

# ***Drosophila* AHR limits tumor growth and stem cell proliferation in the intestine**

Minghua Tsai<sup>1§</sup>, Jiawei Sun<sup>1§</sup>, Cyrille Alexandre<sup>1</sup>, Michael Shapiro<sup>1</sup>, Adrien Franchet<sup>1</sup>, Ying Li<sup>1</sup>, Alex P. Gould<sup>1</sup>, Jean-Paul Vincent<sup>1</sup>, Brigitta Stockinger<sup>1</sup>, Nicola Laura Diny<sup>1\*</sup>

<sup>1</sup> The Francis Crick Institute, NW1 1AT London, UK  
§ equal contribution

\* Corresponding author:  
Nicola L. Diny  
The Francis Crick Institute  
1, Midland Road  
NW1 1AT, London  
[nicola.diny@crick.ac.uk](mailto:nicola.diny@crick.ac.uk)

## **Abstract**

The aryl hydrocarbon receptor (AHR) plays important roles in intestinal homeostasis, limiting tumour growth and promoting differentiation in the intestinal epithelium. Spineless, the *Drosophila* homolog of AHR, has only been studied in the context of development but not in the adult intestine. Here, we show that *spineless* is upregulated in the adult intestinal epithelium after infection with *Pseudomonas entomophila* (*P.e.*). Spineless knockdown increased stem cell proliferation following infection-induced injury. Spineless overexpression limited intestinal stem cell proliferation and reduced survival after infection. In two tumour models, using either *Notch* RNAi or constitutively active Yorkie, Spineless suppressed tumour growth and doubled the lifespan of tumour-bearing flies. At the transcriptional level it reversed the gene expression changes induced in Yorkie tumours, counteracting cell proliferation and altered metabolism. These findings demonstrate a new role for Spineless in the adult *Drosophila* midgut and highlight the evolutionarily conserved functions of AHR/Spineless in the control of proliferation and differentiation of the intestinal epithelium.

## **Introduction**

The aryl hydrocarbon receptor (AHR) is a ligand-activated transcription factor with barrier-protective roles in the intestine<sup>1</sup>. AHR is an environmental sensor of the basic helix-loop-helix Per-Arnt-Sims (bHLH-PAS) family that binds ligands derived from the diet, microbial metabolism or industrial sources. Ligand binding leads to release of AHR from its chaperone complex and nuclear translocation<sup>2</sup>. AHR then dimerizes with AHR nuclear translocator (ARNT) for DNA binding at canonical binding sites. AHR is widely expressed in many intestinal immune cells, stromal cells and the intestinal epithelium<sup>3</sup>, where it is important in the anti-bacterial defence and in limiting tumour growth<sup>1</sup>. Ablation of AHR in the intestinal epithelium of mice leads to increased susceptibility to infection with *Citrobacter rodentium* and increased malignant transformation in an AOM-DSS model<sup>4</sup>. AHR is needed to end the regenerative response of the intestinal epithelium after injury to allow the epithelial barrier to return to its mature state<sup>5</sup>. Treatment of mice with AHR ligand-rich diet was shown to be beneficial in tumour models and epithelial healing after injury with DSS<sup>4,5</sup>.

Spineless is the closest *Drosophila* homolog to AHR and binds the same DNA sequence<sup>6,7</sup>. Amino-acid identity between AHR and Spineless is 41% overall but substantially higher in the PAS domains and reaches 70% in the DNA-binding site<sup>8</sup>. The lowest similarity is found in the ligand-binding domain. This is in line with the idea that Spineless is a ligand-independent transcription factor that cannot bind prototypic AHR ligands like dioxin and does not require ligands for nuclear translocation<sup>9,10</sup>. Moreover, *Drosophila* is unaffected by dioxin<sup>11</sup>. Akin to the AHR-ARNT dimer in vertebrates, Spineless forms a heterodimer with the bHLH-PAS family member Tango and this heterodimer then translocates to the nucleus to bind to dioxin-response elements<sup>7</sup>. Another line of evidence for strong evolutionary conservation between these homologs comes from a study showing that AHR could rescue the developmental phenotypes of Spineless mutants in *Drosophila*<sup>11</sup>. The authors also demonstrated that AHR/Spineless functions in *Drosophila* are highly dependent on gene dosage of Tango or ARNT and that dioxin treatment could enhance AHR functions in murine AHR-transgenic *Drosophila*. Thus, *Drosophila* Spineless might be a useful model to study evolutionarily conserved AHR functions.

The functions of Spineless have been studied extensively in *Drosophila* development. Spineless was first identified for controlling antenna development, with mutants causing the aristapedia phenotype<sup>12</sup>. Spineless has since been shown to function together with Tango to control antennal identity and the development of tarsal segments of the leg<sup>7,8,13,14</sup>. Spineless also plays important roles in regulating dendrite morphology in neurons<sup>15</sup>, the development of sternopleural bristles<sup>16</sup>, and photoreceptor specification in the retina<sup>17-21</sup>. Few studies have focused on the function of Spineless in adult flies and the role of Spineless in the adult intestine has not been studied.

The *Drosophila* midgut consists of a single layer of epithelial cells surrounded by a basement membrane and visceral muscle. Similar to the mammalian intestine, the epithelium regenerates from intestinal stem cells (ISC), which give rise to transient enteroblasts (EB) and further differentiate into mature absorptive enterocytes<sup>22</sup>. The *Drosophila* intestine also contains a secretory lineage, the enteroendocrine cells. ISC are characterized by expression of Escargot (Esg) and Delta (DI) and suppression of Notch (N) signalling for their maintenance<sup>23,24</sup>. Upon asymmetric division and differentiation into EB, cells downregulate Delta, activate Notch signalling and induce Suppressor of Hairless (Su(H)). Differentiation into Pou domain protein 1 (Pdm1)-positive enterocytes is driven by Jak/Stat signalling, Notch activation and a downregulation of Esg. Prospero (Pros) is induced during differentiation into enteroendocrine cells. A range of local, systemic, and environmental stimuli are integrated through multiple signalling pathways (including Notch, Jak/Stat, Egfr and Hippo) to govern ISC proliferation and differentiation and to maintain the epithelial barrier and its function<sup>22</sup>. Many of these pathways have also been shown to interact with AHR to regulate stem cell maintenance and differentiation in the mammalian intestinal epithelium<sup>1,4,5,25,26</sup>.

Given the critical roles that AHR plays in the mammalian intestine, we hypothesized that Spineless may have evolutionarily conserved functions in the *Drosophila* midgut. We found that *spineless* is indeed upregulated in the adult midgut after bacterial infection where it limits the regenerative response and functions as a tumour suppressor in two independent models.

## Results

### Spineless limits intestinal stem cell proliferation after *Pseudomonas entomophila* infection

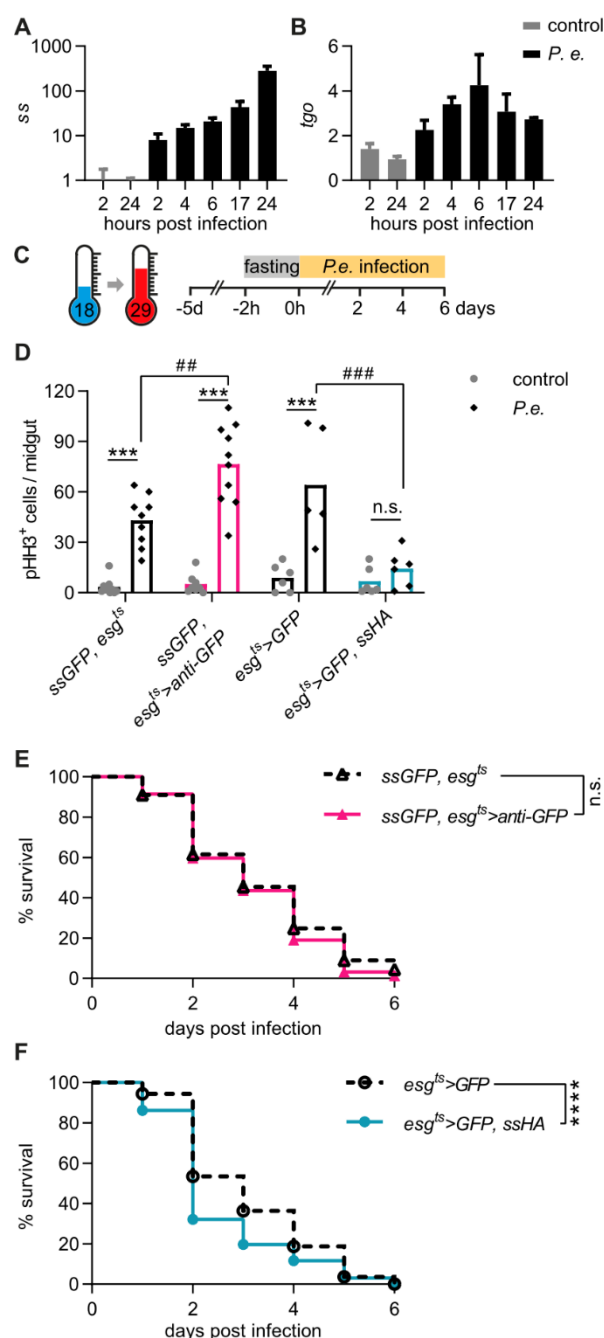
Given the critical functions of AHR in the mammalian intestine, we sought to determine if Spineless has similar evolutionarily conserved roles in the *Drosophila* midgut. We infected flies with the enteropathogenic bacterium *Pseudomonas entomophila* as a model of intestinal damage and regeneration. Bacterial infection induced expression of antimicrobial defence genes *DiptB*, *Nuox*, *Upd3* (Fig S1A-C). *Spineless* was expressed at low levels in the steady state intestine and was induced over 200-fold 24h following bacterial infection (Fig. 1A). Expression of the Spineless binding protein

*tango* (*tgo*) remained largely unchanged after infection (Fig 1B). We analysed published RNA-sequencing data of different intestinal cell populations from Dutta and colleagues<sup>27</sup> which confirmed low expression of *spineless* in the steady-state midgut and strong induction after *P. entomophila* infection (Fig. S1D). Of note, *spineless* was only induced in intestinal stem cells (ISC) and enteroblasts (EB), but not in enterocytes or enteroendocrine cells. This suggests a potential role for Spineless in the progenitor compartment following bacterial infection.

We generated new lines to conditionally knockdown or overexpress Spineless to study its function in midgut progenitors. *ssGFP* flies were created by targeted insertion of GFP at the C-terminus of the endogenous *spineless* locus (Fig S1E). Homozygous *ssGFP* flies showed no apparent phenotype and *ssGFP* was clearly visible in the antenna imaginal disk (Fig. S1F). Flies expressing a membrane-anchored anti-GFP nanobody (*UAS-anti-GFP*) were created to prevent *ssGFP* from translocating to the nucleus (Fig. S1G). To achieve cell type-specific Spineless knockdown we utilized the GAL4/UAS system. GAL4 can be expressed under a cell type-specific promoter to induce expression of transgenes with a GAL4 binding site (UAS). We used the promoter of *rotund*, which is expressed in the imaginal disk, to induce anti-GFP. *ssGFP*, *rotund>anti-GFP* flies showed aristapedia and leg phenotypes typical of Spineless mutants (Fig. S1H)<sup>7,8</sup>. To overexpress Spineless, *UAS-ss* flies with a C-terminal 3xHA tag were generated (Fig S1I). We made use of the *escargot* (*esg*)-*Gal4* driver and the temperature sensitive repressor *tubGal80<sup>ts</sup>* to manipulate Spineless expression in ISC and EB of the adult midgut only. Flies were reared at 18°C and switched to 29°C 5 days prior to infection (Fig. 1C). Infection with *P. entomophila* induces epithelial damage followed by a wave of ISC proliferation that leads to epithelial regeneration<sup>28-31</sup>. ISC proliferation was induced in controls following *P. entomophila* infection. Spineless knockdown further increased ISC proliferation while Spineless overexpression abolished the ISC proliferative response (Fig. 1D). We hypothesized that this could reduce the ability

**Figure 1: *Spineless* limits intestinal stem cell proliferation after *Pseudomonas entomophila* infection**

A, B) Gene expression was determined by qPCR from isolated midguts of *w1118* flies at different timepoints after *P. entomophila* infection. Data are from one experiment with n=3 samples per timepoint. Gene expression was normalized to *Rpl32* and uninfected controls. C) Schematic of *P. entomophila* infection. D) The number of pH3<sup>+</sup> cells per midgut was quantified at 24h post *P. entomophila* infection. Data are from 3 independent experiments, n=5-10 per genotype. E, F) Survival following *P. entomophila* infection. Data are pooled from 5 independent experiments, n=230-300 flies per genotype. *P.e.*, *Pseudomonas entomophila*.



of Spineless overexpressing flies to survive infection. While knockdown of Spineless did not affect survival following infection with *P. entomophila* (Fig. 1E), overexpression of Spineless accelerated death following infection (Fig. 1F). This suggests that overexpression of Spineless limits ISC proliferation after *P. entomophila* infection, which impairs epithelial regeneration and thereby leads to reduced survival.

### Spineless overexpression reduces survival through ISC- and EB-specific effects

In the midgut, *esg>GFP* labels ISC and EB and therefore allows quantification of these cells as well as measurement of cell size and GFP intensity. In naïve flies, Spineless overexpression increased *esg>GFP* positive cells on fluorescent images of the midgut compared to wildtype flies (Fig. 2A, Fig. S2A), suggesting that it affects intestinal progenitor dynamics. In wildtype flies, *P. entomophila* infection resulted in a noticeable increase in the number of *esg>GFP* positive cells and an increase in their cell size. Spineless overexpression blocked the increase in cell number and size after *P. entomophila* infection, in line with the lack of increased ISC proliferation measured by pHH3<sup>+</sup> cell quantification (as seen in Fig. 1D).

We then used flow cytometry to quantify the intestinal cell populations. *Esg>GFP* positive cells are readily detectable by flow cytometry after gating on live, single cells (Fig. S2B, C). EB can be distinguished from ISC based on their higher GFP fluorescent intensity and larger cell size as has been previously reported<sup>32</sup> (Fig. S2D, E). We confirmed higher *Delta* expression in ISC than EB populations by qPCR from FACS sorted cells (Fig. S2F). This approach corroborated that Spineless overexpression increased ISC and EB numbers at steady state and lead to a decrease in the number of ISC and EB in infected versus naïve Spineless overexpressing flies (Fig. 2B, C). It also confirmed an increase in cell size in wildtype flies after *P. entomophila* infection and a reduction in cell size in naïve or infected flies overexpressing Spineless compared to wildtype flies (Fig. 2D, E).

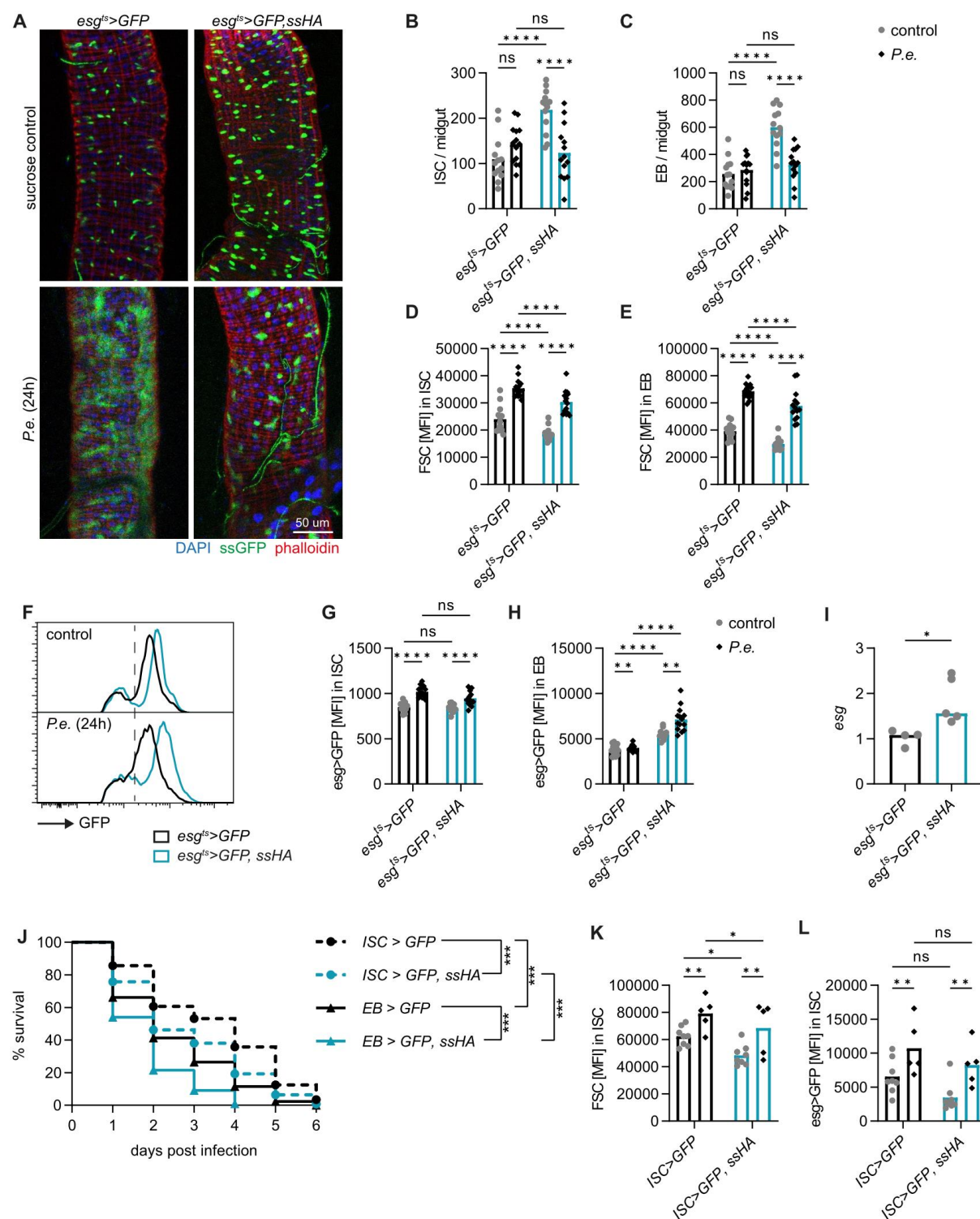
Within the ISC and EB populations, *esg>GFP* fluorescence increased after infection in wildtype controls (Fig. 2F-H). *Escargot* was shown to maintain stemness in ISC and EB<sup>33,34</sup>. In EB, *Escargot* also enhances Notch signalling by inhibiting Amun which may promote differentiation<sup>33</sup>. Spineless overexpression did not affect *esg>GFP* expression in ISC (Fig. 2G), but increased fluorescence intensity in both naïve and infected EB (Fig. 2H). We confirmed that *escargot* gene expression was increased by qPCR in FACS-sorted EB from naïve flies (Fig. 2I). This suggests that Spineless directly affects *escargot* expression in EB.

We next sought to determine whether Spineless overexpression in ISC or EB was responsible for the reduced survival following *P. entomophila* infection. Overexpression of Spineless using cell type-specific drivers for either ISC or EB reduced survival after infection in both cases (Fig. 2J). Spineless overexpression in ISC reduced ISC cell size but had no significant effect on *esg>GFP* fluorescent intensity (Fig. 2K, L), in line with the results observed with the *esg<sup>ts</sup>* driver. Spineless overexpression in EB did not affect EB cell size or *Su(H)>GFP* expression (Fig. S2G, H). Taken together, these results suggest that Spineless reduces survival following bacterial infection by limiting ISC proliferation and possibly by altering EB maturation through *escargot*.

### Spineless prevents intestinal tumour formation in the *Notch*<sup>RNAi</sup> tumour model

Limiting the proliferation of ISC is detrimental in the context of infection-induced damage and regeneration but could be beneficial in the context of tumours. We therefore hypothesized that Spineless may prevent tumour formation in the midgut. Loss of Notch signalling (through *Notch*<sup>RNAi</sup>) results in the proliferation of neoplastic ISC-like cells that fail to differentiate and form multi-layered tumours<sup>23,24,35</sup>. Spineless knockdown had a small but significant effect on survival in the *Notch*<sup>RNAi</sup> tumour model, reducing median survival from 21 to 20 days (Fig 3A, B). Tumour growth induced by *Notch*<sup>RNAi</sup> can be accelerated by infecting flies with *P. entomophila*, which induces a wave of stem cell proliferation<sup>35,36</sup>. We therefore infected flies with a low dose of *P. entomophila* for 24h and then followed their survival. Indeed, bacterial infection reduced the median survival of tumour flies from 21 to 7 days (Fig. 3B-D). Spineless knockdown flies succumbed even faster to tumours, with a

maximum survival of only 16 days compared to 37 days for control tumour flies after *P. entomophila* infection (Fig. 3D).



**Figure 2: *Spineless* overexpression reduces survival through ISC- and EB-specific effects**

A) Representative images of *P. entomophila* infected midguts at 24h post infection. B-H) Midguts from uninfected controls and at 24h post *P. entomophila* infection were analysed by flow cytometry. B, C) Quantification of ISC and EB numbers per midgut. D, E) Mean fluorescent intensity for FSC (cell size) in ISC and EB populations. F) Representative flow cytometry plots depicting GFP fluorescent intensity in GFP<sup>+</sup> cells. G, H) Geometric mean fluorescent intensity of GFP in ISC and EB populations. B-E, G, H) Data are pooled from 2



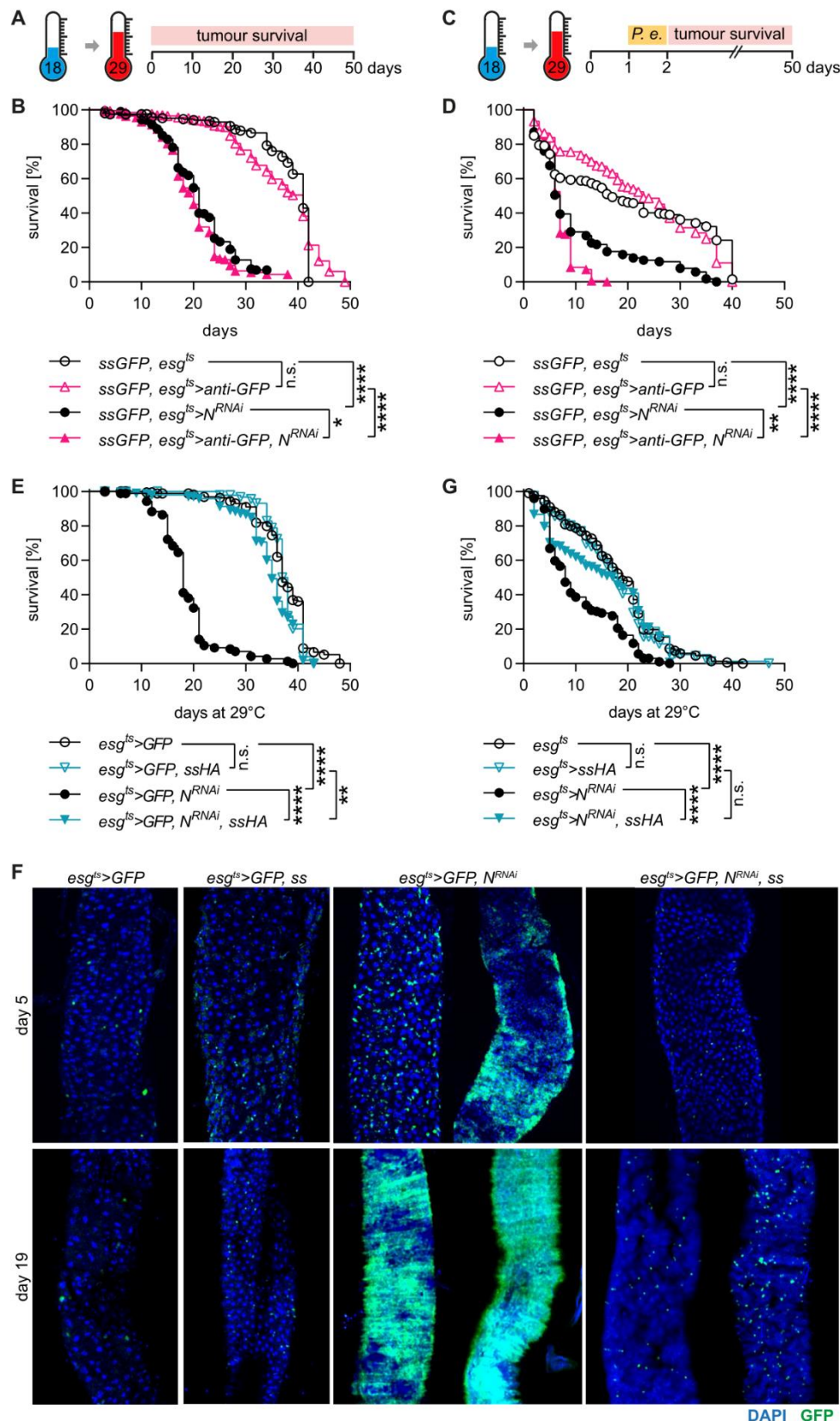
independent experiments, n=13-14 samples per genotype. I) *Escargot* (*esg*) expression was determined by qPCR in FACS sorted EB from naïve flies and gene expression was normalized to *Rpl32*. Data are pooled from 2 independent experiments, n=4-5 per genotype. J-L) *P. entomophila* infection in flies overexpressing *spineless* specifically in ISC (*esg-Gal4*, *Su(H)-Gal80*, *tub-Gal80<sup>ts</sup>*) or in EB (*Su(H)-Gal4*, *tub-Gal80<sup>ts</sup>*). J) Survival following *P. entomophila* infection. Data are pooled from 2 independent experiments, n=295-339 flies per genotype. K, L) Fluorescent intensity of FSC and GFP in ISC populations from uninfected controls and at 24h post *P. entomophila* infection. Data are from one experiment with n=5-8 samples per genotype.

Spineless overexpression completely blocked the development of tumours (Fig 3E, F). The median survival was increased from 18 days to 35 days by overexpressing Spineless in the *Notch<sup>RNAi</sup>* tumour model, which nearly matched the median survival of control flies (38 days). In the *Notch<sup>RNAi</sup>* model increased numbers of *esg>GFP* positive cells or clonal tumours could already be seen by day 5 (Fig 3F). By day 19, the *esg>GFP* positive tumours took over most of the midgut in surviving flies. Spineless overexpressing flies showed no signs of tumour development at day 5 and at most a slight increase in the number of *esg>GFP* positive cells by day 19. Spineless overexpression also increased survival when the *Notch<sup>RNAi</sup>* tumour model was combined with low-dose *P. entomophila* infection (Fig. 3G). Thus, Spineless can block tumour development in the *Drosophila* midgut, likely by inhibiting ISC proliferation.

### Spineless delays tumour formation in the *yki<sup>act</sup>* tumour model

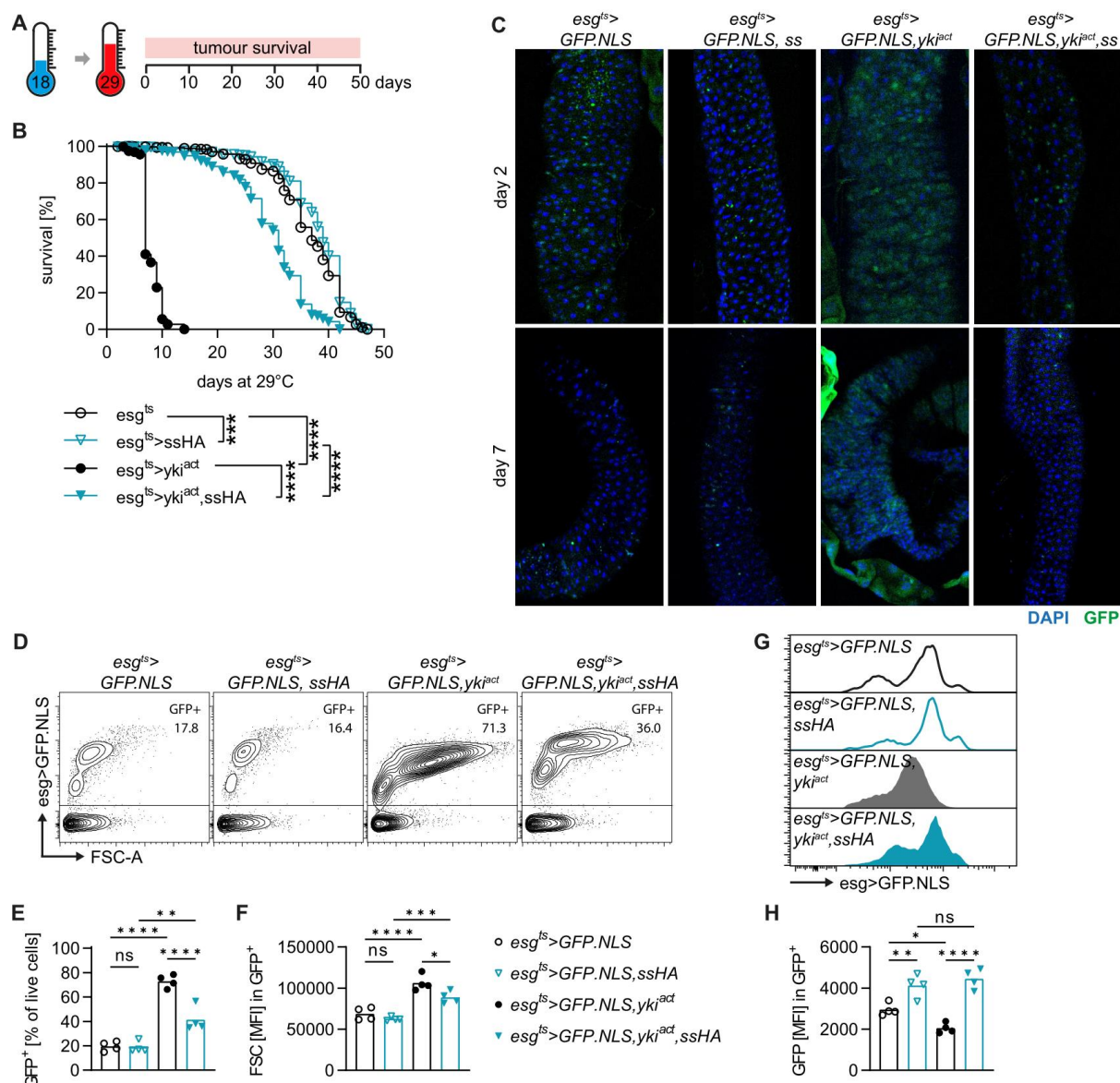
We used a second tumour model to confirm the effect of Spineless. The transcriptional coactivator yorkie can regulate ISC proliferation during midgut epithelial regeneration<sup>31,37</sup>. Mutation of 3 serine phosphorylation sites to alanine leads to a constitutively active form of yorkie (*yki<sup>act</sup>*) that is no longer subject to control by the Hippo pathway and leads to the formation of intestinal tumours<sup>38,39</sup>. Following temperature shift to 29°C, *esg<sup>ts</sup>>yki<sup>act</sup>* flies had a median survival of 7 days (Figure 4A, B). Spineless overexpression increased median survival to 31 days, while control flies without *yki<sup>act</sup>* had a median survival of 37-39 days. This suggests that in this tumour model Spineless overexpression can also significantly delay tumour onset. The expansion of tumour cells in *esg<sup>ts</sup>>yki<sup>act</sup>* flies was already visible after 2 days by microscopy but not visible in spineless overexpressing flies even by day 7 (Fig. 4C).

Next, we sought to use flow cytometry to profile the GFP<sup>+</sup> cells in the midgut at an early timepoint of *yki<sup>act</sup>* tumour development. GFP<sup>+</sup> cells were already visible 24h after temperature shift, but the distinction between ISC and EB populations was more apparent and the GFP intensity was higher after 48h (Figure S3A, B). We therefore chose the 48h timepoint. The expansion of tumour cells with EB-like fluorescent characteristics was clearly visible at this time (Fig. 4D). The frequency of GFP<sup>+</sup> cells increased more than 3-fold in *esg<sup>ts</sup>>yki<sup>act</sup>* flies and was suppressed by spineless overexpression, although not to baseline levels (Fig. 4E). Tumour cells showed an increased size (FSC) compared to GFP<sup>+</sup> cells from tumour-free flies, which was partially rescued by spineless overexpression (Fig. 4F) and decreased intensity of GFP expression which was fully rescued by spineless overexpression (Fig. 4G, H). Together, these findings demonstrate that Spineless can strongly suppress or delay tumour formation in two independent models.



**Figure 3: *Spineless* prevents tumour formation in the *Notch<sup>RNAi</sup>* tumour model**

A) Schematic of *Notch<sup>RNAi</sup>* tumour model. B) Survival of *spineless* knockdown and controls in the *Notch<sup>RNAi</sup>* tumour model. Data are pooled from 2 independent experiments, n=189-297 flies per genotype. C) Schematic of *Notch<sup>RNAi</sup>* tumour model with 24h low-dose *P. entomophila* infection. D) Survival of *spineless* knockdown and controls in the *Notch<sup>RNAi</sup>* tumour model with 24h low-dose *P. entomophila* infection. Data are from two experiments with n=166-265 flies per genotype. E) Survival of *spineless* overexpression and controls in the *Notch<sup>RNAi</sup>* tumour model. Data are pooled from 2 independent experiments, n=173-202 flies per genotype. F) Representative fluorescent images of controls and *spineless* overexpressing flies at different timepoints of the *Notch<sup>RNAi</sup>* tumour model. G) Survival of *spineless* overexpression and controls in the *Notch<sup>RNAi</sup>* tumour model with 24h low-dose *P. entomophila* infection. Data are from two experiments with n=187-232 flies per genotype.



**Figure 4: *Spineless* delays tumour formation in the *yki<sup>act</sup>* tumour model**

A) Schematic of *yki<sup>act</sup>* tumour model. B) Survival of *spineless* overexpression and controls in the *yki<sup>act</sup>* tumour model. Data are pooled from 2 independent experiments, n=409-592 flies per genotype. C) Representative fluorescent images of controls and *spineless* overexpressing flies at different timepoints after induction of the *yki<sup>act</sup>* tumour model. D-H) Flow cytometric analysis of midguts from day 2 of tumour induction at 29°C. Data are from n=4 samples, each pooled of 24-30 midguts. D) Representative flow cytometry plots of GFP expressing cells as a percentage of live, single cells. E, F) Quantification of GFP<sup>+</sup> cells and cell size of GFP<sup>+</sup> cells. G) Representative flow cytometry plots of GFP intensity within GFP<sup>+</sup> cells. H) Quantification of GFP fluorescent intensity within GFP<sup>+</sup> cells.

### Spineless affects cell metabolism, proliferation, and differentiation pathways

To understand how *Spineless* blocks the development of tumours, we sorted ISC and EB populations from flies with or without *yki<sup>act</sup>* tumours and with or without *spineless* overexpression to analyse their transcriptome (Fig. S4A, B). We chose 48h post temperature shift to analyse the transition from normal to tumour cells and to be able to distinguish and isolate the ISC and EB populations by FACS. Given the differences in size and GFP intensity between the genotypes, the gates were adjusted for each genotype to best fit the populations (Fig. S4B). Principle component analysis indicated distinct clustering of ISC and EB populations (mostly along PC2) as well as genotypes (mostly along PC1) (Fig.



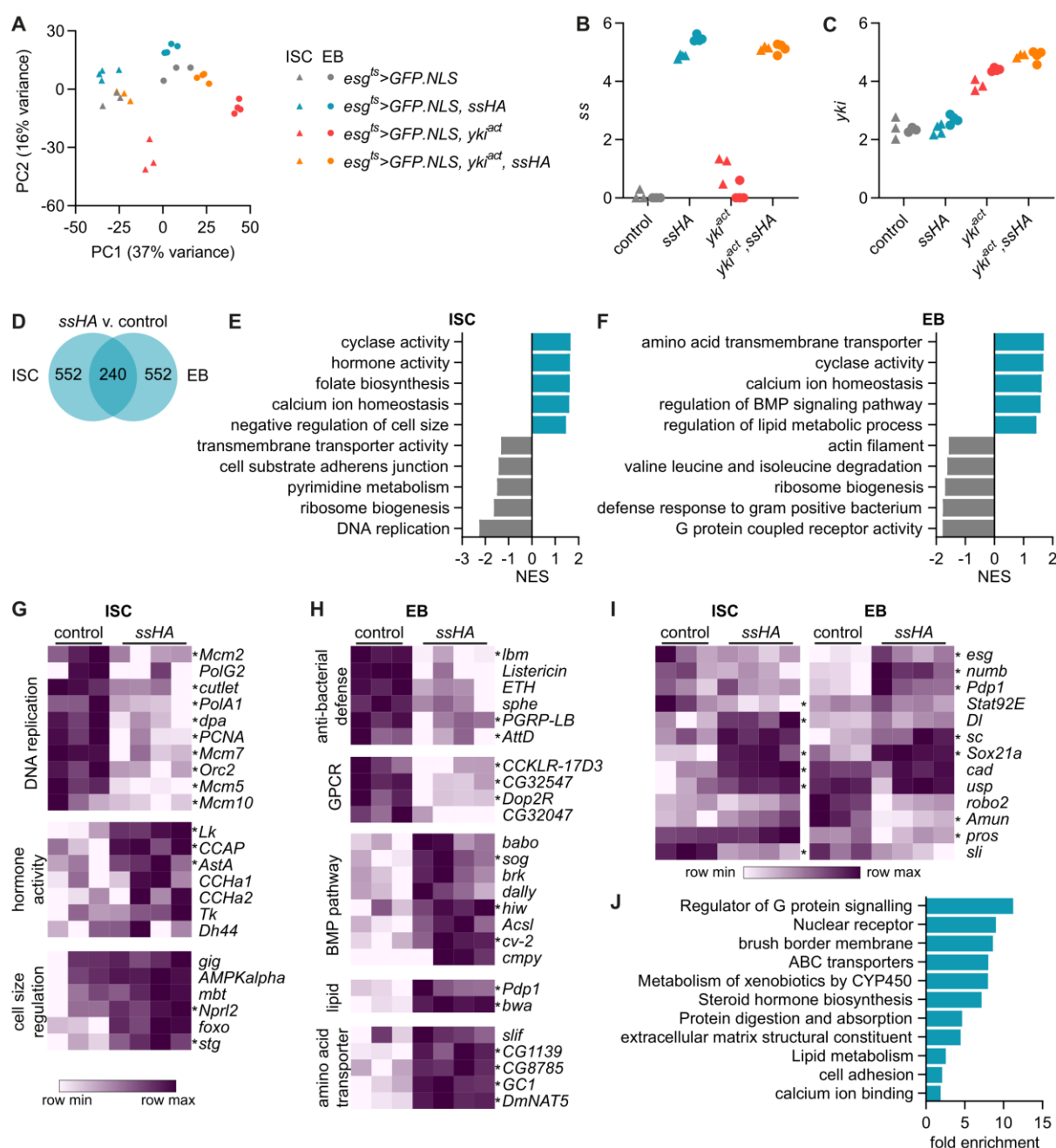
5A). Samples from Spineless overexpressing flies grouped away from controls in the opposite direction of *yki<sup>act</sup>* tumour cells and *yki<sup>act</sup>,ssHA* cells clustered closer to controls than to *yki<sup>act</sup>* tumour cells. This was also apparent in the number and overlap of differentially expressed genes across different comparisons in both the ISC and EB populations (Fig. S4C, D) and broadly reflects the ability of Spineless to suppress tumour growth as seen in the survival experiments. Importantly, *yki<sup>act</sup>* and *ssHA* expression were not reduced when expressed together as compared to flies that only overexpressed *yki<sup>act</sup>* or *ssHA* alone (Fig. 5B, C), confirming that the inhibition of tumour growth in *yki<sup>act</sup>,ssHA* flies is not merely the result of reduced *yki<sup>act</sup>* expression.

We first focused our attention on the gene expression changes driven by overexpression of Spineless. 792 genes were differentially expressed between Spineless overexpression and control samples in both the ISC and EB populations, with 240 genes in common (Fig. 5D, Fig. S4E). As expected, *spineless* was highly overexpressed in both cell types (Fig. S4F, G). Using gene set enrichment analysis, we found that Spineless reduced expression of genes relating to DNA replication in ISC (Fig. 5E), which correlates well with our earlier findings that Spineless suppressed cell proliferation after *P. entomophila* infection (Fig. 1). Pathways relating to DNA replication were only enriched in ISC, but not EB populations which do not proliferate (Fig. 5F). In ISC, Spineless increased expression of genes relating to hormone activity, including several hormones secreted by enteroendocrine cells (*AstA*, *Tk*, *CCh1*, *CCh2*)<sup>40</sup> (Fig. 4G). This suggests that Spineless overexpression may result in an increase in *esg>GFP<sup>+</sup>* enteroendocrine cells, a population that has previously been reported<sup>41</sup>. Spineless overexpression also increased expression of genes relating to negative regulation of cell size, such as *Nprl2*, *foxo* and *stg* (Fig. 5E, G). This is in line with our earlier findings that ISC from Spineless overexpressing flies were smaller on flow cytometry (Fig. 2D). Spineless overexpression increased expression of genes relating to calcium ion homeostasis and cyclase activity and reduced expression of genes relating to ribosome biogenesis in both ISC and EB (Fig. 4E, F).

In EB, Spineless overexpression reduced expression of anti-bacterial defence genes, GPCR activity and actin filament (Fig. 5F, H). Genes relating to lipid metabolism, amino acid transporters and the BMP pathway were increased. The BMP signalling pathway has been shown to antagonize the response to injury and return ISC to a quiescent state after injury-induced proliferation<sup>42-44</sup>. Genes relating to negative regulation of the BMP pathway were also enriched in Spineless overexpressing ISC, although the pathway did not reach significance (Fig. S4H, I). Inhibition of the BMP pathway by Spineless in ISC and EB may partially explain the lack of stem cell proliferation in response to infection we observed earlier.

In addition to annotated GO pathways, we also sought to directly analyse the expression of genes with known critical roles in the midgut (Fig. 5I). *Stat92E* is a key transcription factor downstream of JAK-STAT signalling. It is normally increased during midgut regeneration after injury to promote stem cell proliferation<sup>45</sup>. Spineless overexpressing ISC showed decreased expression of *Stat92E*, which may in part explain their reduced proliferative capacity. The mammalian genes *Cdx2* and *Rxra* are directly regulated by AHR in the intestinal epithelium<sup>5</sup>. Spineless increased the expression of their homologs *caudal* (*cad*) and *ultraspiracle* (*usp*) in ISC. The sequencing data confirmed increased *esg* expression in EB but not ISC from Spineless overexpressing flies, in line with our earlier flow cytometry and qPCR data (Fig. 2G-I). *Esg* has been reported to maintain stemness in ISC and EB<sup>33,34</sup>. In EB, Escargot also enhances Notch signalling by inhibiting Amun which may promote differentiation. Increased expression of the Notch ligand *Delta* (*Dl*) in ISC and the transcription factor *Sox21a* in ISC and EB could promote increased differentiation of EB to enterocytes<sup>22,46-48</sup> and the enteroendocrine marker *pros* was decreased in EB. However, many of the differentially expressed genes promote differentiation into enteroendocrine cells. *Scute* (*sc*) overexpression in ISC and EB leads to an increase in Pros<sup>+</sup> enteroendocrine cells<sup>49</sup>, *numb* facilitates enteroendocrine cell fate specification by limiting Notch signalling<sup>50</sup>, and *Pdp1* is a transcription factor in enteroendocrine cells with binding sites in the promoters of hormones<sup>51</sup>. Spineless increased expression of *sc*, *numb* and *Pdp1* in EB. *Sli* encodes Slit, a ligand for the receptor Robo2. The Slit/Robo2 pathway forms part of a negative feedback loop that limits commitment to the enteroendocrine lineage<sup>52</sup>. We found decreased expression of *sli* in ISC and

of *robo2* in EB from spineless overexpressing flies. Taken together, the RNA sequencing data suggest that spineless changes expression of a range of key factors involved in epithelial cell differentiation and may promote cell fate specification of enteroendocrine cells.



**Figure 5: Spineless alters cell metabolism, proliferation and differentiation in midgut progenitors**

A) Principle component analysis of sequenced ISC and EB samples. B, C) Expression of *ss* and *yki* in sequenced samples is depicted in  $\log_{10}(\text{raw counts} + 1)$ . D) Number of differentially expressed genes between *esg<sup>ts</sup>>GFP.NLS, ssHA* and *esg<sup>ts</sup>>GFP.NLS* samples ( $\text{padj} < 0.05$ ,  $|\text{FC}| > 2$ ). E, F) Gene set enrichment analysis comparing E) ISC and F) EB from *esg<sup>ts</sup>>GFP.NLS, ssHA* to *esg<sup>ts</sup>>GFP.NLS* samples. Selected pathways are shown, the full list is in Table S1. G, H) Examples of genes from the leading edge of enriched pathways shown in E) and F). I) Key genes involved in midgut stem cell maintenance and differentiation. G-I) Genes with significant differential expression ( $\text{padj} < 0.05$ ,  $|\text{FC}| > 2$ ) are denoted by \*. J) Gene ontology analysis using DAVID of AHR-regulated mouse genes with homology to fly genes with differential expression between *esg<sup>ts</sup>>GFP.NLS, ssHA* and *esg<sup>ts</sup>>GFP.NLS* samples. Selected pathways are shown, the full list is in Table S3. Homologous genes are shown in Figure S5 and Table S2.

### Spineless and AHR regulate common target genes in the intestinal epithelium

To determine if AHR and Spineless had evolutionarily conserved target genes in the intestinal epithelium, we generated a list of homologous genes that were regulated by AHR in mouse epithelium and by Spineless in the *Drosophila* midgut. This yielded 213 mouse genes and 260 homologous fly genes (Fig. S5). We then used gene ontology analysis of the mouse genes to determine which pathways were regulated by AHR/Spineless. Target genes were enriched for pathways critical for mature epithelial cells, such as brush border maintenance, protein digestion and absorption and the extracellular matrix. Several enriched pathways were similar to those enriched in *ssHA* compared to control, such as lipid metabolism and calcium ion binding. These data suggest that AHR and Spineless control over 200 evolutionarily conserved target genes, many of which have key functions in the intestinal epithelium.

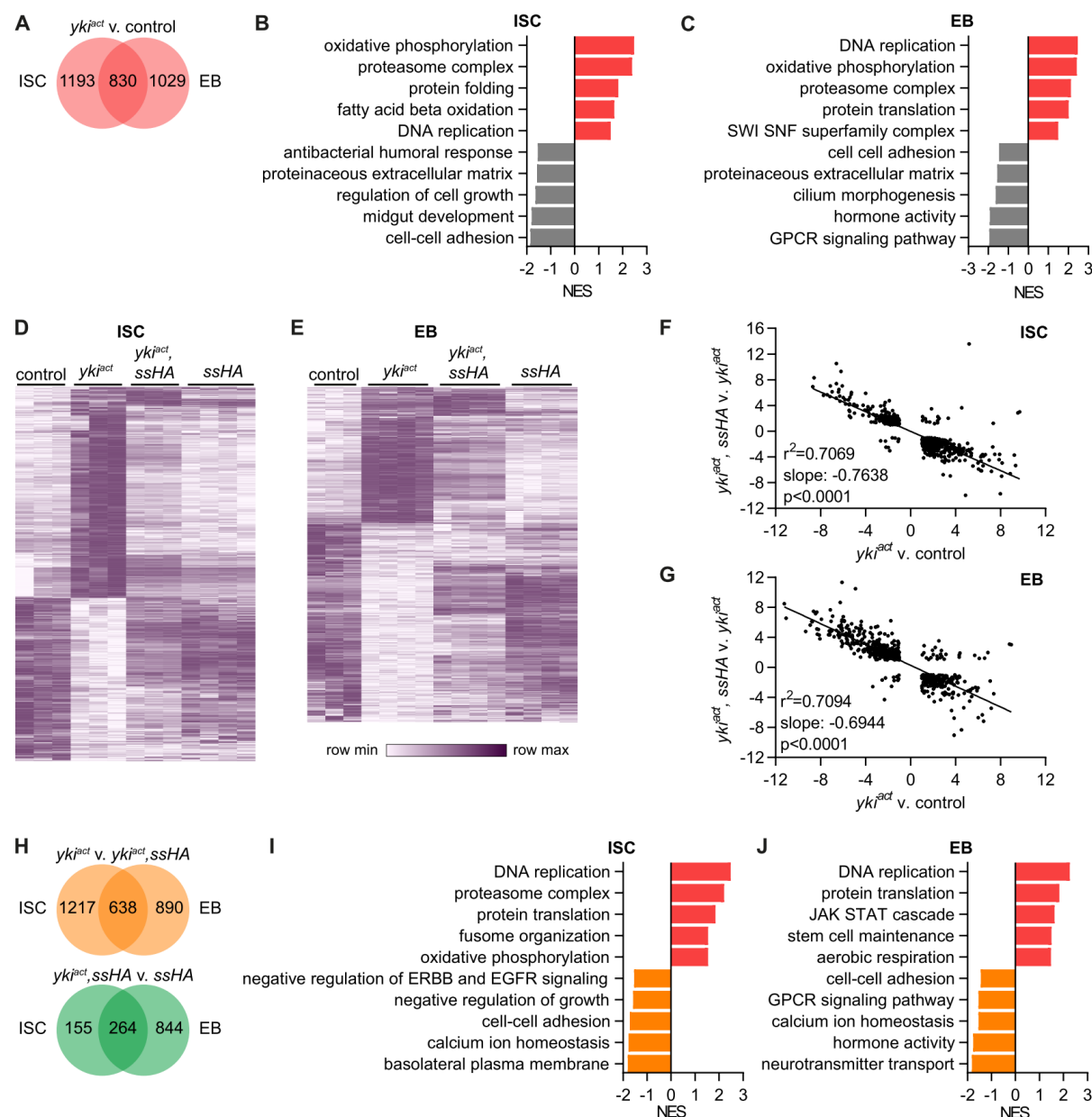
### *yki<sup>act</sup>* induces proliferation and changes cellular metabolism in the midgut

Next, we analysed the pathways that were differentially expressed in *yki<sup>act</sup>* tumour samples. Tumour cells clustered farthest from control cells on the PCA (Fig. 5A) and correspondingly had the highest number of differentially expressed genes, 2023 in ISC and 1859 in EB, 830 of which were in common (Fig. 6A). Tumour ISC were enriched for pathways relating to cell proliferation such as DNA replication and showed an altered metabolism with increased expression of oxidative phosphorylation and fatty acid beta oxidation pathways (Fig. 6B). Pathways relating to cell-cell adhesion, midgut development, regulation of cell growth, the extracellular matrix and antibacterial response were all downregulated in tumour ISC. In EB, tumour samples continued to upregulate pathways of cell proliferation and altered metabolism (Fig. 6C). They also upregulated genes in the SWI SNF superfamily complex such as *Iswi*, *Acf*, *HDAC1* and *Nurf-38*, which are involved in chromatin remodelling. Tumour EB downregulated pathways relating to cell-cell adhesion and the extracellular matrix similar to ISC and also reduced expression of genes involved in cilium morphogenesis, GPCR signalling and hormone activity. This shows that *yki<sup>act</sup>* tumour cells reduced expression of genes critical for the normal function of epithelial cells in exchange for genes driving proliferation and altered metabolism.

### Spineless reverses effects of *yki<sup>act</sup>* tumour on gene expression

From the PCA it was apparent that *yki<sup>act</sup>* tumour samples clustered separately from controls and concurrent spineless overexpression reversed this effect so that *yki<sup>act</sup>,ssHA* samples clustered closer to controls. This reversal was also visible on the level of individual genes. Hierarchical clustering of all genes with differential expression in *yki<sup>act</sup>* vs. control ISC showed that Spineless reversed the expression of most of those genes (Fig. 6D). This effect was less pronounced in EB (Fig. 6E). For genes that were differentially expressed in both comparisons (*yki<sup>act</sup>* v. control and *yki<sup>act</sup>,ssHA* vs. *yki<sup>act</sup>*), the effect of *yki<sup>act</sup>* was almost perfectly reversed by Spineless in ISC and EB (Figure 6F, G). This effect was also visible in the number of differentially expressed genes between *yki<sup>act</sup>,ssHA* vs. *yki<sup>act</sup>*, which were similar to the numbers between *yki<sup>act</sup>* tumours and controls (Fig. 6H). In contrast, there were only 419 differentially expressed genes in ISC between *yki<sup>act</sup>,ssHA* vs. *ssHA*, suggesting that the tumours had a limited effect on gene expression in the presence of Spineless. In EB, the difference was much larger with 1108 genes. On the level of pathways the comparison of *yki<sup>act</sup>* vs. *yki<sup>act</sup>,ssHA* was similar to that of *yki<sup>act</sup>* vs. controls (Fig. 6I, J). Concurrent Spineless overexpression in *yki<sup>act</sup>* tumours suppressed the cell proliferation and oxidative phosphorylation pathways and increased pathways relating to normal epithelial function such as cell-cell adhesion, regulation of growth, basolateral membrane, hormone activity and GPCR signalling. These results clearly show that Spineless can largely reverse the effects of *yki<sup>act</sup>* on gene expression and as a result restore normal cell function and metabolism pathways in midgut progenitors.

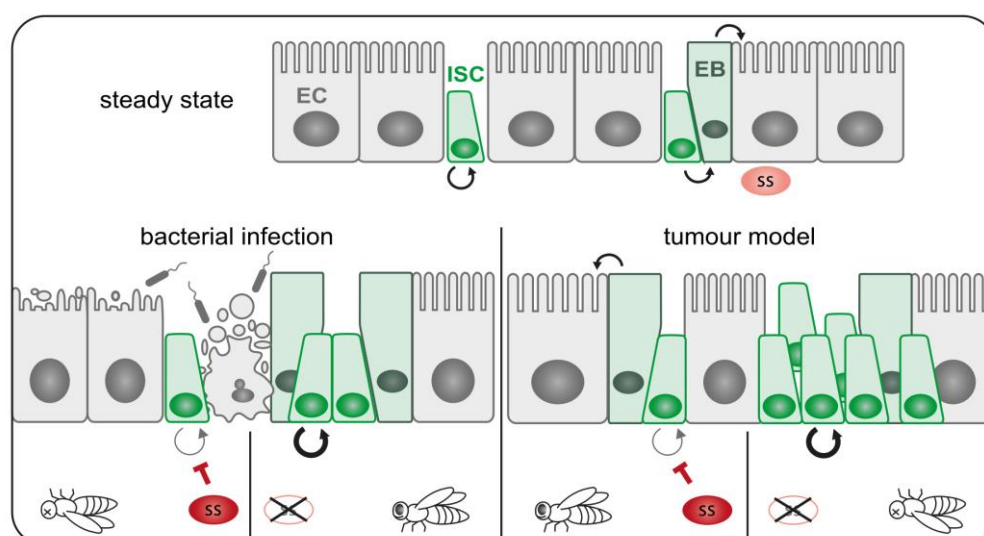
Taken together, our results demonstrate a critical role for Spineless in limiting stem cell proliferation, promoting epithelial cell differentiation, and acting as a tumour suppressor in the adult *Drosophila* intestine (Figure 7). This shows that Spineless has functional activity in the adult intestine and adds to a growing body of evidence that Spineless fulfils roles beyond development.



**Figure 6: Spineless reverses effects of *yki<sup>act</sup>* tumour on gene expression**

A) Number of differentially expressed genes between *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* and *esg<sup>ts</sup>>GFP.NLS* samples ( $\text{padj}<0.05$ ,  $|\text{FC}|>2$ ). B, C) Gene set enrichment analysis comparing B) ISC and C) EB from *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* to *esg<sup>ts</sup>>GFP.NLS* samples. Selected pathways are shown, the full list is in Table S1. D, E) Hierarchical clustering of genes differentially expressed between *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* and *esg<sup>ts</sup>>GFP.NLS* samples. D) 2023 genes with differential expression in ISC are shown. E) 1859 genes with differential expression in EB are shown. F, G) Simple linear regression analysis of F) 946 genes in ISC and G) 741 genes in EB that are differentially expressed in both comparisons of *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* to *esg<sup>ts</sup>>GFP.NLS* and of *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>,ssHA* to *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>*. H) Number of differentially expressed genes between *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* to *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>,ssHA* samples or between *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>,ssHA* to *esg<sup>ts</sup>>GFP.NLS,ssHA* samples ( $\text{padj}<0.05$ ,  $|\text{FC}|>2$ ). I, J) Gene set enrichment analysis comparing I) ISC and J) EB from *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>* to *esg<sup>ts</sup>>GFP.NLS,yki<sup>act</sup>,ssHA* samples. Selected pathways are shown, the full list is in Table S1.





**Figure 7: Schematic of Spineless function in the intestine**

At steady state ISC undergo self-renewal in symmetric divisions or asymmetric divisions into ISC and EB, which in turn give rise to mature enterocytes. Spineless is expressed at low levels during steady state. Bacterial infection leads to epithelial damage and stem cell proliferation to regenerate the epithelium. Spineless overexpression blocks ISC proliferation and reduces survival while spineless knockdown increases ISC proliferation. In tumour models, spineless overexpression blocks tumour growth and promotes differentiation thereby prolonging fly survival. Spineless knockdown accelerates tumour growth and decreases lifespan. EB, enteroblast; EC, enterocyte; ISC, intestinal stem cell; ss, spineless.

## Discussion

Our *Drosophila* study shows that Spineless regulates several conserved pathways in the intestine that are known to be regulated by AHR in mammals. The BMP gradient increases from crypt to villus and drives terminal differentiation of goblet cells and enterocytes in the mammalian intestine<sup>53</sup>. Genes in the BMP pathway were increased by Spineless overexpression in EB, suggesting that Spineless promotes differentiation of EB. Interactions between AHR and the BMP pathway have previously been suggested<sup>54,55</sup>, but have not been explored in the intestine. Spineless overexpression also increased expression of *cad* in ISC. Its homolog *Cdx2* is a direct target of AHR in mouse epithelial cells<sup>5</sup>. Both *cad* and *Cdx2* function in regionalization of the intestine<sup>51,56</sup>. *Cad* has also been shown to repress antimicrobial genes<sup>57</sup>, which could explain reduced expression of anti-bacterial defence genes following Spineless overexpression. *Cad* overexpression results in reduced ISC proliferation and epithelial regeneration after injury<sup>58</sup>, similar to what we observed when overexpressing Spineless during *P. entomophila* infection. Likewise, *usp*, the homolog of the direct AHR target *Rxra*<sup>59</sup> was also increased by Spineless. This suggests that AHR/Spineless have evolutionarily conserved target genes with critical functions in epithelial homeostasis.

This work identified novel tumour suppressing functions for Spineless. In mice, AHR was shown to suppress colorectal cancer development in an epithelial cell-intrinsic manner<sup>4</sup>. We found that Spineless knockdown resulted in accelerated death from *Notch*<sup>RNAi</sup> tumours and Spineless overexpression delayed tumour growth and drastically prolonged survival in two independent tumour models. This suggests that Spineless likely suppresses tumour growth on a fundamental level. Our results support two tumour-suppressing mechanisms for Spineless: the ability to restrain stem cell proliferation as observed following *P. entomophila* infection and to promote epithelial differentiation as highlighted by the transcriptomic data. Evidence for Spineless suppressing proliferation was also visible on the mRNA level, where it reversed *yki*<sup>act</sup>-induced pathways relating to DNA replication and cell division. Spineless largely reversed the effects of *yki*<sup>act</sup> on gene expression and thereby restored

normal cell function and metabolism in midgut progenitors. AHR was shown to directly affect transcription downstream of the Hippo pathway by restricting chromatin accessibility to Yap/Tea transcriptional targets<sup>5</sup> and to interact with chromatin remodelling complexes<sup>60</sup>. It is possible that Spineless similarly alters chromatin accessibility to repress *yki* target genes. Previous work found that Spineless binding is associated with chromatin opening during butterfly wing metamorphosis<sup>61</sup>. The role of AHR in cancer in general is still debated and its effect seems to vary between tumour cells and immune cells<sup>62,63</sup>. In the mammalian intestine, several studies find beneficial functions for AHR as a tumour suppressor<sup>1</sup>, suggesting that the *Drosophila* model system could be used to further study tumour suppressing functions of AHR/Spineless.

Others have shown that Spineless may affect movement, the oxidative stress response, or long-term memory formation in adult flies<sup>64,65</sup>. Our work adds to this growing body of evidence that Spineless fulfils roles beyond development in adult *Drosophila*. In a study by Sonowal and colleagues *Drosophila* healthspan was increased by indoles in a Spineless-dependent manner<sup>66</sup>. This result is surprising given that the ligand-binding domain of AHR is not conserved in Spineless and Spineless nuclear translocation is independent of dioxin<sup>6,11</sup>. In our study, Spineless knockdown or overexpression had no apparent effect on organismal survival in the absence of tumours.

Spineless function is most likely controlled on the transcriptional level but which factors control its expression or if there are additional protein complexes that can retain Spineless in the cytoplasm similar to AHR is not known. Ligand-dependent activation of AHR in vertebrates might offer an evolutionary advantage by allowing a rapid response and the integration of environmental signals. Despite these differences in activation, AHR and Spineless bind the same DNA sequence and we show here that they regulate evolutionarily conserved target genes and pathways critical for the maintenance of intestinal epithelial homeostasis.

## Material and Methods

### Fly Stocks and manipulation of midgut progenitor cells

The following transgenic lines were used: *rotund-Gal4*, *Su(H)-Gal4*, *Su(H)-Gal80*, *esg-Gal4*, *tub-Gal80<sup>ts</sup>*, *UAS-N<sup>RNAi</sup>* (#GD27228, Vienna *Drosophila* Resource Centre), *UAS-yki<sup>act</sup>* (*w<sup>\*</sup>*; *UAS-yki.S111A.S168A.S250A.V5*; #228817 Bloomington *Drosophila* Stock Center)<sup>38</sup>.

The following fly lines were generated in this study: *UAS-ssHA*, *ssGFP*, *UAS-anti-GFP*.

*Esg<sup>ts</sup>* refers to *tub-GAL80<sup>ts</sup>*, *esg-GAL4* which was used to express transgenes in midgut ISC and EB populations. We used the following driver to limit transgene expression to ISC (*esg-Gal4*, *Su(H)-Gal80*, *tub-Gal80<sup>ts</sup>*) or EB (*Su(H)-Gal4*, *tub-Gal80<sup>ts</sup>*) only. Drivers were crossed to *w<sup>1118</sup>* wildtype flies or *UAS-GFP* as control. To knockdown *spineless* in midgut progenitor cells, we used flies with *ssGFP* (homozygous), *tub-GAL80<sup>ts</sup>*, *esg-GAL4*, *UAS-anti-GFP*. Flies with *ssGFP* (homozygous), *tub-GAL80<sup>ts</sup>*, *esg-GAL4* without *UAS-anti-GFP* served as control.

Crosses were set up at 18°C to activate Gal80<sup>ts</sup>, thus restricting the expression of the Gal4-induced transgenes. Adult female offspring were selected at 0-4 days of age and shifted to 29°C to induce expression of transgenes. During incubation at 29°C, flies were transferred onto fresh food every 3-4 days.

### Genetic modification of flies

To generate flies expressing Ss fused to GFP (*ssGFP*), the *ss* locus was modified by CRISPR/Cas9-stimulated homologous recombination. DNA encoding eGFP followed by a lox-3Pax3-CHE-lox cassette was inserted just before the stop codon as described<sup>67</sup>. To generate *UAS-anti-GFP*, we fused the CD8 ORF (NCBI Ref.: NP\_001074579) to GBP (vhhGFP4)<sup>68</sup> separated by a Gly/Ser linker and inserted this new ORF in pUAST. To obtain flies allowing over-expression of HA tagged version of Ss (*UAS-ssHA*), we

inserted 2 HA tags (before the stop codon) in the ss cDNA (isoform A)<sup>20</sup>. The resulting DNA was inserted in pUAST before p-element-mediated transformation.

### Bacterial infection

*Pseudomonas entomophila* (stock kindly provided by Bruno Lemaitre) was grown in LB medium at 29°C for 24h. Bacterial culture was centrifuged at 3000xg for 15 min and pellets resuspended in 5% sucrose solution for a final concentration of OD<sub>600</sub>=200. 1ml concentrated bacteria solution was added to filter paper placed on top of standard fly food for infection. Flies were shifted to 29°C for 5 days and starved in an empty vial for 2 hours prior to infection. Survival was recorded daily.

For tumour survival experiments with low-dose *P. entomophila* infection flies were first shifted to 29°C for 1 day, then infected with bacterial solutions concentrated to OD<sub>600</sub>=70 in 5% sucrose and returned to normal fly food 24h later.

### Tumour survival experiments

Crosses were set up at 18°C to activate GAL80ts, thus restricting the expression of the Gal4-induced transgenes. Adult female offspring were selected at 0-4 days of age and shifted to 29°C to induce expression of transgenes. Every 2-3 days during incubation at 29°C, flies were transferred onto fresh food and survival was recorded.

### Statistical analysis

Statistical analyses were performed using GraphPad Prism (version 9). Kaplan-Meier survival curves were plotted and analysed using log-rank (Mantel-Cox) test. A p value <0.05 was considered significant. The p values of multiple comparisons of survival curves were adjusted using the Bonferroni method. 2-way ANOVA was used to analyse grouped comparisons. All data points and “n” values reflect biological replicates (either from single or from pooled flies).

### Immunofluorescence staining and microscopy

*Drosophila* midguts were dissected and fixed with 4% (w/v) Formaldehyde (ThermoFisher) in PBS at room temperature for 30 minutes, permeabilised with PBS 0.2% Triton x-100 (PBST) at room temperature for 30 minutes and blocked with 10% (w/v) Bovine Serum Albumin (SigmaAldrich) in PBST (PBSA) at room temperature for 30 minutes. Primary antibody rabbit anti-Phospho-histone H3 (Ser10) (Cell signalling, #9701) was diluted 1:1000 in PBSA and added at 4°C overnight. Stained tissue was washed with PBS the next day and stained with secondary antibodies donkey anti-rabbit A555 (LifeTech, A21429) 1:1000 diluted in PBSA at room temperature for 3 hours followed by 1:10000 PBS-diluted DAPI staining (5mg/mL in H<sub>2</sub>O, SigmaAldrich, D9542) for 10 minutes before PBS washing. Ovaries and posterior abdomen were removed, and the remaining midguts were then mounted with antifade (ThermoFisher, P36934) on 21-well glass slides (1 gut/well). Images were acquired on a Zeiss LSM 710 confocal microscope and were further processed in ImageJ (FIJI, version 2.1.0). Proliferating cells were manually counted under Zeiss AxioImager M1 epifluorescence microscope using 20x objective. pHH3-positive cells were counted from 3-5 whole female midguts per experiment. Images within stacks were collected at 3-5µm z-interval, 5-7 images per stack were taken to cover the complete depth of samples acquired from R2 of the midgut<sup>69</sup>.

### Flow Cytometry

The cell isolation protocol was adapted from Dutta *et al*<sup>70</sup>. 96-well v-bottom plate was prepared with 40µL/well of digestion buffer on ice (30µL PBS, 10µL of 4mg/mL Elastase, 0.4µL of 5mg/mL DNase I). 2 midguts/well were digested at 27°C with shaking for 1h, followed by pipetting 20 times for mechanical separation. Cells were washed in wash buffer (PBS, 2mM EDTA, 0.2% BSA), centrifuged and incubated with Live/Dead nIR stain (ThermoFisher) at 4°C for 30 minutes, washed in PBS and fixed with 4% PFA at room temperature for 30 minutes. Samples were washed in PBS, resuspended in wash buffer with count bright beads (Invitrogen, C36950) to determine absolute cell numbers per midgut

and filtered through 40µm filters. Samples were acquired on a BD Fortessa instrument (BD Biosciences) and analysed using FlowJo v10 (TreeStar). Samples were gated on single cells using FSC-A/FSC-H and SSC-A/SSC-H and to exclude debris on FSC-A/SSC-A. Dead cells were excluded based on Live/Dead near-IR staining and autofluorescence (405nm laser, 450/50 filter) before gating on GFP<sup>+</sup> cells.

### Cell sorting

Midguts were dissected and digested as described above for flow cytometry. 10-30 midguts per sample were digested in 100-300µl digestion buffer. Live cells were sorted through a 70µm nozzle on a BD Fusion instrument (BD Biosciences) using BD Diva software. ISC and EB were gated according to Fig. S2B-D. Total live GFP-negative cells (mostly enterocytes) were sorted as a control.

### RNA isolation and qPCR

Entire midguts (2 midguts per sample) or sorted cells (2000-20,000) were vortexed in TriReagent and RNA was extracted according to the manufacturers protocol, using Glycogen to precipitate the RNA pellet. RNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (ThermoFisher). The cDNA served as a template for the amplification of genes of interest and housekeeping genes by real-time quantitative PCR, using TaqMan Gene Expression Assays (Applied Biosystems), universal PCR Master Mix (Applied Biosystems) and the QuantStudio 7 System (Applied Biosystems). The following primer/probes were used: *Ribosomal protein L32* (Dm02151827\_g1), *spineless* (Dm02134622\_m1), *tango* (Dm02373281\_s1), *DptB* (Dm01821557\_g1), *unpaired 3* (Dm01844142\_g1), *Dual oxidase* (Dm01800981\_g1), *NADPH oxidase* (Dm01826191\_g1), *Delta* (Dm02134951\_m1), *escargot* (Dm01841264\_s1). mRNA expression was determined using the  $\Delta C_T$  method by normalizing to *Rpl32* gene expression. To determine expression of *spineless* isoforms, genes were amplified using Power SYBR Green PCR Master Mix (ThermoFisher). Two sets of primers were used for the ss-A isoform: forward 1 (GCGAGGAGTTGGTTCCAATG), reverse 1 (ACTGCGAGTACTGCGTGTA), 242bp product and forward 2 (GCGAGGAGTTGGTTCCAATG), reverse 2 (CGGATGCGGATGATGGTACG), 268bp product. For isoform ss-C/D the following primers were used: forward 1 (GCGAGGAGTTGGTTCCAATG), reverse 1 (CTGCTGAAGCCGATCCATTC), 395bp product and forward 2 (GCGAGGAGTTGGTTCCAATG), reverse 2 (CAATCACCAGAGGAGCGGA), 456bp product. The primers for isoform ss-C/D also generate products of 686bp and 747bp, respectively for ss-A.

### RNA sequencing and data analysis

ISC and EB populations were sorted by FACS as described above. DAPI was used as an additional staining to exclude dead cells. Cells were gated as shown in Figure S4A, B. RNA from 5,000-30,000 cells was isolated using the Qiagen RNeasy Plus Micro Kit and eluted in 15µl water. RNA quality and concentration was analysed on a Bioanalyzer and only samples with RIN>7 were used for sequencing. NEBNext Low Input RNA libraries were prepared manually following the manufacturer's protocol (NEBNext® Single Cell/Low Input RNA Library Prep Kit for Illumina® Instruction Manual Version 5.0\_5/20, NEB #E6420L). Samples were normalized to 1ng total RNA material per library in 8µl of nuclease-free water. RNA samples underwent reverse transcription, and the resulting cDNA was amplified by 10 cycles of PCR (according to the manufacturer recommendation for 1 ng input DNA). Amplified cDNA was subjected to two consecutive bead clean-ups with a 0.6X and 0.9X ratio of SPRISelect beads [B23318; Beckman Coulter] to sample volume. cDNA was fragmented enzymatically to target an insert size of ~200bp. Adaptors (diluted to 0.6µM) were ligated to the cDNA fragments and adaptor-ligated samples were cleaned up with SPRISelect beads (ratio: 0.8x). For the amplification of the sequencing library, 25µl of Q5 Master Mix was added, plus 10µl of a unique index (NEBNext Multiplex Oligos for Illumina [NEB #E6609]). Libraries were amplified by 8 PCR cycles. Final libraries were cleaned up with SPRISelect beads (ratio: 0.9x). The quality of the purified libraries was assessed using an Agilent D1000 ScreenTape Kit on an Agilent 4200 TapeStation. Libraries were sequenced to a depth of at least 25M reads on an Illumina NovaSeq 6000 run in 101-8-8-101 configuration.



Sequencing runs were concatenated into single gzipped fastq files. These were then aligned to genome BDGP6 using nf-core/rnaseq 3.1<sup>71</sup>. The resulting counts file salmon.merged.gene\_counts.tsv was used to create a SummarizedExperiment (<https://bioconductor.org/packages/SummarizedExperiment>) which was then analysed using DESeq2<sup>72</sup> to produce tables of differentially expressed genes and rnk files for Gene set enrichment analysis (GSEA). PCA plots were made using the DESeq2 function varianceStabilizingTransformation. R computations were carried out using R version 4.2.3 (2023-03-15), "Shortstop Beagle". Datasets have been deposited to GEO under accession number GSE229388. Hierarchical clustering of genes was conducted with Morpheus (Broad Institute), using one minus pearson correlation with average linkage method. GSEA was conducted using GSEA 4.2.2 (Broad Institute) with standard settings. The gene sets were taken from gsean (<https://bioconductor.org/packages/release/bioc/html/gsean.html>) using the GO\_dme data set and these were translated into Ensembl gene ids using the biomaRt package<sup>73</sup>. Leading edge analysis was used to remove overlapping pathways and to identify the underlying genes. The full list of pathways is in Table S1.

We used the Homologous Gene Database (<https://ngdc.cncb.ac.cn/hgd/>) to obtain a list of homologous proteins between *Drosophila* and mouse. This list was then filtered on genes regulated by AHR and Spineless. Fly genes with differential expression between *ssHA* and controls with  $|FC| > 2$  in either ISC or EB were retained. For mouse genes, we used previously published RNA sequencing data from wildtype and AHR knockout intestinal organoids (GSE133092)<sup>5</sup>. Mouse genes with differential expression between AHR knockout and controls with  $|FC| > 2$  in either stem cell or differentiated conditions were retained. This yielded a list of 213 mouse genes and 260 homologous fly genes that are regulated by AHR and Spineless (Table S2). Gene ontology of mouse genes was analysed using DAVID (<https://david.ncifcrf.gov/>) and is listed in Table S3.

## Author Contributions

M.T. and J.S. performed and analysed experiments. C.A. designed and created genetically modified flies and designed experiments. M.S. carried out computational analysis of RNA sequencing data. A.F. and Y. L. performed experiments. A.G, J.P.V. and B.S. supervised the project. N.L.D. designed, performed and analysed experiments, supervised the project and wrote the manuscript.

## Acknowledgments

This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC2016), the UK Medical Research Council (CC2016), and the Wellcome Trust (CC2016) and a Sir Henry Wellcome Fellowship 220497/Z/20/Z to N.L. Diny as well as a Wellcome Investigator Award 210556/Z/18/Z to B. Stockinger. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

We would like to acknowledge the Science technology platforms at the Francis Crick Institute. We thank the Fly Facility, especially Joachim Kurth for help with microinjections and fly maintenance, the Flow Cytometry Facility, the Advanced Sequencing Facility and the Light Microscopy Facility. We are thankful to Robert Johnston for providing the *ss* cDNA plasmid, to Nalle Pentimikko for discussion and reagents and to Oscar Diaz, Anke Liebert, and Murali Rao Maradana for critical reading and discussion of the manuscript.

## Declaration of interests

The authors declare no competing interests.

# References

1. Stockinger B, Shah K, Wincent E. AHR in the intestinal microenvironment: safeguarding barrier function. *Nat Rev Gastroenterol Hepatol* 2021;18:559-70.
2. McGuire J, Okamoto K, Whitelaw ML, Tanaka H, Poellinger L. Definition of a dioxin receptor mutant that is a constitutive activator of transcription: delineation of overlapping repression and ligand binding functions within the PAS domain. *J Biol Chem* 2001;276:41841-9.
3. Diny NL, Schonfeldova B, Shapiro M, Winder ML, Varsani-Brown S, Stockinger B. The aryl hydrocarbon receptor contributes to tissue adaptation of intestinal eosinophils in mice. *J Exp Med* 2022;219.
4. Metidji A, Omenetti S, Crotta S, et al. The Environmental Sensor AHR Protects from Inflammatory Damage by Maintaining Intestinal Stem Cell Homeostasis and Barrier Integrity. *Immunity* 2018;49:353-62 e5.
5. Shah K, Maradana MR, Joaquina Delas M, et al. Cell-intrinsic Aryl Hydrocarbon Receptor signalling is required for the resolution of injury-induced colonic stem cells. *Nat Commun* 2022;13:1827.
6. McMillan BJ, Bradfield CA. The aryl hydrocarbon receptor sans xenobiotics: endogenous function in genetic model systems. *Mol Pharmacol* 2007;72:487-98.
7. Emmons RB, Duncan D, Estes PA, et al. The spineless-aristapedia and tango bHLH-PAS proteins interact to control antennal and tarsal development in *Drosophila*. *Development* 1999;126:3937-45.
8. Duncan DM, Burgess EA, Duncan I. Control of distal antennal identity and tarsal development in *Drosophila* by spineless-aristapedia, a homolog of the mammalian dioxin receptor. *Genes Dev* 1998;12:1290-303.
9. Kudo K, Takeuchi T, Murakami Y, Ebina M, Kikuchi H. Characterization of the region of the aryl hydrocarbon receptor required for ligand dependency of transactivation using chimeric receptor between *Drosophila* and *Mus musculus*. *Biochim Biophys Acta* 2009;1789:477-86.
10. Butler RA, Kelley ML, Powell WH, Hahn ME, Van Beneden RJ. An aryl hydrocarbon receptor (AHR) homologue from the soft-shell clam, *Mya arenaria*: evidence that invertebrate AHR homologues lack 2,3,7,8-tetrachlorodibenzo-p-dioxin and beta-naphthoflavone binding. *Gene* 2001;278:223-34.
11. Cespedes MA, Galindo MI, Couso JP. Dioxin toxicity in vivo results from an increase in the dioxin-independent transcriptional activity of the aryl hydrocarbon receptor. *PLoS One* 2010;5:e15382.
12. Struhl G. Spineless-aristapedia: a homeotic gene that does not control the development of specific compartments in *Drosophila*. *Genetics* 1982;102:737-49.
13. Kozu S, Tajiri R, Tsuji T, Michiue T, Saigo K, Kojima T. Temporal regulation of late expression of Bar homeobox genes during *Drosophila* leg development by Spineless, a homolog of the mammalian dioxin receptor. *Dev Biol* 2006;294:497-508.
14. Emmons RB, Duncan D, Duncan I. Regulation of the *Drosophila* distal antennal determinant spineless. *Dev Biol* 2007;302:412-26.
15. Kim MD, Jan LY, Jan YN. The bHLH-PAS protein Spineless is necessary for the diversification of dendrite morphology of *Drosophila* dendritic arborization neurons. *Genes Dev* 2006;20:2806-19.
16. Tsubota T, Saigo K, Kojima T. Hox genes regulate the same character by different strategies in each segment. *Mech Dev* 2008;125:894-905.
17. Voortman L, Anderson C, Urban E, et al. Temporally dynamic antagonism between transcription and chromatin compaction controls stochastic photoreceptor specification in flies. *Dev Cell* 2022;57:1817-32 e5.
18. Miller AC, Urban EA, Lyons EL, Herman TG, Johnston RJ, Jr. Interdependent regulation of stereotyped and stochastic photoreceptor fates in the fly eye. *Dev Biol* 2021;471:89-96.
19. Anderson C, Reiss I, Zhou C, et al. Natural variation in stochastic photoreceptor specification and color preference in *Drosophila*. *Elife* 2017;6.
20. Johnston RJ, Jr., Desplan C. Interchromosomal communication coordinates intrinsically stochastic expression between alleles. *Science* 2014;343:661-5.

21. Wernet MF, Mazzone EO, Celik A, Duncan DM, Duncan I, Desplan C. Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature* 2006;440:174-80.
22. Miguel-Aliaga I, Jasper H, Lemaitre B. Anatomy and Physiology of the Digestive Tract of *Drosophila melanogaster*. *Genetics* 2018;210:357-96.
23. Micchelli CA, Perrimon N. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* 2006;439:475-9.
24. Ohlstein B, Spradling A. The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* 2006;439:470-4.
25. Liu Z, Li L, Chen W, et al. Aryl hydrocarbon receptor activation maintained the intestinal epithelial barrier function through Notch1 dependent signaling pathway. *Int J Mol Med* 2018;41:1560-72.
26. Wisniewski PJ, Nagarkatti M, Nagarkatti PS. Regulation of Intestinal Stem Cell Stemness by the Aryl Hydrocarbon Receptor and Its Ligands. *Front Immunol* 2021;12:638725.
27. Dutta D, Dobson AJ, Houtz PL, et al. Regional Cell-Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult *Drosophila* Midgut. *Cell Rep* 2015;12:346-58.
28. Vodovar N, Vinals M, Liehl P, et al. *Drosophila* host defense after oral infection by an entomopathogenic *Pseudomonas* species. *Proc Natl Acad Sci U S A* 2005;102:11414-9.
29. Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B. *Drosophila* intestinal response to bacterial infection: activation of host defense and stem cell proliferation. *Cell Host Microbe* 2009;5:200-11.
30. Jiang H, Patel PH, Kohlmaier A, Grenley MO, McEwen DG, Edgar BA. Cytokine/Jak/Stat signaling mediates regeneration and homeostasis in the *Drosophila* midgut. *Cell* 2009;137:1343-55.
31. Shaw RL, Kohlmaier A, Polesello C, Veelken C, Edgar BA, Tapon N. The Hippo pathway regulates intestinal stem cell proliferation during *Drosophila* adult midgut regeneration. *Development* 2010;137:4147-58.
32. Tauc HM, Tasdogan A, Pandur P. Isolating intestinal stem cells from adult *Drosophila* midguts by FACS to study stem cell behavior during aging. *J Vis Exp* 2014.
33. Loza-Coll MA, Southall TD, Sandall SL, Brand AH, Jones DL. Regulation of *Drosophila* intestinal stem cell maintenance and differentiation by the transcription factor Escargot. *EMBO J* 2014;33:2983-96.
34. Korzelius J, Naumann SK, Loza-Coll MA, et al. Escargot maintains stemness and suppresses differentiation in *Drosophila* intestinal stem cells. *EMBO J* 2014;33:2967-82.
35. Patel PH, Dutta D, Edgar BA. Niche appropriation by *Drosophila* intestinal stem cell tumours. *Nat Cell Biol* 2015;17:1182-92.
36. Apidianakis Y, Pitsouli C, Perrimon N, Rahme L. Synergy between bacterial infection and genetic predisposition in intestinal dysplasia. *Proc Natl Acad Sci U S A* 2009;106:20883-8.
37. Staley BK, Irvine KD. Warts and Yorkie mediate intestinal regeneration by influencing stem cell proliferation. *Curr Biol* 2010;20:1580-7.
38. Oh H, Irvine KD. In vivo analysis of Yorkie phosphorylation sites. *Oncogene* 2009;28:1916-27.
39. Kwon Y, Song W, Droujinine IA, Hu Y, Asara JM, Perrimon N. Systemic organ wasting induced by localized expression of the secreted insulin/IGF antagonist ImpL2. *Dev Cell* 2015;33:36-46.
40. Chen J, Kim SM, Kwon JY. A Systematic Analysis of *Drosophila* Regulatory Peptide Expression in Enteroendocrine Cells. *Mol Cells* 2016;39:358-66.
41. Guo X, Yin C, Yang F, et al. The Cellular Diversity and Transcription Factor Code of *Drosophila* Enteroendocrine Cells. *Cell Rep* 2019;29:4172-85 e5.
42. Guo Z, Driver I, Ohlstein B. Injury-induced BMP signaling negatively regulates *Drosophila* midgut homeostasis. *J Cell Biol* 2013;201:945-61.
43. Tian A, Wang B, Jiang J. Injury-stimulated and self-restrained BMP signaling dynamically regulates stem cell pool size during *Drosophila* midgut regeneration. *Proc Natl Acad Sci U S A* 2017;114:E2699-E708.

44. Tracy Cai X, Li H, Safyan A, Gawlik J, Pyrowolakis G, Jasper H. AWD regulates timed activation of BMP signaling in intestinal stem cells to maintain tissue homeostasis. *Nat Commun* 2019;10:2988.
45. Bausek N. JAK-STAT signaling in stem cells and their niches in *Drosophila*. *JAKSTAT* 2013;2:e25686.
46. Ohlstein B, Spradling A. Multipotent *Drosophila* intestinal stem cells specify daughter cell fates by differential notch signaling. *Science* 2007;315:988-92.
47. Zhai Z, Kondo S, Ha N, et al. Accumulation of differentiating intestinal stem cell progenies drives tumorigenesis. *Nat Commun* 2015;6:10219.
48. Meng FW, Biteau B. A Sox Transcription Factor Is a Critical Regulator of Adult Stem Cell Proliferation in the *Drosophila* Intestine. *Cell Rep* 2015;13:906-14.
49. Bardin AJ, Perdigoto CN, Southall TD, Brand AH, Schweisguth F. Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* 2010;137:705-14.
50. Salle J, Gervais L, Boumard B, Stefanutti M, Siudeja K, Bardin AJ. Intrinsic regulation of enteroendocrine fate by Numb. *EMBO J* 2017;36:1928-45.
51. Hung RJ, Hu Y, Kirchner R, et al. A cell atlas of the adult *Drosophila* midgut. *Proc Natl Acad Sci U S A* 2020;117:1514-23.
52. Biteau B, Jasper H. Slit/Robo signaling regulates cell fate decisions in the intestinal stem cell lineage of *Drosophila*. *Cell Rep* 2014;7:1867-75.
53. Beumer J, Puschhof J, Yengej FY, et al. BMP gradient along the intestinal villus axis controls zoned enterocyte and goblet cell states. *Cell Rep* 2022;38:110438.
54. Wang Q, Kurita H, Carreira V, et al. Ah Receptor Activation by Dioxin Disrupts Activin, BMP, and WNT Signals During the Early Differentiation of Mouse Embryonic Stem Cells and Inhibits Cardiomyocyte Functions. *Toxicol Sci* 2016;149:346-57.
55. Liu X, Li X, Tao Y, et al. TCDD inhibited the osteogenic differentiation of human fetal palatal mesenchymal cells through AhR and BMP-2/TGF-beta/Smad signaling. *Toxicology* 2020;431:152353.
56. Choi YJ, Hwang MS, Park JS, Bae SK, Kim YS, Yoo MA. Age-related upregulation of *Drosophila* caudal gene via NF-kappaB in the adult posterior midgut. *Biochim Biophys Acta* 2008;1780:1093-100.
57. Ryu JH, Kim SH, Lee HY, et al. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in *Drosophila*. *Science* 2008;319:777-82.
58. Wu K, Tang Y, Zhang Q, et al. Aging-related upregulation of the homeobox gene caudal represses intestinal stem cell differentiation in *Drosophila*. *PLoS Genet* 2021;17:e1009649.
59. Hall BL, Thummel CS. The RXR homolog ultraspiracle is an essential component of the *Drosophila* ecdysone receptor. *Development* 1998;125:4709-17.
60. Gialitakis M, Tolaini M, Li Y, et al. Activation of the Aryl Hydrocarbon Receptor Interferes with Early Embryonic Development. *Stem Cell Reports* 2017;9:1377-86.
61. van der Burg KRL, Lewis JJ, Martin A, Nijhout HF, Danko CG, Reed RD. Contrasting Roles of Transcription Factors Spineless and EcR in the Highly Dynamic Chromatin Landscape of Butterfly Wing Metamorphosis. *Cell Rep* 2019;27:1027-38 e3.
62. Murray IA, Patterson AD, Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat Rev Cancer* 2014;14:801-14.
63. Xue P, Fu J, Zhou Y. The Aryl Hydrocarbon Receptor and Tumor Immunity. *Front Immunol* 2018;9:286.
64. Williams EG, Mouchiroud L, Frochaux M, et al. An evolutionarily conserved role for the aryl hydrocarbon receptor in the regulation of movement. *PLoS Genet* 2014;10:e1004673.
65. Kuzin BA, Nikitina EA, Cherezov RO, et al. Combination of hypomorphic mutations of the *Drosophila* homologues of aryl hydrocarbon receptor and nucleosome assembly protein family genes disrupts morphogenesis, memory and detoxification. *PLoS One* 2014;9:e94975.
66. Sonowal R, Swimm A, Sahoo A, et al. Indoles from commensal bacteria extend healthspan. *Proc Natl Acad Sci U S A* 2017;114:E7506-E15.
67. Poernbacher I. CS, Kurth J., Nojima H., Baena-Lopez A., Alexandre C., Vincent J.P. Lessons in genome engineering: opportunities, tools and pitfalls. *bioRxiv* 2019.



68. Saerens D, Pellis M, Loris R, et al. Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies. *J Mol Biol* 2005;352:597-607.
69. Buchon N, Osman D. All for one and one for all: regionalization of the *Drosophila* intestine. *Insect biochemistry and molecular biology* 2015;67:2-8.
70. Dutta D, Xiang J, Edgar BA. RNA expression profiling from FACS-isolated cells of the *Drosophila* intestine. *Current protocols in stem cell biology* 2013;27:2F. .1-2F. .12.
71. Ewels PA, Peltzer A, Fillinger S, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* 2020;38:276-8.
72. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
73. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009;4:1184-91.