proximal-distal patterning in the Drosophila wing. Development *121*, 589-599.
Pallavi, S.K., and Shashidhara, L.S. (2005). Signaling interactions between squamous and columnar epithelia of the Drosophila wing disc. J Cell Sci *118*, 3363-3370.
Rogulja, D., and Irvine, K.D. (2005). Regulation of cell proliferation by a morphogen gradient. Cell *123*, 449-461.
Romero-Pozuelo, J., Demetriades, C., Schroeder, P., and Teleman, A.A. (2017). CycD/Cdk4 and Discontinuities in Dpp Signaling Activate TORC1 in the Drosophila Wing Disc. Dev Cell *42*, 376-387 e375.
Roy, S., Huang, H., Liu, S., and Kornberg, T.B. (2014). Cytoneme-mediated contact-dependent transport of the Drosophila decapentaplegic signaling protein. Science *343*, 1244624.
Sato, M., and Saigo, K. (2000). Involvement of pannier and u-shaped in regulation of decapentaplegic-dependent wingless expression in developing Drosophila notum. Mech Dev *93*, 127-138.
Schwank, G., Tauriello, G., Yagi, R., Kranz, E., Koumoutsakos, P., and Basler, K. (2011). Antagonistic growth regulation by Dpp and Fat drives uniform cell proliferation. Dev Cell *20*, 123-130.
St Pierre, S.E., Galindo, M.I., Couso, J.P., and Thor, S. (2002). Control of Drosophila imaginal disc development by rotund and roughened eye: differentially expressed transcripts of the same gene encoding functionally distinct zinc finger proteins. Development *129*, 1273-1281.
Strigini, M., and Cohen, S.M. (1997). A Hedgehog activity gradient contributes to AP axial patterning of the Drosophila wing. Development *124*, 4697-4705.
Sudarsan, V., Anant, S., Guptan, P., VijayRaghavan, K., and Skaer, H. (2001). Myoblast diversification and ectodermal signaling in Drosophila. Dev Cell *1*, 829-839.
Tabata, T., and Kornberg, T.B. (1994). Hedgehog is a signaling protein with a key role in patterning Drosophila imaginal discs. Cell *76*, 89-102.
Tabata, T., and Takei, Y. (2004). Morphogens, their identification and regulation. Development *131*, 703-712.
Tamori, Y., Suzuki, E., and Deng, W.M. (2016). Epithelial Tumors Originate in Tumor Hotspots, a Tissue-Intrinsic Microenvironment. PLoS Biol *14*, e1002537.
Teleman, A.A., and Cohen, S.M. (2000). Dpp gradient formation in the Drosophila wing imaginal disc. Cell *103*, 971-980.
Terriente Felix, J., Magarinos, M., and Diaz-Benjumea, F.J. (2007). Nab controls the activity of the zinc-finger transcription factors Squeeze and Rotund in Drosophila development. Development *134*, 1845-1852.
Terriente, J., Perea, D., Suzanne, M., and Diaz-Benjumea, F.J. (2008). The Drosophila gene zfh2 is required to establish proximal-distal domains in the wing disc. Dev Biol *320*, 102-112.

Thacker, S.A., Bonnette, P.C., and Duronio, R.J. (2003). The contribution of E2F-regulated transcription to Drosophila PCNA gene function. Curr Biol *13*, 53-58.
Tripura, C., Chandrika, N.P., Susmitha, V.N., Noselli, S., and Shashidhara, L.S. (2011). Regulation and activity of JNK signaling in the wing disc peripodial membrane during adult morphogenesis in Drosophila. Int J Dev Biol *55*, 583-590.
Verghese, S., and Su, T.T. (2016). Drosophila Wnt and STAT Define Apoptosis-Resistant Epithelial Cells for Tissue Regeneration after Irradiation. PLoS Biol *14*, e1002536.
Wartlick, O., Mumcu, P., Kicheva, A., Bittig, T., Seum, C., Julicher, F., and Gonzalez-Gaitan, M. (2011). Dynamics of Dpp signaling and proliferation control. Science *331*, 1154-1159.
Weinkove, D., Neufeld, T.P., Twardzik, T., Waterfield, M.D., and Leevers, S.J. (1999). Regulation of imaginal disc cell size, cell number and organ size by Drosophila class I(A) phosphoinositide 3-kinase and its adaptor. Curr Biol *9*, 1019-1029.
Wu, J., and Cohen, S.M. (2002). Repression of Teashirt marks the initiation of wing development. Development *129*, 2411-2418.
Zecca, M., Basler, K., and Struhl, G. (1995). Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development *121*, 2265-2278.
Zecca, M., Basler, K., and Struhl, G. (1996). Direct and long-range action of a wingless morphogen gradient. Cell *87*, 833-844.
Zecca, M., and Struhl, G. (2002). Subdivision of the Drosophila wing imaginal disc by EGFR-mediated signaling. Development *129*, 1357-1368.
Zhang, P., Pei, C., Wang, X., Xiang, J., Sun, B.F., Cheng, Y., Qi, X., Marchetti, M., Xu, J.W., Sun, Y.P.*, et al.* (2017). A Balance of Yki/Sd Activator and E2F1/Sd Repressor Complexes Controls Cell Survival and Affects Organ Size. Dev Cell *43*, 603-617 e605.

Figures

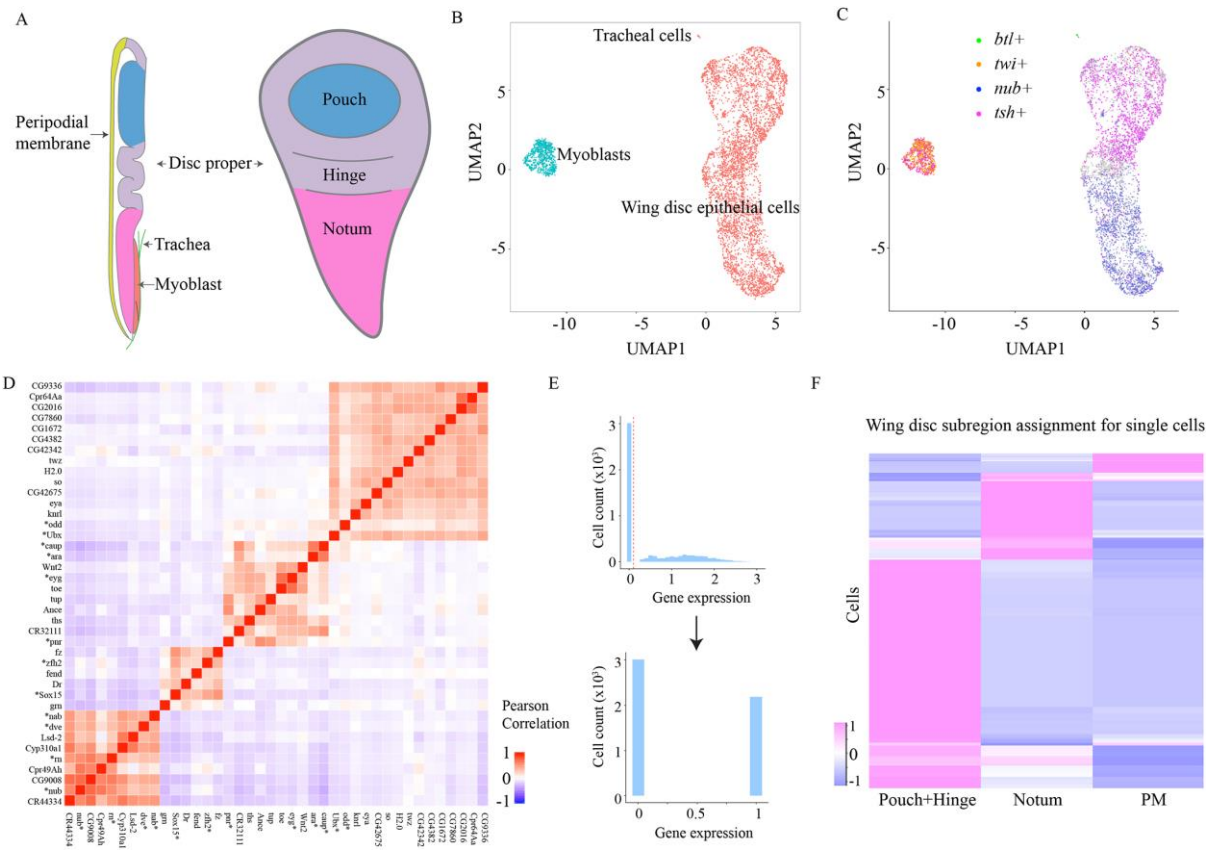

Figure 1 Single cells from wild type wing imaginal discs can be mapped to distinct sub-regions.

(A) A schematic view of a wing imaginal disc composed of two layers of epithelial cells (the squamous peripodial membrane and the columnar disc proper cells), myoblasts and tracheal cells. The disc proper is further divided into the notum, the hinge and the pouch region along the proximal-distal (PD) axis.

(B-C) UMAP projections of wing imaginal disc cells labeled by cell types in (B) and colored by marker genes (C). *Twist (twi)* is specifically expressed in myoblasts. *Breathless (btl)* labels tracheal cells. *Nubbin (nub)* and *teashirt (tsh)* are PD identity markers in the disc proper cells. n=5200 cells.

(D) Plot of Pearson correlation coefficients of the normalized expression levels of landmark genes used for spatial region mapping. The landmark genes form four clusters representing the pouch, hinge, notum and peripodial membrane sub-regions. n=4498 cells (Myoblasts and tracheal cells are removed from this analysis). * marks well-established wing disc region-specific genes including *nub*, *rn*, *dve*, *nab*, *Sox15*, *zfh2*, *pnr*, *eyg*, *ara*, *caup*, *Ubx* and *odd*, which were used as seed genes to identify other landmark genes.

(E-F) The expression profiles of all landmark genes are manually binarized for cell spatial mapping (E). Mattews correlation coefficients (MCC) for 500 randomly-selected wing disc proper cells (n=3 repeats) are plotted as heatmap in (F). The region assignment for each cell was based on Mattews correlation coefficients (MCC).

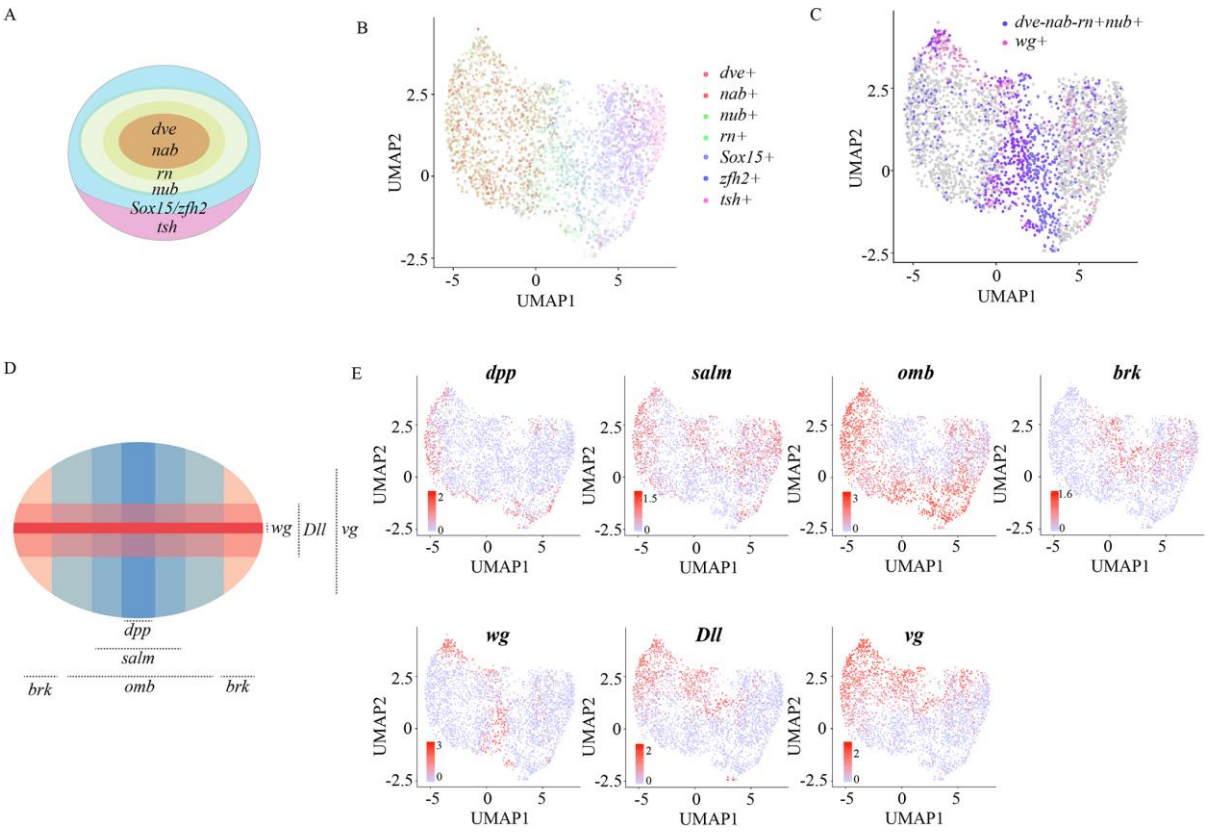

Figure 2 Wild type wing disc cell clustering from the pouch/hinge region recapitulates the patterning processes.

(A) A schematic view of the expression domains of PD identity marker genes in the pouch and hinge region. The schematic drawing is not true to the scale.

(B-C) UMAP projections of pouch and hinge cells colored by PD identity marker genes (B-C). n=2906 cells.

(D) A schematic view of the expression domains of genes responding to the Dpp and Wg gradients in the pouch and hinge region. The schematic drawing is not true to the scale.

(E) UMAP projections of pouch and hinge cells colored by the expression levels of *dpp*, *salm*, *omb*, *brk*, *wg*, *Dll* and *vg*. n=2906 cells.

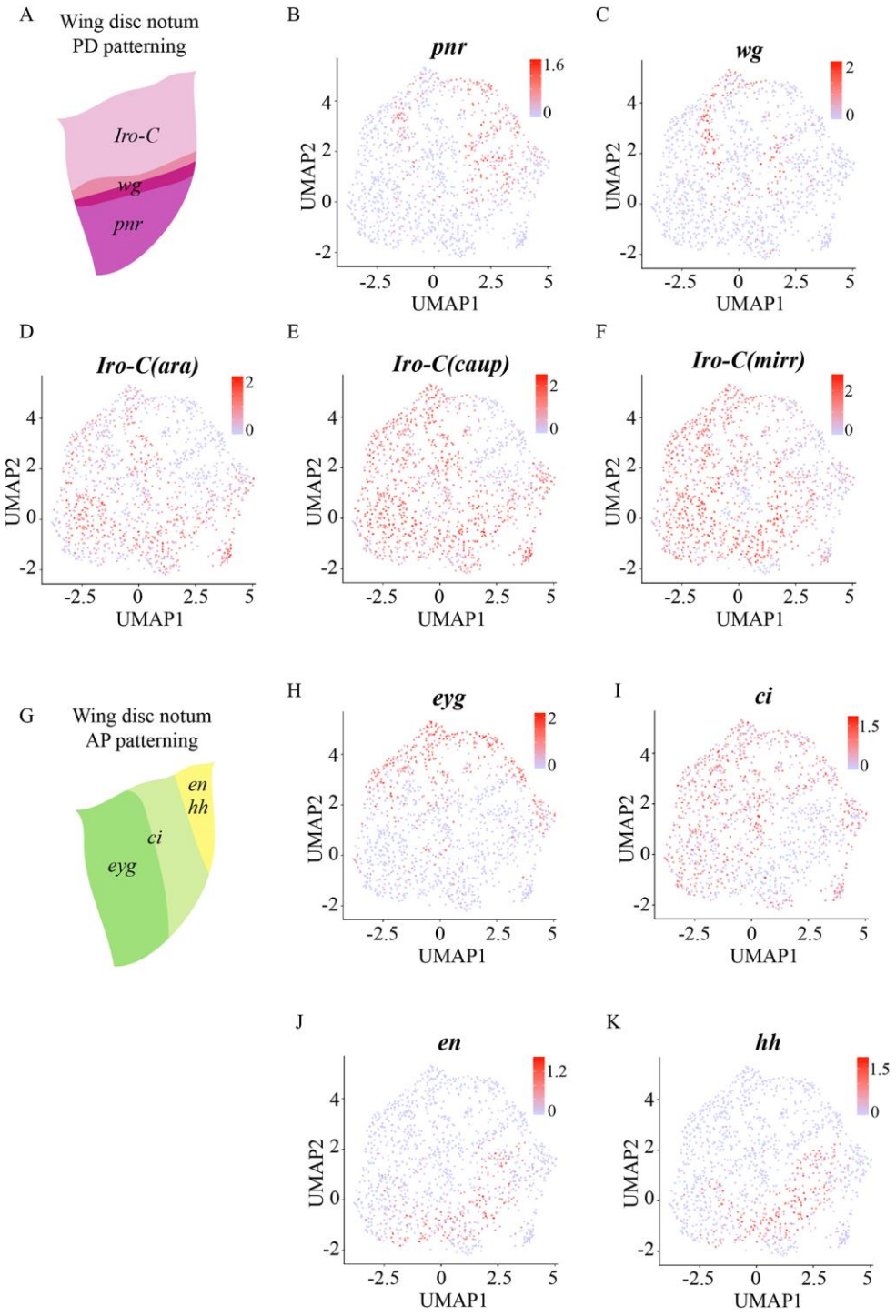

Figure 3 Wild type wing disc cell clustering from the notum region recapitulates the patterning processes.

(A) A schematic view of the expression domains of PD identity marker genes in the notum. The schematic drawing is not true to the scale.

(B-F) UMAP projections of notum cells colored by the expression levels of *pnr* (B), *wg* (C), *ara* (D), caup (E) and *mirr* (F). n=1231 cells.

(G) A schematic view of the expression domains of AP identity marker genes in the notum. The schematic drawing is not true to the scale.

(H-K) UMAP projections of notum cells colored by the expression levels of *eyg* (H), *ci* (I), *en* (J) and *hh* (K). n=1231 cells.

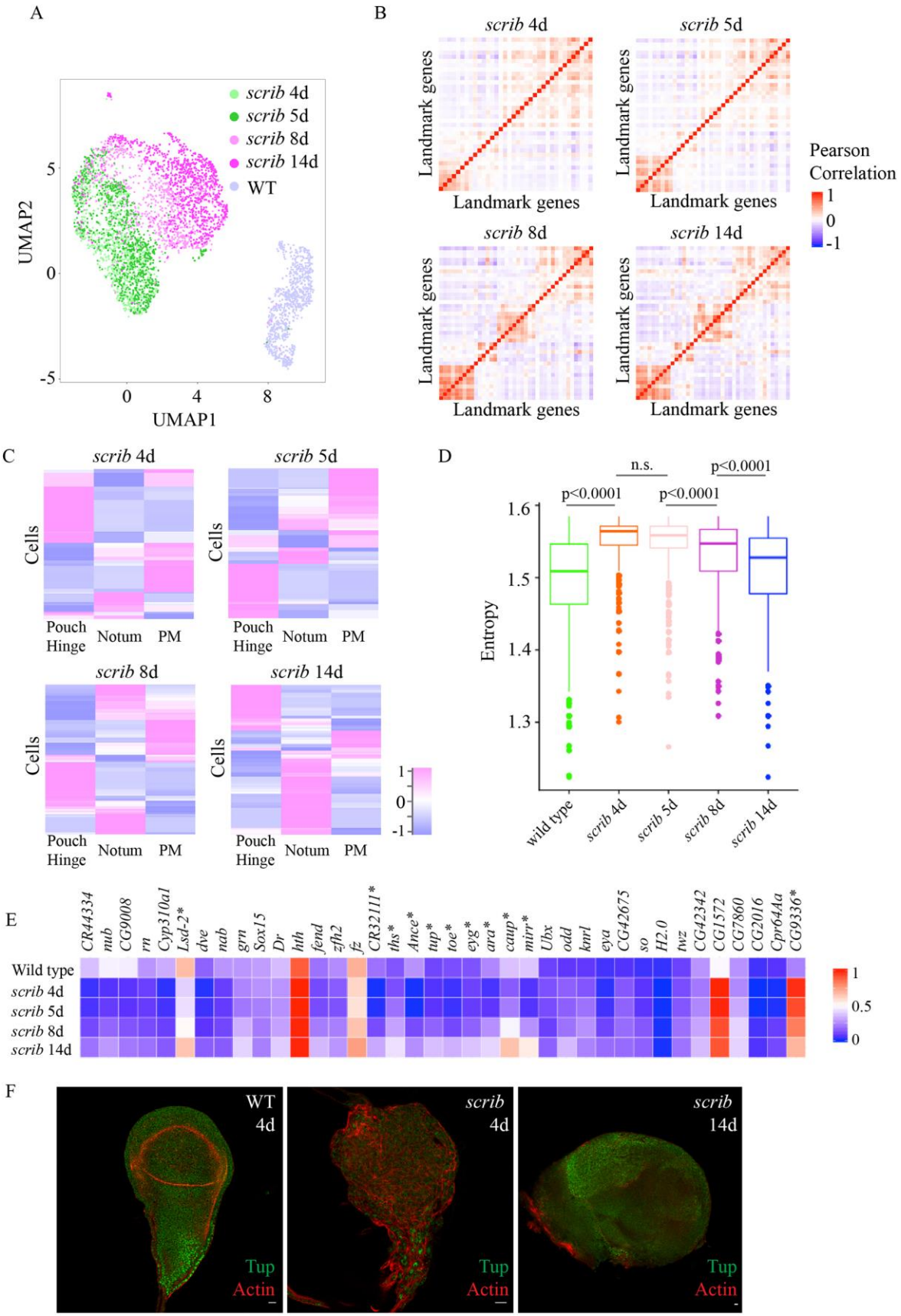

Figure 4 The *scrib* mutant imaginal discs exhibit partial differentiation.

(A) The UMAP projection of single cells from the wild type imaginal disc (4-day AEL) and the *scrib* mutant imaginal discs (4-day, 5-day, 8-day and 14-day AEL) colored by sample origin. n= 948 cells for each sample (random selected for each sample except the 14-day AEL *scrib* mutant imaginal discs). Myoblasts and tracheal cells are removed from this analysis.

(B) Plot of Pearson correlation coefficients of the normalized expression levels of landmark genes used for single cell spatial mapping in the *scrib* mutant imaginal discs (4-day, n=3440, 5-day, n=5614, 8-day, n=2622 and 14-day AEL, n=948). Myoblasts and tracheal cells are removed from this analysis. The landmark genes used are the same ones used for Figure 1D except that the expression of *Cpr49Ah*, *pnr*, *Wnt2* and *CG4382* was not detected in the 4-day scrib mutant imaginal discs and the expression of *pnr* and *CG4382* were not detected in the 5-day scrib mutant imaginal discs.

(C) Assignment of the *scrib* mutant single cells into wing disc sub-regions based on MCC scores (500 randomly-selected cells per sample, n=3 repeats).

(D) Boxplot of single cell entropies calculated from MCC scores for each sample (500 randomly-selected cells per sample, n=3 repeats). Statistical analysis was performed by Wilcoxon rank sum test.

(E) Heatmap plot of the percentage of cells expressing the specific landmark gene in each sample. We ranked 7846 expressed genes by the change of percentage of cells expressing the specific gene between 4-day and 14-day *scrib* mutant imaginal discs and * mark genes that fall into the top 5% genes with significant changes.

(F) A 4-day wild type imaginal disc, a 4-day and a 14-day *scrib* mutant imaginal disc stained for notum identify marker Tailup (Tup, green) and Actin (red). Scale bar: 10μm.

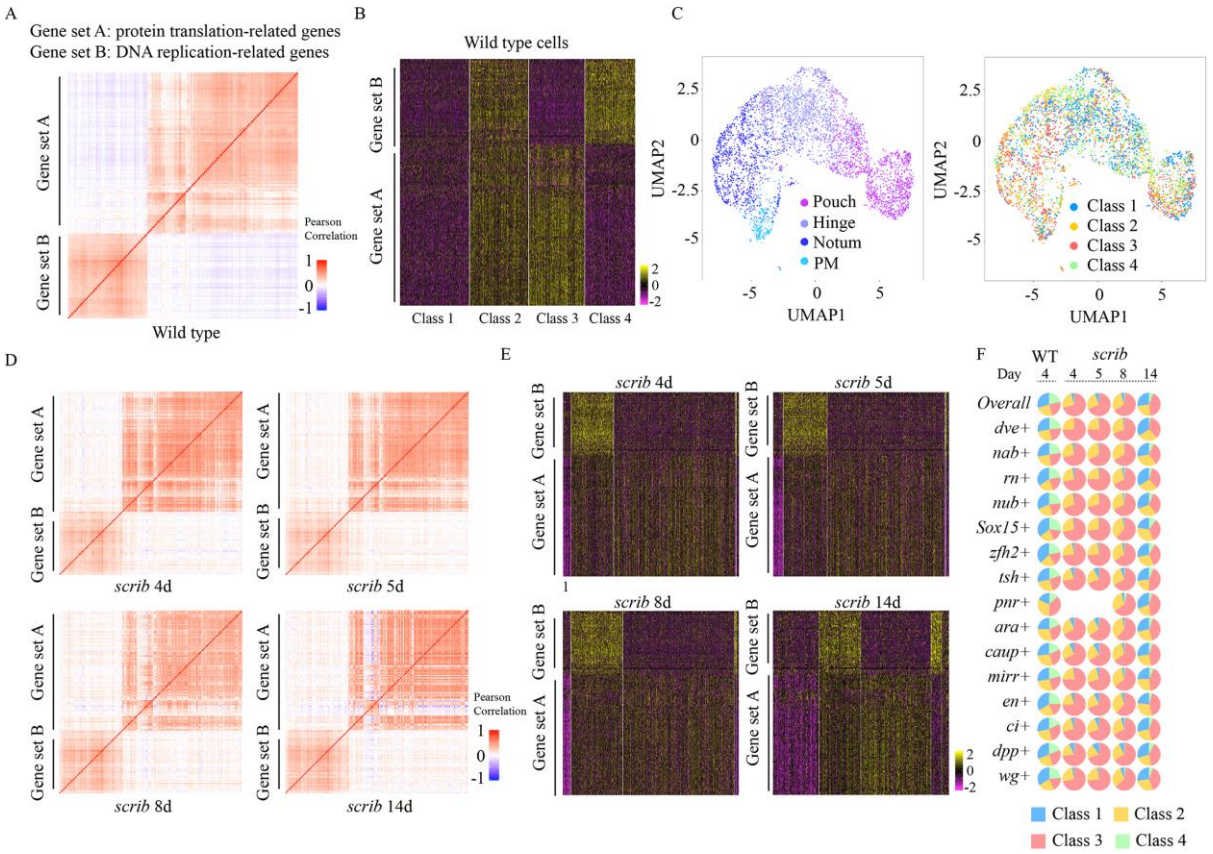

Figure 5 Analysis of single cell proliferation and growth states in the wildtype and the *scrib* mutant imaginal discs.

(A) Plot of Pearson correlation coefficients of the normalized expression levels of a protein translation gene set and a DNA replication gene set in wild type wing imaginal discs (4-day AEL). n=5200 cells.

(B) Hierarchical clustering of single cells from the wild type imaginal disc (4-day AEL) using the protein translation gene set and the DNA replication gene set. n=4498 cells, myoblasts and tracheal cells are removed from this analysis.

(C) UMAP projections of wing disc cells colored by spatial region identity and proliferation/growth classes defined in (B). n=4498 cells, myoblasts and tracheal cells are removed from this analysis.

(D) Plot of Pearson correlation coefficients of the normalized expression levels of the protein translation gene set and the DNA replication gene set in the *scrib* mutant wing imaginal discs

of different stages. 4-day, n=3440, 5-day, n=5614, 8-day, n=2622 and 14-day AEL, n=948, myoblasts and tracheal cells are removed from this analysis.

(E) Hierarchical clustering of the *scrib* mutant single cells into cell classes defined in 5B through k-nearest neighbor algorithm. 4-day, n=3440, 5-day, n=5614, 8-day, n=2622 and 14-day AEL, n=948, myoblasts and tracheal cells are removed from this analysis.

(F) Plot of the distribution of 4 classes of cells defined in 5B and 5E in cell populations expressing specific patterning genes for both wildtype and the *scrib* mutant single cells. WT, 4-day, n=4498 cells; *scrib*, 4-day, n=3440, 5-day, n=5614, 8-day, n=2622 and 14-day AEL, n=948; myoblasts and tracheal cells are removed from this analysis.

Table S1 Summary of 10x single cell RNA sequencing results for 5 samples.

Click here to Download Table S1

Table S2 List of differentially expressed genes in two sub-populations of myoblasts.

Identified through FindMarkers function in Seurat, Wilcox test, Logfc.threshold=0.2.

Click here to Download Table S2

Table S3 List of landmark genes used for single cell spatial mapping in wing imaginal discs.

Click here to Download Table S3

Table S4 List of gene sets used as proxies for single cell proliferation and growth state analysis in wing imaginal discs.

Click here to Download Table S4

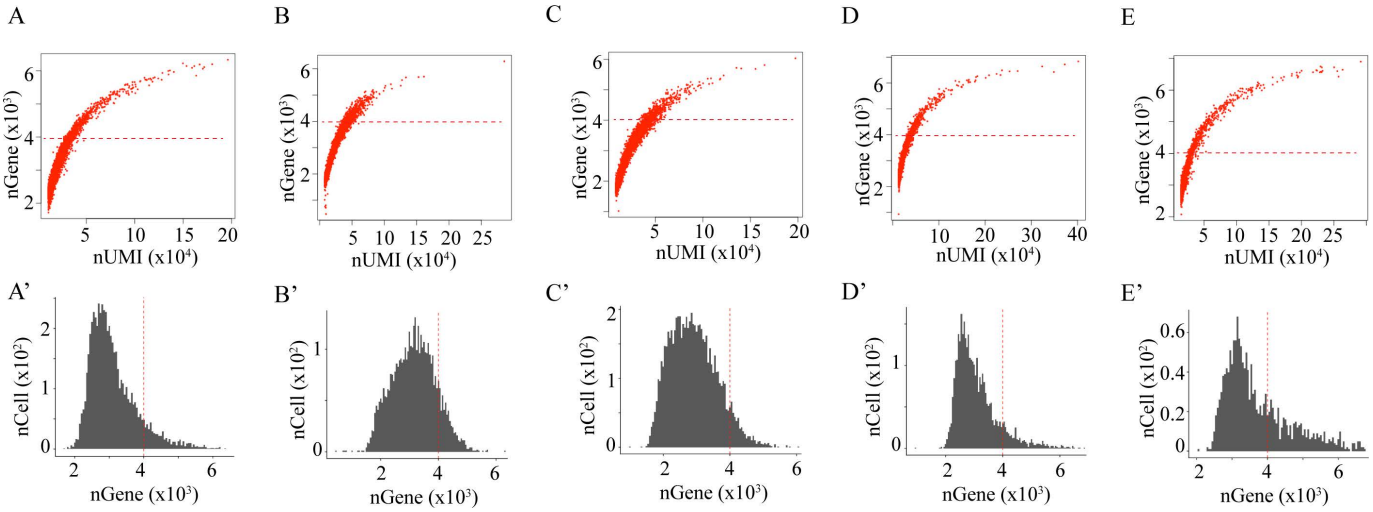

Figure S1 Single cells are filtered to remove possible cell doublets or clusters before downstream analysis.

For single cells in each sample (5 in total), the total genes detected per cell and the total unique molecular identifiers (UMI) per cell are plotted in the upper panel (A-E) and the histogram of genes detected per cell are plotted in the lower panel (A'-E'). We applied a stringent filter and removed cells in which over 4000 genes are detected to avoid possible cell doublets or cell clusters.

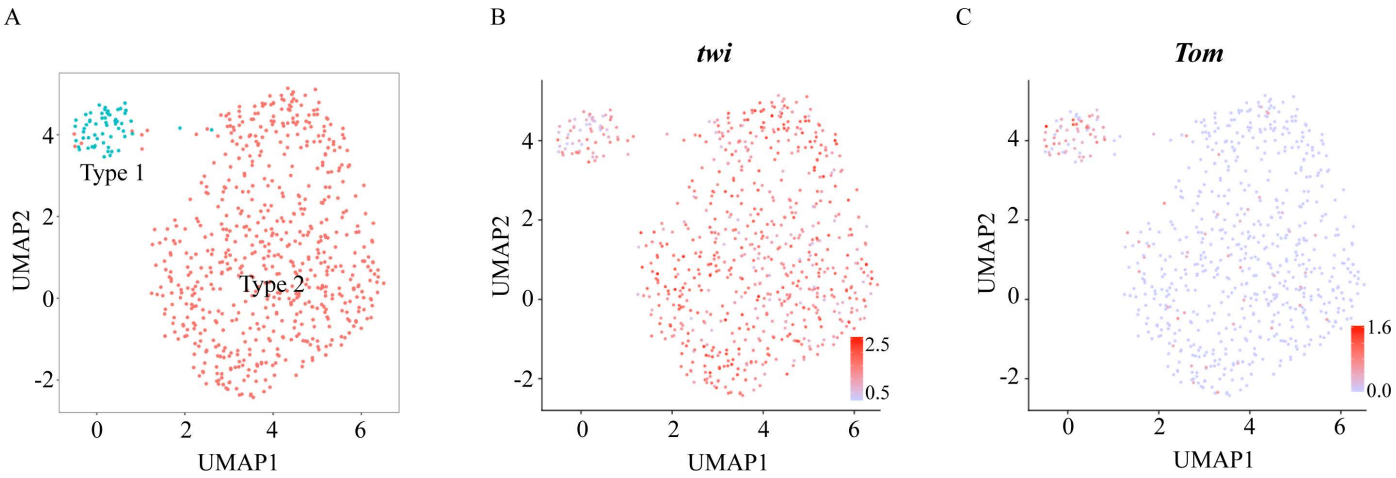

Figure S2 The myoblasts contain two distinct sub-populations of cells.

(A-C) The UMAP projection of myoblasts colored by cell type (A), the expression levels of *twist (twi)* (B) and *Tom* (C). Tom is a Notch pathway component. n=689 cells. A list of differentially expressed genes in these two subpopulations of myoblasts is provided in Supplemental table 2.

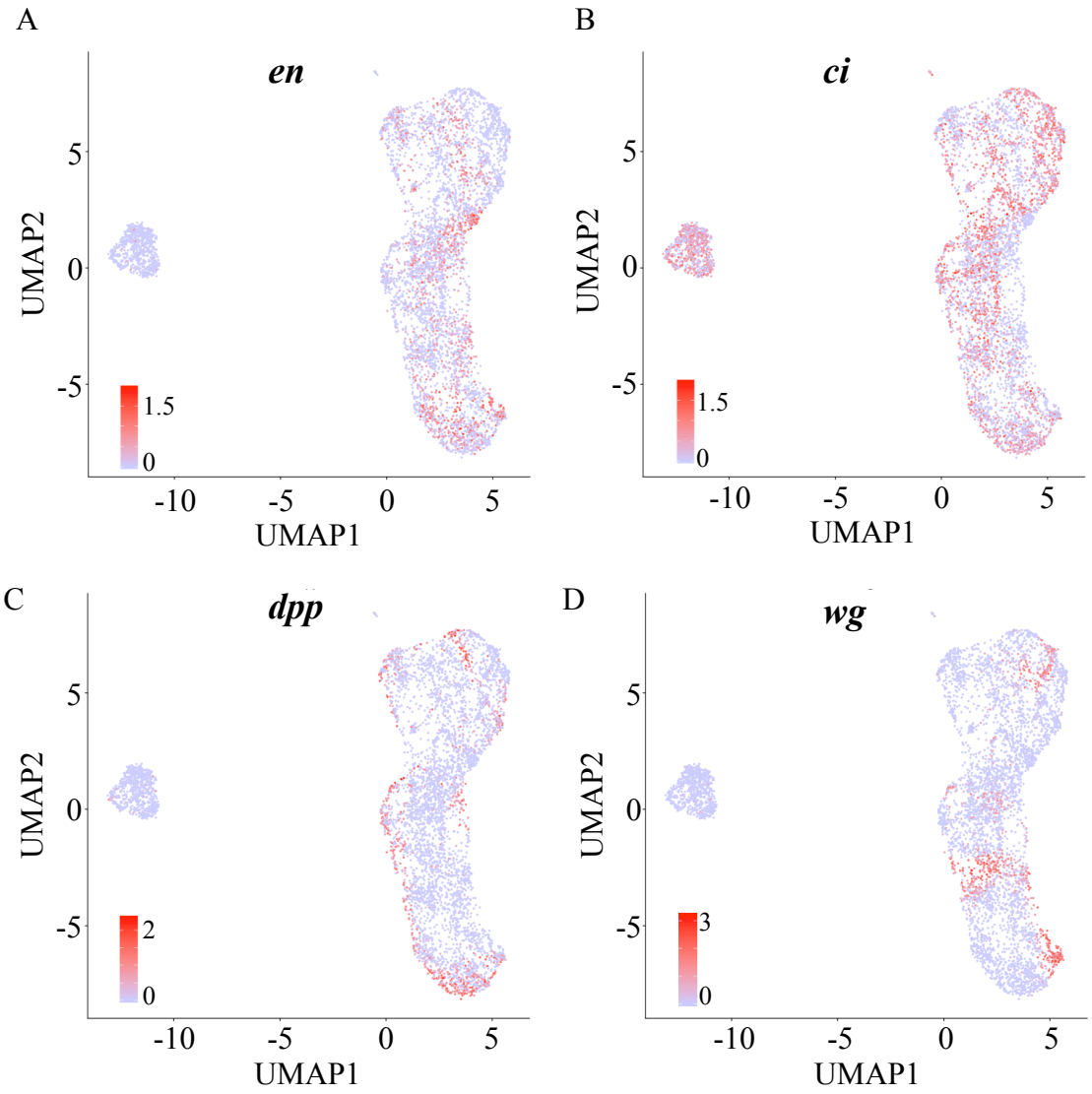

Figure S3 Examination of wing imaginal disc patterning marker genes.

(A-D) UMAP projections of wild type wing disc cells colored by the expression levels of patterning marker genes including *engrailed (en)*(A), *cubitus interruptus (ci)*(B), *decapentaplegic (dpp)* (C) and *wingless (wg)* (D). Note that wing disc cells do not cluster well based on their anterior-posterior identity (marked by *en* and *ci*) before assignment into sub-regions. Neither Dpp-producing nor Wg-producing cells form a single cluster. n=5200 cells.

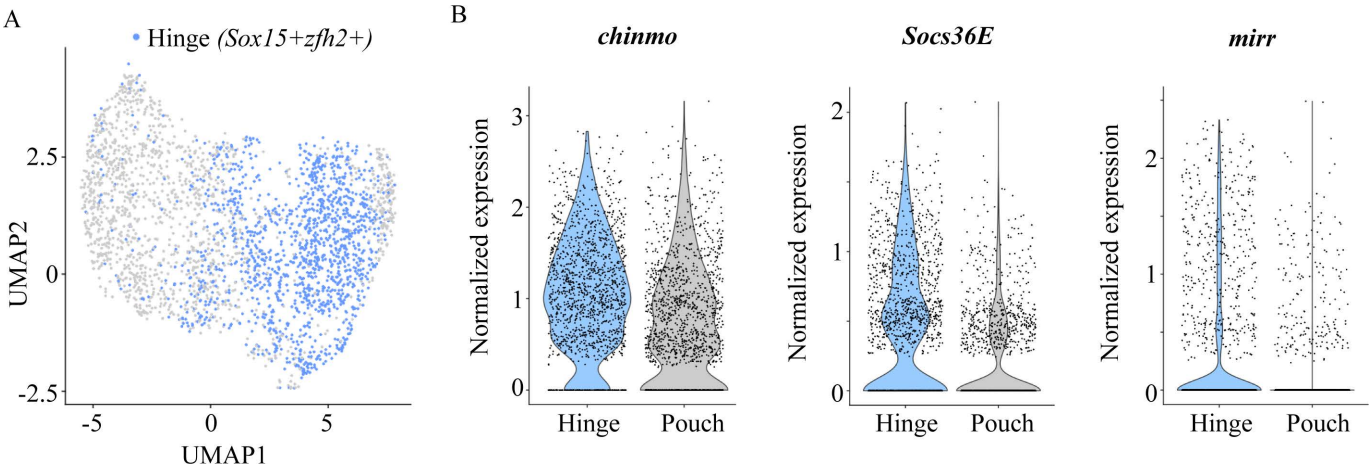

Figure S4 The wing disc hinge cells exhibit high JAK-STAT signaling activity.

(A) The UMAP projection of pouch and hinge cells colored by the hinge identity genes *Sox15* and *Zfh2*. n=2906 cells.

(B) Violin plot of *chinmo, Socs36E* and *mirr* expression levels in the hinge region (*Sox15*+ or *Zfh2*+) and the pouch region (*Sox15*- *Zfh2*-). n=2906 cells.

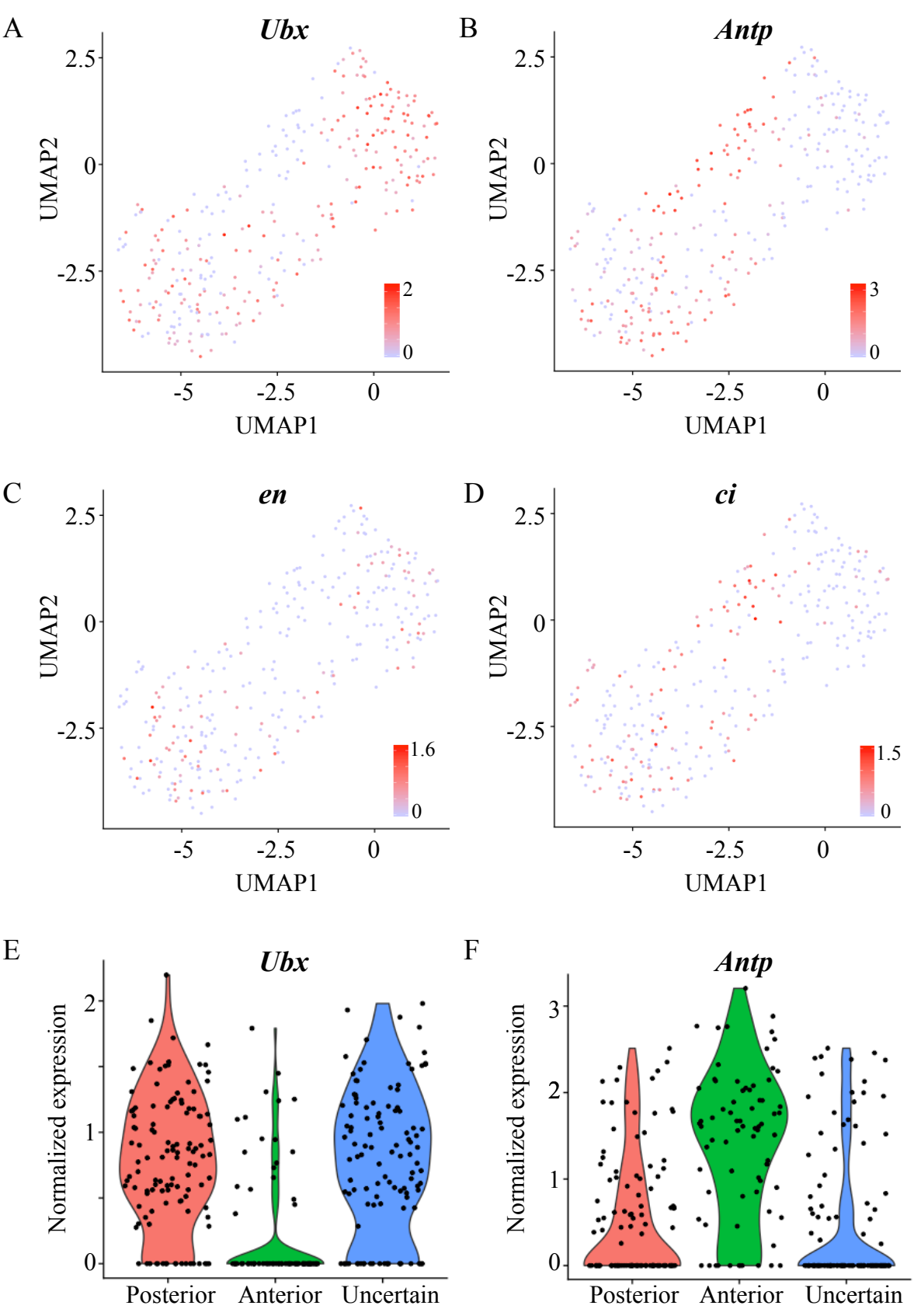

Figure S5 The peripodial membrane (PM) cells contain two cell populations marked by *Ubx* and *Antp*.

(A-D) UMAP projections of PM cells colored by the expression levels of *Ubx* (A), *Antp* (B), *en* (C) and *ci* (D). n=303 cells.

(E-F) Violin plot of *Ubx* (E) and *Antp* (F) expression in the posterior region (*en*+ or *hh*+), the anterior region (*ci*+) and an uncertain region (*en-hh-ci-*, n=112 cells). n=303 cells.

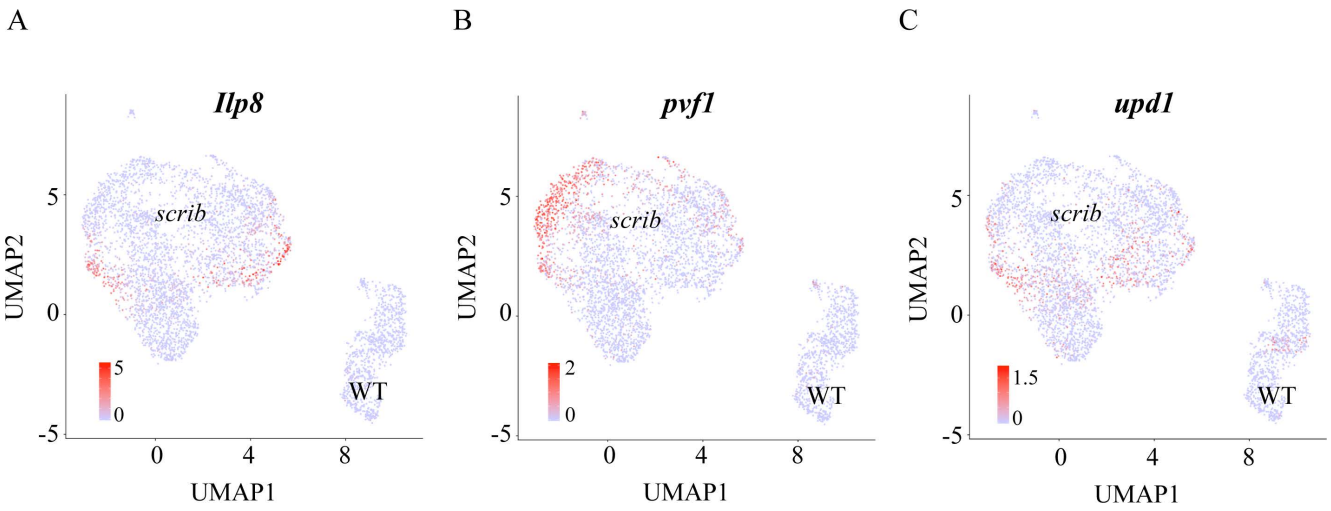

Figure S6 The *scrib* mutant cells express known tumor-specific markers.

UMAP projections of single cells from the wild type imaginal disc (4-day AEL) and the *scrib* mutant imaginal discs (4-day, 5-day, 8-day and 14-day AEL) colored by the expression levels of *Ilp8* (A), *pvf1* (B) and *upd1* (C). n= 948 cells for each sample.

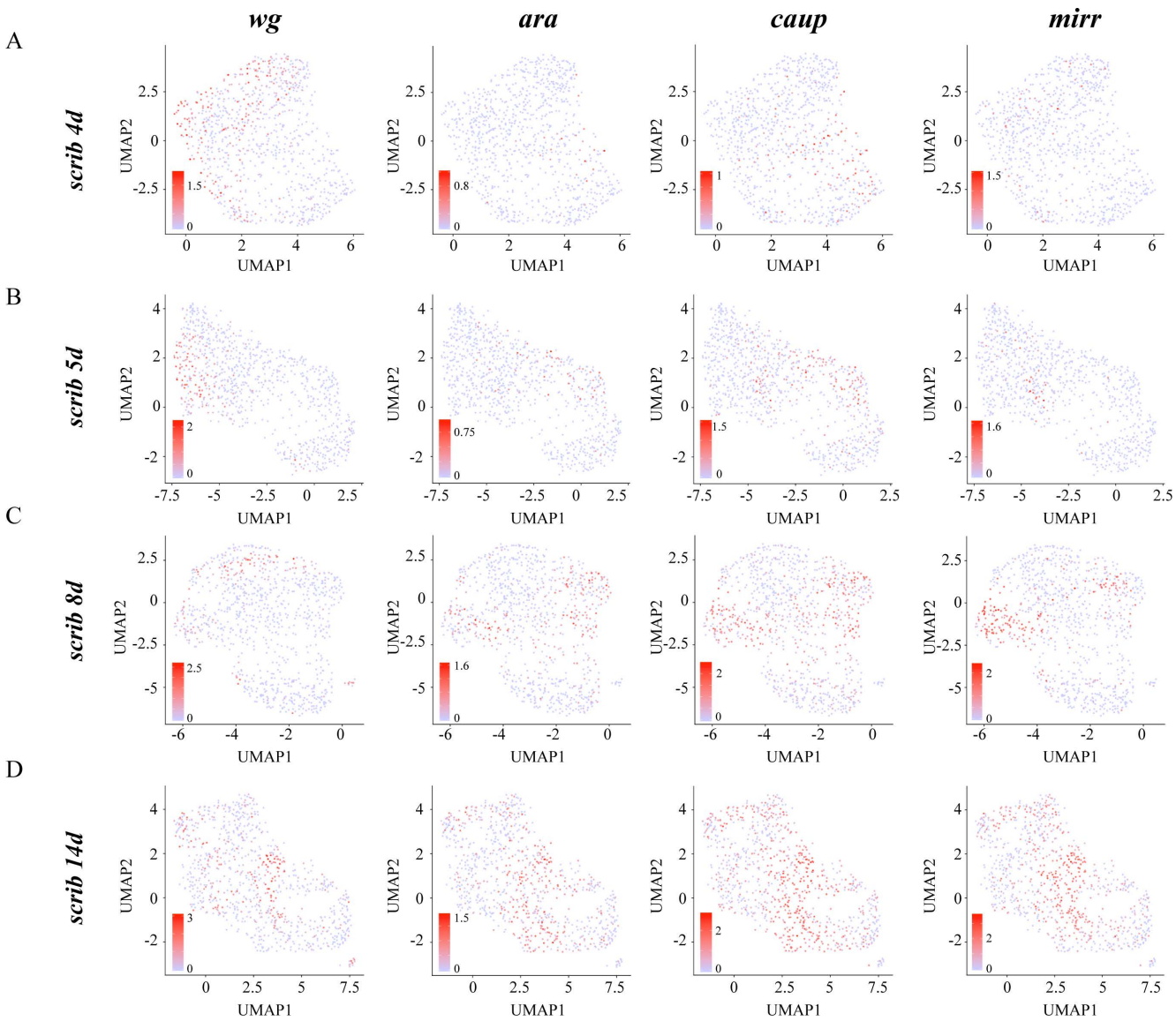

Figure S7 More cells acquire notum identity over time in the *scrib* mutant imaginal discs

UMAP projections of the *scrib* mutant cells (4-day, 5-day, 8-day and 14-day AEL) colored by the expression levels of *wg*, *ara, caup* and *mirr*. n= 948 cells for each sample.

Table S1 Summary of 10x single cell RNA sequencing results for 5 samples.

Click here to Download Table S1

Table S2 List of differentially expressed genes in two sub-populations of myoblasts.

Identified through FindMarkers function in Seurat, Wilcox test, Logfc.threshold=0.2.

Click here to Download Table S2

Table S3 List of landmark genes used for single cell spatial mapping in wing imaginal discs.

Click here to Download Table S3

Table S4 List of gene sets used as proxies for single cell proliferation and growth state analysis in wing imaginal discs.

Click here to Download Table S4

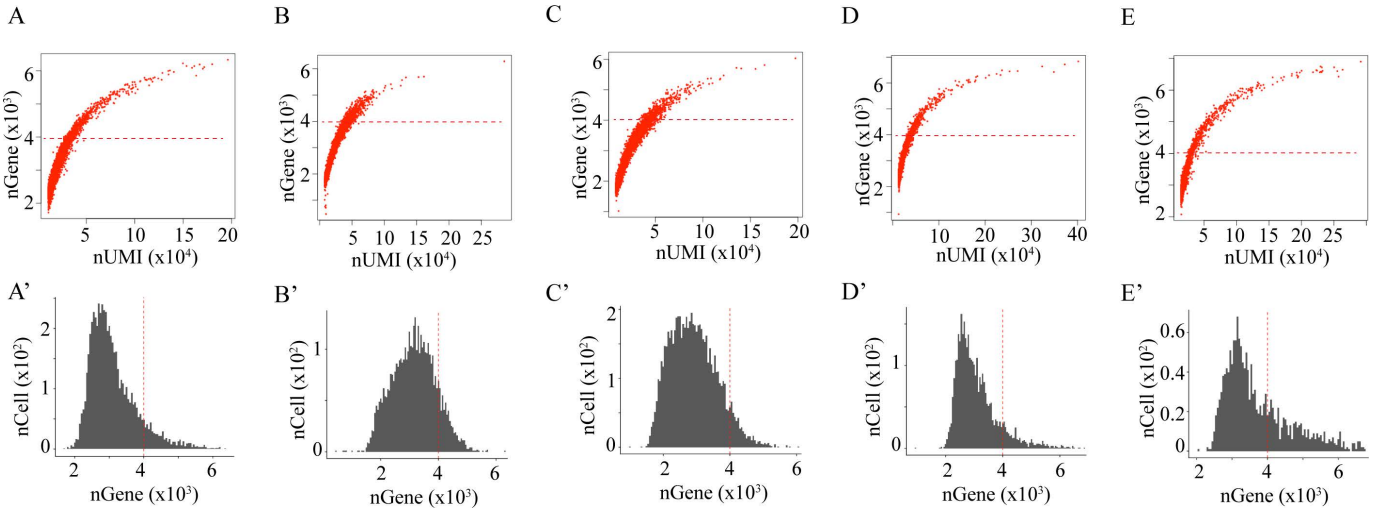

Figure S1 Single cells are filtered to remove possible cell doublets or clusters before downstream analysis.

For single cells in each sample (5 in total), the total genes detected per cell and the total unique molecular identifiers (UMI) per cell are plotted in the upper panel (A-E) and the histogram of genes detected per cell are plotted in the lower panel (A'-E'). We applied a stringent filter and removed cells in which over 4000 genes are detected to avoid possible cell doublets or cell clusters.

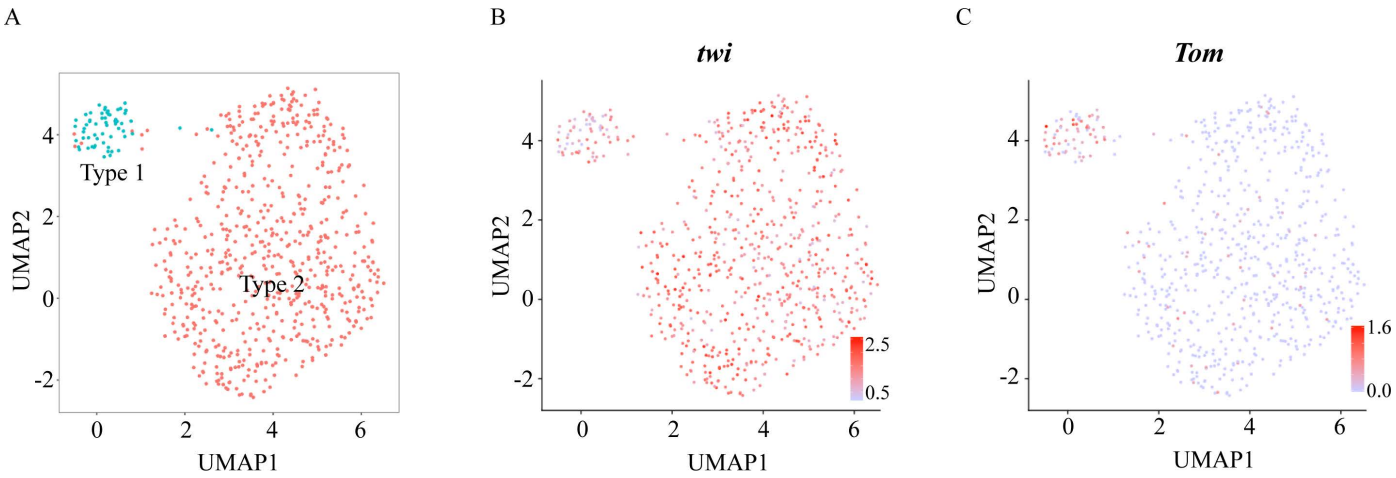

Figure S2 The myoblasts contain two distinct sub-populations of cells.

(A-C) The UMAP projection of myoblasts colored by cell type (A), the expression levels of *twist (twi)* (B) and *Tom* (C). Tom is a Notch pathway component. n=689 cells. A list of differentially expressed genes in these two subpopulations of myoblasts is provided in Supplemental table 2.

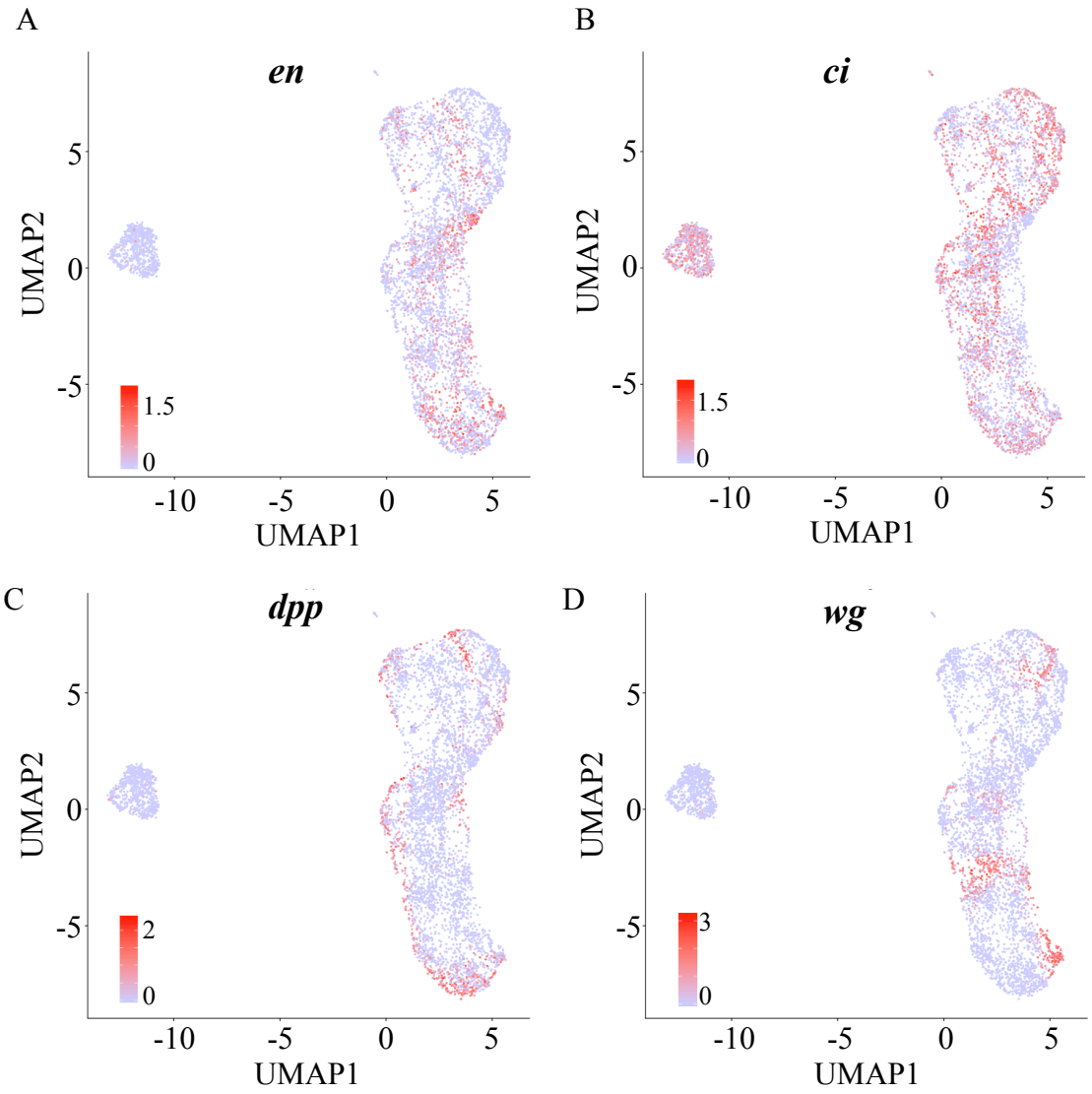

Figure S3 Examination of wing imaginal disc patterning marker genes.

(A-D) UMAP projections of wild type wing disc cells colored by the expression levels of patterning marker genes including *engrailed (en)*(A), *cubitus interruptus (ci)*(B), *decapentaplegic (dpp)* (C) and *wingless (wg)* (D). Note that wing disc cells do not cluster well based on their anterior-posterior identity (marked by *en* and *ci*) before assignment into sub-regions. Neither Dpp-producing nor Wg-producing cells form a single cluster. n=5200 cells.

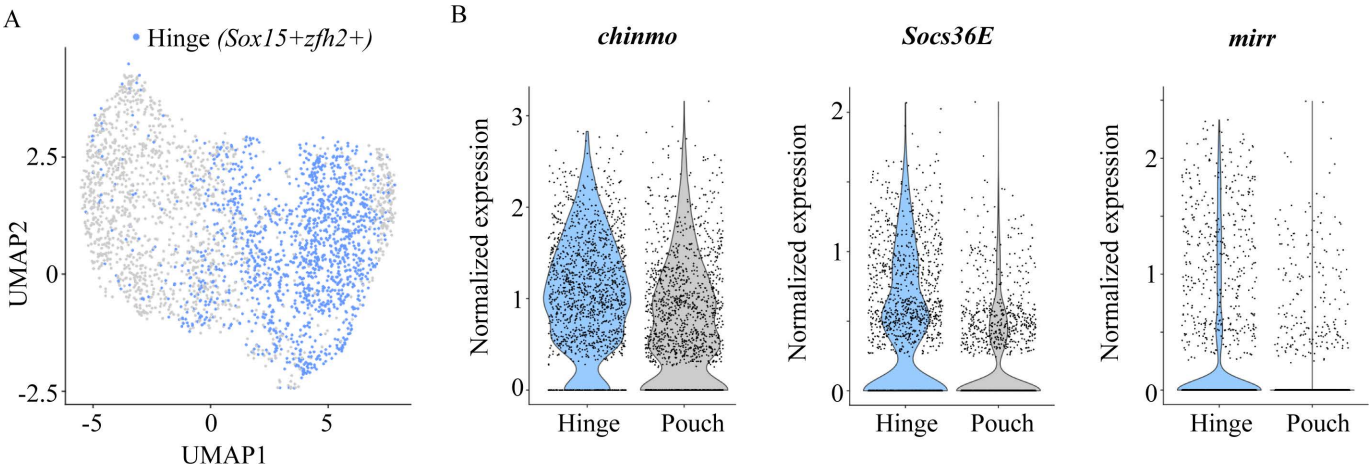

Figure S4 The wing disc hinge cells exhibit high JAK-STAT signaling activity.

(A) The UMAP projection of pouch and hinge cells colored by the hinge identity genes *Sox15* and *Zfh2*. n=2906 cells.

(B) Violin plot of *chinmo, Socs36E* and *mirr* expression levels in the hinge region (*Sox15*+ or *Zfh2*+) and the pouch region (*Sox15*- *Zfh2*-). n=2906 cells.

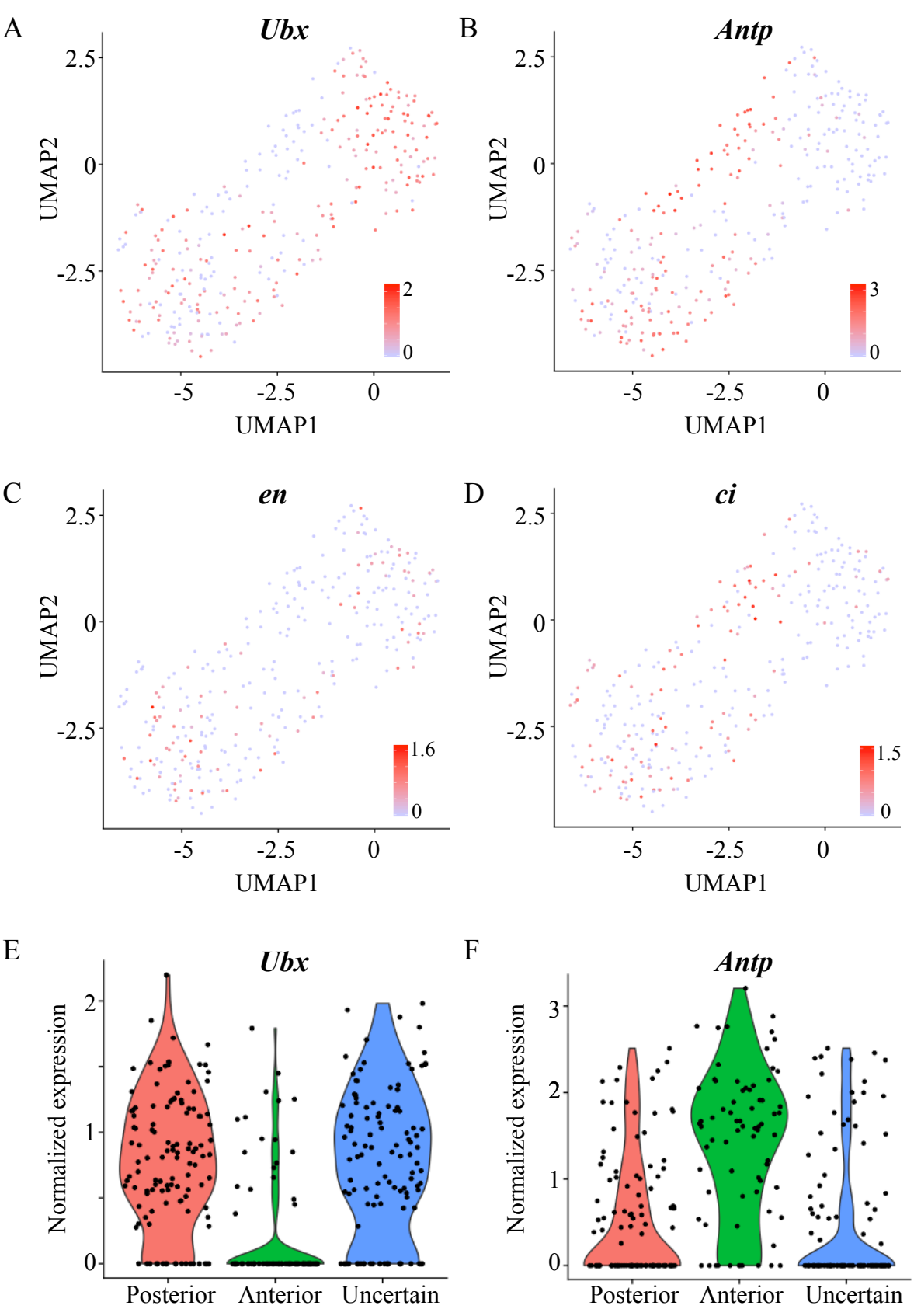

Figure S5 The peripodial membrane (PM) cells contain two cell populations marked by *Ubx* and *Antp*.

(A-D) UMAP projections of PM cells colored by the expression levels of *Ubx* (A), *Antp* (B), *en* (C) and *ci* (D). n=303 cells.

(E-F) Violin plot of *Ubx* (E) and *Antp* (F) expression in the posterior region (*en*+ or *hh*+), the anterior region (*ci*+) and an uncertain region (*en-hh-ci-*, n=112 cells). n=303 cells.

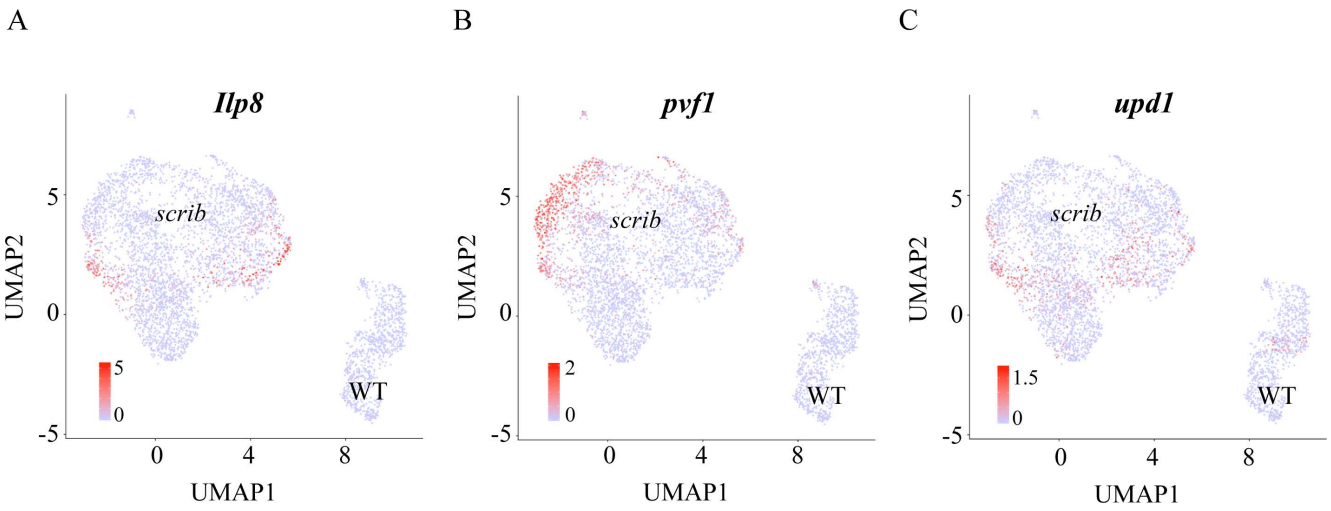

Figure S6 The *scrib* mutant cells express known tumor-specific markers.

UMAP projections of single cells from the wild type imaginal disc (4-day AEL) and the *scrib* mutant imaginal discs (4-day, 5-day, 8-day and 14-day AEL) colored by the expression levels of *Ilp8* (A), *pvf1* (B) and *upd1* (C). n= 948 cells for each sample.

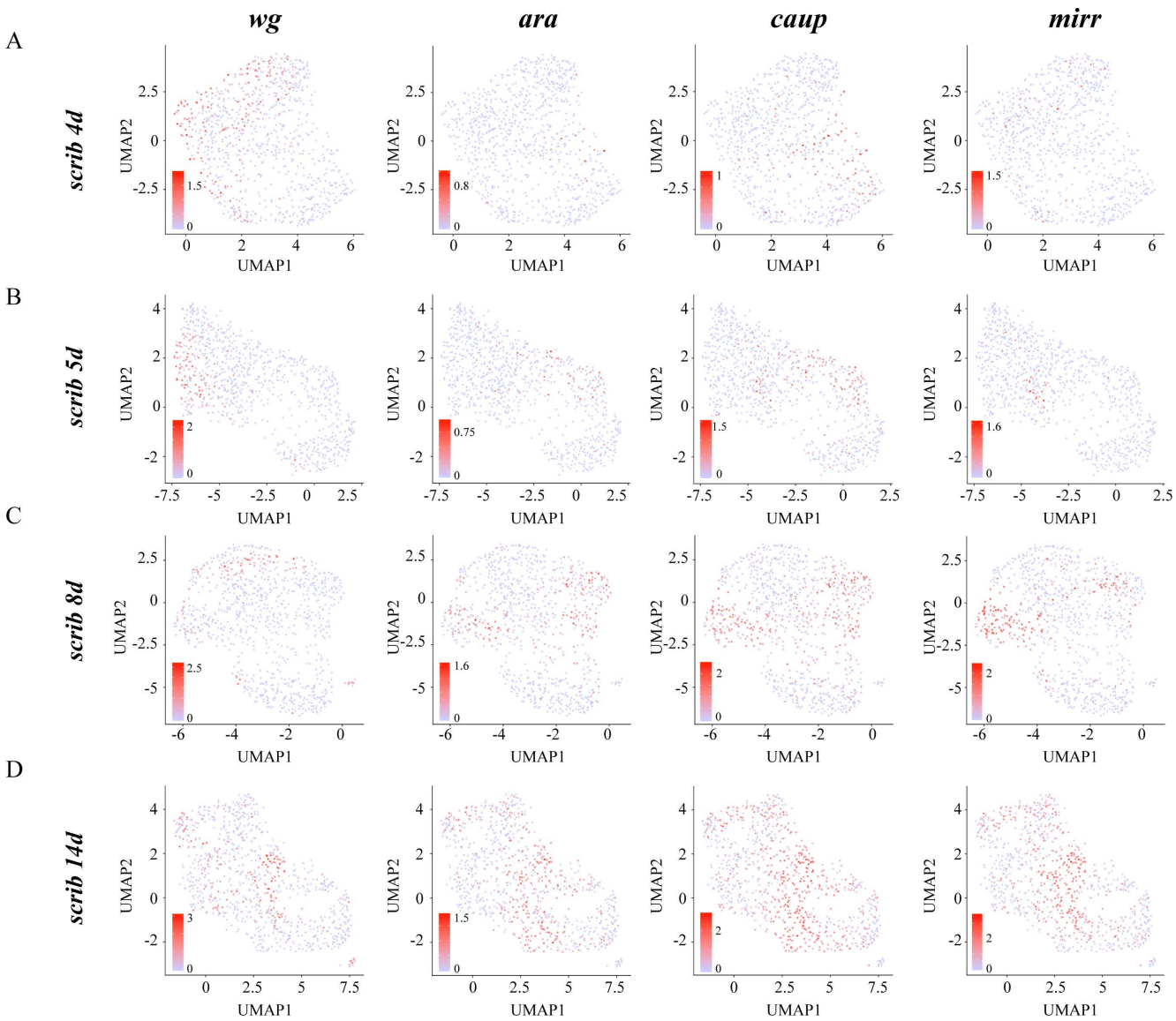

Figure S7 More cells acquire notum identity over time in the *scrib* mutant imaginal discs

UMAP projections of the *scrib* mutant cells (4-day, 5-day, 8-day and 14-day AEL) colored by the expression levels of *wg*, *ara, caup* and *mirr*. n= 948 cells for each sample.