

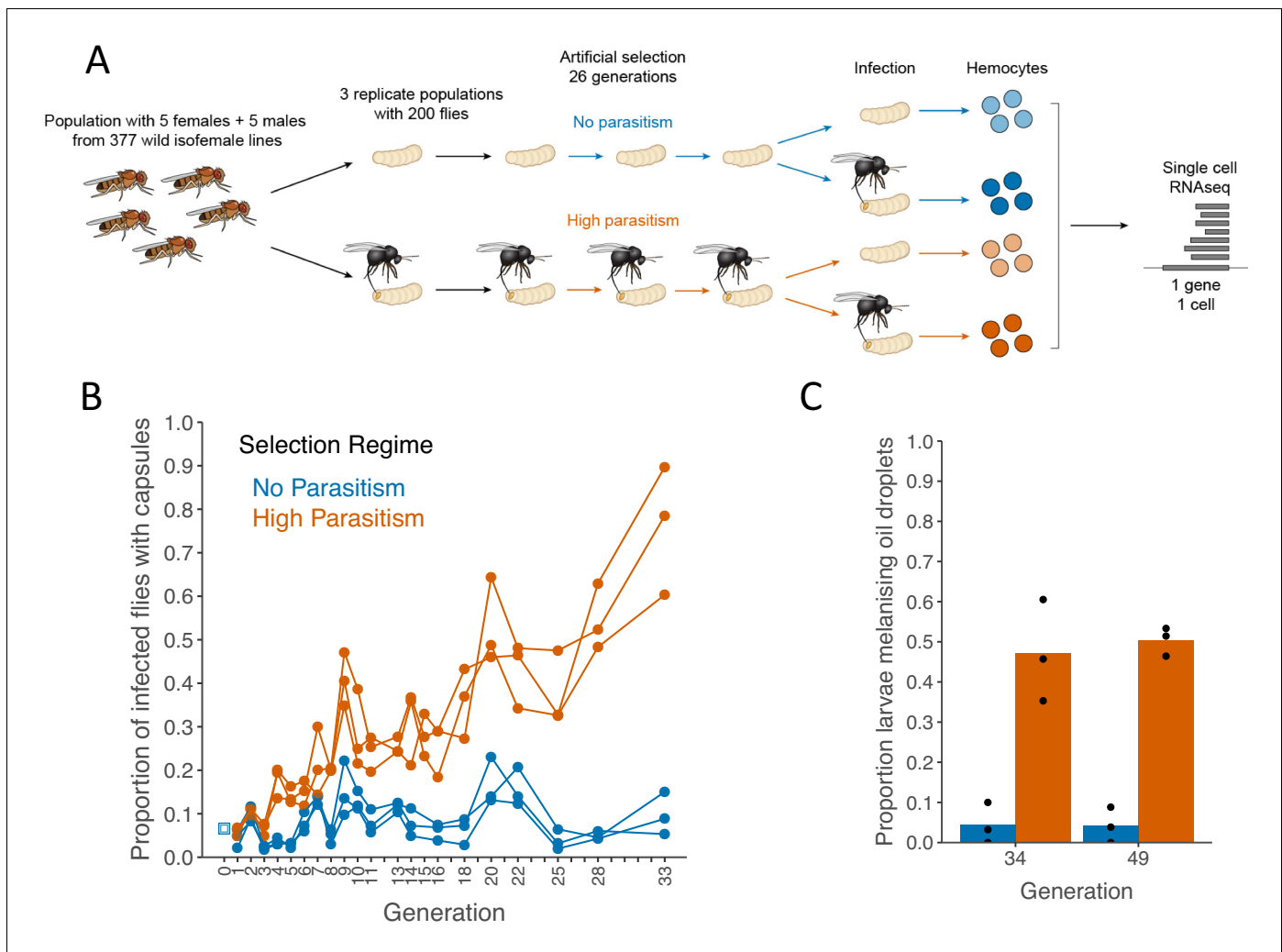


---

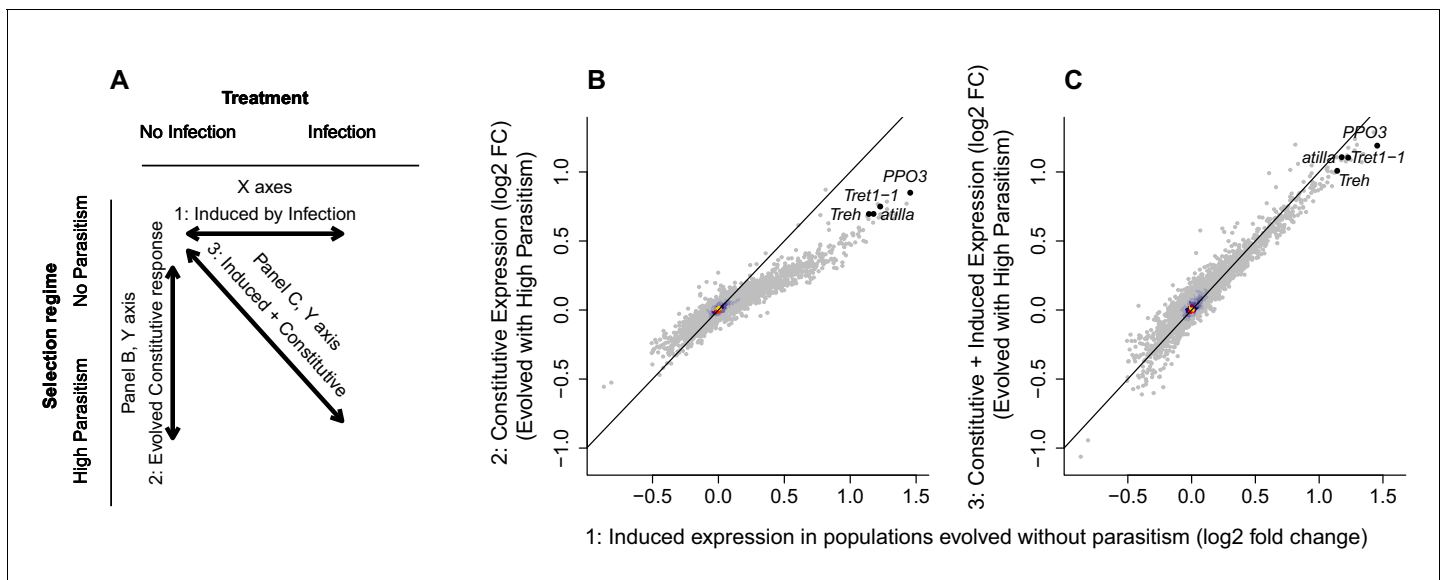
## Figures and figure supplements

Constitutive activation of cellular immunity underlies the evolution of resistance to infection in *Drosophila*

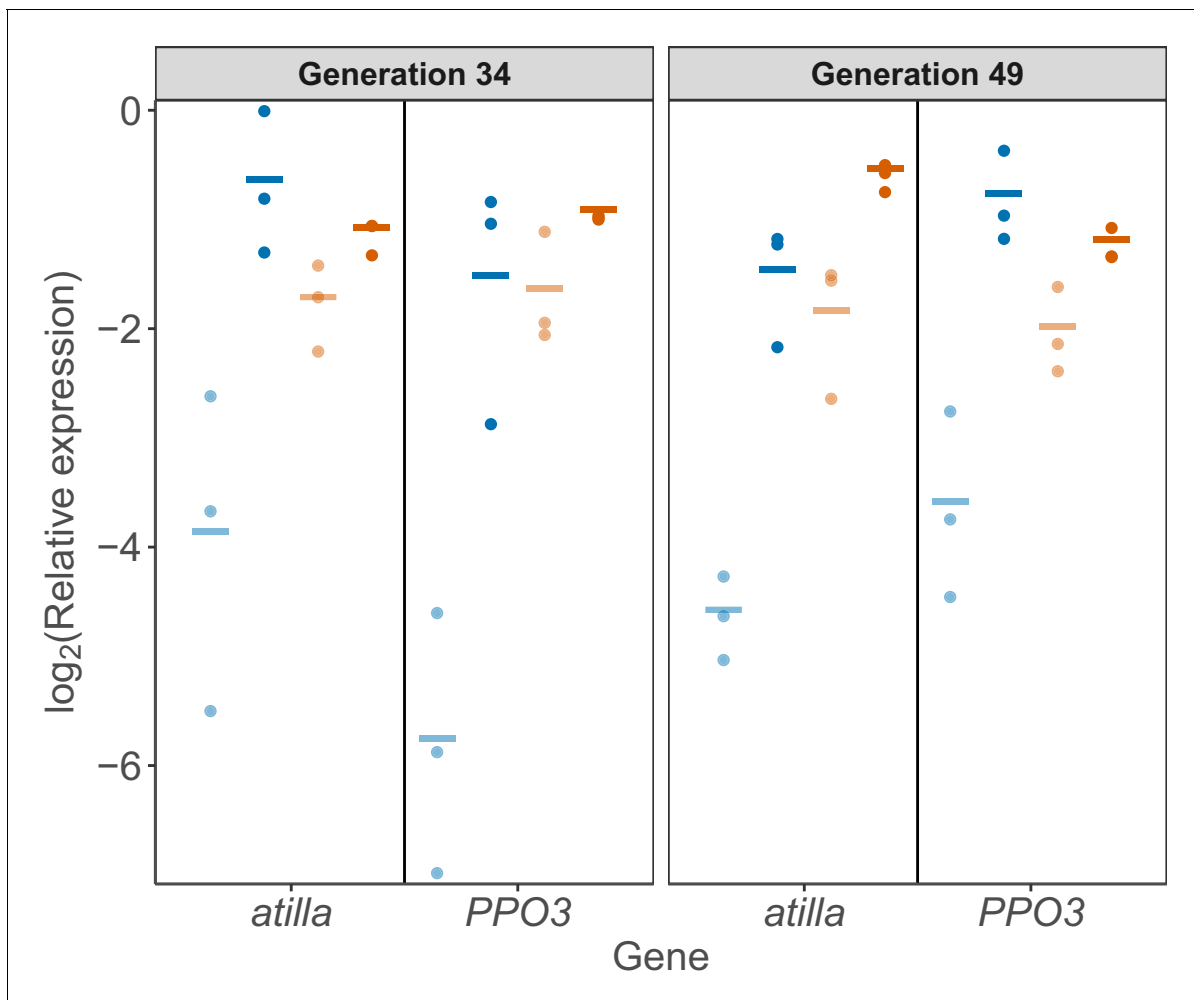
**Alexandre B Leitão et al**



**Figure 1.** Encapsulation rate during selection for parasitoid resistance. (A) Schematic of the experiment (B) The proportion of infected larvae that gave rise to adult flies with visible capsules. An outbred population of *D. melanogaster* (source, blue square) was used to create six populations that were parasitized every generation with *L. boulardi* (high parasitism, orange) or maintained without infection (no parasitism, blue). (C) Proportion of larvae encapsulating oil droplets at generations 33 and 49. Dots represent the proportion calculated from 15 to 40 injected larvae in triplicate populations from each selection regime. Bar heights represent the mean per selection regime.

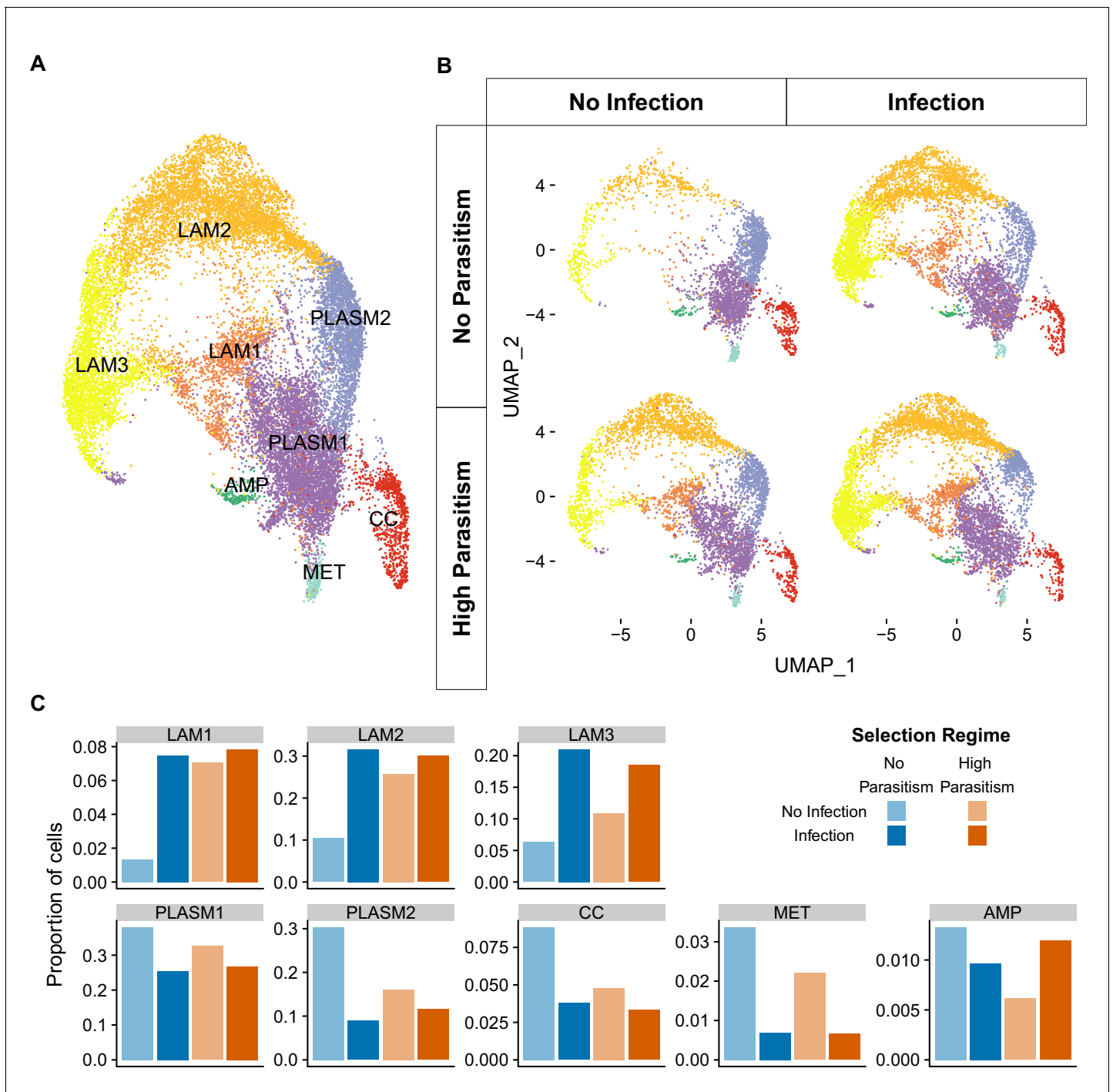


**Figure 2.** Changes in gene expression following selection for resistance and parasitoid infection. (A) Summary of the comparisons shown in the other panels. (B) The change in gene expression following infection in populations maintained without parasitoid infection (x axis) compared to the constitutive gene expression following 26 generations of selection for resistance to the parasitoid wasp *L. bouleari* (y axis). (C) The change in gene expression following infection in populations maintained without parasitoid infection (x axis) compared to the combined induced and constitutive change in gene expression following selection for resistance and infection (y axis). The black diagonals indicate the 1:1 line. The genes *atilla*, *PPO3*, *Trehalase* and *Tret1-1* are highlighted. Colour represents the density of overlain points. Relative expression is expressed as log<sub>2</sub>(fold change).

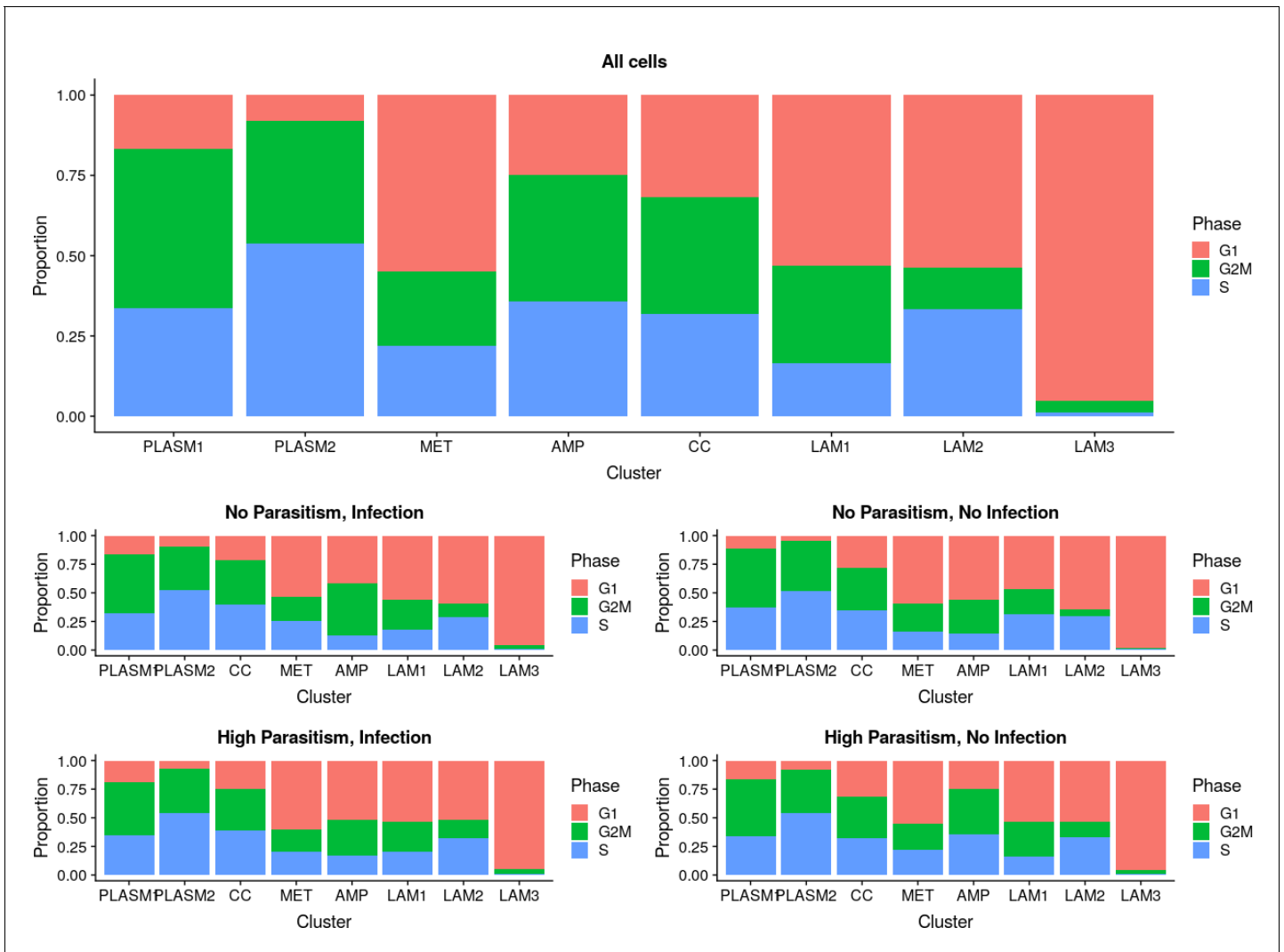


**Figure 2—figure supplement 1.** Gene expression in circulating hemocytes measured by qPCR. *atilla* and *PPO3* expression normalized by the housekeeping gene Rpl32 in populations evolved under no parasitism (blue) or high parasitism (orange). Hemocyte samples for RNA extraction were collected from 50 to 70 larvae 24-hr post infection (dark colours) and from larvae with no infection (light colours). Each dot represents the average  $\Delta Ct$  of two biological replicas per population and lines represent the average per selection regime. *atilla*, Selection regime x Treatment:  $\chi^2 = 15.45$ , d.f = 1,  $p = 8.49 \times 10^{-5}$ ; *PPO3*, Selection regime x Treatment:  $\chi^2 = 17.64$ , d.f = 1,  $p = 2.67 \times 10^{-5}$ .

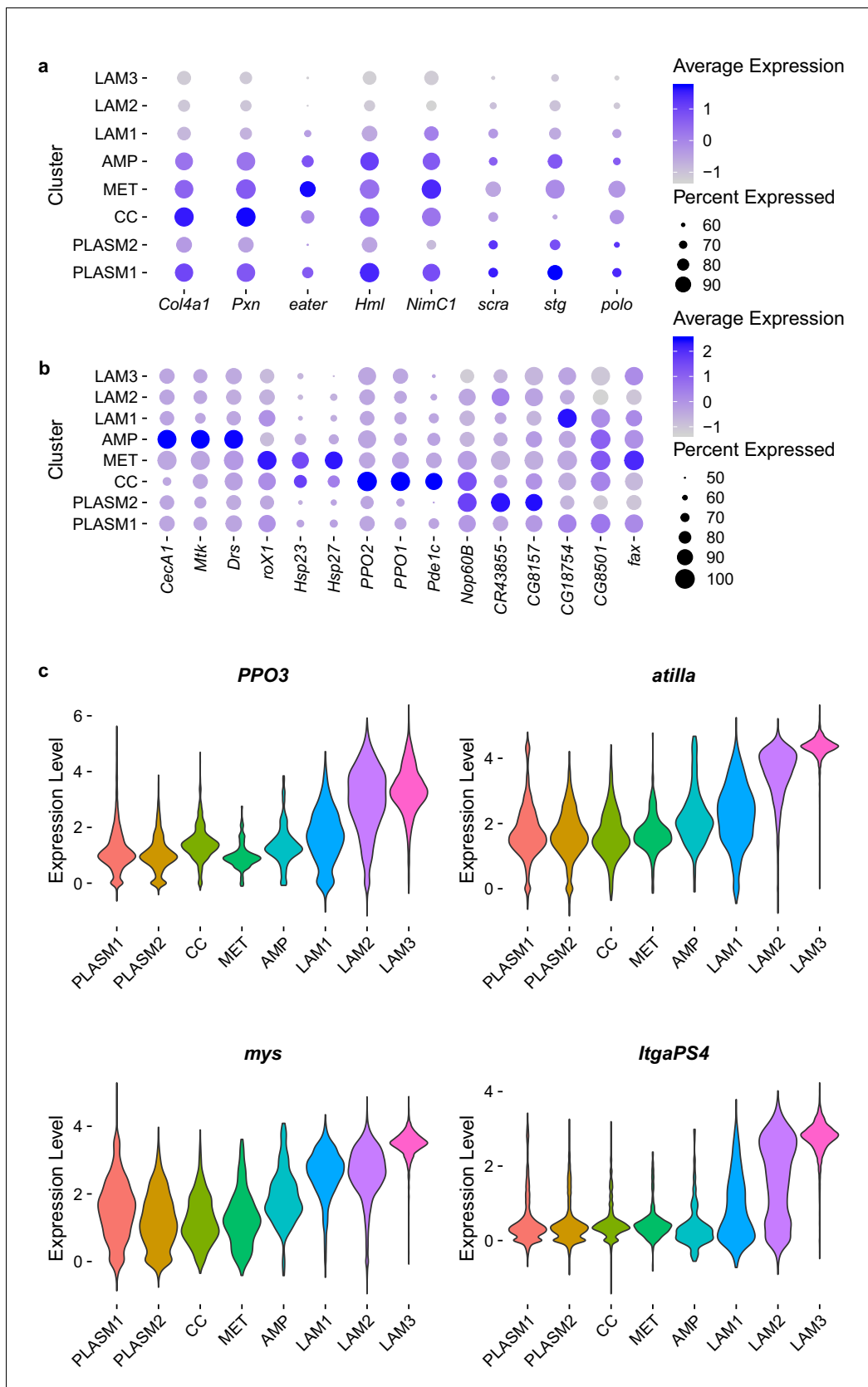




**Figure 3.** Cell states of *Drosophila melanogaster* larval hemocytes. Two-dimensional UMAP projections and cell state classification. (A) All 19,344 cells combined. PLASM: plasmatocytes; MET: metabolic plasmatocyte cluster; AMP: antimicrobial peptide plasmatocyte cluster; LAM: lamellocytes; CC: crystal cells. (B) Cells from populations that had undergone 26 generations of selection with no parasitism (top panels) or high parasitism (bottom panels), and were either uninfected (left panels) or 48 hr post infection (right panels). 64–333 million reads were sequenced in each of the eight 10X libraries that we generated (**Supplementary file 3**). The number of cells detected per library ranged from 787 to 3463 (**Figure 3—figure supplement 4**). (C) Proportion of different cell states after infection and selection.



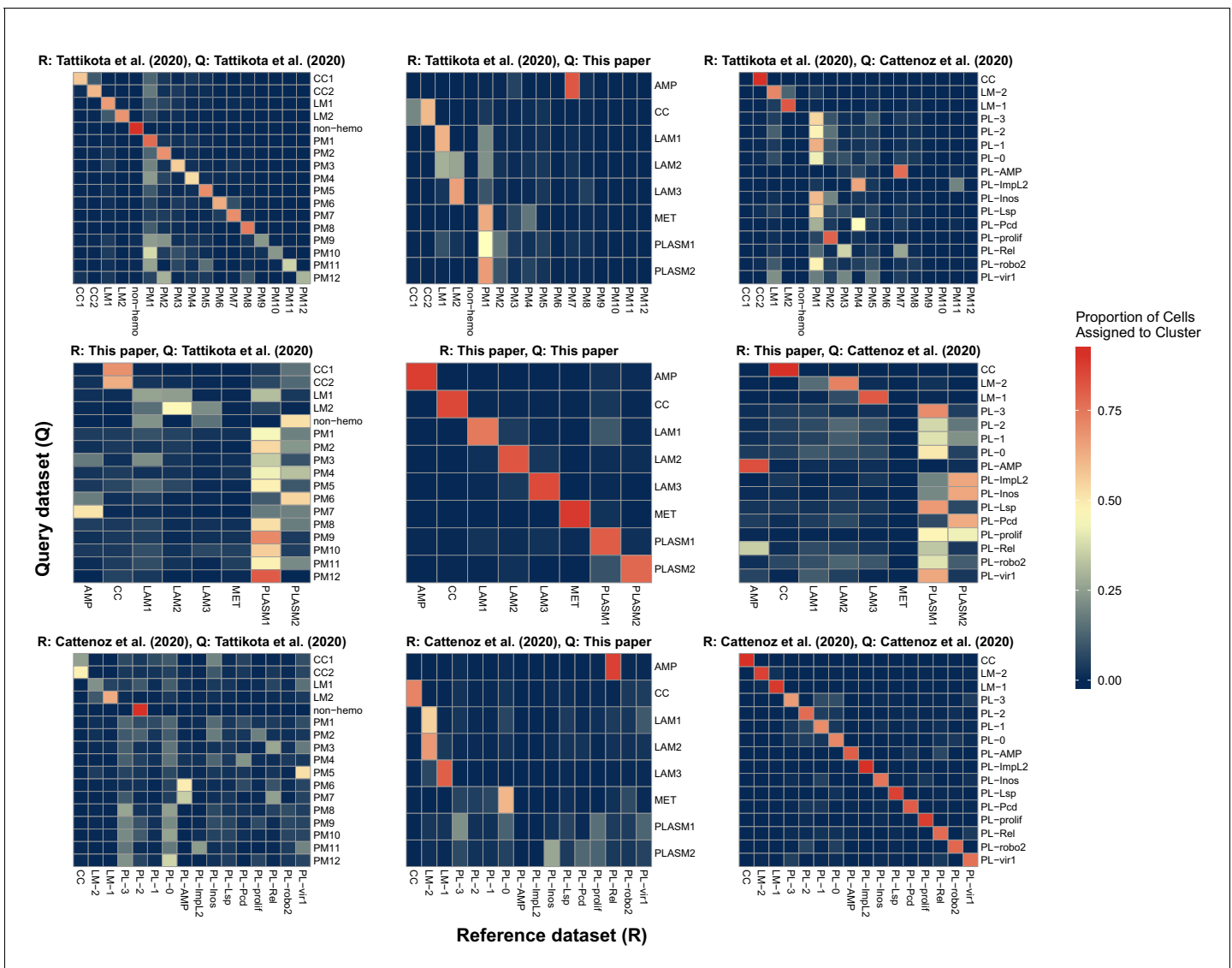
**Figure 3—figure supplement 1.** Proportions of cells in G1, G2M, and S phase in the plasmatocyte, lamellocyte, and crystal cell clusters.



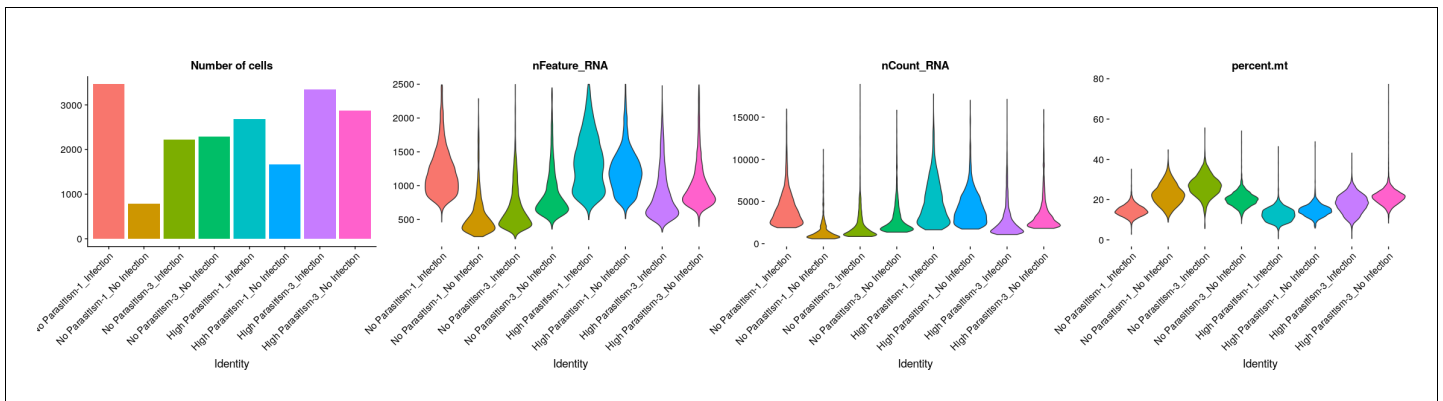
**Figure 3—figure supplement 2.** Expression levels ( $\log_e$ ) of marker genes. (A) Relative level of expression of described plasmacyte and cell cycle genes and (B) Relative level of expression most significant marker genes for the plasmacyte and crystal cell clusters. The MAST test statistic was used  
 Figure 3—figure supplement 2 continued on next page

Figure 3—figure supplement 2 continued

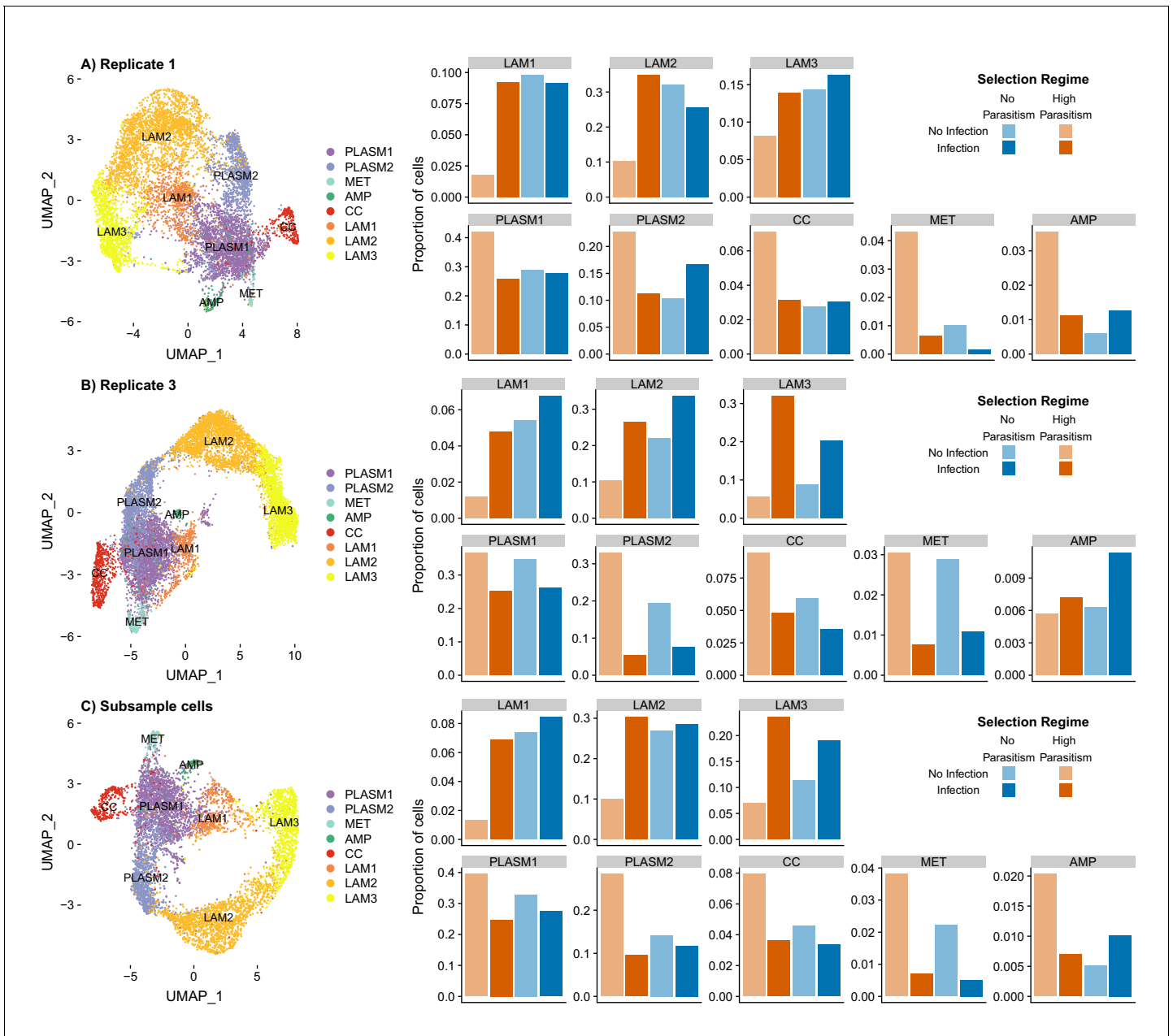
to identify significantly upregulated markers in each cluster comparing the cluster of interest with cells in all other clusters. Percent of cells expressing genes are indicated by circle size. (C) Distribution of gene expression levels for described lamellocyte genes grouped by cluster identities.



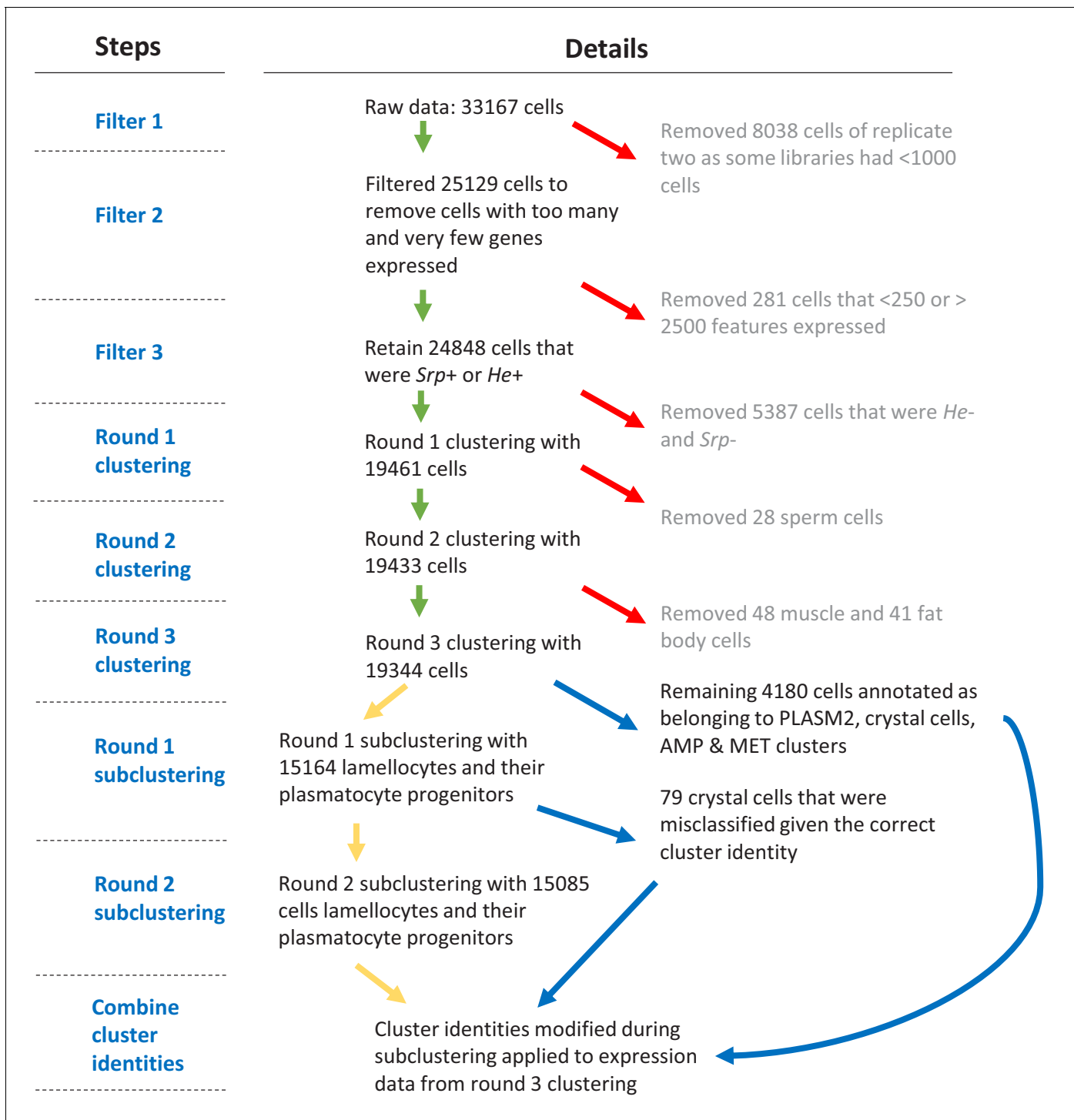
**Figure 3—figure supplement 3.** Comparison of scRNA-seq data sets from this paper to those from two previously published studies (*Cattenoz et al., 2020*; *Tattikota et al., 2020*). A reference-based label transfer was used to predict cell cluster identity in the query data set. The scale bar indicates the proportion of cells in query clusters (y axis) that are predicted to belong to a given reference cluster (x axis).



**Figure 3—figure supplement 4.** Number of cells and genes detected, unique molecular identifier count and percent mitochondrial gene expression per cell grouped by sample. HP: high parasitism, NP: No parasitism, I: Infection and NI: no infection. Numbers following HP or NP correspond to replicate.

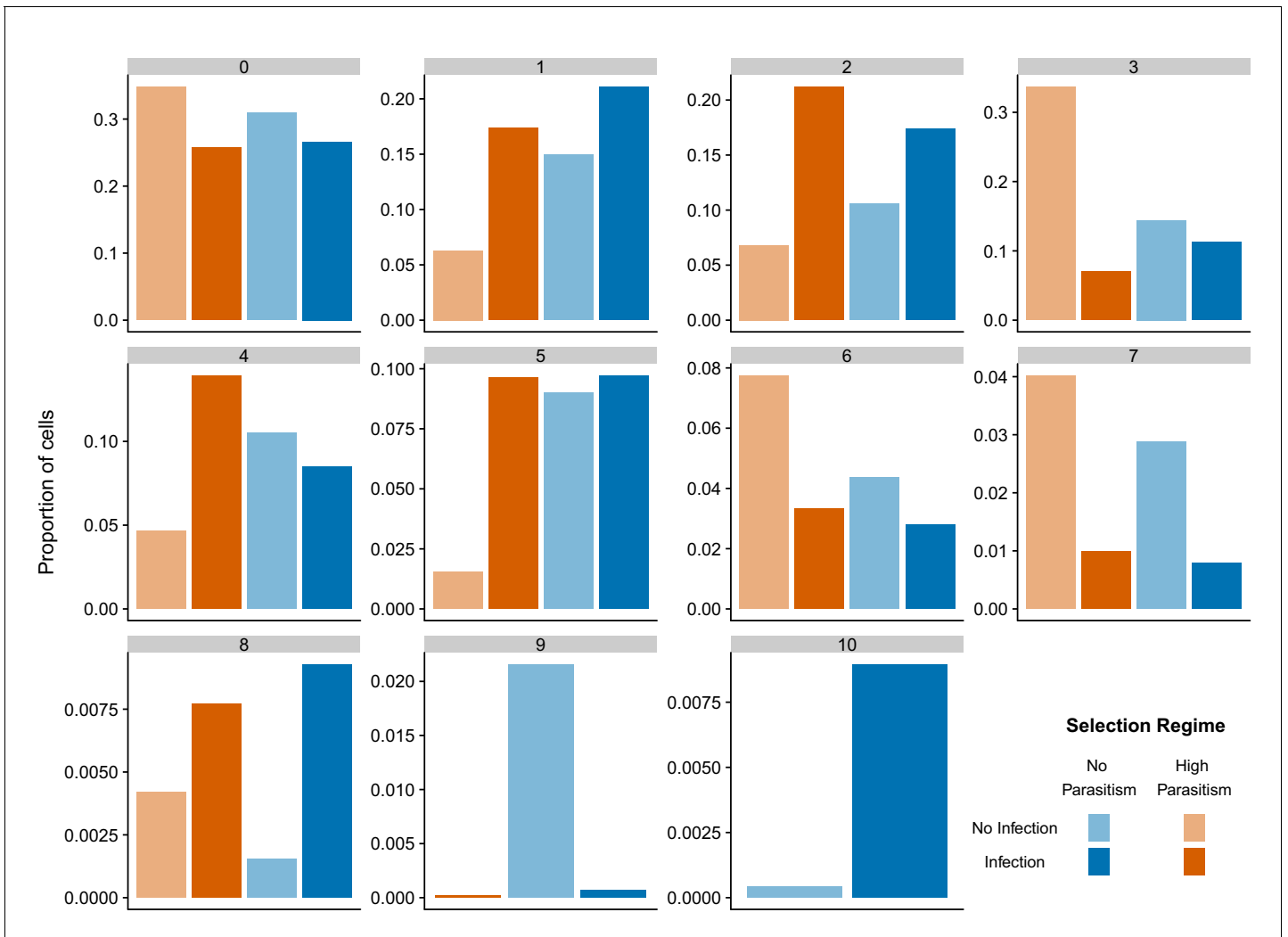


**Figure 3—figure supplement 5.** Impact of cell number and library-specific differences on data integration, cell cluster classification and estimating cluster proportions. (A) Clustering 8,596 cells from replicate one, (B) clustering 10,784 cells from replicate three and (C) clustering 6,296 cells where 787 cells were randomly subsampled without replacement in each of eight libraries used for data integration.

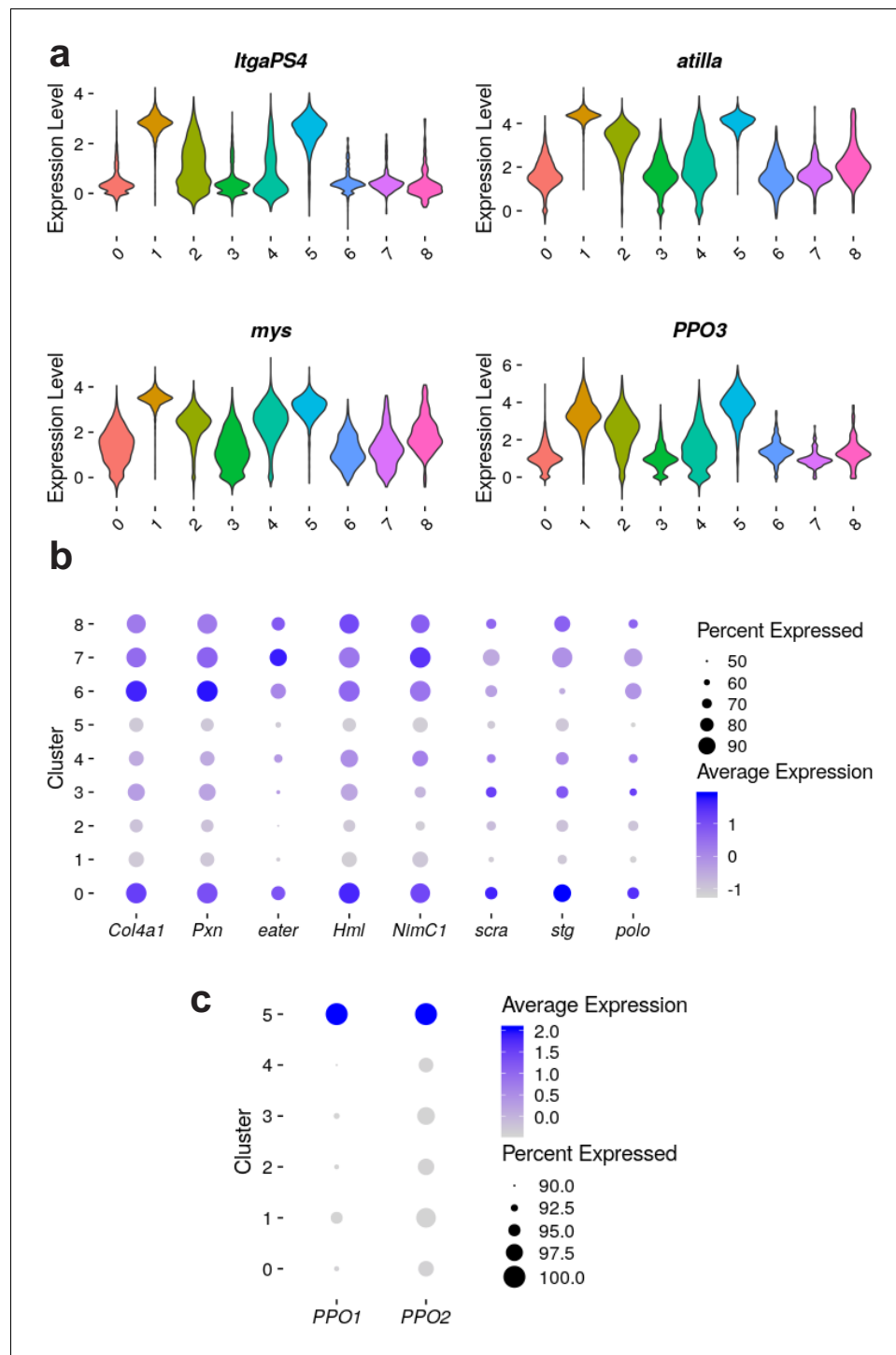


**Figure 3—figure supplement 6.** Flowchart of data filtering and clustering steps. Only cells expressing >250 features and <2500 features and cells that were either *He+* or *Srp+* were kept. Cells expressing high levels of sperm-cell, muscle, or fat body marker genes were identified during three rounds of clustering and removed. 19,344 cells remained after filtering. Three rounds of subclustering identified plasmatocyte progenitor and lamellocyte clusters and these identities replaced those attained from round three clustering for relevant cells.

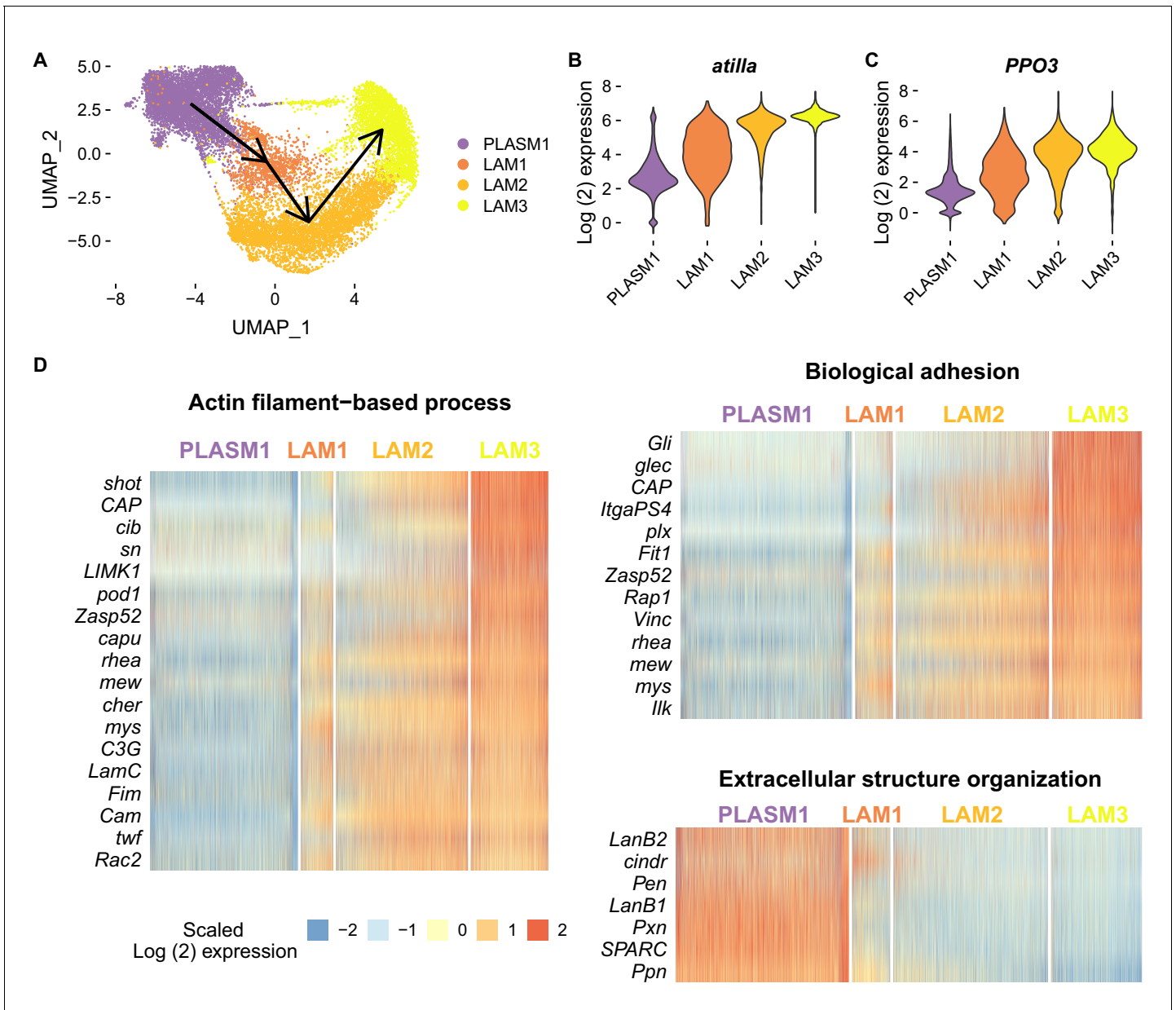




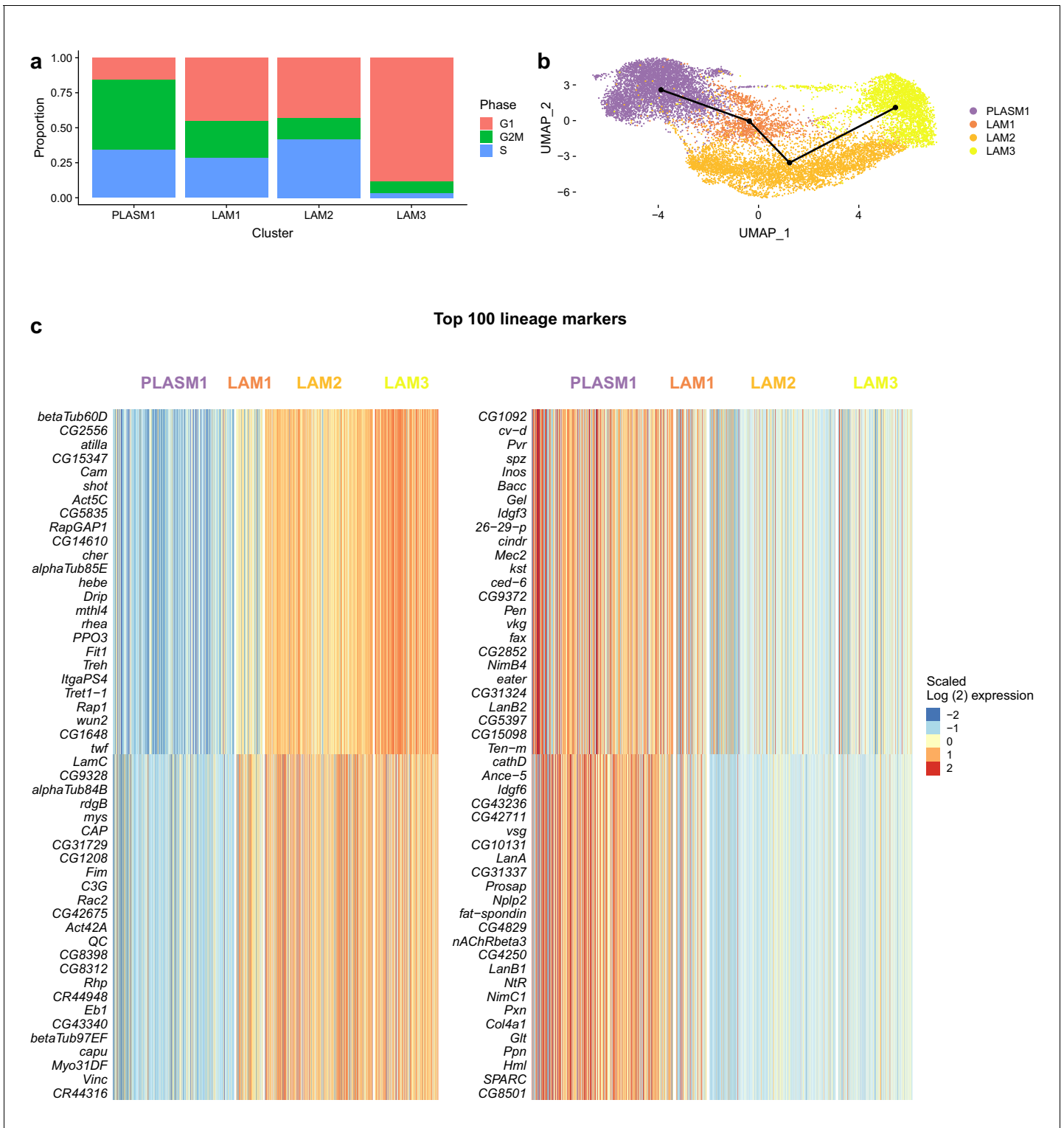
**Figure 3—figure supplement 7.** Proportion of different cell states after infection and selection following data integration with 7716 anchors, the maximum number that can be detected.



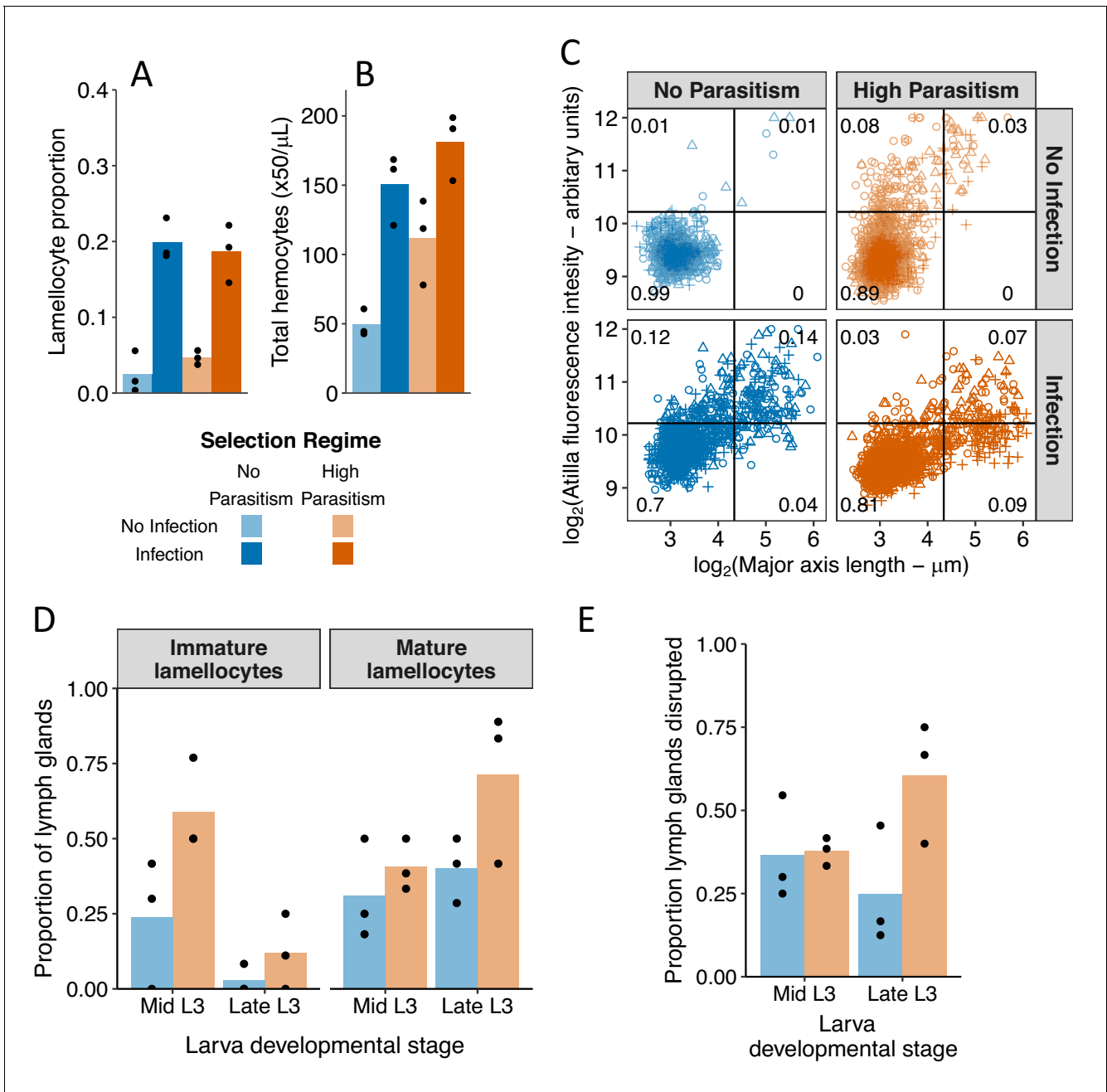
**Figure 3—figure supplement 8.** Expression levels ( $\log_e$ ) of marker genes from early rounds of clustering. (A) Distribution of gene expression levels of described lamellocyte marker genes, for clusters identified from round three clustering. (B) Relative level of expression of described plasmacyte and cell cycle marker genes, for clusters identified from round three clustering. (C) Relative level of expression of two known crystal cell markers from round one of subclustering, using lamellocytes and their plasmacyte progenitors. Percent of cells expressing genes are indicated by circle size (B and C).



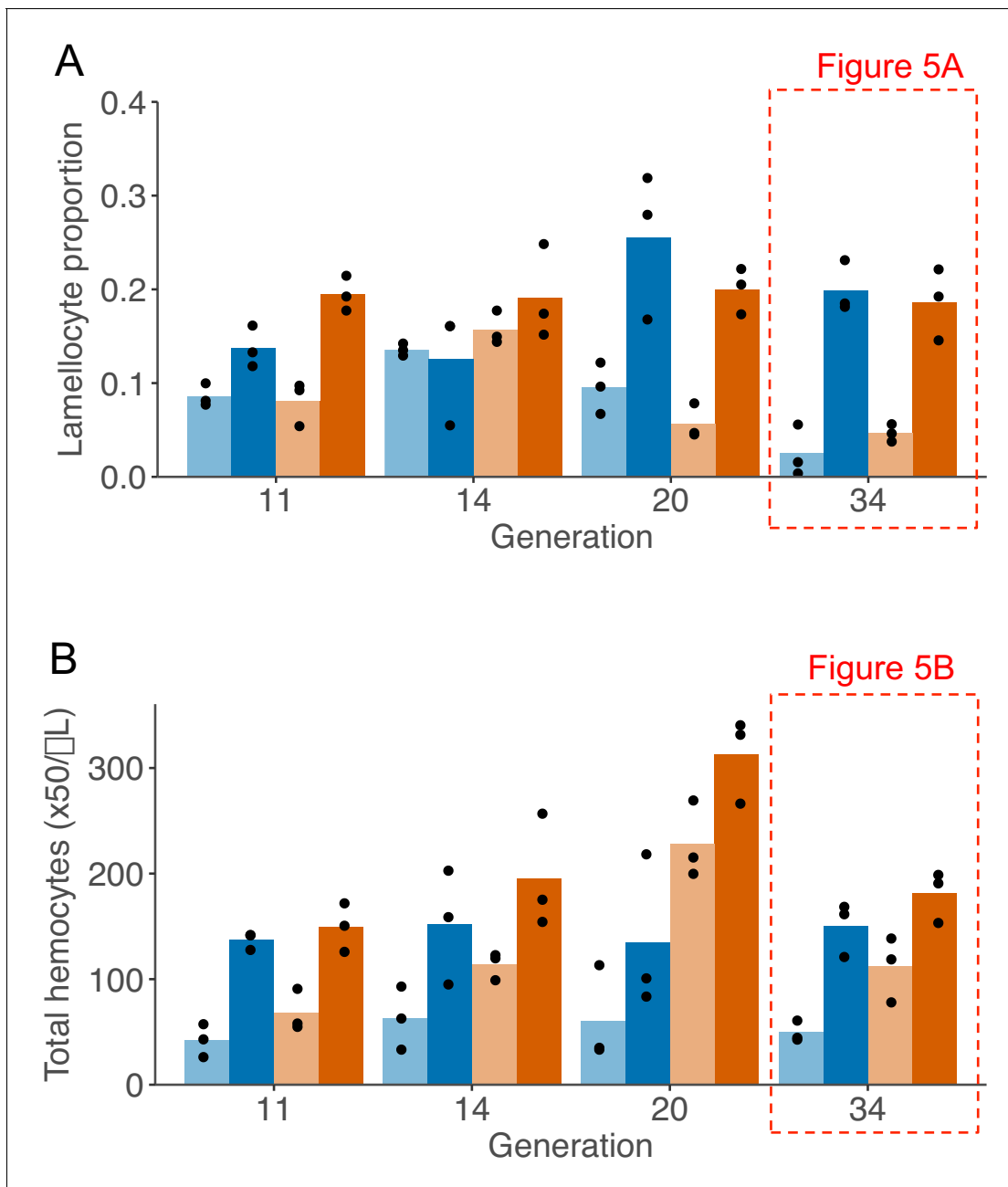
**Figure 4.** Changes in cell state following infection and adaptation to high rates of parasitism. (A) Trajectory of lamellocyte differentiation. Plasmatocyte progenitors and lamellocytes were subclustered, the trajectories inferred from multi-dimensional principle components analysis, and the lineage projected onto the two-dimensional UMAP plot. (B–C) Log<sub>2</sub> expression levels of *atilla* and *PPO3*. Relative expression levels were estimated from normalized and scaled unique molecular identifier counts. (D) Heatmaps of genes that are upregulated in either the starting or terminal clusters of lamellocyte differentiation and belonging to significantly enriched gene ontology categories (biological function). Log<sub>2</sub> fold change gene expression levels were centred on their mean value and divided by their standard deviation.



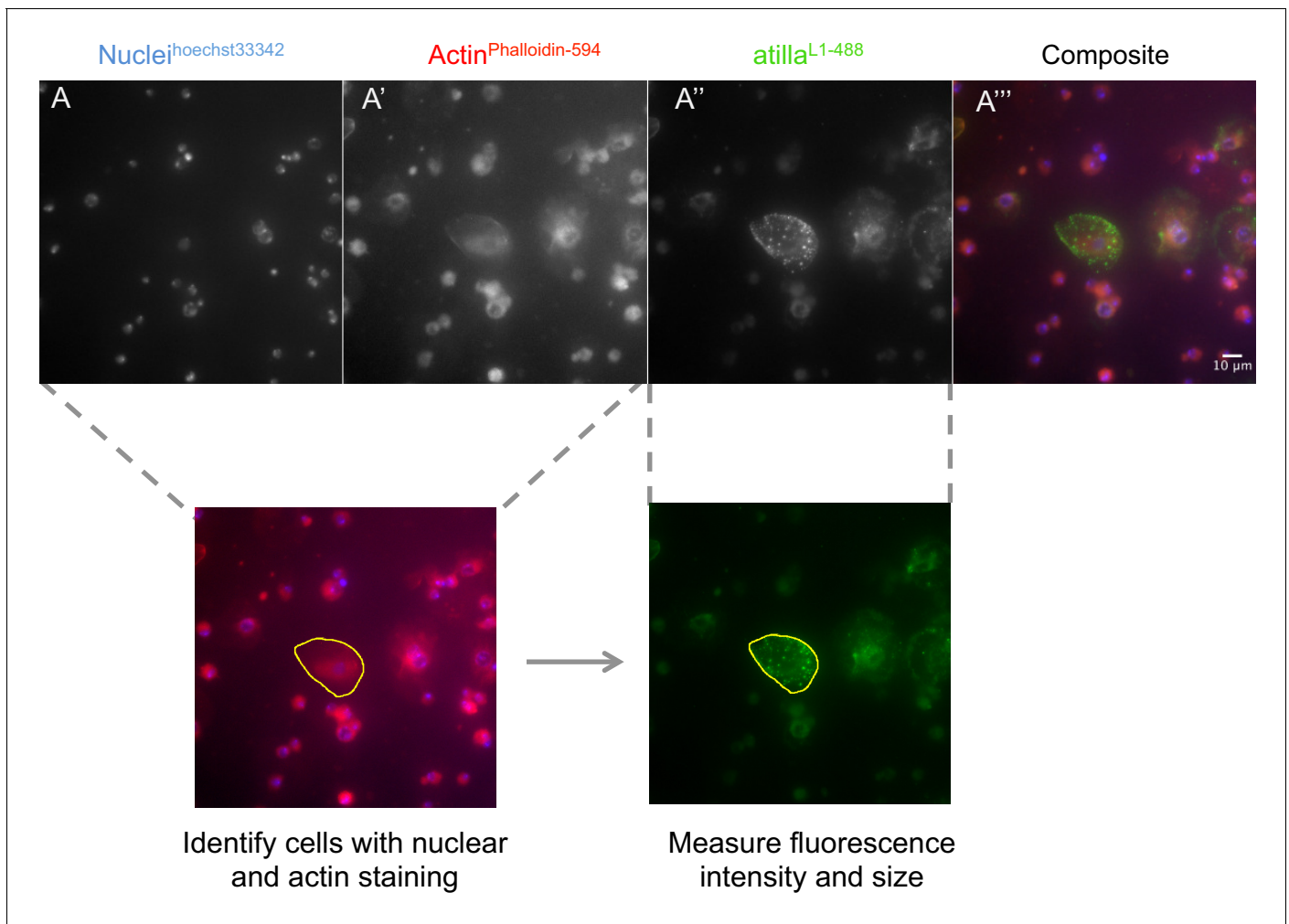
**Figure 4—figure supplement 1.** Trajectory of lamellocyte differentiation following subclustering. (A) Proportions of cells in G1, G2M, and S phase in the plasmacyte progenitor and lamellocyte clusters. (B) Trajectory of lamellocyte differentiation inferred following cell cycle correction and regressing out the difference between the G2M and S phases. Trajectories were inferred from multi-dimensional principle components analysis, and the lineage projected onto the two-dimensional UMAP plot. (C) Heatmap of top 100 markers predicting lamellocyte differentiation as ranked by variable importance. Genes are grouped by mean cluster expression level and cells (columns) were ordered by expression of the 101st marker. Log<sub>2</sub>-fold change gene expression levels were centered on their mean value and divided by their standard deviation.



**Figure 5.** Changes in circulating and lymph gland hemocytes following infection and adaptation to high parasitism rates. (A) The proportion hemocytes that were morphologically identified as lamellocytes and (B) the concentration of total circulating hemocytes in no infection conditions (light colours) and 48 hr post infection (dark colours). Samples were collected from populations that evolved for 34 generation with high parasitism (orange bars) or no parasitism (blue bars) (C) Cell size and atilla expression. The expression of atilla was measured as the fluorescence intensity resulting from immunohistochemistry staining. Cell size was measured along the major axis. Replicated populations from each selection regime are represented by different symbols and the proportion of cells in each quarter of the plot is shown. (D) The proportion of lymph glands from uninfected larvae containing immature lamellocytes (*mys*<sup>+</sup>  $\alpha$ -PS4<sup>+</sup>) and mature lamellocytes (*mys*<sup>+</sup>  $\alpha$ -PS4<sup>+</sup>). (E) Proportion of disrupted lymph glands in uninfected larvae. The lymph gland anterior lobe was considered disrupted when it was either partially or completely disrupted.

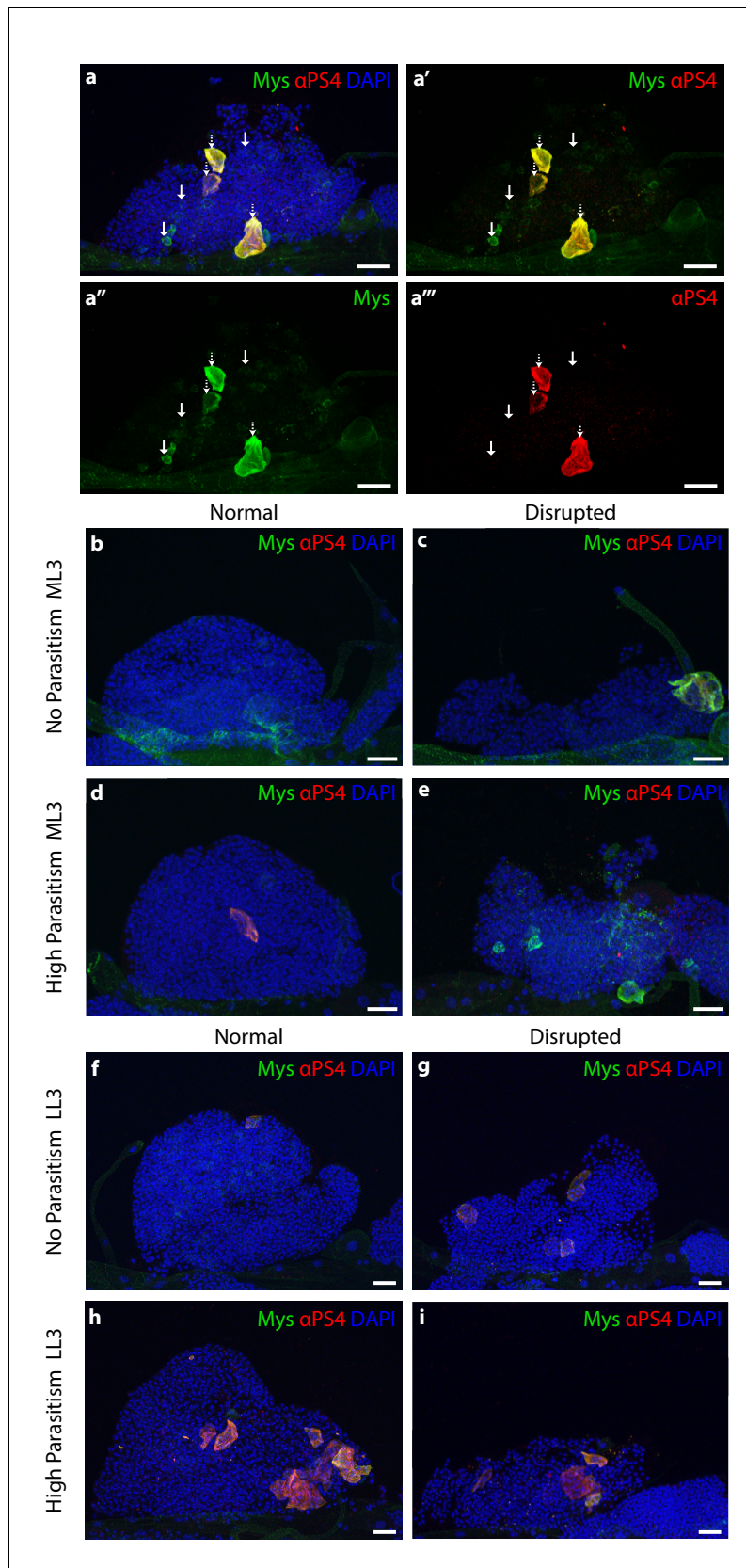


**Figure 5—figure supplement 1.** Changes in number of circulating hemocytes during artificial selection. Proportion of lamellocytes (A) and concentration of total circulating hemocytes (B) in population evolved with no parasitism (blue bars) and with high parasitism (orange bars). Samples were collected as late 3rd instar larvae in homeostasis (light colours) or 48-hr post infection (dark colours). Each dot represents the mean counts of each population calculated from 4 to 10 replicas and bar height represents the mean of the triplicate lines. Generation 34 is represented in **Figure 5A–B**.



**Figure 5—figure supplement 2.** Hemocyte size and atilla staining intensity in circulating hemocytes. Hemocytes were stained for nuclei (A, hoechst 33342), Actin (A', Phalloidin- Alexa594) and atilla (A'', L1 antibody + Alexa488). Composite images were created for phalloidin and hoechst staining to identify cells with selection tool (yellow line) and fluorescence intensity was measured in L1 staining.



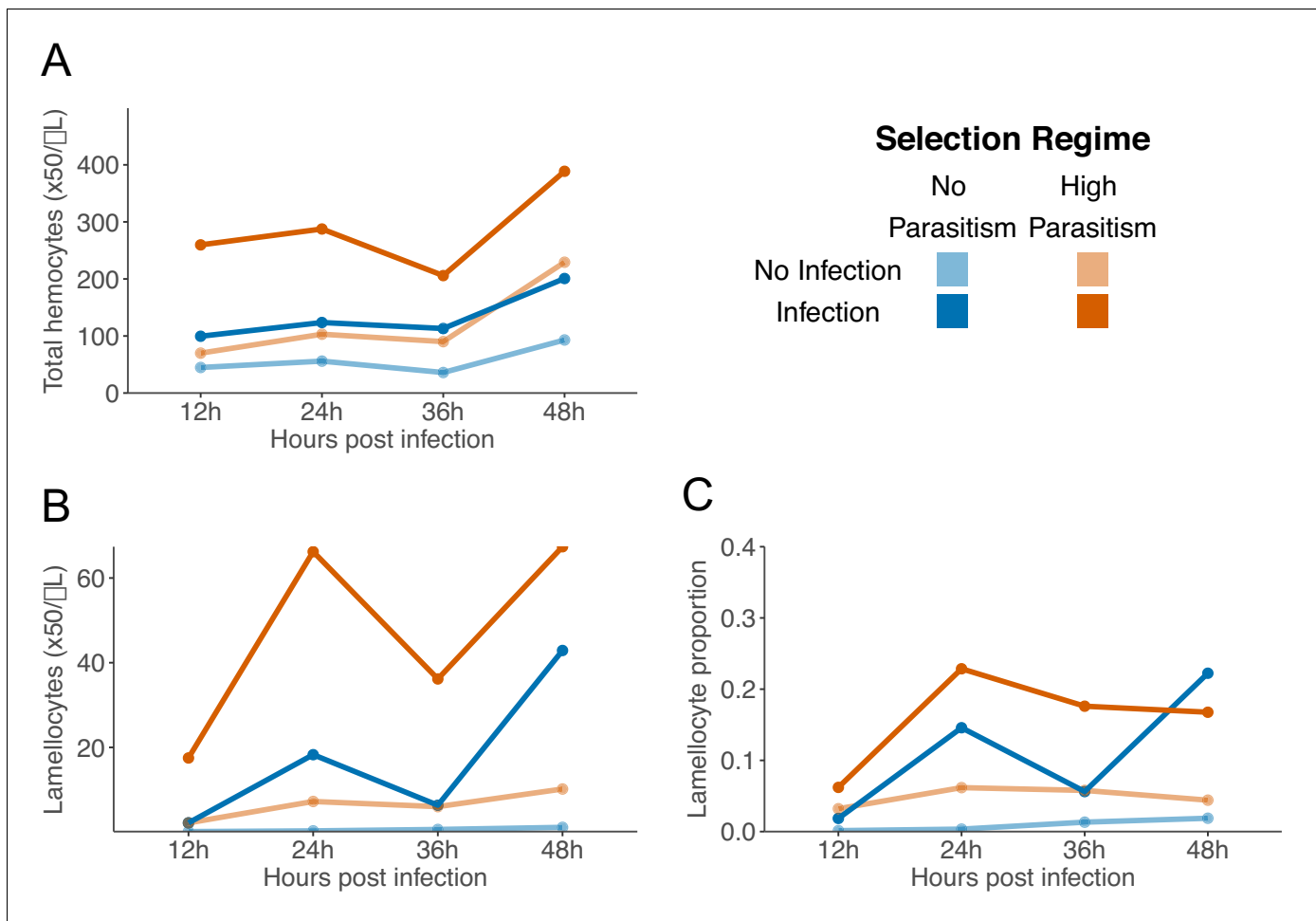


**Figure 5—figure supplement 3.** Lamellocytes differentiation state and morphology of anterior lymph gland lobes. (a) Anterior lobe of a lymph gland from a larva from No Parasitism selection regime. Immature lamellocytes  
 Figure 5—figure supplement 3 continued on next page

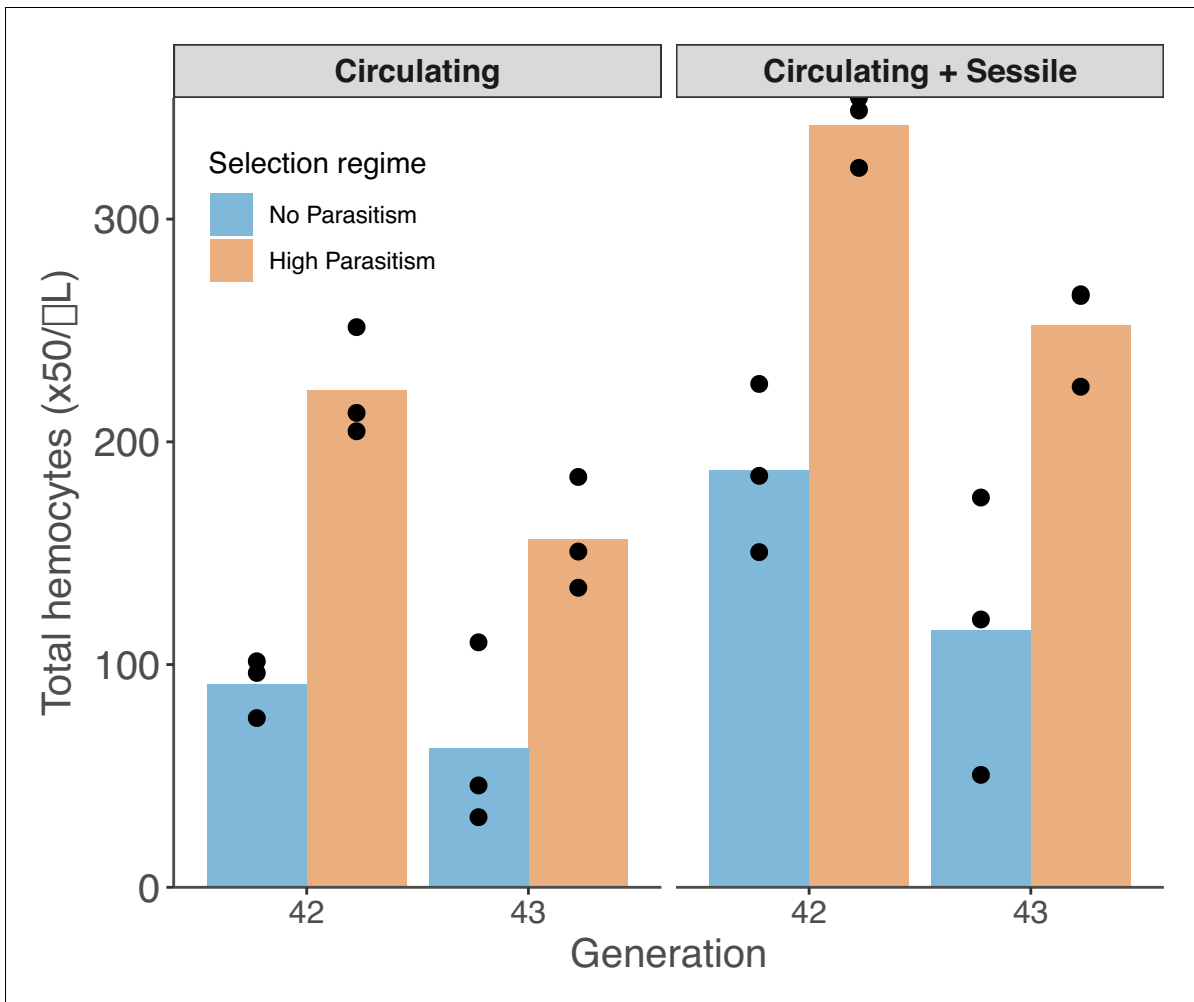


*Figure 5—figure supplement 3 continued*

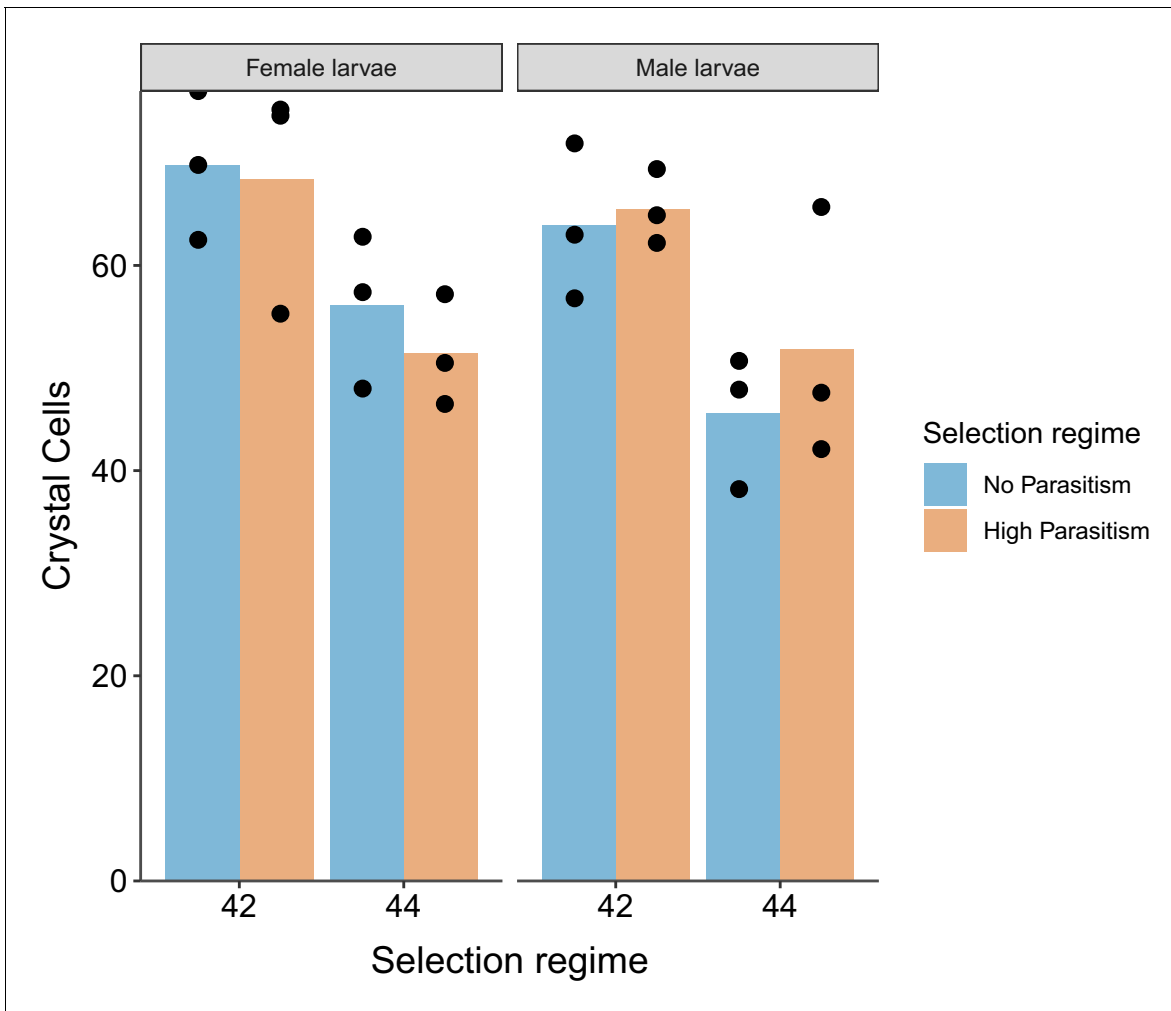
are stained with Myspheroid (Mys, green, with arrow). Mature lamellocytes are double stained with Mys and Integrin alpha PS4 subunit ( $\alpha$ -PS4, red, white-dotted arrow). (b,c) No Parasitism anterior lobe in mild larval stage 3; normal (b) and disrupted lobe (c). (d,e) High parasitism anterior lobe in mild larval stage 3; normal (d) and disrupted lobe (e). (f,g) No parasitism anterior lobe in late larval stage 3; normal (f) and disrupted (g). (h,i) High parasitism anterior lobe in late larval stage 3; normal (h) and disrupted (i). All samples were stained with Mys (green),  $\alpha$ -PS4 (red). Nuclei were stained with DAPI. Scale bars = 20  $\mu$ m.



**Figure 5—figure supplement 4.** Hemocyte concentrations during development of 3rd instar larvae. Total number of hemocytes (A), lamellocytes (B) and lamellocyte proportions (C) in larvae 12-, 24-, 36-, and 48-hr post infection (dark colours) and corresponding controls with no infection (light colors). Points represent the mean of three biological replicas from each triplicate line of No parasitism (blue) and high parasitism (orange) selection regimes. Restrictions on working hours mean the 12 hr and 36 hr timepoints were performed separately to the 24 hr and 48 hr timepoints.



**Figure 5—figure supplement 5.**  $\zeta$  Concentration of hemocytes from dissected larvae without manipulation (Circulating) and after vortexing to displace sessile hemocytes into circulation (Circulating + Sessile). Points represent the mean of four biological replicas for each triplicate line from populations evolved with no parasitism (blue) and high parasitism (orange) selection regimes. Bar heights represent the mean of triplicate lines.



**Figure 5—figure supplement 6.** Sessile crystal cells numbers. Number of sessile crystal cells in dorsal side of the A7 abdominal segment of larvae in male and female larvae. Dots represent the average number of crystal cells calculated from 10 larvae of each triplicate population from both selection regimes. Bar heights represent the average of each selection regime, no parasitism (blue) and high parasitism (Orange).