



Published in final edited form as:

Nat Cell Biol. 2021 September ; 23(9): 939–952. doi:10.1038/s41556-021-00734-6.

Identification of rare post-mitotic cell states induced by injury and required for whole-body regeneration in *Schmidtea mediterranea*

Blair W. Benham-Pyle^{1,2,*}, Carolyn E. Brewster¹, Aubrey M. Kent¹, Frederick G. Mann Jr.^{1,2}, Shiyuan Chen¹, Allison R. Scott¹, Andrew C. Box¹, Alejandro Sánchez Alvarado^{1,2,*}

¹Stowers Institute for Medical Research, Kansas City, MO, USA.

²Howard Hughes Institute for Medical Research, Kansas City, MO, USA.

Abstract

Regeneration requires coordination of stem cells, their progeny, and distant differentiated tissues. Here, we present a comprehensive atlas of whole-body regeneration in *Schmidtea mediterranea* and identify wound-induced cell states. Analysis of 299,998 single-cell transcriptomes captured from regeneration-competent and regeneration-incompetent fragments identified transient regeneration-activated cell states (TRACS) in the muscle, epidermis, and intestine. TRACS were stem-cell-division-independent with distinct spatiotemporal distributions and RNAi depletion of TRACS-enriched genes produced regeneration defects. Muscle expression of *notum*, *folliculin*, *evi/wls*, *glypican-1*, and *junctional-1* was required for tissue polarity. Epidermal expression of *agat-1/2/3*, *cyp3142a1*, *zfhx3*, and *atp1a1* was important for stem cell proliferation. Finally, expression of *spectrinβ* and *atp12a* in intestinal basal cells and *Irrk2*, *cathepsinB*, *myosin1e*, *polybromo-1*, and *talin-1* in intestinal enterocytes regulated stem cell proliferation and tissue remodeling, respectively. Our results identify cell types and molecules important for regeneration, indicating that regenerative capacity can emerge from coordinated transcriptional plasticity across all three germ layers.

Robust injury and tissue repair mechanisms provide critical fitness advantages to organisms and regenerative capacity is widely, but heterogeneously, distributed across the animal kingdom. Invertebrate species can regenerate entire body plans from dissociated cells (1) or small tissue fragments (2–5), while fish and amphibians have the capacity to regenerate appendages and organs after damage or amputation (6–8). Even in mammals with relatively limited regenerative potential, numerous organ systems can replace damaged tissue or re-populate ablated cellular compartments (9–11). Regeneration of adult tissues requires coordinated signaling between differentiated and proliferating cells. In many tissues,

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: <https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>

*Correspondence to: asa@stowers.org and bbp@stowers.org.

Author Contributions Statement: Conceptualization and data interpretation (BWBP, ASA), data analysis (BWBP, CEB, SC), acquisition of data (BWBP, AMK, ARS, ACB), cloning of planarian gene transcripts (FGM), writing of original manuscript (BWBP), supervision and funding acquisition (ASA), and revision and editing of manuscript (all authors).

Competing Interests Statement: The authors declare no competing interests.

dedicated resident stem cells are maintained by multicellular niches (12) and injury results in increased functional plasticity in both niche cells and stem cells, facilitating the replacement of missing cell types and re-establishment of tissue homeostasis (13). During whole-body regeneration, pluripotent adult stem cells can mobilize to sites of injury, proliferate, and differentiate in accordance with local patterning cues to specify and re-build missing tissues (3, 5, 14, 15). While mechanisms coordinating stem cell proliferation and differentiation near the wound site have been described (3, 15–17), less is known about the transcriptional states that occur in distant tissues during regeneration or how these distant cells support tissue repair.

The free-living planarian *Schmidtea mediterranea* is ideal for the study of regenerative plasticity across an animal due to their extraordinary ability to repair and regenerate any organ system from tiny fragments (2). T.H Morgan showed that tissue fragments only 1/279th of the intact animal could regenerate all missing tissues (18) and surface area to volume models predict that regeneration requires fragments containing ~10,000 cells (19, 20). In addition, asexual *S. mediterranea* produce progeny of a fixed size (~1.2mm long) by transverse fission, independent of parent length (21), indicating there may be a minimal fragment size required for regeneration. The small number of cells required for planarian regeneration, combined with robust methods for single cell transcriptomics and RNAi, makes possible an un-biased identification and characterization of regeneration-induced transcriptional states. Here, we report a comprehensive atlas of successful and un-successful planarian regeneration, as well as the identification and characterization of post-mitotic transient regeneration-activated cell states (TRACS).

A single cell reconstruction of planarian regeneration

To characterize regenerative signaling across a whole animal and determine which signaling mechanisms are dependent upon stem cells, we generated a single cell reconstruction of successful and unsuccessful planarian regeneration. We first determined the smallest *S. mediterranea* tissue fragment competent to regenerate and the impact of stem cell depletion on regenerative capacity. Tissue biopsies ranging from 0.75mm – 1.50mm in diameter were taken from the tail region of large animals treated with control, lethal, or sub-lethal doses of ionizing radiation (Extended Data Fig. 1A). Biopsies were taken at 3, 7, and 10 days after sub-lethal irradiation, producing tissue fragments containing increasing numbers of *piwi*⁺ stem cells (Extended Data Fig. 1A, B). Surprisingly, only biopsies taken from un-irradiated animals were competent to regenerate (Extended Data Fig. 1B). Biopsies taken from sub-lethally irradiated animals, independent of time since irradiation, had reduced regenerative capacity but comparable survival rates to un-irradiated biopsies (Extended Data Fig. 1B), whereas biopsies taken from lethally irradiated animals died within 21 days of amputation. Based on these results, we chose to characterize regenerative signaling in three different biological contexts: (1) Biopsies from un-irradiated animals competent to regenerate, (2) Biopsies taken from sub-lethally irradiated animals where stem cells were depleted, and (3) Biopsies from lethally irradiated animals where stem cells were ablated (Fig. 1A).

Split-pool ligation-based single-cell RNA sequencing (SPLiT-seq) (22) was used to capture single-cell transcriptomes from 1.00mm biopsies taken from un-irradiated, sub-lethally

irradiated, and lethally irradiated animals over 2 weeks of regeneration (Fig. 1A, Extended Data Fig. 1C). Fluorescently tagged linker molecules and image cytometry were used to visualize SPLiT-seq barcoding reagents in dissociated planarian cells (Extended Data Fig. 2A,B) and find optimal dye-based sort conditions to ensure that only intact barcoded cells were sequenced (Extended Data Fig. 2C–F). Biopsies were then dissociated, barcoded, and sequenced 0, 1, 2, 4, 7, 10, and 14 days post amputation (dpa), resulting in twenty-one samples from three regeneration time courses. After filtering for transcriptome quality, the final dataset contained 299,998 single-cell transcriptomes with an average of 14,285 transcripts per sample, 1,981 UMIs/cell, and 429 genes/cell (Extended Data Fig. 2G,H, Table S1). Clustering of cell states identified 89 global clusters, most of which could be assigned to known tissue classes based on previously reported marker genes, comparison to transcriptional profiles from published planarian atlases (23, 24), and expression patterns of enriched genes (Fig. 1B–D, Extended Data Fig. 3–5, 6A–B, Table S2). Once assigned to known tissue classes, cell states were split into tissue-level data subsets and re-analyzed to identify additional subcluster diversity (Extended Data Figs. 2–5, Tables S3–S14), producing 211 tissue subclusters.

To validate the reconstruction, we quantified the relative abundance of all tissue subclusters during successful and unsuccessful regeneration (Fig. 1E, Fig. 2C, Extended Data Fig. 6C–L). Unsurprisingly, stem cell clusters marked by high expression of the stem cell marker *piwi-1* (25) were enriched in biopsies taken from un-irradiated animals (Fig. 2A–D) and RNAi-mediated knockdown of genes enriched in irradiation-sensitive sub-clusters resulted in strong tissue maintenance and regeneration defects, as expected for regulators of stem cell function (Fig. 2D–J, Tables S16–S17, see also (26–32)). Together, these results demonstrated that relative abundance of tissue subclusters across samples could identify functionally important cell states and RNAi of enriched genes could be used to test physiological function. Therefore, to identify post-mitotic cell states important for whole-body regeneration, we next identified amputation-induced states in the differentiated tissue sub-clusters.

Transient Regeneration-Activated Cell States

We postulated that cell states important for organizing regeneration would be present primarily at early time points (24 – 96 hours after amputation). Indeed, there were several cellular states identified that were enriched in the un-irradiated dataset or in only early timepoints (Fig. 1E, 3A–C). To identify amputation-induced cell states, the proportion of cells originating from each time point after amputation was calculated for all differentiated tissue subclusters (Fig. 3D–L). Because several wound-induced genes important for regeneration respond to injury even in the absence of stem cells (33), we performed this calculation on both the full dataset and the un-irradiated time course (Table S15). On average, transcriptomes collected 0 – 4dpa were ~45% of cells in a subcluster (44.70% \pm 14.01%) in the full dataset, or ~ 50% (51.26% \pm 18.65%) in the un-irradiated time course (Fig. 3M). However, thirteen tissue subclusters were disproportionately occupied (2 standard deviations above the mean) by cells from early timepoints (Fig. 3D–L (black dotted lines) and Fig. 3M (red dotted lines)). These subclusters were termed transient regeneration-activated cell states (TRACS) and represented 5.04% of captured cells.

We next calculated the proportion of cells in each tissue belonging to TRACS. Surprisingly, TRACS were not uniformly distributed across cell lineages and were instead enriched in the intestine, epidermis, muscle, and kidney-like protonephridia (Fig. 3N). The relative abundance of excretory tissues in planaria has been shown to be highly dependent on culture conditions and water salinity, so we chose to focus on the muscle, epidermal, and intestinal cell states for characterization. In an effort to characterize TRACS likely to be upstream of stem cell activation, we focused on those that occurred in both the un-irradiated regeneration time course and in biopsies where stem cells were depleted or ablated (Fig. 3D–F, black arrows). These subclusters were muscle subcluster 16 (M16, Fig. 3F, Fig. 4A), epidermal subclusters 2 and 20 (E2/E20, Fig. 3E, Fig. 4B), and intestinal subclusters 9 and 12 (I9/I12, Fig. 3D, Fig. 4C).

Muscle subcluster 16 was highly enriched at early time points after amputation (Fig. 4A) and analysis of high enriched genes compared to other muscle clusters identified *notum* as a highly specific M16 marker (Fig. 4D). Interestingly, *notum* is known to be activated in muscle at anterior wounds and is required for polarity re-establishment during whole-body regeneration (34). Indeed, visualization of *notum*⁺/*collagen*⁺ cells *in vivo* confirmed their induction 24 hours post amputation at anterior facing wounds and quantitation of *notum*⁺/*collagen*⁺ cells indicated that M16 represents ~50% muscle cells near the wound site (Fig. 4E,F, Extended Data Fig. 7A,B). In addition to *notum*, M16 was marked by expression of several regulators of polarity re-establishment after injury (Extended Data Fig. 8E), including *folliculin* (35,36), *evi/wls* (37), and *glypican-1* (38). Thus, M16 is a wound-induced muscle cell state transiently induced 24 hours post amputation and marked by dynamic expression of position control genes (PCGs).

Epidermal subclusters 2 and 20 (E2, E20, Fig. 4B) represented distinct cell types in the epidermal lineage and had distinct temporal dynamics. E2 expressed markers of mature epidermal cells and was marked by increased expression of *hadrian* (Fig. 4D). Visualization of *hadrian*⁺ cells confirmed that they were enriched in the surface epithelium at 24 – 48 hours post amputation (Fig. 4E,G, Extended Data Fig. 7A). Epidermal cluster 20, on the other hand, strongly expressed markers of late epidermal progenitors, such as *agat-1*, *agat-2*, and *agat-3* (39–41). E20 was marked by increased expression of *actin-3* (Fig. 4D,E,H, Extended Data Fig. 9E) and included a smaller subset of epidermal cells. Visualization and quantitation of *agat-1*⁺/*actin-3*⁺ cells confirmed that *actin-3*⁺/*agat-1*⁺ cells increased 24 – 48 hours after amputation and remained slightly elevated 4 – 7 days post amputation (Fig. 4E, H, Extended Data Fig. 7A). While *hadrian*⁺ epidermal cells were observed near and far from the wound-site, *actin-3*⁺/*agat-1*⁺ cells were enriched adjacent to the wound and in blastema tissue (Extended Data Fig. 7A).

Like the TRACS we identified in the epidermis, intestinal subclusters 9 and 12 (I9, I12, Fig. 4C) were related to distinct cell types in the intestine and had unique temporal dynamics. The planarian intestine has three main cell types: secretory goblet cells, absorptive phagocytes/enterocytes, and recently identified ‘outer’ or ‘basal’ intestinal cells (23, 42). Intestinal subcluster 9 expressed markers of basal/outer intestinal cells (*tdo2*, *HAL*) and was marked by increased expression of *lectin2b* (Fig. 4D,E,I, Extended Data Fig. 10E). Intestinal subcluster 12, on the other hand, expressed markers associated with intestinal

enterocytes, and increased expression of methyltransferases (*polybromo-1*, *mettrans*) or regulators of endocytosis (*myosin1E*) and lysosomal degradation (*cathepsinB*, *cathepsinL*) after amputation. In particular, intestinal cluster 12 was marked by the uncharacterized planarian gene *26267*, which is typically expressed only in peripheral phagocytic cells (Fig. 4D,E,J, Extended Data Figs. 7A, 10E). Because of its expression during injury-induced gastrovascular remodeling, we refer to gene *26267* as Injury Remodeling Activated 1 gene (*ira-1*). Visualization and quantitation of *prcn*⁺/*lectin2b*⁺ cells confirmed that the percentage of *prcn*⁺ intestinal cells also *lectin2b*⁺ increased 2 – 4 days after amputation (Fig. 4E, I, Extended Data Fig. 7A). Interestingly, we observed a high level of variability in the proportion of outer basal cells expressing *lectin2b* between individual animals (Fig. 3I) and when *lectin2b* was highly expressed, it was expressed in basal cells throughout the intestine, rather than enriched near the wound site (Extended Data Fig. 9A). Intestinal cluster 12, in contrast, was a very rare cell state among enterocytes (Fig.4J). Visualization and quantitation of *prcn*⁺/*ira-1*⁺ cells confirmed that these cells were only ~4% of intestinal cells 1 – 4 days post amputation (Fig. 3E,J) and were enriched in gut branches near the wound or adjacent to the regenerating pharynx (Extended Data Fig. 7A).

TRACS occur in the absence of cell divisions

Transient regeneration-activated cell states (TRACS) could be induced in pre-existing post-mitotic cells or could arise as a product of stem cell divisions. Indeed, all the TRACS we characterized in the muscle, epidermis, and intestine occurred at different rates in unirradiated biopsies compared to those taken from irradiated animals (Extended Data Figs. 8A–D, 9A–D, 10A–D). Trajectory inference methods can be used to predict connectivity between cell states in single-cell sequencing datasets, so we used partition-based graph abstraction (PAGA, 43) to predict connectivity between our observed cell states and the stem cell compartment. Amputation-specific cell states did not connect directly to the stem cell progenitor subcluster (Fig. 5A–C). Instead, the connectivity path from the stem cell compartment (black arrows) to muscle subcluster 16 (Fig. 5A), epidermal subclusters 2 and 20 (Fig. 5B), and intestinal subcluster 12 (Fig. 5C) passed through several cell states that were not amputation-induced or only present at later stages of regeneration. Intestinal subcluster 9 was connected to the stem cells via intestinal subcluster 1 (Fig. 5C), another wound-induced cell state (Extended Data Fig. 10D). Therefore, PAGA analysis of subcluster connectivity indicated that most TRACS were more likely to arise from pre-existing post-mitotic tissues than from proliferating stem cells.

To test experimentally if generation TRACS required stem cell divisions, we treated animals with 0.15% colchicine after amputation (Fig. 5D), which results in a mitotic arrest and accumulation of H3P⁺/*piwi-1*⁺ cells (Extended Data Fig. 7C–D, see also 44). Despite this mitotic arrest, TRACS in the muscle, epidermis, and intestine were induced at comparable or higher levels after amputation (Fig. 5E, F). *Notum*⁺ muscle cells were slightly reduced in colchicine-treated animals, as predicted by our single cell sequencing (Extended Data Fig. 8A–D), but this reduction was not statistically significant. While E20 was reduced in sequenced biopsy samples taken from irradiated animals (Extended Data Fig. 9A–D), *agat-1*⁺/*actin-3*⁺ cells (E20) became more plentiful in colchicine-treated animals (Fig. 5E,F). Therefore, the irradiation sensitivity of E20 in our single-cell reconstruction was most likely

due to loss of *agat-I*⁺ progenitors prior to amputation, rather than a failure to induce the cell state. These results reveal that TRACS are more related to pre-existing post-mitotic tissues than to stem cells and do not require stem cell divisions to occur *in vivo*. Thus, we identified subsets of differentiated cells in the muscle, epidermis, and intestine that transiently change their function after amputation, independent of stem cell proliferation.

Muscle TRACS express genes required for polarity

Muscle cells constitutively express regionally restricted position control genes (PCGs), which are required to maintain and regenerate the adult body plan (45). As such, re-establishment of positional information in muscle cells is an important step required for regeneration of missing tissues in planaria (34,46–48). Since muscle subcluster 16 (MC16) was marked by increased expression of the polarity gene *notum*, we postulated that the cell state could represent wound-induced muscle express additional genes important for tissue polarity. Genes enriched in M16 relative to other muscle subclusters were highly expressed in the muscle and wound-induced in all three single-cell regeneration time courses and in a bulk RNAseq dataset (49) of planarian regeneration (Extended Data Fig. 8E,F). M16-enriched genes included several previously identified wound-induced genes (*follistatin*, *notum*, *glypican-1*, *evi/wls*) that produce regeneration defects upon RNAi depletion. A targeted RNAi screen of M16-enriched genes re-capitulated previously reported regeneration defects for *follistatin*, *notum*, *evi/wls*, and *glypican-1* (34–38). In addition, *grp78*, *Ca-trans ATPase*, and *Tubulin-β* were identified as regulators of tissue homeostasis (Extended Data Fig. 8G–J) and RNAi inhibition of *junctophilin-1* resulted in small blastemas and increased incidence of cyclopia, reminiscent of *evi/wls* RNAi-treated animals (Fig. 6A–C). *junctophilin-1* was expressed in body wall musculature and was enriched adjacent to the wound site (Fig. 6A, Extended Data Fig. 8K). To determine if *junctophilin-1* was indeed a regulator of tissue patterning, we compared regeneration defects in *junctophilin-1* RNAi-treated animals to those produced after *notum*, *follistatin*, *glypican-1*, and *evi/wls* RNAi depletion. Visualization of the central nervous system (*PC-2*) and intestinal branches (*gata4/5/6*) 14 days post amputation in RNAi-treated animals revealed absent or small anterior cephalic ganglia in all RNAi treatments, including *junctophilin-1* (Fig. 6D). We next visualized *notum*⁺ anterior poles and *wnt1*⁺ posterior poles in RNAi treated animals (Fig. 6E). *Follistatin* and *glypican-1* RNAi-treated animals had an altered number or distribution of *notum*⁺ cells in the regenerated anterior pole, while *evi/wls* and *junctophilin-1* RNAi-treated animals had a normal number and distribution of *notum*⁺ (Fig. 6F–H). In addition, all evaluated RNAi conditions except *evi/wls* knockdown resulted in an expanded range of *Wnt1*⁺ cells in the posterior pole and *follistatin* RNAi-treated animals also had an increased number of *Wnt1*⁺ cells (Fig. 5I–K).

Altogether, our results indicate that muscle subcluster 16 represents a subset of muscle cells adjacent to the anterior wound site that co-express genes important for anterior pole specification and inhibition of posterior-tissue specification by *Wnt1*. While several genes expressed in muscle cluster 16 have been previously identified as wound-induced position control genes, their co-expression in a subset of wound-adjacent muscle cells had not been shown. Moreover, by characterizing several genes expressed within this cell state, we were able to identify additional regulators of homeostasis and polarity re-establishment.

Epidermal genes regulate stem cell proliferation

In both vertebrate and invertebrate systems, the wounded epidermis expresses and secretes molecules that promote wound closure and tissue regeneration (11, 13, 38, 41). In planaria, the molecules regulating differentiation of *piwi-1*⁺ progenitors into mature epidermal cells are some of the most well characterized amongst all tissue lineages. *zfp1*⁺ zeta neoblasts progress through a *prog2*⁺ early progenitor state and an *agat-3*⁺ late progenitor state, then transition through a *zpu6*⁺ or *vim1*⁺ stage before becoming mature *rootletin*⁺, *PRSS12*⁺, or *laminB*⁺ epidermal cells (39–41). While *zfp1*⁺ zeta neoblasts and *prog2*⁺ early progenitors did not significantly alter transcription after amputation, both *agat-3*⁺ late epidermal progenitors and mature epidermal cells had associated TRACS – epidermal subclusters 20 (EC20) and 2 (EC2), respectively (Fig. 3).

To test the function of epidermal TRACS during regeneration, genes enriched in EC2 and EC20 were depleted by RNAi and phenotypes were observed at homeostasis and after amputation. Surprisingly, a range of homeostatic and regeneration defects were observed following depletion of genes enriched in *agat-3*⁺ epidermal progenitors (E20) (Fig. 7A–E, Extended Data Fig. 9E,F). *Agat-1* RNAi-treated animals had reduced photoreceptor pigmentation in regenerates at 14 days post amputation (Fig. 7B,D). Animals treated with RNAi targeting *agat-3* or *zinc finger homeobox protein 3 (zfhx3)* had no blastema growth or reduced blastema growth (Fig. 7B,D) and RNAi of *cyp3142a1* or *agat-2* produced lesions at homeostasis and either no blastema or lesions after amputation (Fig. 7B,D). Finally, a gene enriched in the mature epidermis, *sodium/potassium-transporting ATPase subunit α (atp1a1)* (Fig. 7B,D), resulted in lesions or a small blastema after amputation. Visualization of the central nervous system and intestine at 7 days post amputation confirmed a reduction in the size of regenerating cephalic ganglia, but minimal defects in anterior fusion of the left and right gut branches (Fig. 7C,E). EC20-enriched genes (e.g., *actin-3*, *agat-1*, *agat-2*, *agat-3*, *cyp3142a1*) were expressed in dispersed cells in the mesenchymal space at homeostasis, but more highly expressed in more densely packed cells around the regenerating blastema, while *atp1a1* was expressed throughout the mature epidermis in both intact and amputated animals (Extended Data Fig. 9G).

Given the increased incidence of lesions in RNAi-treated animals, we visualized *piwi-1*⁺ stem cells and quantitated H3P⁺ mitotic cells during early (2dpa) and late (7dpa) stages of regeneration (Fig. 7F,G, Extended Data Fig. 9H–K). Surprisingly, *zfhx3*, *agat-2*, *cyp3142a1*, and *atp1a1* RNAi-treated animals had reduced proliferation 2 days post amputation and *piwi-1*⁺ stem cells were less plentiful in *agat-2*, *cyp3142a1*, and *atp1a1* RNAi-treated animals (Fig. 7F,G). The reduction in proliferation in RNAi-treated animals was less significant after the wound-induced mitotic burst, but the continued loss of *piwi-1*⁺ stem cells could be observed in *agat-2*, *cyp3142a1*, and *atp1a1* RNAi-treated animals (Extended Data Fig. 9J,K). We also assessed stem cell proliferation in intact animals for genes that produced the most severe phenotypes. Both *agat-1* and *agat-3* RNAi-treated animals had increased mitotic divisions at 7 days post RNAi feeding, while *agat-2* and *atp1a1* RNAi-treated animals had reduced H3P⁺ mitotic cells (Extended Data Fig. 9H,I).

Prior studies have identified regulators of epidermal differentiation expressed in *agat*⁺ progenitor cells. Our results indicate that *agat*⁺ late epidermal progenitors respond to injury and express many genes required for regulation of stem cell proliferation. We have also identified a Na/K ATPase subunit expressed in mature epidermal cells that is required for regulation of stem cell proliferation. Therefore, our results indicate that both mature epidermal cells and *agat*⁺ late progenitor cells are transcriptionally responsive to amputation and required for the regulation of stem cell proliferation in intact and regenerating tissues.

Intestinal genes regulate stem cells and tissue remodeling

The endoderm is a well-established model for stem cell regulation during normal tissue turnover and in response to injury (10, 50, 51). Several types of support and stem cells important for tissue maintenance and injury repair have been identified, including rare or transient cell states that only respond to particular types of injuries (10, 52, 53). Genes enriched in intestinal TRACS (Extended Data Fig. 10E,F) included methyltransferases (*polybromo-1*, *mettrans*), regulators of endocytosis (*myosin1E*), regulators of lysosomal degradation (*cathepsinB*, *cathepsinL*), and many regulators of gut function and metabolism (*lectin2b*, *lipase*, *TDO*, *PEPCK*, *glycoside hydrolase*).

To test if genes upregulated in intestinal TRACS after amputation were required for planarian regeneration, IC9- and IC12-enriched genes were depleted using RNAi. Seven genes (Fig. 8A) produced reproducible defects in either tissue homeostasis or tissue regeneration. RNAi depletion of *atp12a* and *spectrinβ* (both enriched in ‘outer/basal’ cells, Extended Data Fig. 10G), resulted in lesions or inch-worming at 3 days post feeding (Fig. 8B) and lysis before 21 days post feeding (Fig. 8C). Animals treated with *atp12a* RNAi had a drastic reduction in stem cell proliferation 7 days post RNAi feeding, while *spectrinβ* RNAi-treated animals had reduced *piwi-1* expression in posterior tissues (Fig. 8D,E). *Atp1a1* and *spectrinβ* were upregulated in wound-adjacent intestinal tissues, but were also expressed in intestinal, phagocytic, and protonephridial tissues at homeostasis (Extended Data Fig. 10G). Therefore, while we conclude that ‘basal/outer’ intestinal cells alter transcription after amputation (Figs. 3–5) and express genes important for stem cell maintenance (Fig. 8B–E), we have not yet determined if they have a regeneration-specific function.

In contrast to genes in IC9, RNAi depletion of IC12-enriched genes resulted in regeneration defects, but minimal defects in intact animals. RNAi knockdown of *myosin1E*, *polybromo-1*, and *talin-1* resulted in small blastemas and limited or no pharynx regeneration, while RNAi knockdown of *cathepsinB* and *Irrk2* resulted in an increased incidence of cyclopia or un-pigmented photoreceptors (Fig. 8F,G). Visualization of the central nervous system confirmed a reduction in the size of regenerated cephalic ganglia in *myosin1E* and *polybromo-1* RNAi-treated animals (Fig. 8H,I). In addition, visualization and quantitation of *notum*⁺ cells revealed a reduction in either the number or distribution of *notum*⁺ cells in RNAi-treated animals (Extended Data Fig. 10I,K,L). A reduction or loss of posterior intestinal branches was also observed in *myosin1E* and *talin-1* RNAi-treated animals (Fig. 8H). Thus, inhibition of IC12-enriched genes (*e.g.* *myosin1E*) resulted in defects in regeneration of anterior structures and in gut remodeling. These defects were not due to

a reduction in *piwi-I*⁺ stem cells or the number of proliferating stem cells at 2 days post amputation (Extended Data Fig. 10H,J), but there was a slight reduction in proliferation observed 14 days post amputation in *myosin1E* and *polybromo-1* RNAi-treated animals (Fig. 8J,K). Together, our data indicate that a rare subset of intestinal enterocytes upregulate phagocytic gene sets in response to amputation and a screen of enterocyte enriched genes identified several important for new tissue growth during regeneration.

Discussion

To date, the study of regenerative capacity has largely focused on signaling mechanisms near the injured tissue required for stem cell proliferation and differentiation following injury (3,7,10,13,54–56). Relatively little is known about how post-mitotic cells contribute to whole-body regeneration, particularly cell types that may act on stem cells and their immediate progeny from a greater distance. Here, we present a comprehensive single cell reconstruction of whole-body regeneration over time. Minimal body fragments were taken from un-irradiated, sub-lethally irradiated, and lethally irradiated planarians, allowing us to define the cellular dynamics of successful and un-successful regeneration across an entire animal and to identify post-mitotic transient regeneration-activated cell states (TRACS). We then determined which TRACS were dependent upon stem cell proliferation and identified TRACS-enriched genes important for tissue regeneration and remodeling of pre-existing tissues.

We identified TRACS in the mesoderm (*collagen*⁺/*notum*⁺ MC16), ectoderm (*agat-I*⁺/*actin-3*⁺ EC20, *hadrian*⁺ EC2), and endoderm (*prcn*⁺/*lectin2b*⁺ IC9 and *prcn*⁺/*ira-I*⁺ IC12). A subset of anterior-facing muscle cells induce genes (*notum*, *follistatin*, *evi/wls*, *glypican-1*, *junctophilin-1*) required for re-establishment of the anterior and posterior poles and subsequent tissue regeneration (Fig. 6). *agat-I*⁺/*actin-3*⁺ cells are induced by injury and required for stem cell proliferation (Fig. 7). Finally, ‘basal/outer’ intestinal cells respond to amputation and regulate stem cell proliferation and a rare subset of enterocytes transiently express genes required for tissue remodeling (Fig. 8). Importantly, TRACS were independent of cell divisions (Fig. 5), consistent with observations that some wound-induced genes do not require new proliferation (38, 47).

An increase in transcriptional plasticity in rare cell types is often required for tissue repair, particularly in systems where stem cells have been ablated (52,53). Our results build on these observations and indicate that regenerative capacity may arise from an organism’s ability to produce rare, transient, and functionally distinct cell states within a subset of differentiated tissues after injury. We refer to these as transient regeneration-activated cell states (TRACS). The observation that TRACS are both transient and rare may have important implications for the regulation of growth capacity. Transiency of regenerative signaling may act as a guard against un-checked growth. Indeed, genes expressed in intestinal TRACS (*polybromo-1* and *TDO*) have been implicated in cancer proliferation and tumor immune evasion (57–60). The rarity of TRACS may also account for why regenerative capacity is so unevenly distributed. A rare cell state regulating a subset of regenerative signaling is more easily lost during evolution and – perhaps – more easily re-activated (61). Regenerative capacity of planarian species incapable of head regeneration

has been restored through inhibition of Wnt signaling (62, 63) and we can now postulate that this may have restored rare muscle TRACS (*notum*⁺ MC16) and the ability to re-pattern the body plan. If generally applicable, re-activation of TRACS could unleash regenerative capacity in systems where such capacities are more limited.

The results presented here characterize cellular components required for regeneration at unprecedented molecular resolution, facilitating the discovery of rare cell states (TRACS) and identification of genes required for patterning, stem cell proliferation, and tissue remodeling. These data also support a model in which regenerative capacity may be linked to coordinated plasticity in differentiated cells across all three germ layers.

Methods

Animal husbandry

S. mediterranea animals from the asexual clonal strain CIW-4 (C4) (64) were maintained in 1X Montjuic salts in a planarian re-circulation culture system as previously described (65). Tissue biopsies were taken from large animals fed chunk beef liver three times per week in a recirculation system. Small animals utilized for *in situ* hybridizations and the RNAi screens were fed chunk liver once a week in a recirculation system and starved in static culture at least 4 weeks prior to use. Worms were dosed with gamma irradiation in an MDS Nordion Gammacell 40 Exactor low dose-rate research irradiator. All animal husbandry was performed in accordance with ethical guidelines of the Stowers Institute for Medical Research.

Tissue biopsies

Animals were anesthetized using cold chloretone solution (0.1 – 0.2% w/v chloretone in 1X Montjuic salts) for 4 – 5 minutes or until immobile, then rinsed with cold Holtfreter's solution (3.5g/L NaCl, 0.2g/L NaHCO₃, 0.05g/L KCl, 0.2g/L MgSO₄, 0.1g/L CaCl₂, pH 7.0 – 7.5) and placed on Whatman #3 filter paper saturated with cold Holtfreter's to minimize movement (66). Limited re-use biopsy punches (World Precision Instruments #504528 [0.5mm], #504529 [0.75mm], #504646 [1.0mm], #504530 [1.2mm], #504647 [1.5mm]) were used to extract round tissue fragments from the center of the animal's posterior tail tissue and the fragments were deposited into cold Holtfreter's solution. Tissue fragments were stored at 20°C overnight and Holtfreter's solution was exchanged for 1X Montjuic + 50µg/mL gentamycin at 24 hours post amputation.

Generation of fixed planarian single-cell suspensions

Animals or biopsies were rinsed in calcium-magnesium free buffer with 1% BSA (CMFB) (67), finely chopped with a single edge stainless steel blade, then agitated on a benchtop rocker in CMFB for 20 – 30 minutes with vigorous pipetting every 3 – 5 minutes. After maceration, cells were filtered through a 30µm Filcon cell-strainer and centrifuged for 5 minutes at 300g. Cells were re-suspended in 1X CMFB + 1:1000 LIVE/DEAD Dye (Life Technologies L34975) and incubated on ice for 20 minutes, then centrifuged for 5 minutes at 300g and resuspended in 1mL PBS + RNase Inhibitors (Invitrogen AM2694, Qiagen Beverly Inc Y9240L) and fixed as previously described for split-pool ligation-based RNA-

seq (22). Cells were centrifuged for 10 minutes at 1000g, re-suspended in 1mL 5 μ g/mL Hoechst 33342 in 0.5X PBS + RI, incubated at RT for 10 minutes, and stored on ice prior to FACS.

Image cytometry

Image cytometry was used to determine the penetrance and localization of SPLiT-seq reagents in dissociated and fixed planarian cells. Linker oligonucleotides tagged with Atto dyes were purchased from IDT to visualize round 2 and round 3 barcoding reagents. In addition, biotin tagged barcoded molecules were visualized using a FITC-tagged streptavidin (Extended Data Fig. 2A). At each successive step of oligonucleotide barcoding, samples were analyzed by imaging flow cytometry to assess by signal intensity of labeled oligonucleotides whether reagent penetration had occurred in dissociated intact cells and/or in cellular debris.

Data was acquired on an ImageStream^X Mk II instrument at 60X running at the low/sensitive flow rate. Channels 1, 2, 7, 10, 11 and 12 were used to collect signal from BF, FITC, Hoechst, Atto 565, Atto 633, and live/dead780 respectively. Bandpass filters for probes were 528/55 (Ch2), 457/45 (Ch7), 610/30 (Ch10), 702/85 (Ch11), and 762/35 (Ch12). Following imaging, spectral compensation was performed using files collected from single-stained samples for each fluorophore and spillover coefficients were calculated in IDEAS software v6.2.189. Analysis for the presence of probe intensity above background levels was also performed in IDEAS by setting gating regions, first on Hoechst positive and brightfield area dot plots to filter out non-cellular events lacking a DNA signal, then by setting gating regions on negative control samples (no probe or no secondary stain) and reporting the frequency of events above the level of negative control samples.

Methods for split-pool ligation-based RNAseq (SPLiT-seq) of fixed planarian cells

Dissociated planarian cells stained with the Near-IR LIVE/DEAD Dye and Hoechst 33342 were sorted with an influx sorter using a 100 μ m tip. Hoechst was excited with 100mW of 355nm UV and collected behind a 450/20 bandpass. Near-IR LIVE/DEAD Dye was excited with 631nm and collected behind a 750LP filter. Sorting was done in purity mode at approximately 12,000 – 15,000 eps. For each sample, 400,000 – 500,000 single Hoechst⁺/DeadDye⁻ cells were collected for barcoding. Split-pool ligation-based barcoding of dissociated cells was performed as previously described (22), then centrifuged at 1000g for 10 minutes in a swinging bucket centrifuge and re-suspended in 1X PBS + SUPERase[•]In RNase Inhibitor (Invitrogen AM2694). All oligonucleotides and primers sequences were ordered from IDT with the sequences previously reported (22). Single, intact Hoechst⁺ cells were sorted with an influx sorter using a 100 μ m tip and 150,000 – 200,000 cells were re-suspended in 400 μ L 1X PBS + SUPERase[•]In RNase Inhibitor. Cells were split into 8 sub-libraries, with 50 μ L of suspension going into each tube. 50 μ L lysis buffer (20mM Tris, pH 8.0, 400mM NaCl, 100mM EDTA, pH 8.0, 4.4% SDS) and 10 μ L proteinase K (20mg/mL) was added to the barcoded cells. Lysates were incubated at 55°C for 2 hours with 200rpm agitation, then stored at –70°C until library preparation.

Library preparation and sequencing

Cell lysates were thawed and sequencing libraries generated according to the SPLiT-seq protocol with some modifications. (22, v3.0). Incubation plus agitation steps were performed at room temperature with 500rpm agitation on a thermomixer (with the final, 42°C, template switching incubation agitated at 300rpm). Following template switch clean-up, all steps were conducted in 96-well plate format. Monitored cDNA amplification was stopped once signal left exponential phase (10 cycles), and a SPRI bead clean-up performed using two additional (200µL) ethanol washes. Amplified cDNA was checked for quality and quantity using an Agilent 2100 Bioanalyzer and Invitrogen Qubit Fluorometer and normalized to 600pg per sample for sequencing library preparation using Illumina Nextera XT library preparation kit. The PCR amplification program was modified to incorporate an initial 3-minute 72°C hold for increased yield. Resulting short fragment libraries were assessed for quality and quantity, pooled equal molar in batches of eight, and sequenced on a total of nine High-Output flow cells (three for each regeneration time course) of an Illumina Next Seq 500 instrument using NextSeq Control Software 2.2.0.4 with the following paired read lengths: 66 bp Read1, 6 bp I7 Index and 94 bp Read2. Following sequencing, Illumina Primary Analysis version NextSeq RTA 2.4.11 and bcl2fastq2 v2.20 were run to demultiplex reads for all libraries and generate FASTQ files.

Data processing and alignment of SPLiT-seq dataset and public Drop-seq dataset

Pooled SPLiT-seq libraries were demultiplexed into individual timepoint samples using a custom python script (available on github upon publication). In brief, Read2 contains three rounds of cell barcodes which were designed during library prep and pooling step (UMI: 1–10bp, cell barcode1: 87–94bp, cell barcode2: 49–56bp, cell barcode3: 11–18bp). The UMI and cell barcodes were extracted from raw reads and cell barcode 1 was used to assign each read to the well/sample of origin. Reads were assigned a sample, UMI, and 24bp cell barcode allowing up to one mismatch per cell barcode round, with quality score more than 10. Each timepoint's samples were then processed using Drop-seq tools software (v2.4.0) (<http://mccarrolllab.org/dropseq/>) for barcode extraction, UMI collapse, gene annotation and cell expression matrix generation. All samples were aligned using STAR aligner with the following parameters (--outFilterMatchNmin 0 --outFilterMismatchNmax 10 --outFilterMismatchNoverLmax 0.3 --outFilterMismatchNoverReadLmax 1). Uniquely aligned reads were annotated using Drop-seq tools function TagReadWithGeneExon based on planarian gene interval refFlat file. Final UMI expression matrix was collapsed using Drop-seq tools function DigitalExpression with parameter "EDIT_DISTANCE=1". Raw and processed data files can be retrieved from the GEO database under accession number GSE146685. The public Drop-seq reference datasets used for tissue annotation and LabelTransfer were downloaded via GEO accession GSE111764 and processed using the same functions and parameters used with SPLiT-seq datasets as described above.

Quality filtering, data normalization, and clustering of cells

Low quality cells from the raw sequencing ($n\text{UMI} < 500$ and/or $n\text{Gene} < 75$) were excluded from subsequent analysis. Utilizing R's Seurat 3 Package (68) version 3.1.1, we used SCTransform (69) with the default settings to normalize and variance stabilize the data.

The 3,000 most highly variable genes were used for principal component analysis of the SCTransformed data. 150 principal components (PCs) were calculated using Seurat's RunPCA function. RunUMAP was then used to generate both two-dimensional and three-dimensional UMAP embeddings (70). FindNeighbors was run using 150 PCs to generate a shared nearest neighbor graph and FindClusters (all default settings) was used to cluster the data, resulting in 89 global clusters designated "0" through "88" based on cluster size.

Identifying enriched transcripts and tissue annotation

We used Seurat's FindAllMarkers with default parameters to identify highly variable transcripts that serve as markers for each of the 89 global clusters, resulting in a total of 3,558 markers comprised of 1,015 unique transcripts (some transcripts were markers for more than one cluster). Gene annotations were sourced from the Sánchez Alvarado lab's Rosetta Stone Transcript Mapper (71). The most significant 20 – 40 transcripts enriched in each global cluster were mapped to dd_Smed_v6 or dd_Smed_v4 using the same mapper and their expression and tissue specificity was characterized in published planarian single-cell datasets (23, 24). Most global clusters were easily assigned to a previously described tissue lineage due to a shared gene expression profile and, in many cases, a cluster could be matched 1-to-1 to a cluster or subcluster in a previously published dataset or to a planarian ontology term (72, Table S15).

In addition to manually annotating each global cluster, we used Seurat's TransferData function (69) to classify cells in our dataset by tissue using one of the prior planarian single-cell studies (23) as a reference for tissue annotation. Raw data (GSE111764) was aligned using STAR aligner using the same parameters outlined above and a gene expression matrix generated, again using Drop-seq_tools-1.13. Only cells which could be unambiguously identified in the published Fincher metadata using a combination of cell barcode and the type of tissue sample were retained. Tissue annotations from the published metadata (23) were added to each cell barcode. FindIntegrationAnchors (dims = 1:30) was used to identify "anchor" cells, then IntegrateData (dims = 1:30) was used to integrate the published dataset with our full SPLiT-seq dataset. PCA and UMAP embedding were performed with 30 PCs. FindTransfer Anchors (dims = 1:30) found transfer anchors using the published data (23) as the reference and our dataset as the query. TransferDataset used the identified transfer anchors and the published metadata's tissue annotations to predict tissue annotations for all cells in our SPLiT-seq dataset. The manual and TransferData tissue annotations were combined to produce a predicted tissue type for each cluster, which was then added to the gene expression matrix metadata for each cell in that cluster.

Interestingly, global cluster 1 contained very few significantly enriched genes and most enriched genes were also highly expressed in cells from the epidermis, muscle, or phagocytic cell lineages. In prior work, clusters lacking specific markers were labeled an artifact and excluded from subsequent analysis (23). However, we found that cells in cluster 1 were not uniformly distributed across all samples. Instead, they were more common in early time points in all three regeneration time courses, suggesting that the transcriptional state might be a wound-induced state characterized by a lack of markers. We therefore included cluster 1 in the full dataset and labeled it 'non-differentiated,' but cells from this

cluster were not assigned to any tissue subset or included in analysis of cells/subclusters from known tissue lineages.

Subclustering of tissue subsets

The full dataset was broken into subsets by tissue type for tissue-specific subclustering. For each tissue subset, the raw RNA counts were normalized with SCTransform prior to performing PCA. Following normalization, PCA was performed with the number of PCs determined on a tissue-by-tissue basis. The global clusters combined into each tissue subset and the PC parameters for tissue level subclustering are summarized in the Table S3. As described for the full dataset, we generated both two- and three-dimensional UMAP embeddings and a shared nearest neighbor graph using the same number of PCs as in the principal component analysis, then clustered using the default settings. We then used FindAllMarkers with default settings to find markers for each of the tissue-level clusters (“subclusters”). This resulted in a total of 211 subclusters (18 phagocytic/Cathepsin⁺, 22 epidermis, 19 intestine, 20 muscle, 44 neural, 36 parenchymal, 8 pharyngeal, 17 protonephridia, 12 stem cell, and 15 non-differentiated). The tissue subcluster assignment of each cell in the dataset was added to the gene expression matrix metadata and is listed in Table S1.

Proportional down-sampling for global data visualization

The number of cells in each sample in the dataset varies from 3,461 to 36,344. While relative abundance of cells in each cluster could be normalized by sample quite easily in tissue proportion heat plots, this was more difficult in UMAP plots. To mitigate this difficulty, we created a proportionately down-sampled dataset with a target maximum of 6000 cells per sample. Both the un-irradiated and sub-lethally irradiated day 0 samples had fewer than 6000 cells, so the entire sample was included. For each of the other samples, we calculated the relative proportion of each global cluster and multiplied by 6000 to determine the number of cells to retain from each cluster. The desired number of cells was then randomly selected from among the cells in that cluster in the given sample. When needed, we rounded all cell counts up to the next equal or greater integer. This sub-sampled dataset was used as the reference object for most global UMAP plots (Fig. 3A–C, Tissue highlight plots, Extended Data Fig. 6A). All other plots (cluster proportion, scaled gene expression, etc.) were produced from calculations on the full dataset containing all 299,998 cells.

PAGA Methods

We chose to use partition-based graph abstraction (PAGA) which uses fewer assumptions and determines connectivity between clusters of cells, making it more robust to incomplete sampling of possible transcriptional states (54). For each tissue we performed pairwise comparison of expression between all tissue subclusters and all of the stem cell subclusters. The stem cell cluster that showed the highest similarity to any tissue subcluster was chosen as the “progenitor” cluster for that tissue and those cells were added to the tissue-specific data subset (stem cells-5 for epidermis and intestine, stem cells-2 for muscle). Tissue subset PCA embeddings (using the number of PCs from tissue subclustering section above) and subcluster assignments were exported from Seurat 3 and read into PAGA. To visualize the connectedness of subclusters/partitions, we used PAGA’s native plotting (Fig. 5A–C). In

these plots, nodes represent subclusters/partitions, with node size indicating the number of cells in that subcluster/partition. Black lines represent the highest-confidence minimal spanning tree (fewest edges needed to connect all nodes) starting from a randomly selected root cell in the stem cell subcluster.

Cloning of *S. mediterranea* transcripts

The Sánchez Alvarado lab consensus transcriptome was used as the reference sequence for designing PCR primers to amplify transcripts of interest. Primer3 was used to design PCR primers to amplify target regions 400 – 500 nucleotides in length from cDNA template. Overhangs homologous to the pPR-T4P vector were added to 5' ends of each primer. PCR products were inserted into linearized vector using Gibson Assembly. Reactions were transformed directly into *Escherichia coli* strain HT115. Insert sequences were verified by sequencing. SMEDIDs for all cloned genes are listed in Tables S16 and S17 and sequence information can be found at <https://planosphere.stowers.org/find/genes>.

Colchicine treatment of *S. mediterranea*

Asexual C4 planarian worms were treated with colchicine (SigmaAldrich C9754) as previously described (44). In brief, colchicine powder was dissolved at room temperature in planarian water at dilutions ranging from 0.0% – 0.25% colchicine, by weight. Planarian water was replaced by colchicine dilutions and animals were then immediately bisected. Regenerating fragments were left in colchicine dilutions for 24 – 48 hours, as needed, washed twice with planarian water, then fixed for subsequent experiments.

In situ hybridizations and antibody staining

For RNA expression analysis in whole planarians, NBT/BCIP and Fluorescent *in situ* hybridizations were performed as previously described (73). Following hybridization, DIG-probes were detected with antibodies in MABT containing 5% horse serum appropriate for NBT/BCIP *in situ* hybridization (Roche 16646822, anti-DIG-AP, 1:1000) or for FISH (Roche 28557000, anti-DIG-POD, 1:1000 or Roche 20052220, anti-FLCN-POD, 1:1000). NBT/BCIP developed animals were stored in 80% glycerol overnight prior to mounting, while FISH animals were soaked overnight in a modified ScaleA2 solution (40% glycerol, 2.5% DABCO, 4M urea) prior to mounting and imaging. Anti-phospho-histone H3 (Ser10) (H3P) antibody (1:500, abcam, ab32107) and Alexa-conjugated goat anti-rabbit secondary antibodies (1:1000, abcam, ab150086) were used to stain mitotic cells.

RNAi food preparation and feeding

Cloned gene vectors transformed into *E. coli* strain HT115 were cultured in 24-well, round-bottom culture plates in 2X YT bacterial growth media with 50µg/mL kanamycin and 10µg/mL tetracycline for 16 – 18 hours at 37°C. Production of dsRNA was induced by adding 6mL 2X YT bacterial growth media with 50µg/mL kanamycin, 10µg/mL tetracycline, and 1mM IPTG. Bacteria were cultured for an additional 4 hours at 37°C following induction. Beef liver was homogenized by adding 800µL 0.2X food coloring in Montjuic Salts to 2g beef liver puree, followed by pipetting until homogeneous. Cultures were centrifuged 10 minutes at 1500g, supernatant removed, and bacterial pellets were

resuspended in 60µL homogenized beef liver. Planarians were fed RNAi food four times with two days between each feeding. Water exchanges were performed after each feeding and dish exchanges along with full water exchanges performed 24 hours prior to each new feeding.

Phenotype quantitation

Following RNAi feedings, animals were moved to clean dishes with fresh Montjuic solution + 50µg/mL gentamycin. In the primary RNAi screen, animals were imaged 7 days post feeding and all surviving animals were bisected through the pharyngeal region to produce anterior and posterior fragments. Regeneration phenotypes were visually inspected and scored 7 days post amputation and animals were imaged and scored 14 days post amputation. For the secondary RNAi screen, the number of animals was doubled (n = 20). RNAi depletions that had produced lesions by 3 days post feeding in the primary screen were assigned to survival curve assays, where animals were monitored daily after the fourth feeding and lysis events noted until all animals were dead. A log-rank (Mantel-Cox) test was used to compare survival of all RNAi conditions to the *unc-22* RNAi control. Genes that produced regeneration phenotypes or homeostatic defects only after amputation were assigned to regeneration assays. In regeneration assays, animals were imaged 7 days post feeding and 10 animals were bisected to produce anterior and posterior fragments, while 10 were left intact. At 14 days post amputation or 21 days post feeding, homeostatic and regeneration phenotypes were imaged and scored. Mann-Whitney U tests were used to compare RNAi conditions to *unc-22* RNAi controls. Full results of primary and secondary screens are provided in Supplementary Tables S16 and S17.

Microscopy

Colorimetric whole-mount *in situ* hybridization samples and live worm or fragment images were acquired using a Leica M205 microscope using the Leica Application Suite (LASX). Following image acquisition, non-tissue background was subtracted, contrast and image intensity adjusted, and the edited image was converted to grayscale for data presentation. No quantifications were performed on contrast or intensity adjusted images and all raw, original data is available in the original data repository (<http://www.stowers.org/research/publications/libpb-1513>). Confocal images of fluorescent *in situ* hybridization samples were acquired using an LSM-700 inverted confocal microscope with Zeiss Zen Black Software (v8.1) or via high throughput imaging on a Nikon Eclipse Ti with a Yokogawa W1 spinning disk and robotic plate loader. With automated image capture, whole slides were imaged at 4X and objects of interest were automatically detected, re-imaged at 10X, then batch stitched, quantified and aligned using Fiji macros (<https://github.com/jouyun/smc-macros>). Stitching of tiles for whole worm images was performed with Fiji plugins (grid/collection stitching) using custom macros for batch processing. Maximum intensity projections of the stitched z-stacks were generated to visualize gene expression across the entire animal, then rotated, and cropped for data presentation.

Quantitation of fluorescent *in situ* images

TRACS quantitation: Confocal z-stacks of wound-site adjacent tissues (anterior tissues in regenerating tail fragments or posterior tissue in regenerating head fragments) were acquired using an LSM-700 inverted confocal microscope. Fiji ROI manager tool was used to mark all nuclei expressing either the TRACS or tissue marker with an 8µm diameter circular ROI (Extended Data Fig. 7B) and ROI measurements was used to quantify mean intensity of all channels in the confocal image. Raw mean intensity values were plotted and a threshold pixel intensity for a positive or negative cell was set. The same threshold value was used for all treatments in an experiment. The percentage of positive cells was calculated for each tissue section (1 per animal) and an unpaired two-sided t-test was used to compare the percentage of positive cells between treatments.

Cephalic ganglia size: The Fiji straight line tool was used to measure anterior to posterior length of the cephalic ganglia and total body length of each animal and the ratio of cephalic ganglia length/body length was reported. An unpaired two-sided t-test was used to compare RNAi-treated animals to *unc-22* RNAi-treated control animals.

Anterior and posterior poles: Images for each animal were rotated to align AP animal axis with Y axis of confocal image. Fiji multi-point tool was used to mark center of all *notum*⁺ or *wnt1*⁺ cells through confocal z-stack, as well as the most anterior or posterior point of animal body wall. Raw coordinate positions were plotted in GraphPad prism and absolute values of y coordinates used to calculate the distance from the pole. An unpaired two-sided t-test was used to compare the number of *notum*⁺ or *wnt1*⁺ cells per animal in RNAi-treated animals vs. *unc-22* RNAi-treated animals.

H3P quantitation: Imaging and quantitation was performed as detailed previously (44). An unpaired two-sided t-test was used to calculate the statistical difference between the number of H3P⁺ cells in whole animals or amputated fragment. Colchicine-treated animals were compared to 0% colchicine controls, while RNAi-treated animals were always compared to *unc-22* RNAi-treated animals.

Statistics and Reproducibility

No statistical method was used to predetermine sample size in the reported experiments. For RNAi primary and secondary screens (Tables S16 and S17), animals were randomly assigned to treatment conditions and investigators were blinded to allocation during experiments and phenotype assessment. For phenotype characterization after initial screening, the investigators were not blinded to allocation during experiments and outcome assessment. No data were excluded from the analyses. Data analysis and statistical tests were performed using Microsoft Excel (v16.36) and Graph Prism (v8.3.0). Cytometry data collection and analysis was performed with BD FACSDiva Software (v8.0).

The single cell sequencing data presented in the study was a single experiment and was not independently replicated due to resource constraints. Gene expression patterns of subcluster enriched genes shown by *in situ* hybridization (Figures 1D, 2E, Extended Data 3C, F, I; 4D, H; 5C, F, I; 8G, K, 9G, 10G) are representative of 4–5 animals from a single experiment.

Experiments related to the visualization and characterization of TRACS (Figures 4E, 5E, Extended Data 7A) were repeated twice independently with similar results. RNAi phenotyping of survival and regeneration (Figures 6B, C; 7C, D; 8B, C, F, G) were repeated at least three times with similar results. Characterization of the CNS and gut in RNAi treated animals (Figures 6D, 7C, and 8H) was repeated at least twice independently with similar results. Visualization of *notum*⁺ and *Wnt1*⁺ cells in RNAi treated animals (Figure 5E–K, Extended Data 10I) is representative of 2–6 animals from a single experiment. Quantitation of H3P Density (Figures 7F–G, 8D, J, Extended Data 9H–J, 10H) were repeated at least twice independently with similar results.

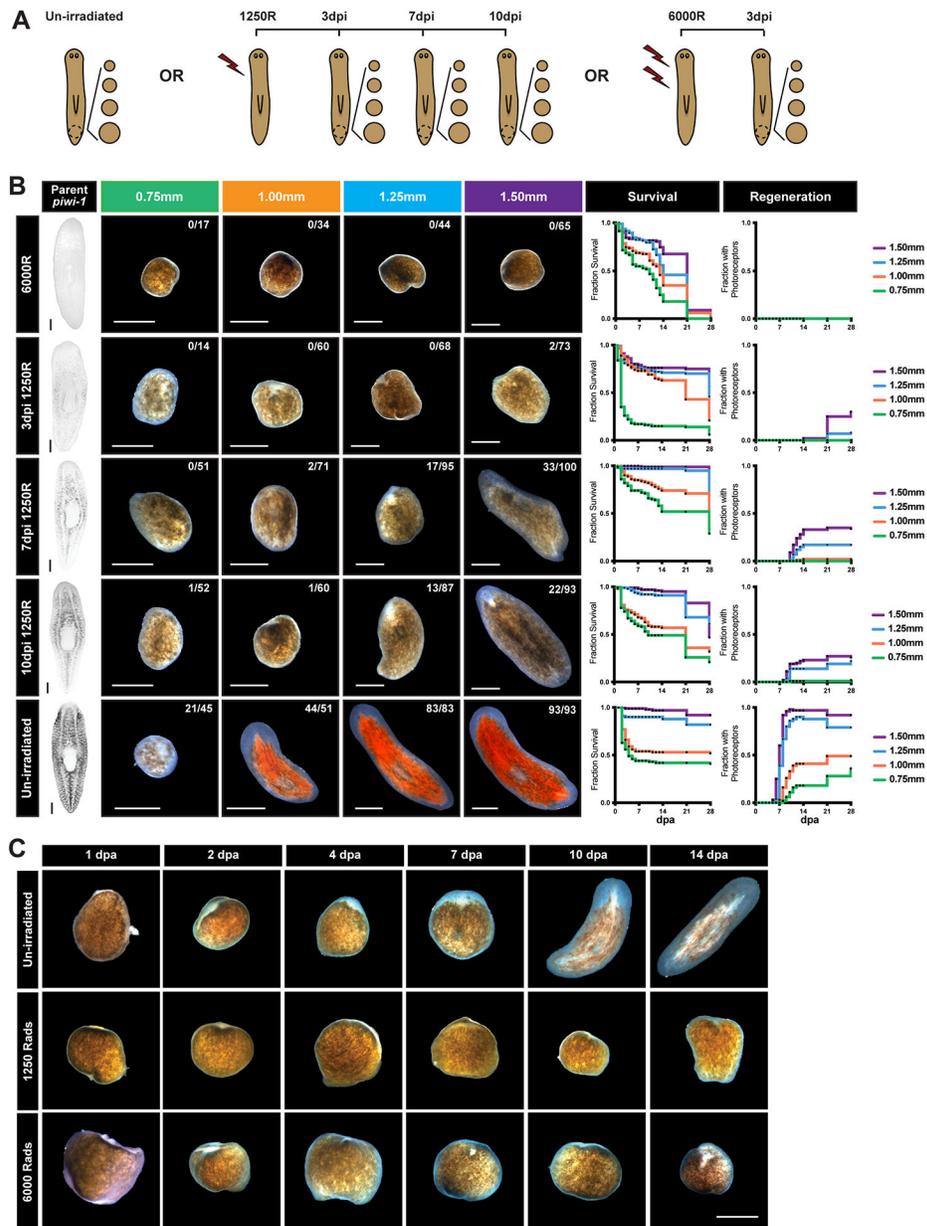
Data Availability

All scRNA-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE146685. scRNA-seq data can also be explored in our shiny app at: https://simrcompbio.shinyapps.io/bbp_app/. Previously published sequencing data that were re-analysed here are available under accession codes GSE111764 (24) and GSE107874 (49). SMEDIDs for all cloned genes are listed in Supplementary Table 15 and sequence information can be found at <https://planosphere.stowers.org/find/genes>. All other data supporting the findings of this study are available from the corresponding author on reasonable request or can be accessed from the Stowers Original Data Repository at <http://www.stowers.org/research/publications/libpb-1513>.

Code Availability:

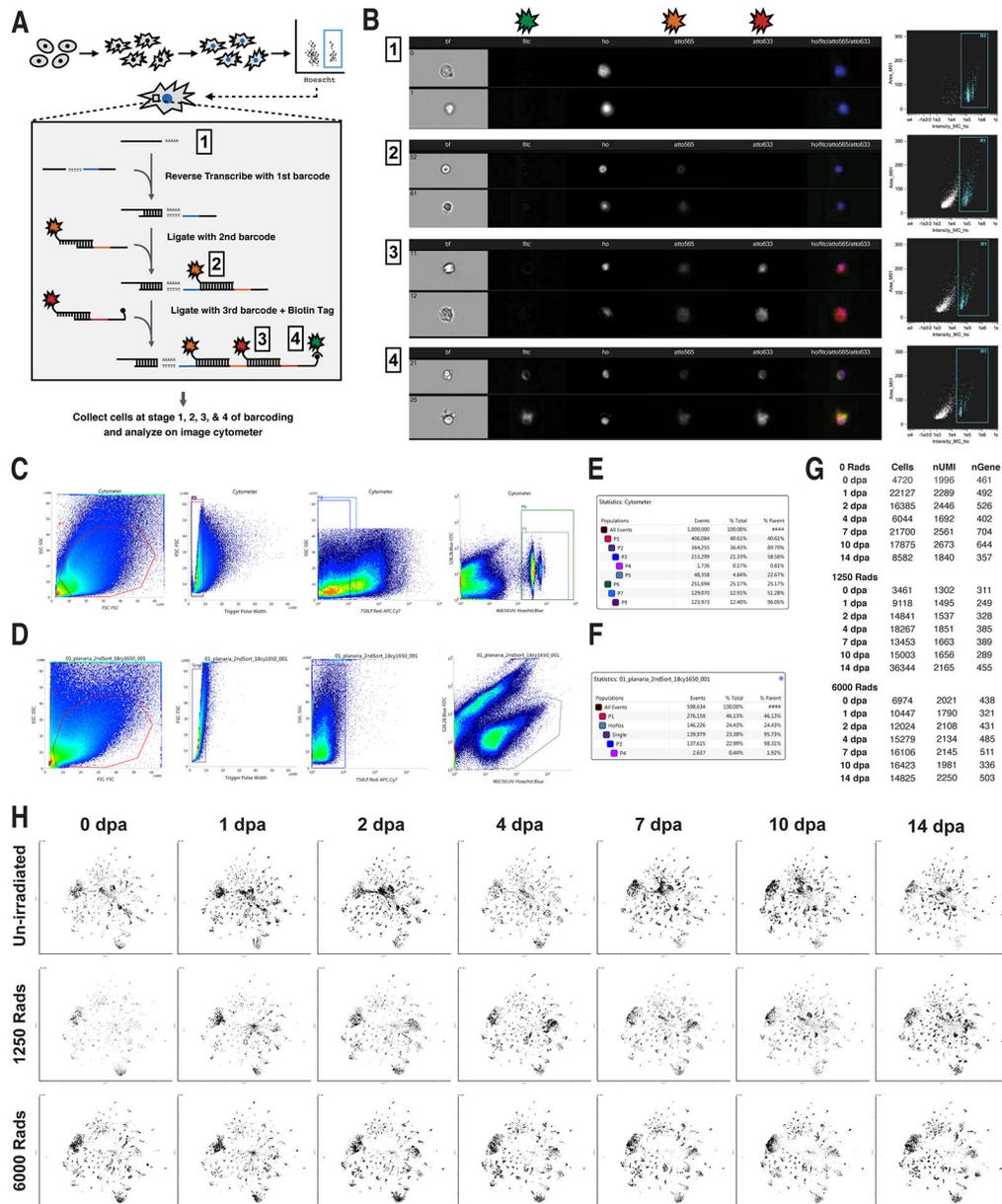
Original scripts used for the analysis and visualization of single-cell sequencing data is available at <https://github.com/0x644BE25/smedSPLiT-seq>.

Extended Data

**Extended Data Fig. 1. Identification of optimal biopsy sizes and treatments for single cell reconstruction.**

(A) Schematic of experimental design. (B) *Piwi-1* expression of parent animals at time biopsy was taken after irradiation treatment (right), as well as representative images 14 days post amputation (dpa), survival curves, and scoring of regeneration of photoreceptor pigmentation of biopsies 0.75mm – 1.50mm taken from comparable parent animals following irradiation treatment. Notation on representative images indicates number of fragments that regenerated photoreceptors by 14dpa out of total surviving at 14dpa (exact n is provided in source data file). (C) Representative images of biopsies from un-irradiated,

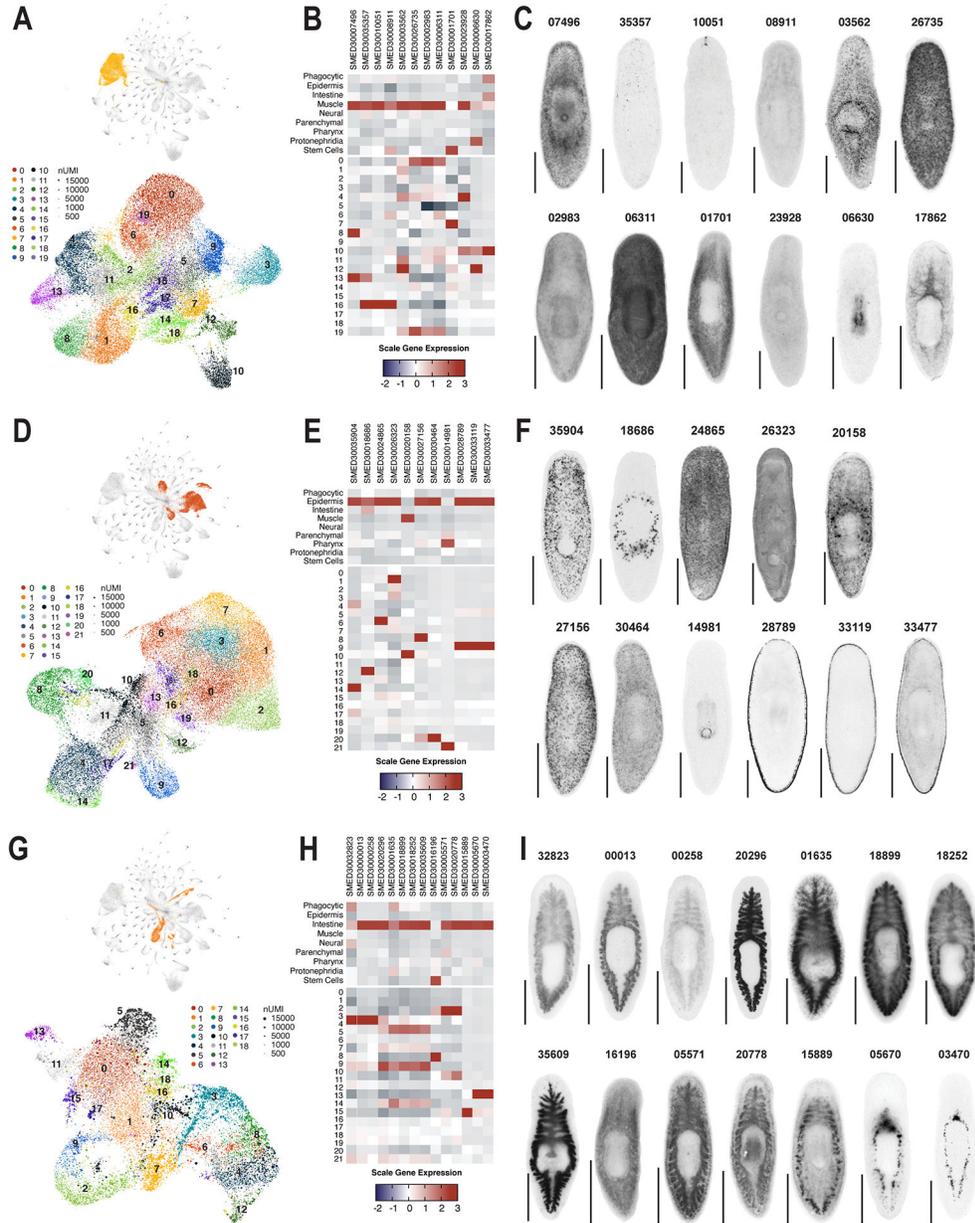
sub-lethally irradiated, and lethally irradiated animals imaged 1, 2, 4, 7, 10, and 14 days post amputation. Scale = 500 μ m.



Extended Data Fig. 2. Optimization and acquisition of a single cell reconstruction of planarian regeneration

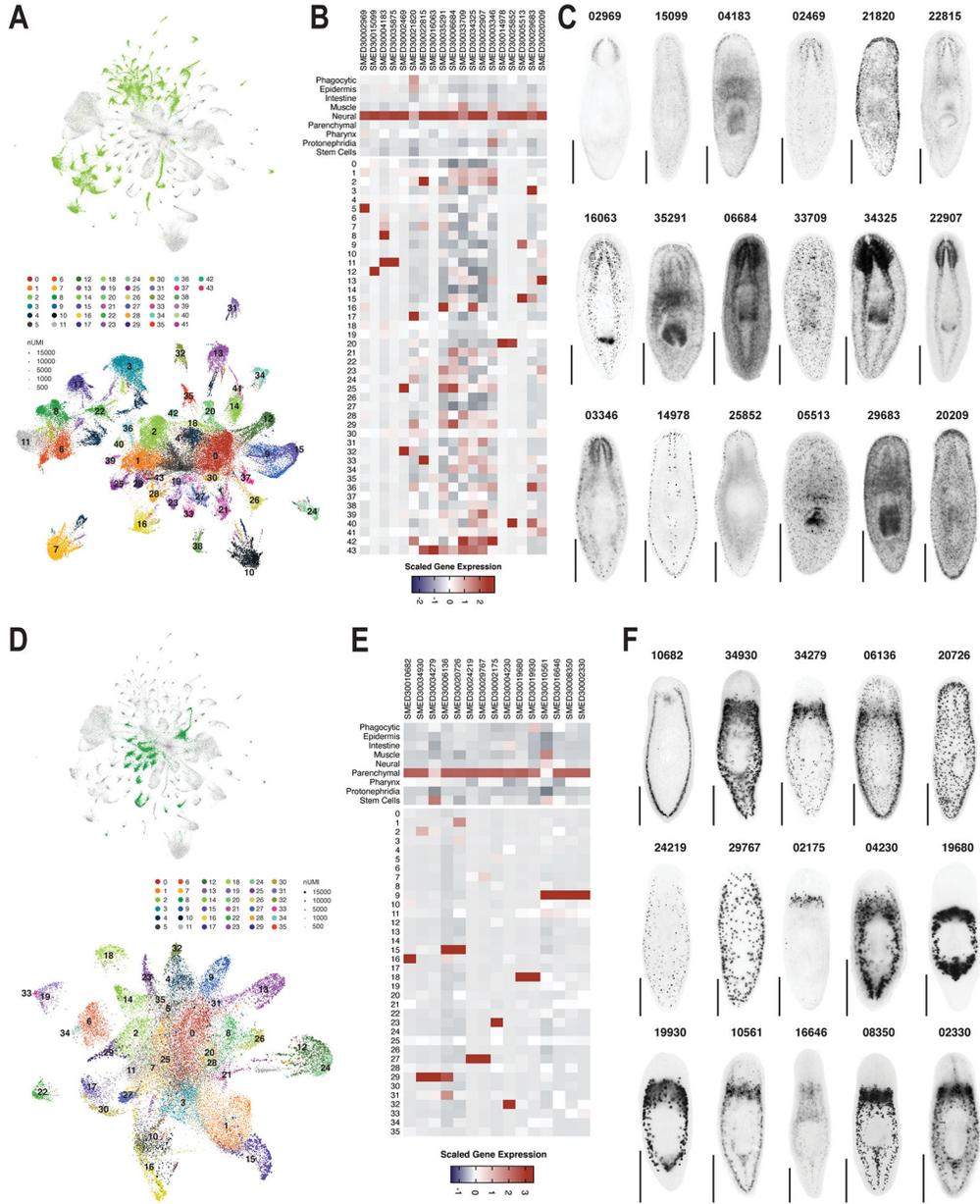
(A) Schematic of experimental design using Atto-conjugated linker molecules to visualize SPLiT-seq reagents after second [2] and third [3] round barcoding, and to detect biotin tagged molecules [4]. (B) Representative images of cells/objects detected in the Hoechst⁺ compartment at all steps of the barcoding process and area vs. Hoechst intensity plots with Hoechst⁺ cell compartment highlighted from each stage of barcoding. Note the accumulation of non-nucleated debris that occurs during the barcoding process that needed to be removed prior to sequencing. As a result, Hoechst⁺ intact cells were sorted following

barcoding using the plot depicted in B [4] as a guide. Gating strategy utilized pre-barcoding (C) and post-barcoding (D). Abundance of sorted population pre-barcoding (E) and post-barcoding (F). (G) Number of cells captured, mean nUMI/cell, and mean nGene/cell for each of the 21 conditions. (H) UMAP embeddings cells sampled from each of the 21 conditions (see materials and methods for sub-sampling methodology) illustrating the change in tissue composition and captured transcriptional states across the dataset.

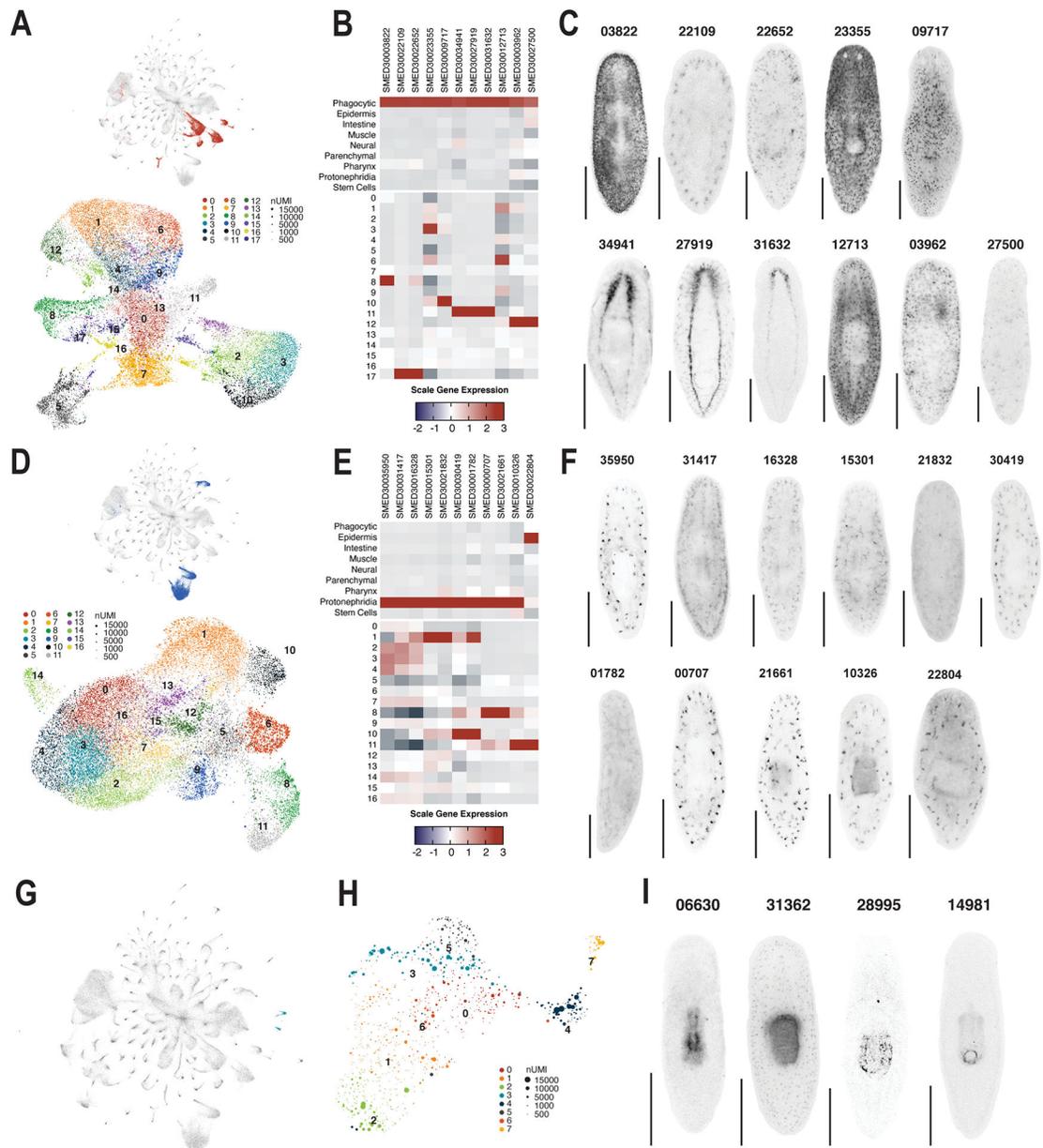


Extended Data Fig. 3. Annotation of muscle, epidermal, and intestinal tissue subclusters. UMAP embedding of global dataset with tissue highlighted and UMAP embedding of tissue cells colored by tissue subcluster ID for muscle (A), epidermis (B), or intestine (C). Scaled mean expression of cluster enriched genes by tissue and tissue subcluster for muscle (D),

epidermis (E), and intestine (F) enriched genes. Whole mount *in situ* hybridization of tissue markers analyzed in D (G), E (H), and F (I). Scale = 500µm.

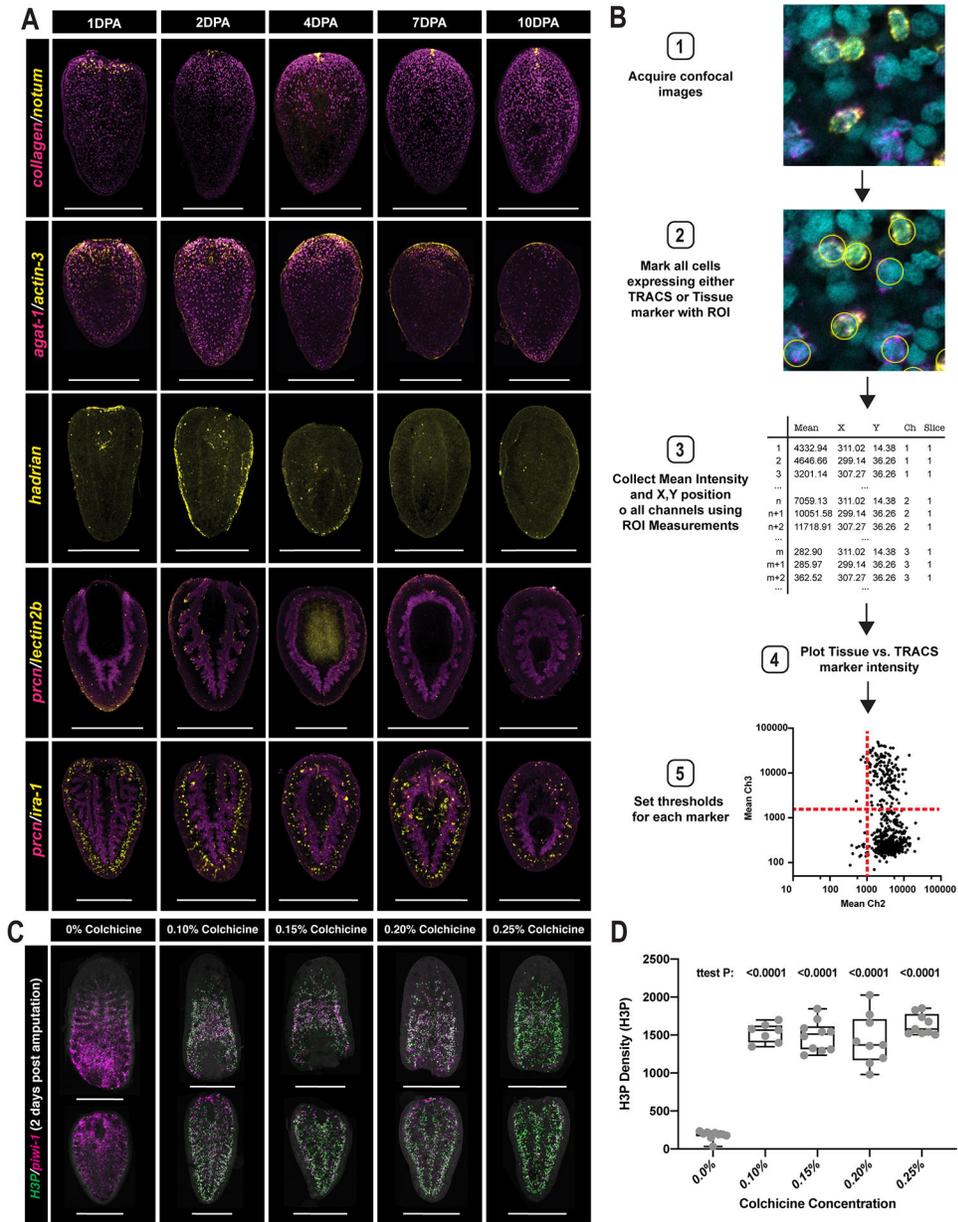


Extended Data Fig. 4. Annotation of neural and parenchymal tissue subclusters. UMAP embedding of global dataset with nervous system (A) or parenchyma (E) highlighted. UMAP embedding of neural (B) or parenchymal (F) cells colored by tissue subcluster ID. Scaled mean expression of cluster neural-enriched (C) or parenchymal-enriched (G) genes by tissue and tissue subcluster. Whole mount *in situ* hybridization of tissue markers analyzed in C (D) or G (H). Scale = 500µm.

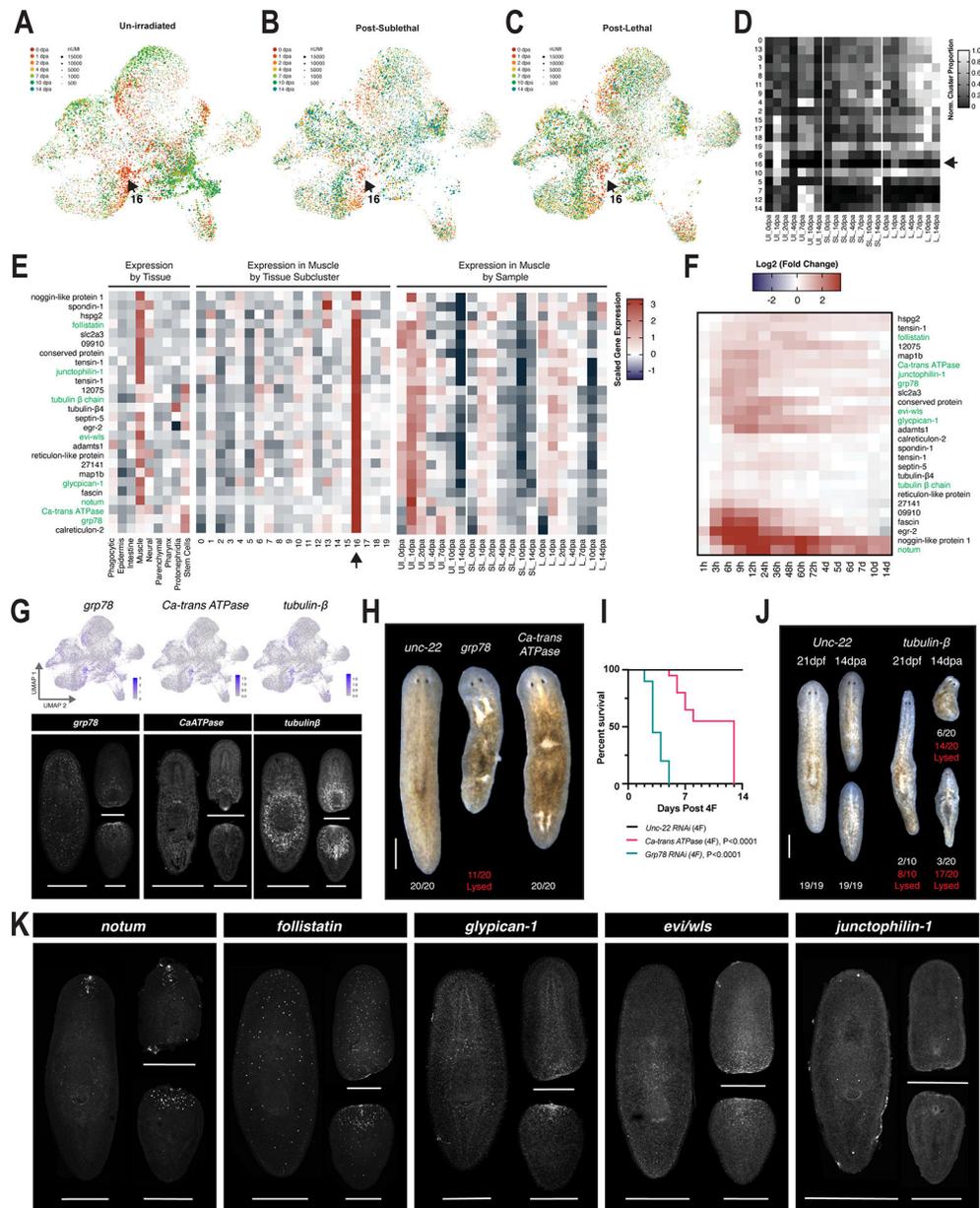


Extended Data Fig. 5. Annotation of phagocytic, protonephridial, and pharyngeal tissue subclusters.

UMAP embedding of global dataset with tissue highlighted and UMAP embedding of tissue cells colored by tissue subcluster ID for phagocytic (A) and protonephridial (D) cells. Scaled mean expression of cluster enriched genes by tissue and tissue subcluster for phagocytic (B) and protonephridia (E) enriched genes. Whole mount *in situ* hybridization of tissue markers analyzed in B (C) and E (F). (G) UMAP embedding of global dataset with pharyngeal clusters highlighted. (H) UMAP embedding of pharyngeal cells. (I) Whole mount *in situ* hybridization of pharynx-enriched genes. Scale = 500 μ m.

**Extended Data Fig. 7. Data supported the visualization and quantitation of TRACS**

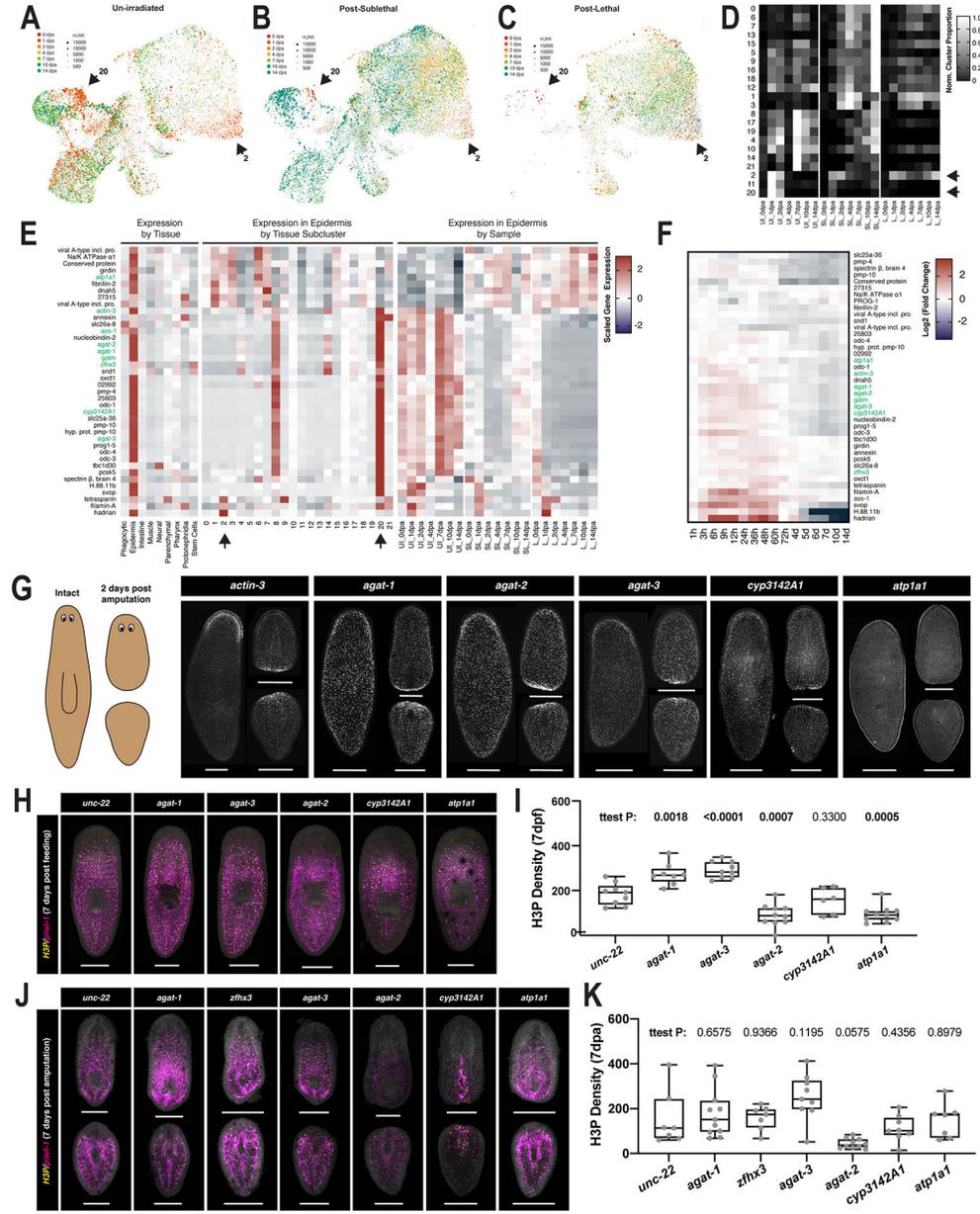
(A) Selected images of whole mount double-fluorescent *in situ* hybridizations of wound-induced genes (yellow) and tissue-specific markers (magenta). (B) Schematic representation of TRACS quantitation strategy. (C) Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and immunohistochemistry of H3P⁺ mitotic cells. (D) Quantitation of density of H3P⁺ cells in control and colchicine-treated animals. $n = 9$ (0%), 7 (0.10%), 10 (0.15%), 9 (0.20%), and 9 (0.25%) biologically independent animals. Data are presented in box plots of min, max, and median, with the bounds of the box at the first and third quartile and whiskers extending from quartile to the minimum or maximum. P values are two-sided unpaired t-tests compared to 0%, with no corrections for multiple comparison. Scale = 500 μ m.



Extended Data Fig. 8. Additional data supporting wound-induced muscle cluster and M16-enriched genes requires for tissue maintenance and regeneration.

UMAP embedding of all muscle cells, colored by time after amputation and split by cells from biopsies taken from un-irradiated (A), sub-lethally irradiated (B), or lethally irradiated (C) animals. (D) Scaled proportion of cells from each muscle subcluster across sampled conditions, normalized to sample in which subcluster had maximum representation. (E) Scaled mean expression of muscle cluster 16 enriched genes (black arrow) by tissue, muscle subcluster, and sample. (F) Gene expression of screened muscle genes in bulk RNA-seq dataset of planarian regeneration (62). (G) UMAP feature plots and gene expression patterns of *grp78*, *CaATPase*, and *tubulin-β* visualized by fluorescent *in situ* hybridization in intact animals and regenerating fragments 2 days post amputation. (H) Representative images RNAi-treated animals 3 days post feeding. (I) Survival of RNAi-treated animals shown in H

(n = 20 for each condition). (J) Representative images of homeostatic (21 days post feeding, 10 animals) and regeneration phenotypes (14dpa, 20 animals) in *unc-22* and *tubulin-β* RNAi-treated animals. (K) Gene expression patterns of M16-enriched genes visualized by fluorescent *in situ* hybridization in intact animals and regenerating fragments 2 days post amputation. P values are log-rank test (I) compared to *unc-22* control. Scale = 500μm.



Extended Data Fig. 9. Additional data supporting amputation-specific epidermal clusters and epidermal genes requires for stem cell proliferation. UMAP embedding of all epidermal cells, colored by time after amputation and split by cells from biopsies taken from un-irradiated (A), sub-lethally irradiated (B), or lethally irradiated (C) animals. (D) Scaled proportion of cells from each epidermal subcluster across sampled conditions, normalized to sample in which subcluster had maximum representation.

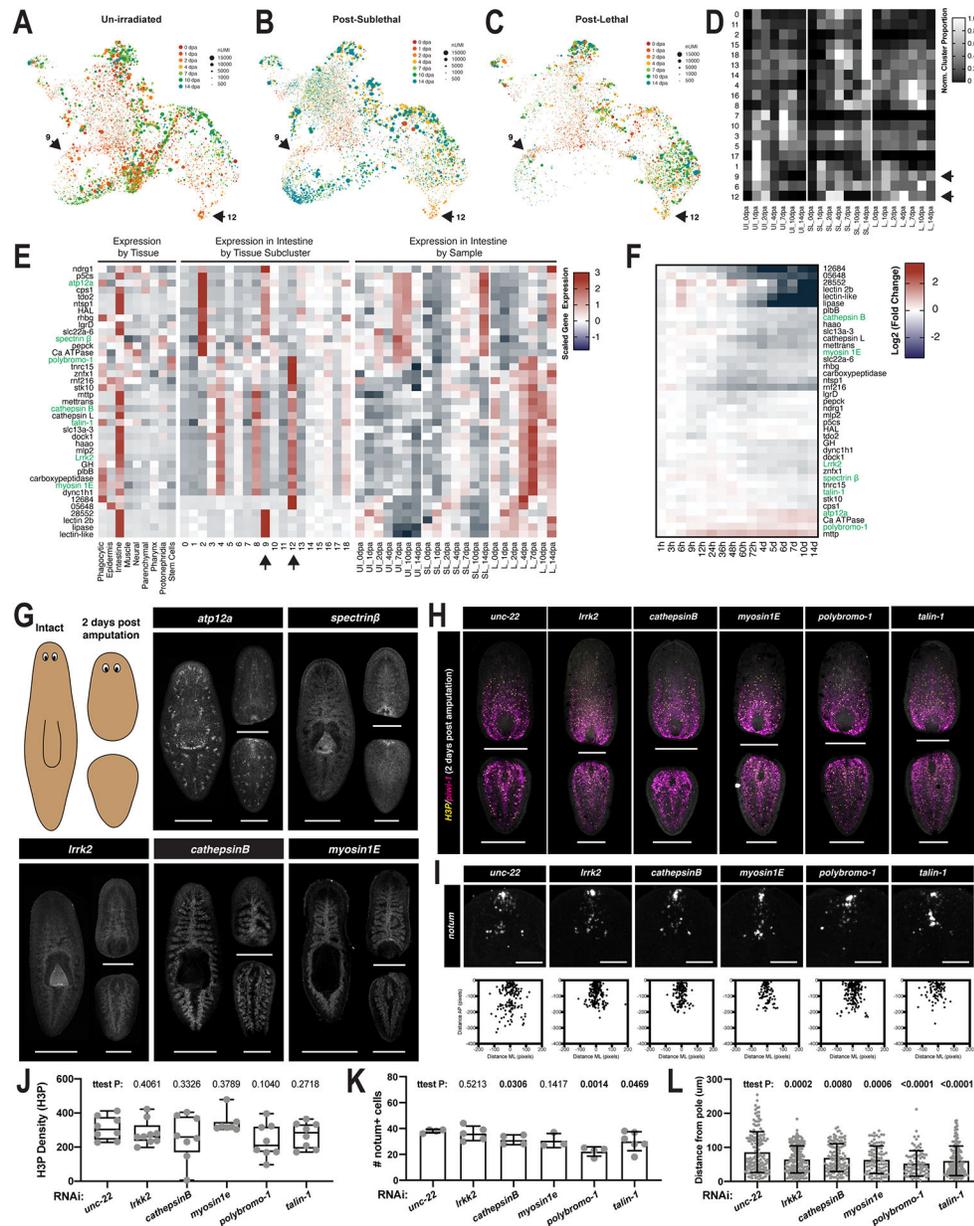
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(E) Scaled mean expression of E2/E20 enriched genes (black arrows) by tissue, muscle subcluster, and sample. (F) Gene expression of screened epidermal genes in bulk RNA-seq dataset of planarian regeneration (62). (G) Gene expression patterns of epidermal genes visualized by fluorescent *in situ* hybridization in intact animals and regenerating fragments 2 days post amputation. Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and immunohistochemistry of H3P⁺ mitotic cells in RNAi-treated animals 7 days post feeding (H), or 7 days post amputation (J). H3P⁺ cell density in RNAi-treated animals 7 days post feeding (I) or 7 days post amputation (K). n (7dpf) = 9 (*unc-22*), 8 (*agat-1*), 9 (*agat-3*), 10 (*agat-2*), 6 (*cyp3142a1*), and 10 (*atp1a1*) biologically independent animals. n (7dpa) = 7 (*unc-22*), 11 (*agat-1*), 7 (*zfhx3*), 9 (*agat-3*), 10 (*agat-2*), 8 (*cyp3142a1*), and 7 (*atp1a1*) biologically independent animals. Data are presented in box plots of min, max, and median, with the bounds of the box at the first and third quartile and whiskers extending from quartile to the minimum or maximum. P values are two-sided unpaired t-tests with no corrections for multiple comparison compared to *unc-22* control. Scale = 500µm.



Extended Data Fig. 10. Additional data supporting amputation-induced intestinal clusters and intestinal genes required for tissue homeostasis and regeneration.

UMAP embedding of all intestinal cells, colored by time after amputation and split by cells from biopsies taken from un-irradiated (A), sub-lethally irradiated (B), or lethally irradiated (C) animals. (D) Scaled proportion of cells from each intestinal subcluster across sampled conditions, normalized to sample in which subcluster had maximum representation. (E) Scaled mean expression of 9/112-enriched genes (black arrows) by tissue, muscle subcluster, and sample. (F) Gene expression of screened intestinal genes in bulk RNA-seq dataset of planarian regeneration (62). (G) Gene expression patterns of intestinal genes visualized by fluorescent *in situ* hybridization in intact animals and regenerating fragments 2 days post amputation. (H) Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and immunohistochemistry of H3P⁺ mitotic cells in RNAi-treated animals 2 days

post amputation. (I) Selected images and raw x,y position of *notum*⁺ anterior pole cells visualized by whole mount *in situ* hybridization. (J) H3P⁺ cell density in RNAi-treated animals 2 days post amputation. n = 8 (*unc-22*), 9 (*Irrk2*), 8 (*cathepsinB*), 7 (*myosin1e*), 8 (*polybromo-1*), and 8 (*talin-1*) biologically independent animals. (K) Number of *notum*⁺ cells in RNAi treated animals. n = 4 (*unc-22*), 5 (*Irrk2*), 4 (*CathepsinB*), 3 (*myosin1e*); 4 (*polybromo-1*) and 6 (*talin-1*) biologically independent animals. (L) Distribution of *notum*⁺ cells in RNAi-treated animals, n = 152 (*unc-22*), 180 (*Irrk2*), 125 (*cathepsinB*), 92 (*myosin1e*); 88 (*polybromo-1*), and 180 (*talin-1*) *notum*⁺ cells.. P values are a two-sided unpaired t-tests, with no adjustments for multiple comparisons. Data are presented in box plots of min, max, and median, with the bounds of the box at the first and third quartile and whiskers extending from quartile to the minimum or maximum (j) or as mean values +/- SD (k,l). Scale = 500µm (G,H) or 150µm (I).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Sean McKinney for assistance with automated confocal imaging. We thank members of the ASA laboratory for discussion and advice. We are grateful to the Stowers cytometry and molecular biology core facilities for technical contributions and methods development.

Funding:

ASA is an investigator of the Howard Hughes Medical Institute (HHMI) and the Stowers Institute for Medical Research. BWBP is a Jane Coffin Childs Memorial Fund Postdoctoral Fellow. FGM is a HHMI Postdoctoral Fellow. This work was supported in part by NIH R37GM057260 to ASA.

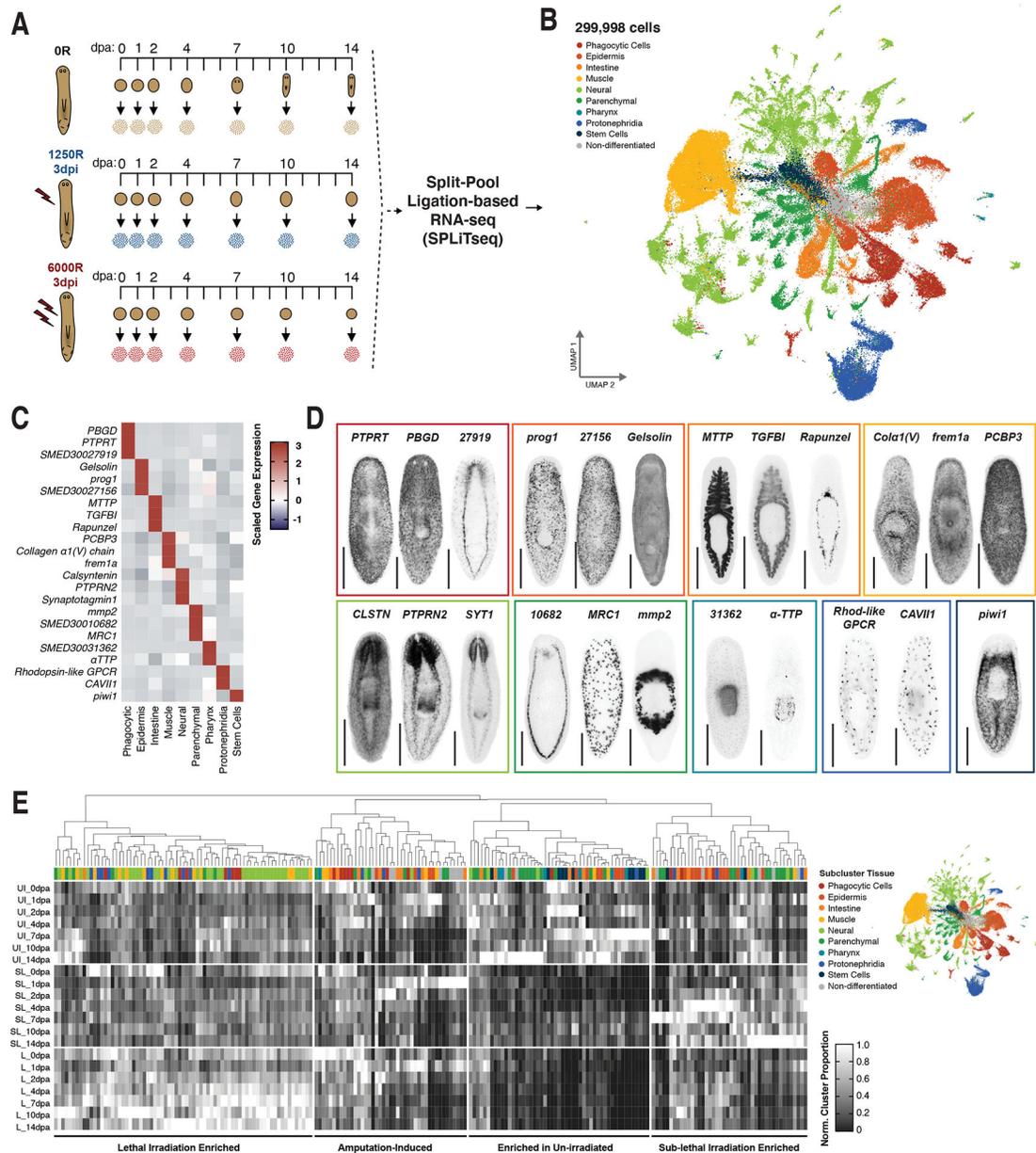
References:

- Gierer A et al. , Regeneration of Hydra from Reaggregated Cells. *Nature New Biol.* 239, 98–101 (1972). [PubMed: 4507522]
- Elliott SA, Sánchez Alvarado A, The history and enduring contributions of planarians to the study of animal regeneration. *Wiley Interdiscip Rev Dev Biol.* 2, 301–326 (2013). [PubMed: 23799578]
- Reddien PW, The Cellular and Molecular Basis for Planarian Regeneration. *Cell.* 175, 327–345 (2018). [PubMed: 30290140]
- Layden MJ, Rentzsch F, Röttinger E, The rise of the starlet sea anemone *Nematostella vectensis* as a model system to investigate development and regeneration. *Wiley Interdiscip Rev Dev Biol.* 5, 408–428 (2016). [PubMed: 26894563]
- Laird DJ, De Tomaso AW, Weissman IL, Stem Cells Are Units of Natural Selection in a Colonial Ascidian. *Cell.* 123, 1351–1360 (2005). [PubMed: 16377573]
- Kragl M et al. , Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature.* 460, 60–65 (2009). [PubMed: 19571878]
- Gerber T et al. , Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science.* 362, eaaq0681 (2018). [PubMed: 30262634]
- Gemberling M, Bailey TJ, Hyde DR, Poss KD, The zebrafish as a model for complex tissue regeneration. *Trends Genet.* 29, 611–620 (2013). [PubMed: 23927865]
- Yanger K et al. , Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes Dev.* 27, 719–724 (2013). [PubMed: 23520387]
- Beumer J, Clevers H, Regulation and plasticity of intestinal stem cells during homeostasis and regeneration. *Development.* 143, 3639–3649 (2016). [PubMed: 27802133]

11. Seifert AW et al. , Skin shedding and tissue regeneration in African spiny mice (*Acomys*). *Nature*. 489, 561–565 (2012). [PubMed: 23018966]
12. Morrison SJ, Spradling AC, Stem Cells and Niches: Mechanisms That Promote Stem Cell Maintenance throughout Life. *Cell*. 132, 598–611 (2008). [PubMed: 18295578]
13. Blanpain C, Fuchs E, Plasticity of epithelial stem cells in tissue regeneration. *Science*. 344, 1242281–1242281 (2014). [PubMed: 24926024]
14. Guedelhofer OC4, Sánchez Alvarado A, Amputation induces stem cell mobilization to sites of injury during planarian regeneration. *Development*. 139, 3510–3520 (2012). [PubMed: 22899852]
15. Chera S, Ghila L, Wenger Y, Galliot B, Injury-induced activation of the MAPK/CREB pathway triggers apoptosis-induced compensatory proliferation in hydra head regeneration. *Dev Growth Differ*. 53, 186–201 (2011). [PubMed: 21338345]
16. Rinkevich Y et al. , Repeated, Long-Term Cycling of Putative Stem Cells between Niches in a Basal Chordate. *Dev Cell*. 24, 76–88 (2013). [PubMed: 23260626]
17. Lengfeld T et al. , Multiple Wnts are involved in Hydra organizer formation and regeneration. *Dev Biol*. 330, 186–199 (2009). [PubMed: 19217898]
18. Morgan TH, *Regeneration*. Macmillan, New York (1901), (available at <https://archive.org/details/regeneration00morggoog>).
19. JR M, Coward SJ, On the minimal size of a planarian capable of regeneration. *Trans Am Microsc Soc*. 93, 386–391 (1974). [PubMed: 4853459]
20. Coward SJ, The relation of surface and volume to so-called physiological gradients in planaria. *Dev Biol*. 18, 590–601 (1968). [PubMed: 5710081]
21. Arnold CP, Benham-Pyle BW, Lange JJ, Wood CJ, Alvarado AS, Wnt and TGFβ coordinate growth and patterning to regulate size-dependent behaviour. *Nature*. 7, 1–5 (2019).
22. Rosenberg AB et al. , Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. 360, 176–182 (2018). [PubMed: 29545511]
23. Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien PW, Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. *Science*. 360, eaaq1736 (2018). [PubMed: 29674431]
24. Plass M et al. , Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science*, eaaq1723 (2018). [PubMed: 29674432]
25. Reddien PW, Oviedo NJ, Jennings JR, Jenkin JC, Sánchez Alvarado A, SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science*. 310, 1327–1330 (2005). [PubMed: 16311336]
26. Wagner DE, Ho JJ, Reddien PW, Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. *Cell Stem Cell*. 10, 299–311 (2012). [PubMed: 22385657]
27. Palakodeti D, Smielewska M, Lu Y-C, Yeo GW, Graveley BR, The PIWI- proteins SMEDWI-2 and SMEDWI-3 are required for stem cell function and piRNA expression in planarians. *RNA*. 14, 1174–1186 (2008). [PubMed: 18456843]
28. Wang Y, Stry JM, Wilhelm JE, Newmark PA, A functional genomic screen in planarians identifies novel regulators of germ cell development. *Genes Dev*. 24, 2081–2092 (2010). [PubMed: 20844018]
29. Bansal D et al. , Cytoplasmic poly (A)-binding protein critically regulates epidermal maintenance and turnover in the planarian *Schmidtea mediterranea*. *Development*. 144, 3066–3079 (2017). [PubMed: 28807897]
30. Zhu SJ, Hallows SE, Currie KW, Xu C, Pearson BJ, A mex3 homolog is required for differentiation during planarian stem cell lineage development. *eLife*. 4 (2015), doi:10.7554/eLife.07025.
31. Ma K-X, Chen G-W, Liu D-Z, cDNA cloning of heat shock protein 90 gene and protein expression pattern in response to heavy metal exposure and thermal stress in planarian *Dugesia japonica*. *Mol Biol Rep*. 39, 7203–7210 (2012). [PubMed: 22391652]
32. Cheng X et al. , A planarian RPS3 homolog is critical to the modulation of planarian regeneration. *Gene*. 691, 153–159 (2019). [PubMed: 30615916]
33. Wenemoser D, Reddien PW, Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. *Dev Biol*. 344, 979–991 (2010). [PubMed: 20599901]

34. Petersen CP, Reddien PW, Polarized notum activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science*. 332, 852–855 (2011). [PubMed: 21566195]
35. Gavino MA, Wenemoser D, Wang IE, Reddien PW, Tissue absence initiates regeneration through follistatin-mediated inhibition of activin signaling. *eLife*. 2, e00247 (2013). [PubMed: 24040508]
36. Roberts-Galbraith RH, Newmark PA, Follistatin antagonizes activin signaling and acts with notum to direct planarian head regeneration. *Proc Natl Acad Sci U S A*. 110, 1363–1368 (2013). [PubMed: 23297191]
37. Adell T, Salo E, Boutros M, Bartscherer K, Smed-Evi/Wntless is required for beta-catenin-dependent and -independent processes during planarian regeneration. *Development*. 136, 905–910 (2009). [PubMed: 19211673]
38. Wenemoser D, Lapan SW, Wilkinson AW, Bell GW, Reddien PW, A molecular wound response program associated with regeneration initiation in planarians. *Genes Dev*. 26, 988–1002 (2012). [PubMed: 22549959]
39. Eisenhoffer GT, Kang H, Sánchez Alvarado A, Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. *Cell Stem Cell*. 3, 327–339 (2008). [PubMed: 18786419]
40. Tu KC et al. , Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *eLife*. 4, e10501 (2015). [PubMed: 26457503]
41. Wurtzel O, Oderberg IM, Reddien PW, Planarian Epidermal Stem Cells Respond to Positional Cues to Promote Cell-Type Diversity. *Dev Cell*. 40, 491–504.e5 (2017). [PubMed: 28292427]
42. Forsthoefel DJ, Cejda NI, Khan UW, Newmark PA, Cell-type diversity and regionalized gene expression in the planarian intestine. *eLife*. 9, e52613 (2020). [PubMed: 32240093]
43. Wolf FA et al. , PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biology*. 20, 59 (2019). [PubMed: 30890159]
44. Guo L, Accorsi A et al. An adaptable chromosome preparation methodology for use in invertebrate research organisms. *BMC Biology*. 16, 25 (2018). [PubMed: 29482548]
45. Witchley JN, Mayer M, Wagner DE, Owen JH, Reddien PW, Muscle cells provide instructions for planarian regeneration. *Cell Rep*. 4, 633–641 (2013). [PubMed: 23954785]
46. Petersen CP, Reddien PW, A wound-induced Wnt expression program controls planarian regeneration polarity. *Proc Natl Acad Sci U S A*. 106, 17061–17066 (2009). [PubMed: 19805089]
47. Gurley KA et al. , Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. *Dev Biol*. 347, 24–39 (2010). [PubMed: 20707997]
48. Adler CE, Sánchez Alvarado A, PHRED-1 is a divergent neurexin-1 homolog that organizes muscle fibers and patterns organs during regeneration. *Dev Biol*. 427, 165–175 (2017). [PubMed: 28461239]
49. Zeng A et al. , Prospectively Isolated Tetraspanin⁺ Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration. *Cell*. 173, 1593–1608.e20 (2018). [PubMed: 29906446]
50. Forsthoefel DJ, Park AE, Newmark PA, Stem cell-based growth, regeneration, and remodeling of the planarian intestine. *Dev Biol*. 356, 445–459 (2011). [PubMed: 21664348]
51. O'Brien LE, Bilder D, Beyond the Niche: Tissue-Level Coordination of Stem Cell Dynamics. *Annu Rev Cell Dev Biol*. 29, 107–136 (2013). [PubMed: 23937350]
52. Ayyaz A et al. , Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. *Nature*. 569, 121–125 (2019). [PubMed: 31019301]
53. Sato T et al. , Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature*. 469, 415–418 (2011). [PubMed: 21113151]
54. Sánchez Alvarado A, Stem cells and the Planarian *Schmidtea mediterranea*. *C R Biol*. 330, 498–503 (2007). [PubMed: 17631444]
55. Bergmann A, Steller H, Apoptosis, stem cells, and tissue regeneration. *Sci Signal*. 3, re8 (2010). [PubMed: 20978240]
56. Lane SW, Williams DA, Watt FM, Modulating the stem cell niche for tissue regeneration. *Nature Biotechnology*. 32, 795–803 (2014).
57. Pan D et al. , A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science*. 359, 770–775 (2018). [PubMed: 29301958]

58. Nargund AM et al. , The SWI/SNF Protein PBRM1 Restrains VHL-Loss-Driven Clear Cell Renal Cell Carcinoma. *CellReports*. 18, 2893–2906 (2017).
59. Platten M, von Knebel Doeberitz N, Oezen I, Wick W, Ochs K, Cancer Immunotherapy by Targeting IDO1/TDO and Their Downstream Effectors. *Front. Immunol* 5, 19961 (2015).
60. Platten M, Wick W, Van den Eynde BJ, Tryptophan Catabolism in Cancer: Beyond IDO and Tryptophan Depletion. *Cancer Res*. 72, 5435–5440 (2012). [PubMed: 23090118]
61. Zattara EE, Fernández-Álvarez FA, Hiebert TC, Bely AE, Norenburg JL, A phylum-wide survey reveals multiple independent gains of head regeneration in Nemertea. *Proc Biol Sci*. 286, 20182524 (2019). [PubMed: 30836873]
62. Liu S-Y et al. , Reactivating head regrowth in a regeneration-deficient planarian species. *Nature*. 500, 81–84 (2013). [PubMed: 23883932]
63. Sikes JM, Newmark PA, Restoration of anterior regeneration in a planarian with limited regenerative ability. *Nature*. 500, 77–80 (2013). [PubMed: 23883929]
64. Newmark PA, Sánchez Alvarado A, Not your father’s planarian: a classic model enters the era of functional genomics. *Nat Rev Genet*. 3, 210–219 (2002). [PubMed: 11972158]
65. Arnold CP et al. , Pathogenic shifts in endogenous microbiota impede tissue regeneration via distinct activation of TAK1/MKK/p38. *eLife*. 5 (2016), doi:10.7554/eLife.16793.
66. Guedelhofer OC4, Sánchez Alvarado A, Planarian immobilization, partial irradiation, and tissue transplantation. *J Vis Exp* (2012), doi:10.3791/4015.
67. Lei K, McKinney SA, Ross EJ, Lee H-C, Sánchez Alvarado A, Cultured pluripotent planarian stem cells retain potency and express proteins from exogenously introduced mRNAs. *bioRxiv*. 107, 573725 (2019).
68. Stuart T et al. , Comprehensive Integration of Single-Cell Data. *Cell*. 177, 1888–1902.e21 (2019). [PubMed: 31178118]
69. Hafemeister C, Satija R, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 20, 1–15 (2019). [PubMed: 30606230]
70. Becht E et al. , Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*. 37, 38–44 (2019).
71. Davies EL et al. , Embryonic origin of adult stem cells required for tissue homeostasis and regeneration. *eLife*. 6, e21052 (2017). [PubMed: 28072387]
72. Nowotarski SH, Davies EL, Robb SMC et al. The planarian anatomy ontology: A resource to connect data within and across experimental platforms. *BioRxiv*. (2020)
73. King RS, Newmark PA, *In situ* hybridization protocol for enhanced detection of gene expression in the planarian *Schmidtea mediterranea*. *BMC Dev Biol*. 13, 8 (2013). [PubMed: 23497040]

**Figure 1.**

A single-cell reconstruction of successful and unsuccessful planarian regeneration. (A) Schematic depicting experimental design of single cell reconstruction. (B) UMAP embedding of captured single cell transcriptomes, colored by tissue annotation. Scaled mean gene expression in single-cell reconstruction (C) and gene expression patterns by whole mount *in situ* hybridization (D) of select tissue-specific markers (See Extended Data Figs. 3–5 for additional markers and tissue-subcluster specificity). (E) Scaled proportion of cells from each tissue subcluster in sampled conditions, normalized to sample in which subcluster had maximum representation. Scale = 500 μ m.

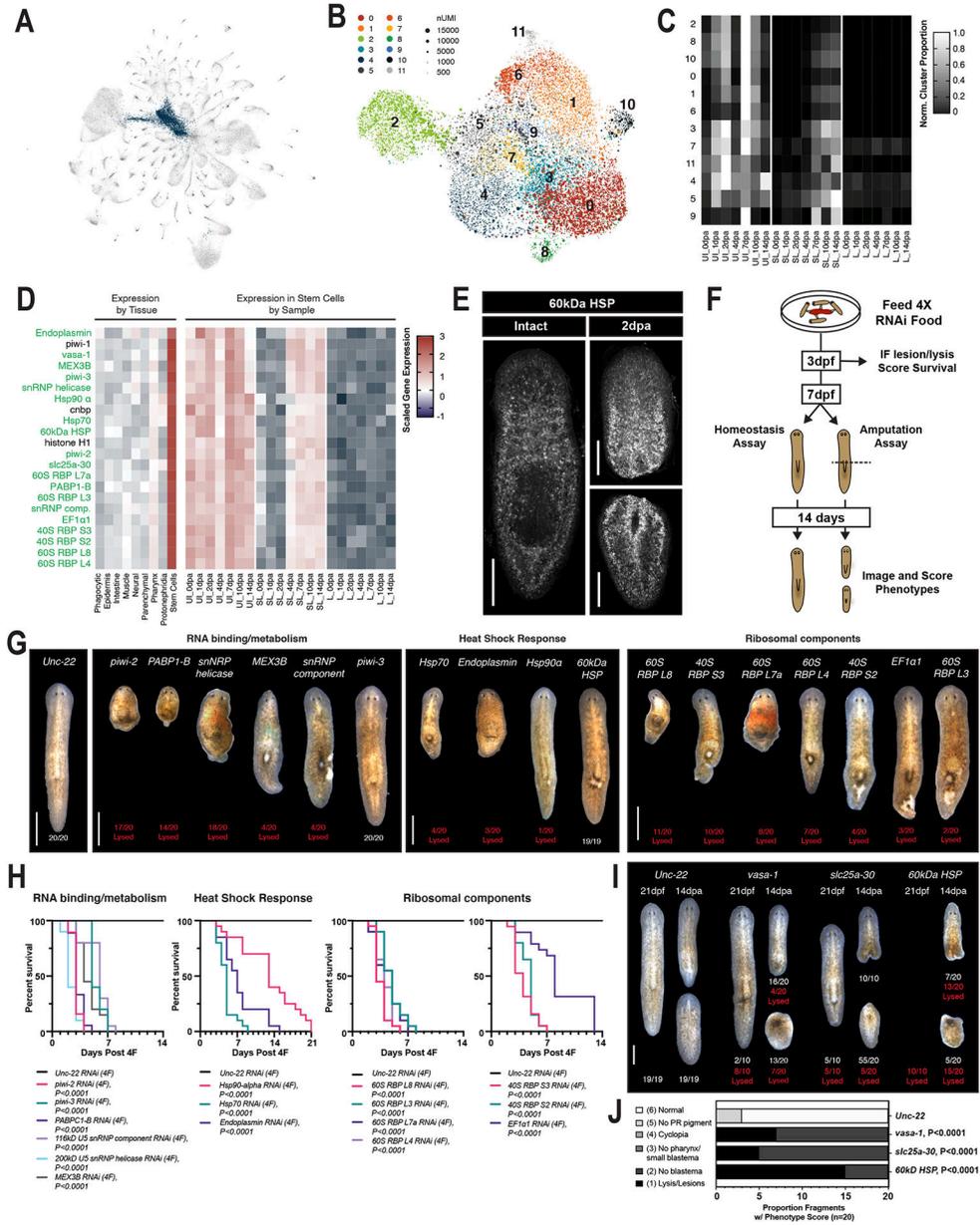


Figure 2. Regulators of tissue homeostasis were identified in irradiation sensitive stem cell clusters. (A) UMAP embedding of global dataset with stem cells highlighted. (B) UMAP embedding of all stem cells, colored by time after amputation. (C) Scaled proportion of cells from each stem cell subcluster across sampled conditions, normalized to sample in which subcluster had maximum representation. (D) Scaled mean expression of stem cell enriched genes by tissue and sample. (E) Z projection of confocal stack of stem cell enriched gene, *60kD HSP*. (F) Schematic representation of RNAi screen design (G) Representative images of RNAi-treated animals 3 days post feeding. (H) Survival of RNAi-treated animals shown in F (n = 19 or 20 as noted in G). Representative images (I) of homeostatic (21 days post feeding, n = 10) and regeneration phenotypes (14dpa, n = 20) and scoring of phenotypes

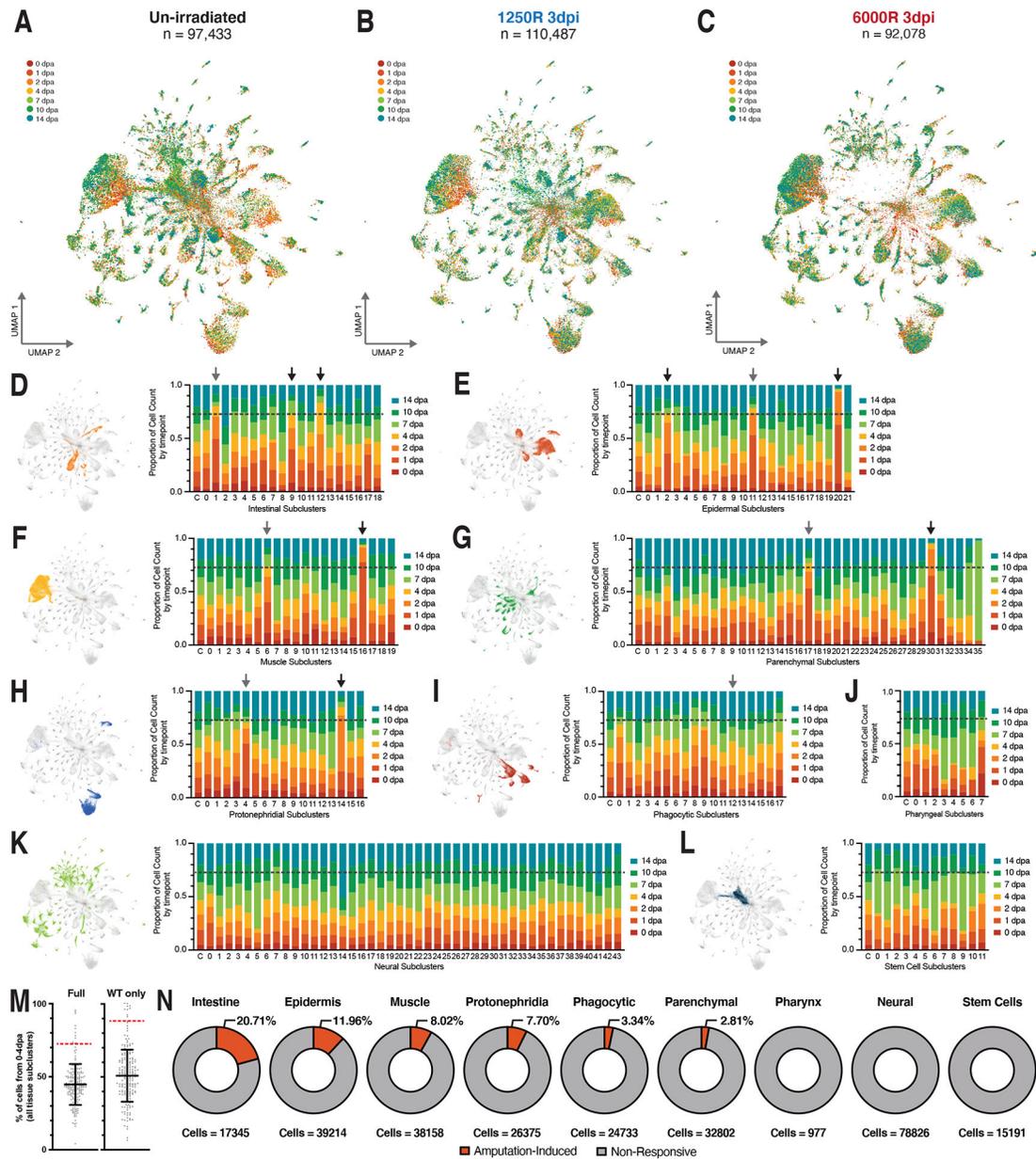
(J) in RNAi-treated animals. Scale = 500 μ m. P values are log-rank (Mantel-Cox) test (H) and two-sided Mann-Whitney U test (J), with no adjustments for multiple comparisons compared to *unc-22* control. Scale = 500 μ m.

Author Manuscript

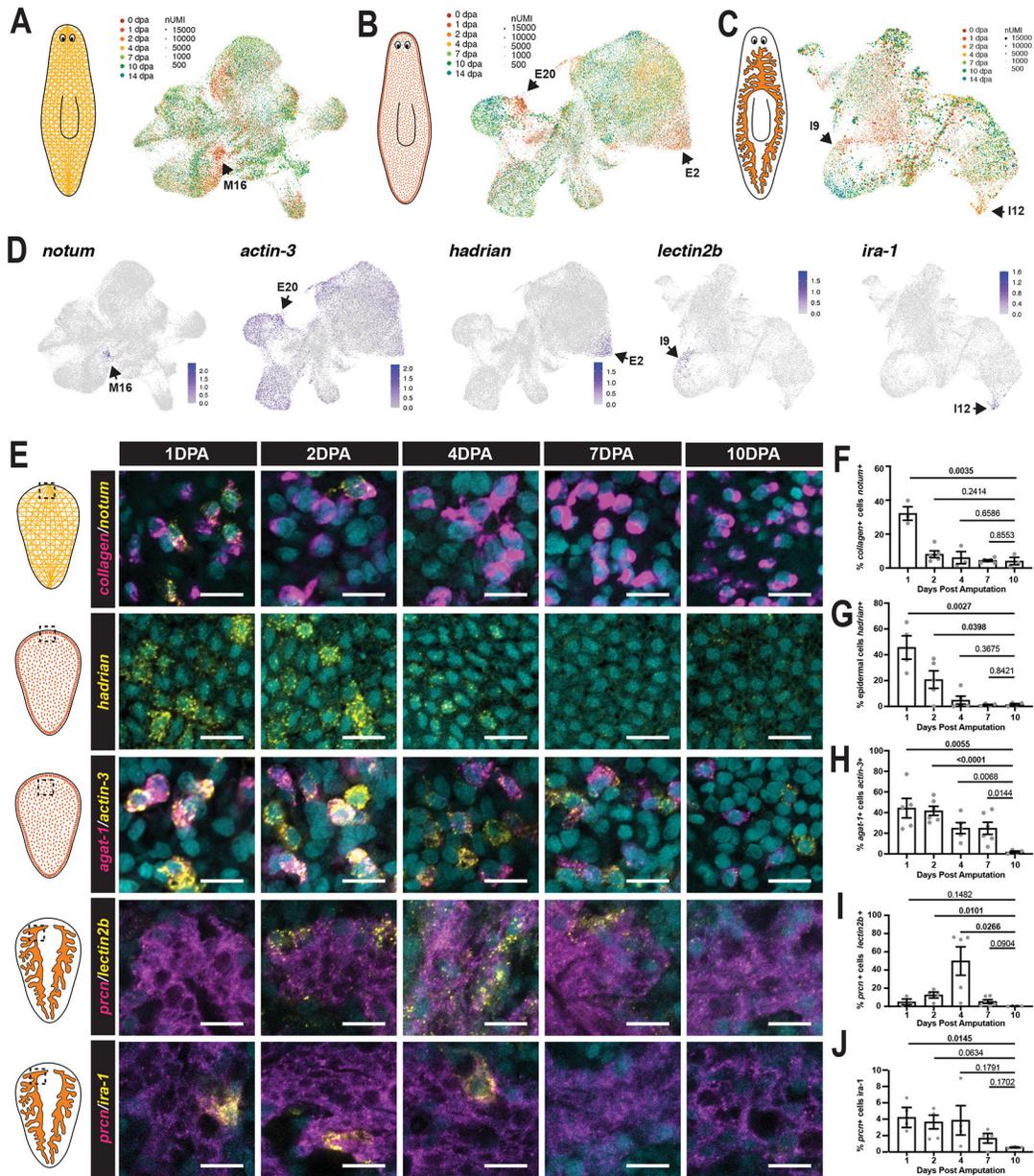
Author Manuscript

Author Manuscript

Author Manuscript

**Figure 3.**

Amputation specific clusters are non-uniformly distributed across tissue lineages. UMAP embedding of captured single cell transcriptomes from biopsies taken from un-irradiated (A), sub-lethally irradiated (B), or lethally irradiated (C) animals, colored by time. Proportion of cells in full dataset that fall into each timepoint within intestinal (D), epidermal (E), muscle (F), parenchymal (G), protonephridial (H), phagocytic (I), pharyngeal (J), neural (K), and stem cell (L) subclusters. Arrows indicate amputation-specific clusters in either un-irradiated or full data (gray), or both (black). (M) % of cells in each cluster (n=196 subclusters) that come from 0 – 4 dpa. Error bars depict mean \pm SD. Red dotted line indicates $+2$ SD. (N) Proportion of cell within each tissue that are in an amputation-specific cluster (Table S15).



3 biologically independent animals. P values are a two-sided unpaired t-test (F-J) compared to 10dpa time, with no adjustments for multiple comparisons. Data are presented as mean values \pm SEM. Scale = 20 μ m.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

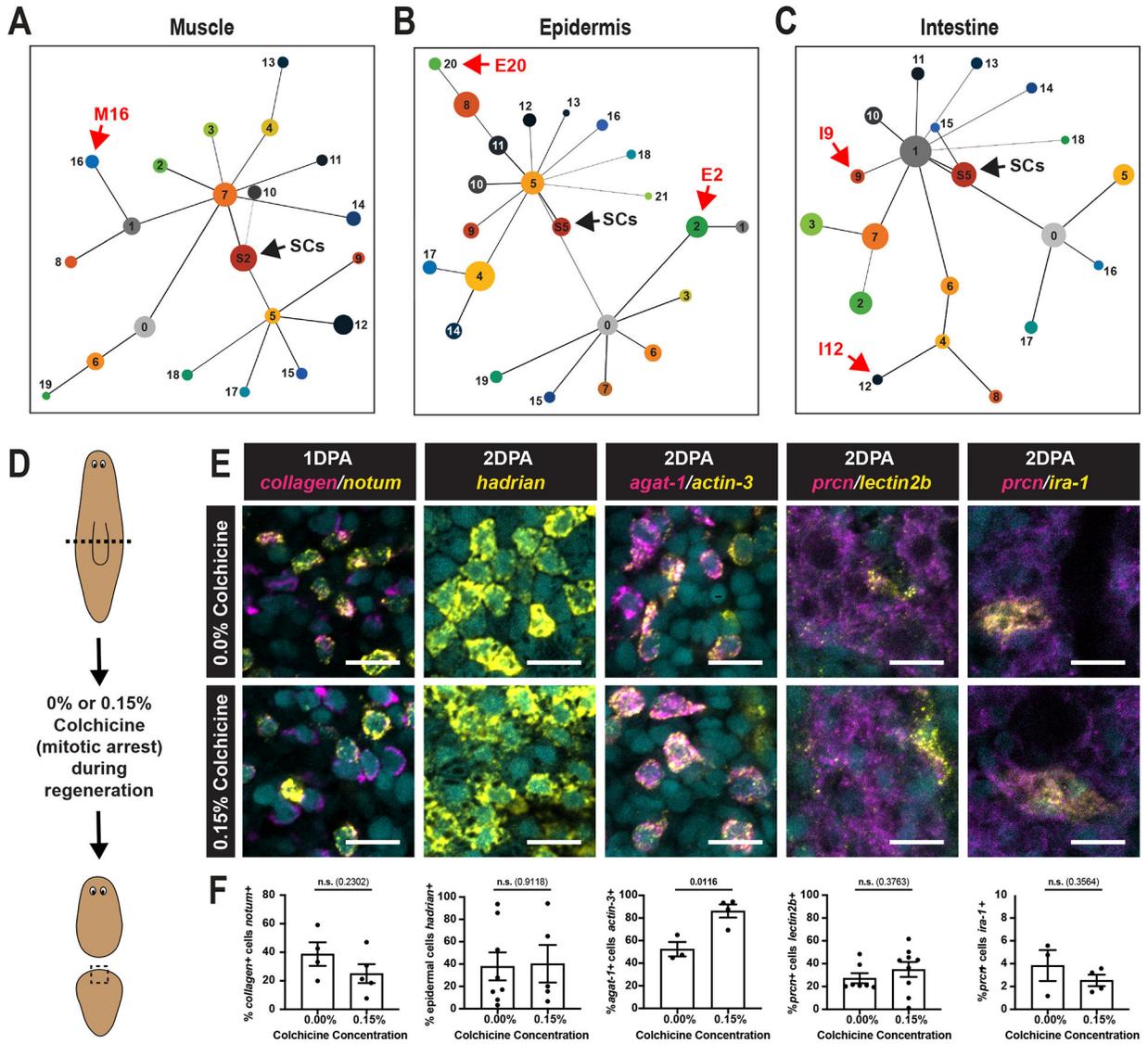


Figure 5.

Amputation-specific cell states arise from pre-existing post-mitotic cells. (A) Native PAGAs of muscle (A), epidermal (B), and intestinal (C) subcluster connectivity with stem cell progenitors (‘SCs,’ black arrow). Black lines represent the highest-confidence minimal spanning tree (fewest edges needed to connect all nodes) and amputation-specific cell states are highlighted by red arrows. (D) Schematic summary of colchicine-based experimental design (E) Visualization of amputation-induced gene (yellow) in surface epithelium or co-expressed with tissue specific marker (magenta) for each amputation-specific cell state following amputation. (F) Quantitation of TRACS markers as a percentage of the relevant tissue or cell type after amputation for MC16, n = 4 and 5 biologically independent animals; EC20, n = 8 and 5 biologically independent animals; EC2, n = 3 and 4 biologically independent animals; IC9, n = 7 and 9 biologically independent animals; and IC12, n = 3 and 4 biologically independent animals. P values are a two-sided unpaired t-test, with no

adjustments for multiple comparisons. Data are presented as mean values \pm SEM. Scale = 20 μ m.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

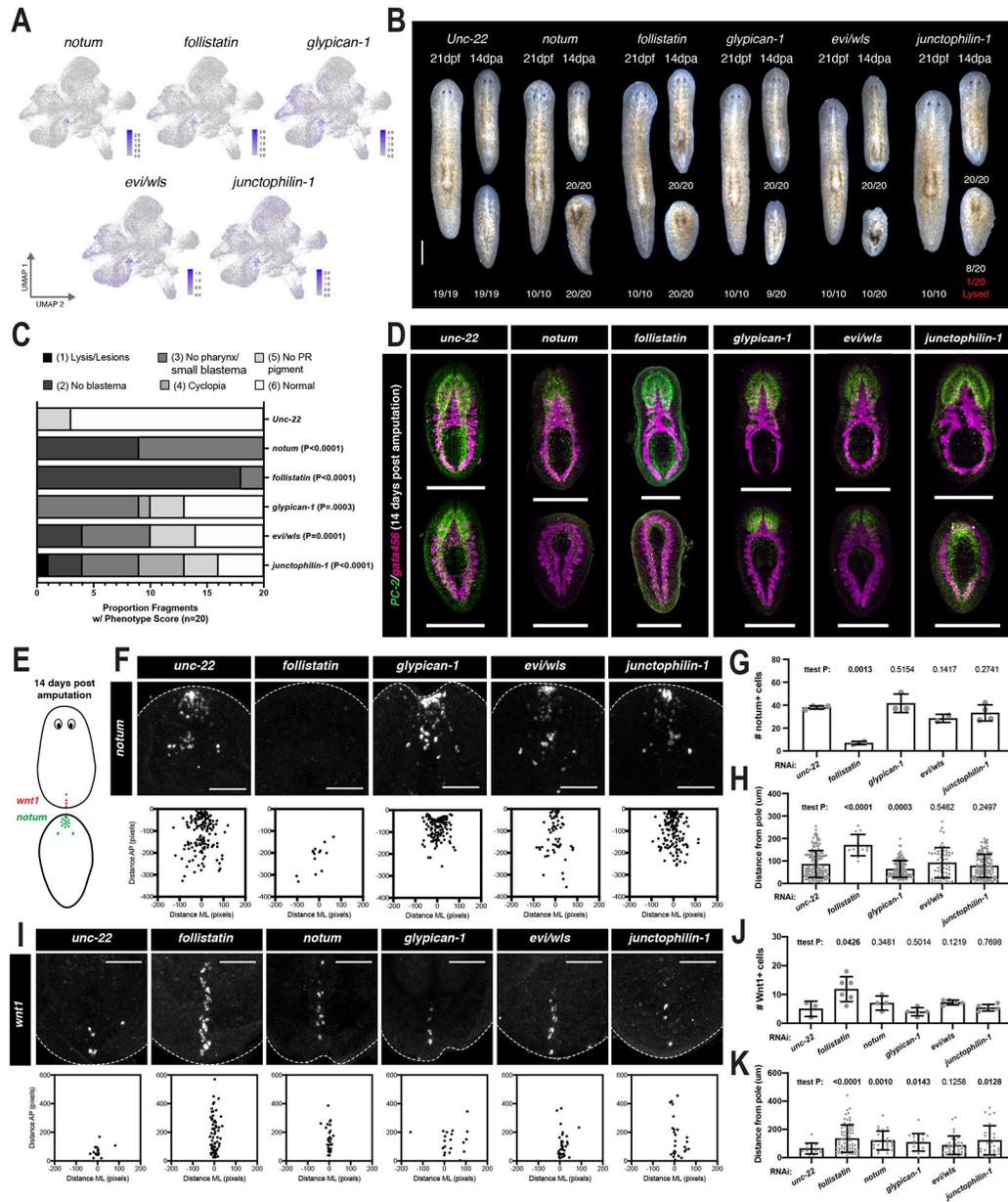
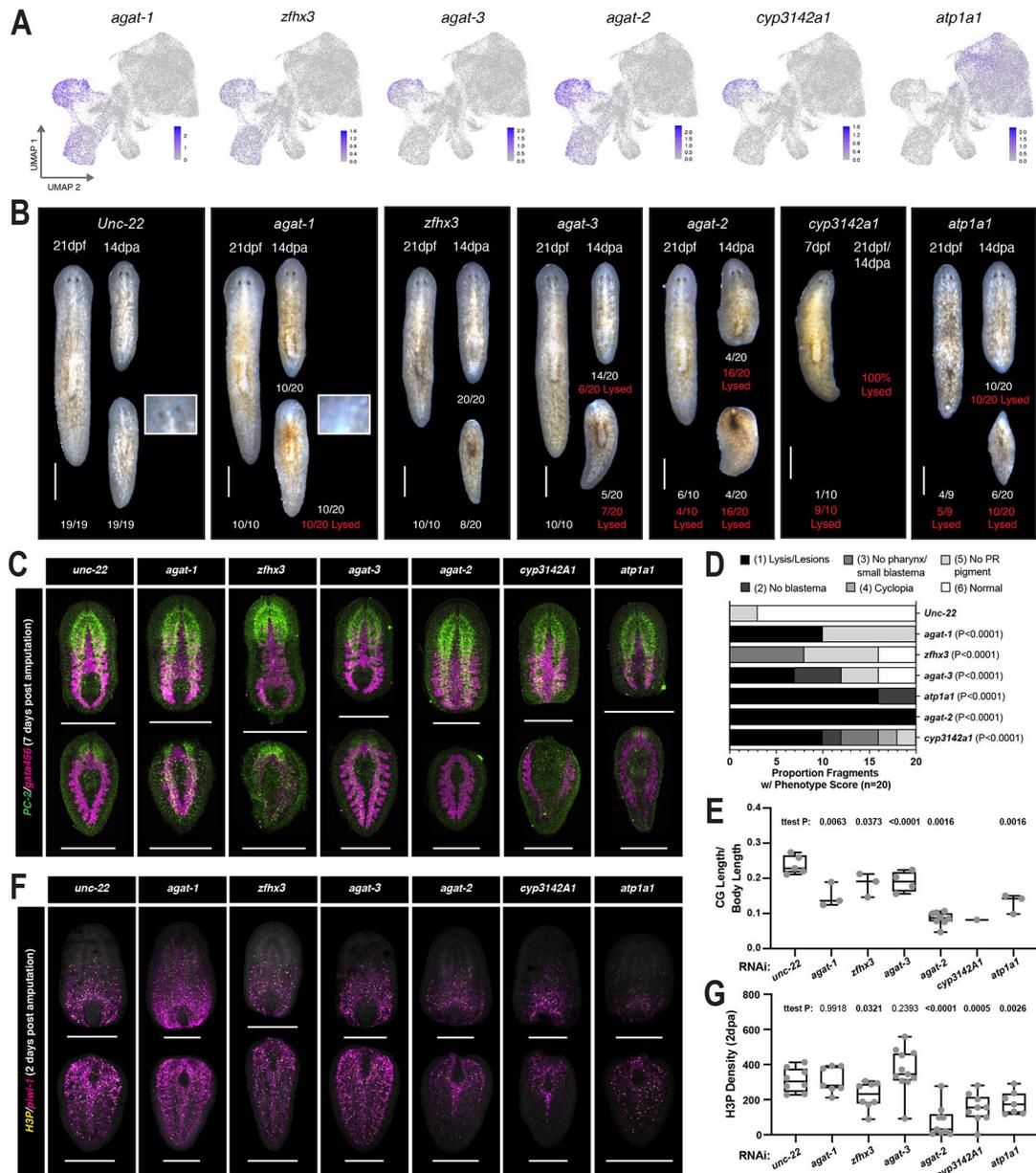


Figure 6. Muscle TRACS express genes required for tissue polarity. (A) UMAP feature plots showing muscle expression of M16-enriched genes that produced penetrant RNAi phenotypes. (B) Representative images of homeostatic (21 days post feeding, 10 animals) and regeneration phenotypes (14dpa, 20 animals) in RNAi-treated animals. Scale = 500µm. (C) Scoring of regeneration phenotypes. (D) Selected images of whole mount dFISH visualization of CNS (*PC2*) and intestine (*gata4/5/6*) in RNAi-treated animals, Scale = 500µm. (E) Schematic depicting visualized anterior and posterior pole cells in regenerating fragments. (F) Selected images and raw x,y position of *notum*⁺ anterior pole cells visualized by whole mount *in situ* hybridization, Scale = 150µm. (G) Number of *notum*⁺ cells in RNAi treated animals. n = 4 (*unc-22*), 2 (*folliculin*), 3 (*glypican-1*), 2 (*evi/wls*); and 4 (*junctophilin-1*) biologically

independent animals. (H) Distribution of *notum*⁺ cells in RNAi-treated animals, n = 152 (*unc-22*), 14 (*follistatin*), 125 (glypican-1), 57 (*evi/wls*); and 133 (*junctophilin-1*) *notum*⁺ cells. (I) Selected images and raw x,y position of *wnt1*⁺ posterior pole cells visualized by whole mount *in situ* hybridization, Scale = 150μm. (J) Number of *wnt1*⁺ cells in RNAi treated animals. n = 3 (*unc-22*), 6 (*follistatin*), 4 (notum), 5 (glypican-1), 5 (*evi/wls*); and 5 (*junctophilin-1*) biologically independent animals. (K) Distribution of *wnt1*⁺ cells in RNAi-treated animals, n = 15 (*unc-22*), 71 (*follistatin*), 28 (notum), 20 (glypican-1), 36 (*evi/wls*); and 27 (*junctophilin-1*) *wnt1*⁺ cells. P values are two-sided Mann-Whitney U test, with no adjustments for multiple comparisons compared to *unc-22* control (C) and two-sided unpaired t-test (G,H,J,K), with no adjustments for multiple comparisons. Data are presented as mean values +/- SD. compared to *unc-22* control.

**Figure 7.**

Epidermal TRACS express gene required for regulation of stem cell proliferation. (A) UMAP feature plots showing epidermal expression of genes that produced penetrant RNAi phenotypes. (B) Representative images of homeostatic (21 days post feeding) and regeneration phenotypes (14 dpa) in RNAi-treated animals. (C) Selected images of whole mount dFISH visualization of CNS (*PC2*) and intestine (*gata4/5/6*) in RNAi-treated animals. (D) Scoring of regeneration phenotypes. n = 20 animals. (E) Quantitation of cephalic ganglia length normalized by total body length in RNAi-treated animals. n = 5 (*unc-22*), 3 (*agat-1*), 3 (*zfhx3*), 4 (*agat-3*), 8 (*agat-2*), 1 (*cyp3142a1*), and 3 (*atp1a1*) biologically independent animals. (F) Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and immunohistochemistry of H3P⁺ mitotic cells in RNAi-treated animals 2 days post

amputation. (G) H3P⁺ cell density in RNAi-treated animals 2 days post amputation. n = 8 (*unc-22*), 7 (*agat-1*), 8 (*zfhx3*), 11 (*agat-3*), 9 (*agat-2*), 9 (*cyp3142a1*), and 7 (*atp1a1*) biologically independent animals. Data are presented in box plots of min, max, and median, with the bounds of the box at the first and third quartile and whiskers extending from quartile to the minimum or maximum. P values are two-sided Mann-Whitney U test, with no adjustments for multiple comparisons compared to *unc-22* control (D) and two-sided unpaired t-tests with no corrections for multiple comparison (E,G) compared to *unc-22* control. Scale = 500 μ m.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

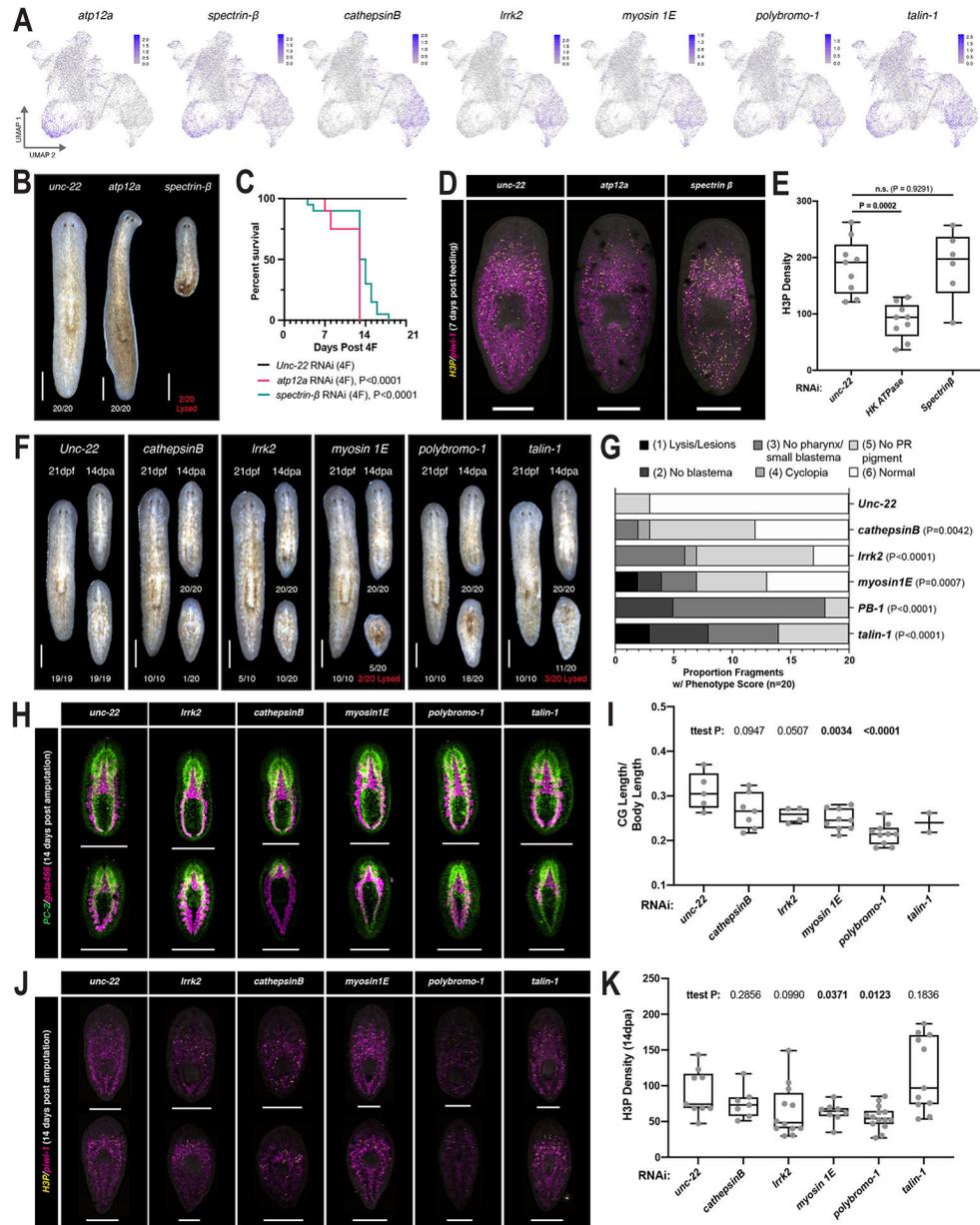


Figure 8. Intestinal TRACS express genes required for stem cell maintenance and tissue remodeling. (A) UMAP feature plots showing intestinal expression of genes that produced penetrant RNAi phenotypes. (B) Representative images RNAi-treated animals 3 days post feeding. (C) Survival of RNAi-treated animals shown in F ($n = 20$ for each condition). Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and H3P⁺ mitotic cells (D) and H3P⁺ cell density (E) in RNAi-treated animals 7 days post amputation. $n = 9$ (*unc-22*), 9 (*atp12a*), and 6 (*spectrinβ*) biologically independent animals. (F) Representative images of homeostatic (21 days post feeding, 10 animals) and regeneration phenotypes (14dpa, 20 animals) in RNAi-treated animals. (G) Scoring of I12-enriched gene RNAi regeneration phenotypes. $n = 20$ animals. (H) Selected images of whole mount dFISH visualization of

CNS (*PC2*) and intestine (*gata4/5/6*) in RNAi-treated animals. (I) Quantitation of cephalic ganglia length normalized by total body length in RNAi-treated animals. n = 5 (*unc-22*), 7 (*cathepsinB*), 4 (*lrrk2*), 9 (*myosin1e*), 10 (*polybromo-1*), and 2 (*talin-1*) biologically independent animals. Selected images of whole mount *in situ* hybridization of *piwi-1*⁺ stem cells and H3P⁺ mitotic cells (J) and H3P⁺ cell density (K) in RNAi-treated animals 14 days post amputation. n = 9 (*unc-22*), 7 (*cathepsinB*), 12 (*lrrk2*), 8 (*myosin1e*), 14 (*polybromo-1*), and 11 (*talin-1*) biologically independent animals. Data are presented in box plots of min, max, and median, with the bounds of the box at the first and third quartile and whiskers extending from quartile to the minimum or maximum. P values are log-rank test (C) and two-sided unpaired t-tests compared to *unc-22* control, with no corrections for multiple comparison (E,I,K). Scale = 500µm.