Title: Identification of a regeneration organizing cell in the *Xenopus* tail

Authors: C. Aztekin[1,2,†], T.W. Hiscock[1,3,†], J.C. Marioni[3,4,5], J.B. Gurdon[1,2], B.D. Simons[1,6,7,*] and J. Jullien[1,2,*]

Affiliations:

1. Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK

2. Department of Zoology, University of Cambridge, Cambridge, UK

3. Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK

4. EMBL-European Bioinformatics Institute, Wellcome Genome Campus, Cambridge, UK

5. Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK

6. Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK

7. Wellcome Trust Centre for Stem Cell Research, University of Cambridge, Tennis Court Road, Cambridge, UK

† These authors contributed equally

* Corresponding authors: jj256@gurdon.cam.ac.uk, bds10@cam.ac.uk

Abstract

Unlike mammals, *Xenopus laevis* tadpoles have a high regenerative potential. To characterize this regenerative response, we performed single-cell RNA sequencing (scRNA-seq) following tail amputation. By comparing naturally-occurring regeneration-competent and incompetent tadpoles, we identified a previously unrecognized cell type that we term the regeneration-organizing cell (ROC). ROCs are present in the epidermis during normal tail development, and specifically relocalize to the amputation plane of regeneration-competent tadpoles, forming the wound epidermis. Genetic ablation or manual removal of ROCs blocks regeneration, whereas transplantation of ROC-containing grafts induces ectopic outgrowths in early embryos. Transcriptional profiling revealed that ROCs secrete ligands associated with key regenerative pathways, signaling to progenitors to reconstitute lost tissue. These findings reveal the cellular mechanism through which ROCs form the wound epidermis and ensure successful regeneration.

One Sentence Summary: Regeneration-organizing-cells play an essential role in coordinating the regeneration of *Xenopus* tail following amputation.

100 words Summary: Some vertebrae show a remarkable, if sometimes restricted, ability to regenerate lost appendages. By utilizing single-cell mRNA sequencing and comparing naturally-occurring regeneration competent and -incompetent *Xenopus laevis* tadpoles, Aztekin *et al*. identified a new cell type, termed regeneration organizing cells (ROCs) that coordinate tail regeneration. Relocation of ROCs from the body to the amputation plane enables specialized wound epidermis formation and subsequent regeneration. ROCs simultaneously express many different ligands that can induce proliferation of different progenitor cell populations. Hence, by signaling to underlying progenitors, ROCs act as a center that orchestrates the growth of a new appendage.

Main Text

Appendage regeneration involves coordinated changes in many cell types, and has been largely characterized by morphological assessments, lineage tracing studies, and low-throughput gene investigations. As a result, regeneration is broadly divided into three essential steps: the formation of a specialized wound epidermis, blastema/regenerative bud formation, and outgrowth via proliferation (*1*, *2*). However, a comprehensive understanding of the changes in cell types, transcriptional dynamics, and cellular mechanisms accompanying these processes is lacking. For example, the first morphological change upon amputation is the formation of the specialized wound epidermis, a ligand-expressing structure covering the wound that is essential in many different regeneration scenarios in different species (*e.g.* zebrafish, axolotl) (*1*, *3*, *4*). However, it is not clear which cell types are present in the specialized wound epidermis, what is its origin, what is the broad catalogue of ligands expressed from it, and why is it crucial for regeneration. To answer such questions and identify essential regulators of regeneration, we focused on *Xenopus* tadpoles, with their naturally occurring regeneration-competent and -incompetent developmental stages, making them an ideal system for comparative studies (*5*).

To assess comprehensively the transcriptional dynamics and cell type changes that occur during regeneration, we took advantage of high-throughput scRNA-seq to target *Xenopus laevis* tails at various stages following amputation in both regeneration-competent and incompetent tadpoles, as well as uninjured (intact) tails at the same developmental stage (Fig. 1A). We sequenced >13,000 cells, with at least 2 biological replicates per condition, with an average of ~2,300 genes detected per cell (table S1). Cells from all samples were pooled and visualized by the dimensionality reduction method, UMAP (*6*) (Fig. 1A). Cluster identity was assigned using multiple known markers and revealed a total of 46 putative cell types (some rare and others uncharacterized) encompassing the immune system, skin, nervous system and somites, emphasizing the cellular heterogeneity of the tail (Fig. 1B, and fig. S1). Biological replicates showed a similar distribution of cell types, confirming the reproducibility of the atlas (fig. S2). Computational inference of cell cycle state

confirmed that progenitor cell populations are mostly positioned in G2/M and S phases, whereas terminally differentiated cells are in G1 (fig. S3). Our comprehensive cell atlas can be viewed using the interactive platform (marionilab.cruk.cam.ac.uk/XenopusRegeneration/).

Having established the atlas, we then questioned what transcriptional and cell type changes are associated specifically with tail regeneration. By comparing samples, we could make a distinction between developmental (Fig. 2A), amputation-specific (Fig. 2B), and regeneration-specific (Fig. 2C) effects. Most cell types were found in all samples (fig. S4A). Consistent with lineage tracing studies (*7*), we found no evidence for the emergence of a multipotent progenitor population during regeneration, nor did we observe "intermediate" cell states reflective of transdifferentiation. Indeed, the only new cell type to emerge following amputation was an uncharacterized motor neuron-like cell type that expressed genes associated with spinal cord injury (*e.g. Fgf10* (*8*)) and metabolic hormones (*e.g. Leptin* (*9*, *10*)) (Fig. 2B). This phenotype was also observed in regeneration-incompetent tadpoles (Fig. 2B). Hence, we considered it an amputation-response, and focused instead on regeneration-specific changes.

Surveying the range of single-cell data, we found that the most significant cell-type change specific to regeneration was related to a previously unidentified cell type of the epidermis (fig. S4A and B), which (for reasons that will become clear) we designated as the regeneration-organizing cell (ROC). Notably, based on the scRNA-seq data, ROCs were found to be present in both intact and regenerating tails, but were observed following tail amputation only in regeneration-competent tadpoles (Fig. 2C). As this cell population distinguishes the amputation response of regeneration–competent tadpoles from incompetent ones, and expresses multiple genes that support regeneration (*e.g. Wnt5a* (*11*), *Fgf10* (*12*), *Fgf20* (*13*), *Msx1* and *Bmpr1a* (*5*)), we hypothesized that ROCs may represent an essential component of the regenerative response.

To assess the function of ROCs, we first investigated their location using marker genes identified by scRNA-seq. Using published *in situ* data from Xenbase, we found that >25 ROC marker genes were expressed along the midline edge of the epidermis (*e.g. Fgf7, Msx2, C3, Wnt3a*), from the posterior trunk towards the tail tip

(table S2). We further confirmed the localization of ROCs using a Lef1 reporter line (pbin7LEF:GFP (*14*)), in combination with TP63 immunolabelling. Whilst Lef1 is expressed in multiple cell types, only ROCs express high levels of both *Lef1* and *Tp63* (Fig. 3A and fig. S5A). Therefore, we identified ROCs as LEF1+/TP63+ cells and confirmed that they are localized to the edge of the epidermis (Fig. 3B and fig. S5B). ROCs are present in this location in both regeneration-competent and incompetent tadpoles and, immediately following amputation, this population is largely removed from the amputation plane, but remains along the posterior trunk (Fig. 3C**,** left hand column). During successful regeneration, ROCs reappear in the amputation plane within 24 hours; however, they remain notably absent at the amputation plane of regeneration-incompetent tadpoles (Fig. 3C and D, fig. S5C). Our *Lef1* data were consistent with published *in situ* data of ROC marker genes seen in the tails of amputated regeneration-competent tadpoles (*Fgf9, Fgf10, Wnt5a, Wnt3a, Msx1, Msx2* (*5*, *15*)*, C3* (*16*)) and lacking in regeneration-incompetent tadpoles (*Msx1, Msx2* (*5*, *15*)). The gene expression profile and location of ROCs at the amputation plane further suggest that they are in fact a single cell type that defines the specialized wound epidermis, which specifically forms in regeneration-competent tadpoles to trigger the regenerative response.

As the presence of ROCs at the amputation plane correlates with regenerative outcome, we tested whether these cells are required for regeneration. We first performed NTR/MTZ based genetic ablation (*17*) of ROCs using F_0-transgenic tadpoles expressing Nitroreductase (NTR) under the control of the Krt.L promoter, a member of the keratin gene family that is expressed in a highly specific manner in ROCs in stages where regeneration is assessed (Fig. 4A, and fig. S6A to D; *Krt.L* is also known as *Krt70.L*). Upon Metronidazole (MTZ) treatment, we were able to specifically ablate ROCs, as confirmed by the disappearance of GFP positive cells in pbin7LEF:GFP /Krt.L:NTR F_0 transgenic tadpoles (Fig. 4A), together with the observation of no apparent gross off-target effects in other tissue types (fig. S6E). Ablation of ROCs in regeneration-competent tadpoles led to drastically reduced tail regeneration (Fig. 4B and fig. S7A-D), demonstrating that ROCs are indeed required for regeneration.

We then eliminated ROCs in a spatially localized manner, by manually removing ROCs in the posterior trunk that remain directly after amputation (Fig. 4C, and fig. S7E). When these regions were removed at the same time as the amputation, we observed a reduction in *Lef1*+ cells at the amputation plane, and correspondingly reduced regeneration (Fig 4D, and fig. S7E). However, when removed 12-16 hours after amputation, *Lef1*+ expression was maintained at the amputation plane, and regeneration could proceed (Fig 4C and D, and fig. S7E), indicating that there is a critical time window during which posterior trunk ROCs are required to initiate regeneration. This suggests that, soon after amputation, existing ROCs may have to relocate to the amputation area to initiate the regenerative response. To test this hypothesis, we traced ROCs following amputation with the *Lef1* reporter, and observed the mobilization of resident ROCs from the posterior trunk towards the amputation area within 2-8 hours (Fig. 4E). Together, these observations suggest that mobilization of ROCs to the amputation area is a necessary step in wound epidermis formation and subsequent regeneration, in contrast to previous suggestions that the wound epidermis is a novel state that differs from normal epithelium and appears upon amputation (*2*).

Next, we asked whether inhibition of pathways that are necessary for wound epidermis formation, which are rapidly upregulated upon injury, interfere with the mobilization of ROCs to the amputation plane. Indeed chemical inhibition of reactive oxygen species (ROS) production (*13*, *18*, *19*) or the TGFβ pathway (*20*) resulted in significantly reduced ROC migration (fig. S8). In contrast, inhibition of the FGF pathway, which is known to not affect wound epidermis formation (*15*), had no effect on ROC migration (fig. S8). Hence, wound-induced ROS production and TGFβ pathway activation are necessary for ROCs to migrate to the amputation plane, where they form the wound epidermis.

To understand the essential role played by ROCs during regeneration, we used our scRNA-seq data to dissect their transcriptional signature. Ligands of signalling pathways that are known to be required for regeneration and increase proliferating cell numbers, including FGF (*15*), BMP (*5*, *21*), WNT (*15*), NOTCH (*5*) and TGFβ (*20*), are simultaneously expressed in ROCs, but not in any other cell type. In contrast, receptors for these pathways are mostly expressed in progenitor cell types (Fig. 5A, and fig. S9). Moreover, we found that

progenitor populations at the amputation site showed an increase in the fraction of cells in G2/M and S phases during regeneration (Fig. 5B). These results suggest that ROCs function as a signalling centre by secreting factors that promote progenitor proliferation in multiple tissues. Such an increase in proliferation can explain how tissue loss can be reconstituted from progenitors of the tail without requiring the emergence of a new multipotent cell state or states.

Having established the central role of ROCs during regeneration, we then asked whether they resemble other cell types associated with early development. ROCs express markers of limb development and appendage growth (fig. S10), including the limb Apical Ectodermal Ridge (AER) regulating transcription factors *Sp8* and *Sp9* (*22*). The AER is a structure formed at limb bud tips, and plays an essential role in limb growth and patterning by sending extracellular signals to underlying tissues (*23*). Moreover, in urodele limb regeneration, a similar structure, the Apical Epithelial Cap (AEC), is shown to be necessary for limb regeneration (*1*). Although ROCs resemble AER and AEC transcriptionally (*e.g., Sp9, Msx2, Wnt5a)*, we were not able to detect their well-known regulators (*e.g. Fgf2, Fgf4, Fgf8, and Cx43(Gja1))* (fig. S10A). Despite this, we hypothesized that ROCs could play an instructive role during tail growth by secreting growth factors and extracellular cues, similar to the AER during limb growth, and aimed to test this using transplantation assays.

In order to isolate potential ROCs for grafting, we investigated where the transcriptional signature of ROCs is first detected during early embryonic development. Marker genes of ROCs first appear at the early tailbud tip (NF stage 23), and later expand posteriorly along the midline edge of the epidermis (table S2). We grafted different sizes of posterior tailbud tissues, which contain ROCs, to the surface of different regions of trunk of a host embryo (Fig. 5C). All grafts induced ectopic outgrowth, regardless of graft size or implantation location. Larger grafts induced tail-like structures, with a corresponding defect in donor tail growth; whereas smaller grafts, composed of only skin layers, induced fin-like structures, without significantly impacting the donor (Fig. 5C, and fig. S11A and B). In contrast, control grafts, in which dissected trunk skin tissues were transplanted, did not result in outgrowths and instead were incorporated into the host trunk (fig. S11C). To further pinpoint the cell type responsible for the outgrowths, we repeated our grafting experiments whilst

simultaneously removing ROCs using the MTZ/NTR system. The ablation of Krt.L expressing cells in the donor graft significantly reduced the length of the ectopic outgrowths (fig. S11D and E), suggesting that ROCs are involved in the outgrowth phenomenon, although other cell types may also contribute.

If these distal growths are induced by the organizing abilities of the grafted ROCs, we would expect the transplanted cells to localize to the tip of the ectopic outgrowths. Indeed, when we tested this using GFP-labelled donor or host embryos, we found that most of the donor tissues were located at the tip of the ectopic structures. Moreover, host cells contributed significantly to the ectopic structures in all grafts, indicating that ROCs can stimulate outgrowth of both donor and host cells (Fig. 5D, and fig. S11F and G). Together, these results suggest that ROCs act as an instructive signalling centre that induces outgrowth during both the development and regeneration of the tail.

Overall, our comprehensive analysis of cell types in the regenerating *Xenopus* tail provides a mechanistic understanding of the initiation and organization of tail regeneration via the re-establishment of a ROC-signalling centre. By acting as the primary source of major growth factors and instructive signals, ROCs promote proliferation of underlying progenitors to regenerate tissue following amputation (Fig. 6). Our data also suggests that ROCs are a single cell type that characterizes the wound epidermis, a structure that is crucial for regeneration in many contexts (*2*). Investigation of other species (*e.g.* neonatal mouse, salamander) will indicate whether a ROC-based mechanism is a conserved feature of specialized wound epidermis formation and appendage regeneration.

Finally, signatures of the specialized wound epidermis formation are absent in non-regenerating animals such as birds, adult mice and adult frogs. However, reintroduction of molecules secreted from the specialized wound epidermis can re-initiate cell cycle entry to some degree in these animals (*24, 25, 26*). The discovery of a single cell type defining the wound epidermis offers a new perspective on cell replacement therapies, suggesting that “organizer grafts” may perhaps one day serve as a potential alternative to full organ replacement in regenerative therapies.

References and Notes

1. E. M. Tanaka, The Molecular and Cellular Choreography of Appendage Regeneration. *Cell*. **165**, 1598–1608 (2016).
2. C. L. Stoick-Cooper, R. T. Moon, G. Weidinger, Advances in signaling in vertebrate regeneration as a prelude to regenerative medicine. *Genes & Development*. **21**, 1292–1315 (2007).
3. C.-H. Chen, K. D. Poss, Regeneration Genetics. *Annu. Rev. Genet.* **51**, 63–82 (2017).
4. J. Li, S. Zhang, E. Amaya, The cellular and molecular mechanisms of tissue repair and regeneration as revealed by studies in Xenopus. *Regeneration*. **3**, 198–208 (2016).
5. C. W. Beck, B. Christen, J. M. W. Slack, Molecular Pathways Needed for Regeneration of Spinal Cord and Muscle in a Vertebrate. *Developmental Cell*. **5**, 429–439 (2003).
6. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 1–51 (2018). https://arxiv.org/pdf/1802.03426.pdf
7. Gargioli, C. & Slack, J. M. W. Cell lineage tracing during Xenopus tail regeneration. *Development* **131,** 2669–2679 (2004).
8. Chen, J. *et al.* Neuron and microglia/macrophage-derived FGF10 activate neuronal FGFR2/PI3K/Akt signaling and inhibit microglia/macrophages TLR4/NF-κB-dependent neuroinflammation to improve functional recovery after spinal cord injury. *Cell Death and Disease* 1–12 (2017).
9. J. Kang *et al.*, Modulation of tissue repair by regeneration enhancer elements. *Nature*. **532**, 201–206 (2016).
10. N. R. Love *et al.*, Genome-wide analysis of gene expression during Xenopus tropicalis tadpole tail regeneration. *BMC Dev Biol*. **11**, 70 (2011).
11. T. Sugiura, A. Tazaki, N. Ueno, K. Watanabe, M. Mochii, Xenopus Wnt-5a induces an ectopic larval tail at injured site, suggesting a crucial role for noncanonical Wnt signal in tail regeneration. *Mechanisms of Development*. **126**, 56–67 (2009).
12. Y. Taniguchi, T. Sugiura, A. Tazaki, K. Watanabe, M. Mochii, Spinal cord is required for proper regeneration of the tail in Xenopus tadpoles. *Develop. Growth Differ.* **50**, 109–120 (2008).
13. N. R. Love *et al.*, Amputation-induced reactive oxygen species are required for successful Xenopus tadpole tail regeneration. *Nature Cell Biology*. **15**, 222–228 (2013).
14. Thi Tran, H., Sekkali, B., Van Imschoot, G., Janssens, S. & Vleminckx, K. Wnt/β-catenin signaling is involved in the induction and maintenance of primitive hematopoiesis in the vertebrate embryo. *Proc Natl Acad Sci USA* 1–6 (2010).
15. G. Lin, J. M. W. Slack, Requirement for Wnt and FGF signaling in Xenopus tadpole tail regeneration. *Developmental Biology*. **316**, 323–335 (2008).
16. A. Tazaki *et al.*, Macroarray-based analysis of tail regeneration inXenopus laevis larvae. *Dev. Dyn.* **233**, 1394–1404 (2005).
17. Martinez-De Luna, R. I. & Zuber, M. E. Rod-Specific Ablation Using the Nitroreductase/Metronidazole System to Investigate Regeneration in Xenopus. *Cold Spring Harb Protoc* 1–9 (2018). doi:10.1101/pdb.prot100974
18. F. Ferreira, G. Luxardi, B. Reid, M. Zhao, Early bioelectric activities mediate redox-modulated regeneration. *Development*. **143**, 4582–4594 (2016).
19. Ferreira, F., Raghunathan, V., Luxardi, G., Zhu, K. & Zhao, M. Early redox activities modulate Xenopus tail regeneration. *Nature Communications* 1–15 (2018). doi:10.1038/s41467-018-06614-2
20. D. M. Ho, M. Whitman, TGF-β signaling is required for multiple processes during Xenopus tail regeneration. *Developmental Biology*. **315**, 203–216 (2008).
21. C. W. Beck, B. Christen, D. Barker, J. M. W. Slack, Temporal requirement for bone morphogenetic proteins in regeneration of the tail and limb of Xenopus tadpoles. *Mechanisms of Development*. **123**, 674–688 (2006).
22. Y. Kawakami, Sp8 and Sp9, two closely related buttonhead-like transcription factors, regulate Fgf8 expression and limb outgrowth in vertebrate embryos. *Development*. **131**, 4763–4774 (2004).
23. F. Petit, K. E. Sears, N. Ahituv, Limb development: a paradigm of gene regulation. *Nature Publishing Group*. **18**, 245–258

(2017).
24. K. Kostakopoulou, A. Vogel, P. Brickell, C. Tickle, 'Regeneration' of wing bud stumps of chick embryos and reactivation of *Msx-1* and *Shh* expression in response to FGF-4 and ridge signals. *Mechanisms of Development* 119–131 (1996).
25. L. Yu, *et al.* BMP signaling induces digit regeneration in neonatal mice. *Development* **137,** 551–559 (2010).
26. G. Lin, Y. Chen, J. M. W. Slack, Imparting Regenerative Capacity to Limbs by Progenitor Cell Transplantation. *Developmental Cell* **24,** 41–51 (2013).
27. S. Hoppler, P. D. Vize, Xenopus Protocols, *Methods in Molecular Biology* 917, (2012).
28. Nieuwkoop, P.D., Faber, J., 1967. Normal Table of Xenopus laevis (Daudin). North-Holland Publishing Company, Amsterdam.
29. S. Pandey, K. Shekhar, A. Regev, A. F. Schier, Comprehensive Identification and Spatial Mapping of Habenular Neuronal Types Using Single-Cell RNA- Seq. *Current Biology*. **28**, 1052–1065.e7 (2018).
30. T. Ilicic *et al.*, Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.*, 1–15 (2016).
31. L. McInnes, J. Healy, & J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 1–51 (2018). https://arxiv.org/pdf/1802.03426.pdf
32. A. T. L. Lun, D. J. McCarthy, J. C. Marioni, A step-by-step workflow for low-level analysis of single-cell RNA-seq data. *F1000Res*. **5**, 2122–68 (2016).
33. A. M. Session *et al.*, Genome evolution in the allotetraploid frog Xenopus laevis. *Nature*. **538**, 336–343 (2016).
34. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **26**, 139–140 (2009).
35. A. T. L. Lun, A. C. Richard, J. C. Marioni, Testing for differential abundance in mass cytometry data. *Cell Death and Disease*, 1–5 (2017).
36. M. Ashburner, *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* **25,** 25–29 (2000).
37. S. Aibar *et al.*, SCENIC: single-cell regulatory network inference and clustering. *Nat Meth*. **14**, 1083–1086 (2017).
38. T. Goldberg *et al.*, A draft network of ligand-receptor-mediated multicellular signalling in human. *Nature Communications*. **6**, 1–11 (2016).
39. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36,** 411–420 (2018).
40. I. Tirosh *et al.*, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 1–10 (2019).
41. N. R. Love *et al.*, pTransgenesis: a cross-species, modular transgenesis resource. *Development*. **138**, 5451–5458 (2011).
42. S. Curado *et al.*, Conditional targeted cell ablation in zebrafish: A new tool for regeneration studies. *Dev. Dyn.* **236**, 1025–1035 (2007).
43. H. Ogino, W. B. McConnell, R. M. Grainger, High-throughput transgenesis in Xenopus using I-SceI meganuclease. *Nat Protoc*. **1**, 1703–1710 (2006).

Acknowledgments

We thank the Cambridge Institute Genomics Core for their support with this work on the 10X-Genomics and sequencing library preparations. The transgenic testes that are used in this study, and pTransgenesis vectors were obtained from the European Xenopus Resource Centre, curated with funding from the Wellcome Trust/BBSRC and maintained by the University of Portsmouth, School of Biological Sciences. Anti-PCNA

and anti-PHH3 antibodies were kind gifts from the Brand Lab. CFP-NTR construct was a kind gift from J. Mumm and M. Zuber. We would like to thank H. Ma and M. Huch for use of their stereoscopes. We would like to thank A. Lun for advice on scRNA-seq analysis; J. Griffiths for assisting with the creation of the website; C. Baker and M. Minarik for their help during revision period; B. Steventon, J. Robert, J. Kaufman, N. McGovern and D. Wagner for general discussion of the single-cell data. We are grateful to A. Philpott, V. Gaggioli and E. Rawlins for their critical reading of the manuscript.

Funding

C.A. is funded by University of Cambridge and Cambridge Trust. J.J. and J.B.G. are funded by a grant from the Wellcome Trust (101050/Z/13/Z). T.W.H., J.C.M and B.D.S are funded as part of a Wellcome Strategic Award to study cell fate decisions (105031/D/14/Z). T.W.H. is also supported by an EMBO Long Term Fellowship (ALTF 606-2018). B.D.S also acknowledges funding from the Royal Society E.P. Abraham Research Professorship (RP\R1\180165) and Wellcome Trust (098357/Z/12/Z).. J.C.M. acknowledges core funding from the European Molecular Biology Laboratory and Cancer Research UK (A17197). This work is funded by a grant from the Wellcome Trust (101050/Z/13/Z) and supported by the Gurdon Institute core grant from Cancer Research UK (C6946/A14492) and the Wellcome Trust (092096/Z/10/Z).

Author Contributions

Conceptualization: C.A., J.J.; **Experiments:** C.A. with assistance from J.J. with grafting; **Computational Analysis:** T.W.H.; **Data interpretation:** C.A., T.W.H., J.J.; **Writing – Original, Draft:** C.A., T.W.H., J.J.**;** **Writing – Review & Editing:** all authors; **Supervision:** J.C.M., B.D.S., J.B.G. contributed to general supervision, and J.J. mainly supervised the project.

Competing interests

The authors declare no competing interests.

Data and materials availability

Materials and codes are freely available and requests should be addressed to J.J. (jj256@gurdon.cam.ac.uk) and B.D.S. (bds10@cam.ac.uk).

Supplementary Materials:

Materials and Methods

Figures S1-S11

Tables S1-S3

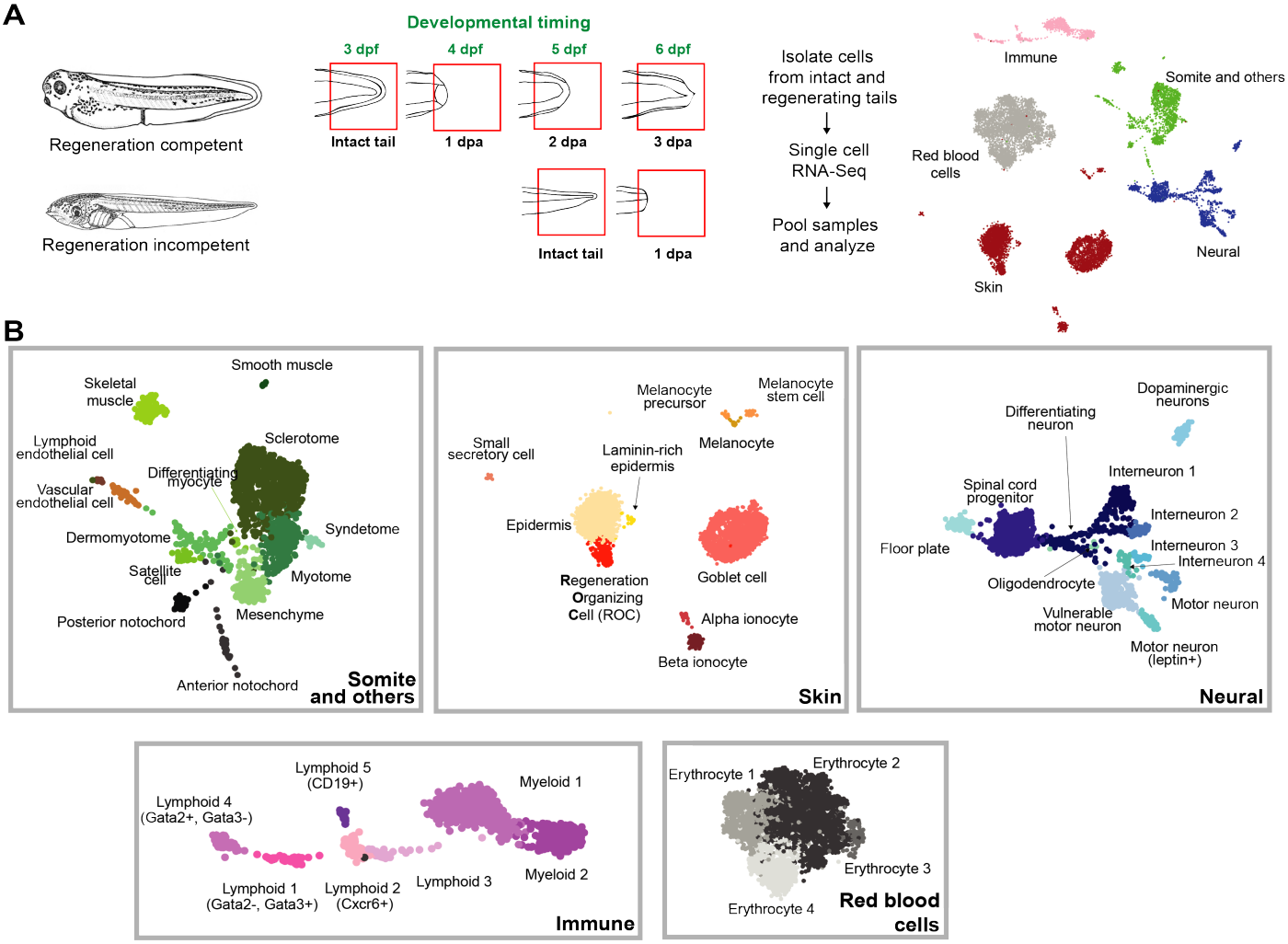

Fig. 1. Pooled transcriptional cell state atlas of the *Xenopus laevis* tail before and after amputation. **(A)** Samples were prepared for single-cell RNA-seq analysis from regeneration-competent and incompetent tadpoles, collecting either intact tails, or tails at various stages following amputation: 1-3 days post amputation (dpa) for regeneration-competent, and 1 dpa for regeneration-incompetent tadpoles. Developmental timing is indicated for each sample (days post-fertilisation, dpf). Samples were processed separately for sequencing and then pooled for UMAP visualisation (Methods). Each dot represents a single cell; colour indicates main tissue group (n≥2 for each sample). **(B)** Cluster identities based on established cell type markers. For details of cluster annotations, see main text, fig. S1 and Methods.

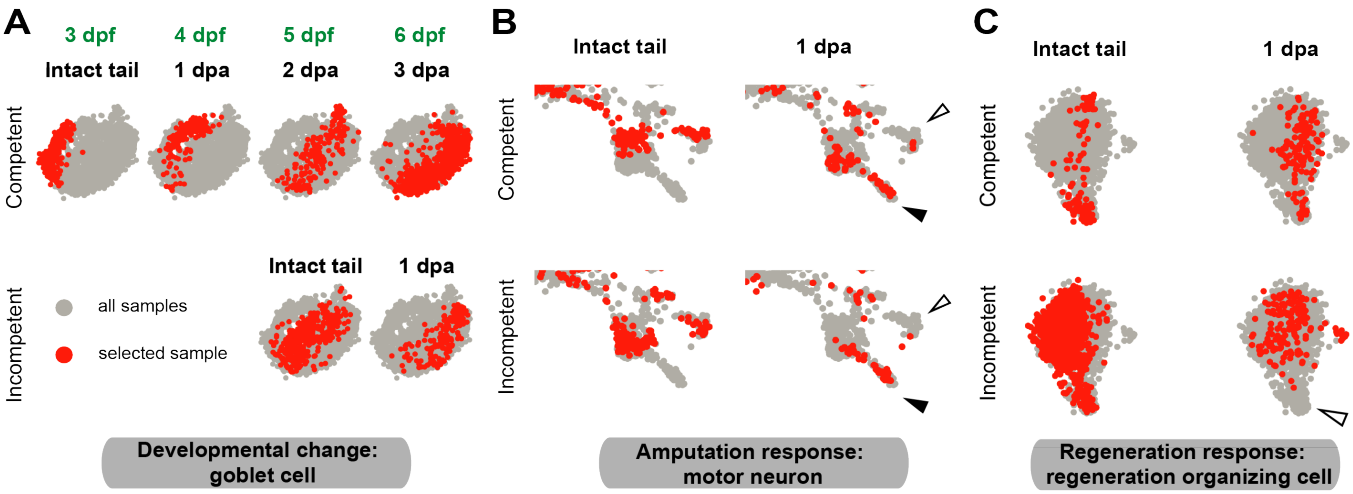

Fig. 2. Comparison of scRNA-seq samples discriminate gene expression and cell state changes that take place during development from those associated with the response to amputation or regeneration. Examples of cell-specific gene expression changes that take place **(A)** during development, **(B)** in response to amputation, and **(C)** in response to regeneration. Grey dots: cells from samples at all respective time points; red dots: cells from selected time point and condition. Black and white filled arrows indicate presence and absence of populations, respectively, when comparing intact tail to 1 dpa samples. Panel A shows a continuous change in the gene expression profile of Goblet cells that takes place during development both in regeneration-competent and incompetent tail; panel B shows gene expression changes that take place in motor neurons in response to amputation, both in regeneration-competent and incompetent tail; and panel C shows differential gene expression changes that take place in epidermis between regeneration-competent and incompetent tail, identifying a cell state change specific to regeneration.

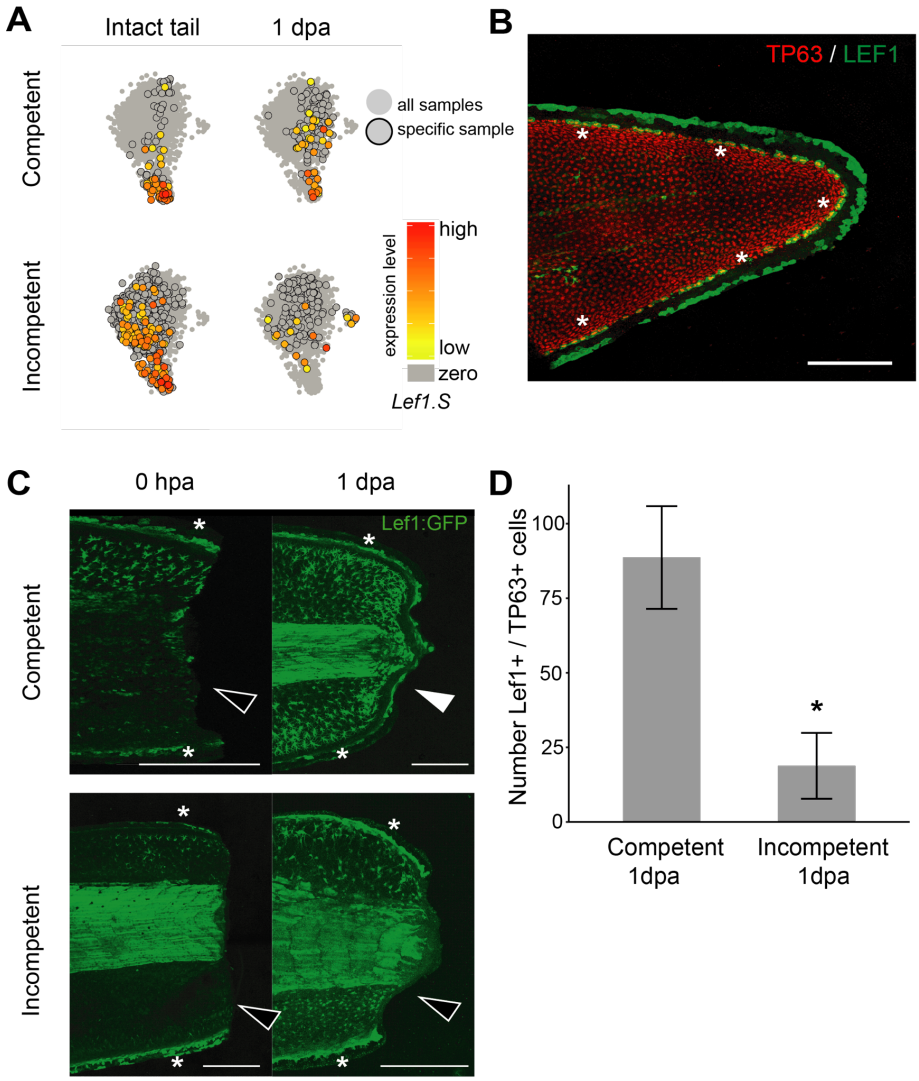

Fig. 3. Regeneration-organizing cells (ROCs) characterize the specialized wound epidermis in regeneration competent tadpole.
(A) ROCs express high *Lef1* mRNA level, and reappear after amputation specifically in regeneration-competent tadpoles. Grey dots: *TP63* positive epidermal clusters; circled dots: selected sample. Relative *Lef1* expression visualized for each cell. **(B)** ROCs (TP63+/LEF1+ cells, denoted by asterisks) are localized along the midline edge of the epidermis in intact tails. Green, pbin7Lef; Red, TP63. Scale bar: 500 µm. **(C)** ROCs (LEF1+) remain along the posterior trunk following amputation (asterisks), but are removed from the amputation plane (empty arrowheads). ROCs specifically reappear in the amputation plane of 1 dpa regeneration-competent tadpoles (filled arrowhead). hpa: hours post-amputation. Green, pbin7LEF:GFP. Scale bars: 250 µm; a total of ≥3 tadpoles per conditions were imaged from 2 biological replicates. **(D)** Quantification of TP63+/LEF1+ cells at the amputation plane (mean ± standard deviation) shows a significant reduction in regeneration-incompetent tadpoles at 1 dpa (n=12 and n=11 for competent and incompetent samples, respectively, both from 2 biological replicates). *: $p < 0.001$.

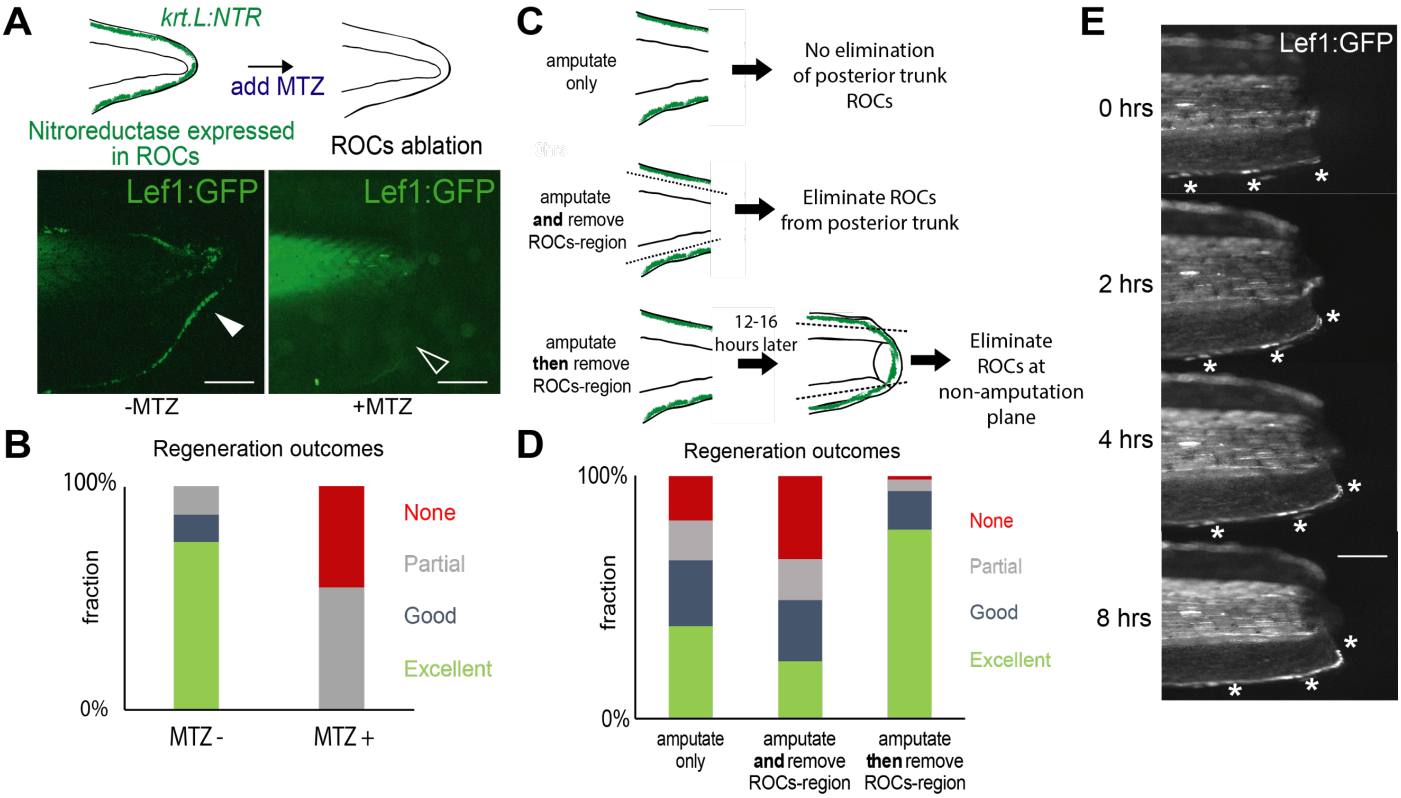

Fig. 4. Regeneration-organizing cells relocation to the amputation area mediates tail regeneration. **(A)** Nitroreductase (NTR)/Metronidazole-mediated ablation of ROCs during regeneration. pbin7LEF:GFP /Krt.L:NTR F_0 transgenic tadpoles (bottom) show successful cell ablation at 3 dpa: GFP positive ROCs are present in control (plain arrowhead) but lost in MTZ treated animal (empty arrowhead). Scale, 1 mm. **(B)** ROC-ablated tadpoles cannot regenerate (n=11 from 2 biological replicates). **(C)** Schematic of ROCs-containing region manual removal protocol. Green colored area indicate ROCs localization. **(D)** Manual removal of posterior trunk ROCs at the same time as tail amputation reduces regeneration (n= 45 from 5 biological replicates), but manual removal of posterior trunk ROCs 12-16 hours post amputation does not negatively affect regeneration (n= 95 from 3 biological replicates). **(E)** Time-lapse images of ROCs relocating to the amputation plane, as assessed by pbin7LEF:GFP. Asterisks denote cells with brighter GFP that can be tracked (n=8 from 3 biological replicates). Scale bar, 500 µm.

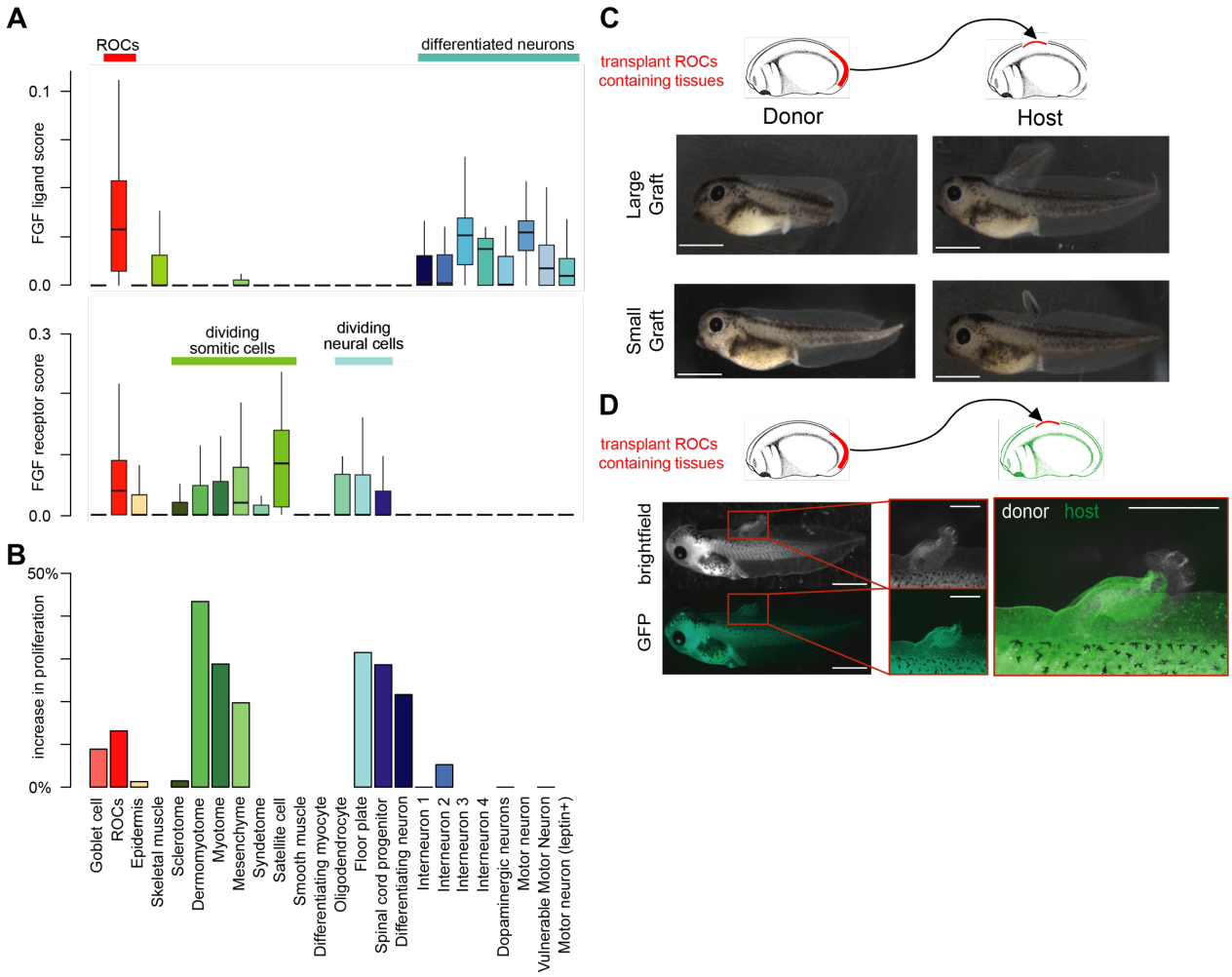

Fig. 5. ROCs act as a signalling centre coordinating progenitor outgrowth during tail regeneration. **(A)** Expression of FGF ligands (Top) and receptor (Bottom) shown for selected cell types as a boxplot (outliers not shown). **(B)** Bar plot indicating the change in the fraction of cells in G2/M and S phases between regeneration-competent 2 dpa and incompetent intact tail samples, all taken at 5 dpf. **(C)** (Left) Removal of large or small ROC-containing tissues causes tail development defects in donors (n= 20), (Right) grafting these regions to the trunk enables tail-enriched or fin-enriched distal growth in hosts, respectively. Matching donor-acceptor pairs are shown 2 days post-grafting (n= 20 from 3 biological replicates). **(D)** Non-labelled grafts to CMV:GFP positive embryos induce outgrowth containing GFP positive cells; donor tissues are at the tip of the ectopic structure (n=12 from 3 biological replicates). Green, CMV:GFP. Scale bars: full tadpoles, 1 mm; zoomed grafts and merged graft images, 500 µm.

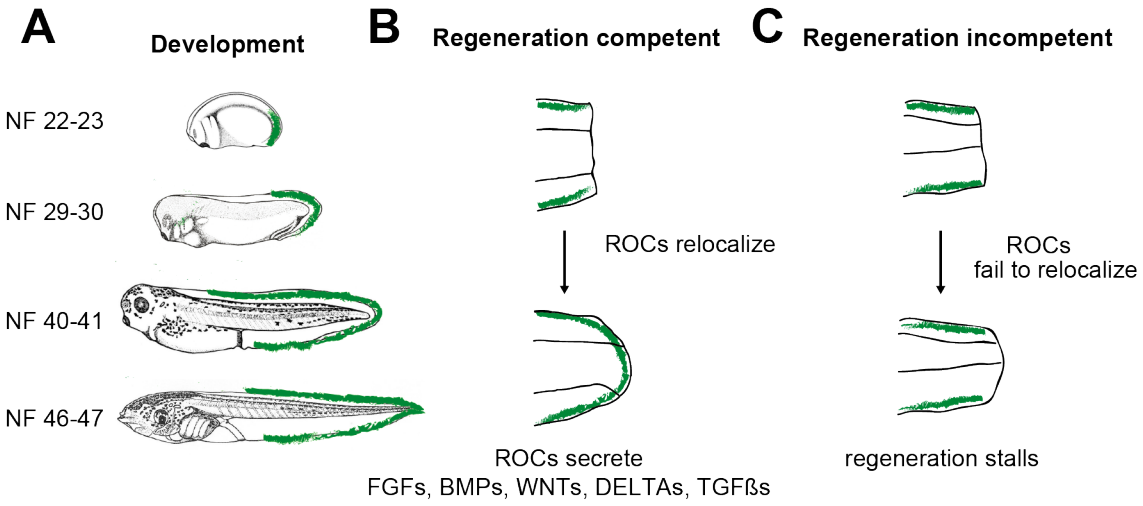

Fig. 6. ROC-based model of tail regeneration.
Transcriptional signature of ROCs first appears in NF stage 22-23 embryos at the tip of the tail-bud, then expand towards the edges of the epidermis midline from the tail tip to the posterior trunk during development (table S2). Relocalization of ROCs to the wound area forms the specialized wound epidermis and is a hallmark of successful tail regeneration.

Science
AAAS

Supplementary Materials for

Identification of a regeneration organizing cell in the *Xenopus* tail

C. Aztekin†, T.W. Hiscock†, J.C. Marioni, J. B. Gurdon, B. D. Simons and J. Jullien

† These authors contributed equally

Correspondence to: J. Jullien (jj256@gurdon.cam.ac.uk), B. D. Simons (bds10@cam.ac.uk)

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S11
Captions for Tables S1 to S3

Other Supplementary Materials for this manuscript include the following:

Table S1 to S3

Materials and Methods

Tadpole Generation and Husbandry

Xenopus laevis embryos and tadpoles were generated and maintained in 0.1X MMR (0.1 M NaCl, 2.0 mM KCl, 1 mM $MgSO_4$, 2 mM $CaCl_2$, 5 mM HEPES (pH 7.8) as previously described (27), and raised to the indicated developmental stages according to Niewkoop and Faber (NF) developmental tables (*28*). Wild-type *Xenopus laevis* were used for experiments unless otherwise stated. Animal experiments were approved by the University Biomedical Services at University of Cambridge and complied with UK Home Office guidelines (Animal Act 1986).

Regeneration Assays

Regeneration competent tadpoles were staged as NF stage 40-41, and regeneration-incompetent tadpoles were staged as NF stage 46-47. Tadpoles were anesthetized by incubation for 30-60 seconds at room-temperature (RT) with 0.1% MS222 (Sigma, E10521) in 0.1X MMR. Anesthetized tadpoles were then transferred to fresh 0.1X MMR for amputation. Amputation was carried out by removing ~30-50% of the tail using a sterile scalpel. To remove posterior trunk ROCs, dorsal and/or ventral fin regions were similarly removed, either at the same time, or 12-16 hours post-amputation. Control tadpoles were subjected to the same protocol for anesthesia. Regeneration scoring was assessed at 7 dpa, and classified as excellent, good, partial or none, as previously described (*27*). Briefly, "Excellent" tails show an elongation indistinguishable from normal tails, except for missing somite segmentation. "Good" tails have either problems with elongation, or are missing fin regeneration. "Partial" tails are much shorter, or have defects in patterning with missing fin regeneration, or show an elongated bulge formation. "None" tails present either a blunt end or a small bulge. We calculated the regeneration index by multiplying the tadpole numbers showing excellent, good, partial, or none regeneration phenotypes by 3, 2, 1, or 0 points, respectively, and then divided the result by the total number of tadpoles. Any sample with less than 3 tadpoles in a replicate were omitted for analysis. Any experiment with visible contamination or unexplained tadpole death was omitted for analysis. A maximum of 10 tadpoles were cultured in each well of a 6-well plate. Each experiment was performed with multiple biological replicates, with each batch derived from a different male or female adult.

Single cell dissociation, library preparation and sequencing

30-40% of intact tails or cells formed posterior to the amputation plane were collected (>20 tadpoles per sample) and dissociated to single cells by first washing with Ca-Mg free 1X MBS ((Barth-HEPES Saline) 10X stock: 88 mM NaCl, 1 mM KCl, 2,4 mM NaHCO3, 0.82mM MgSO4.7H2O, 0.33mM Ca(NO3)2.4H2O, 0.41 mM Cacl2.6H2O, 10 mM HEPES. Add ~3 mL of 10N NaOH to obtain a pH of 7.4 to 7.6), then incubating with 1X Trypsin (Sigma, 59427C) in Ca-Mg free 1X MBS with 0.5 µM EDTA for 5 minutes at room-temperature. Trypsin reaction was diluted with Ca-Mg free 1X MBS after 5 minutes. Before and after trypsinization, physical dispersion was applied by trituration with a pipette. Since samples contain few newly-generated cells at 1dpa, anterior tissues for these samples were also collected. Cells were spun down at 211 g for 5 minutes then resuspended in 1X MBS. Cells were passed through a 35 µm diameter cell strainer then stained with 2 µg/ml propidium iodide (PI) (Sigma, P4170) and 20 µM Hoechst 33342 (Sigma, 2261) in 1X MBS for 5 minutes, and single live cells (PI negative, Hoechst positive) were sorted using a Sony SH800s Cell Sorter. scRNA-seq libraries were generated using 10X Genomics and sequenced on an Illumina HiSeq 4000.

scRNA-seq: Data processing

10X Genomics output files were processed using CellRanger v2.1.0, which aligns, filters and counts unique molecular identifiers (UMIs) to generate a (cell x gene) count matrix. Sequences were mapped to the *Xenopus* laevis 9.1 genome (Xenbase, ftp://ftp.xenbase.org/pub/Genomics/JGI/Xenla9.1/Xla.v91.repeatMasked.fa.gz and ftp://ftp.xenbase.org/pub/Genomics/JGI/Xenla9.1/1.8.3.2/XL_9.1_v1.8.3.2.allTranscripts.gff3.gz). Cells were distinguished from empty droplets using CellRanger; no further cell filtering was performed. Raw counts were normalized by total counts per cell, and converted to TPX (*29*) (transcripts per 10^4).

scRNA-seq: Data visualization

Highly variable genes (Fano factor > 65th percentile, HVGs) were selected for clustering and visualization, with both lowly (mean expression < 5th percentile) and highly expressed (mean expression > 80th percentile)

genes removed to minimize batch effects (*30*). One cell (out of 13,200) showed no expression of any of the HVGs and was therefore removed from subsequent analysis. Data were projected onto two dimensions with UMAP (*6*), using log2-transformed values for highly variable genes, cosine distance as a similarity measure, and parameters k = 20, min_dist = 0.5. Gene expression plots are colored according to expression level (log10-transformed, normalized), with zeros colored gray, and non-zero expression colored by a yellow-to-red colorscale (R package *ggplot2*).

scRNA-seq: clustering

To perform clustering, we first constructed a graph to represent the similarity structure of the data. Rather than using a k-nearest-neighbor graph, we constructed the graph in the same way that UMAP describes similarity during high dimensional data visualization (*31*). Specifically, we built a fuzzy topological representation of the data as a weighted graph, using *fuzzy_simplicial_set* from UMAP, using log2-transformed HVGs, cosine distance as the similarity measure, and k = 10 nearest-neighbors. Having constructed this graph, we performed graphical clustering using the walktrap algorithm (*cluster_walktrap* from R package *igraph*, with steps = 10). The parameter, k, controls the size of clusters generated: smaller values of k results in more clusters of a smaller size. Here, k was manually tuned to generate clusters that were in accordance with manually annotated cell types.

scRNA-seq: annotation of cell-types

Cluster identity was determined by looking for genes uniquely upregulated in different regions of the UMAP projection (*findMarkers* in R package *scran* (*32*)), and comparing to expression of known marker genes. Where possible, clusters were assigned to a well-characterized, functional cell type; heatmaps of key marker genes are provided in fig. S1. However, in 23/46 clusters, an unambiguous annotation was not possible. In these cases, clusters were assigned a broad label (*e.g.* lymphoid), with a numeric label distinguishing between clusters within the same broad label; heatmaps of putative marker genes are provided in fig. S1. Marker gene expression was visualised using heatmaps (R package *ggplot2*), coloured by mean expression (log10-transformed, normalized), and adjusted relative to the cluster with maximum expression of that gene (except in fig. S6B, where we do not perform this adjustment). *Xenopus* laevis is pseudotetraploid, and gene expression is available for each allele from the Large (Gene.L) or Short (Gene.S) chromosomes (*33*). In this work, we did not combine these two expression patterns as they may have different functions and we have recorded cluster-specific allele expression differences. We report both large and short allele expressions if needed. In other figures (such as heatmaps), we report whichever allele has higher maximal expression.

scRNA-seq: differential abundance analysis

Differential abundance analysis was performed using edgeR (*34*), following the method introduced for mass cytometry data, cydar (*35*). Briefly, for all samples (all replicates across all experimental conditions), the number of each cell type was computed, thereby generating a counts matrix, with columns corresponding to cell types and rows corresponding to individual samples. Cell types with low mean abundance (< 5 cells per sample) were removed from the analysis; red blood cells were also removed since we observed high variability in their capture efficiency. Pairwise differential abundance tests were performed between regeneration-competent and incompetent 1 dpa conditions, with edgeR accounting for the replicate structure and counts-based nature of the data. Note that covariates (such as sequencing batch) were not explicitly considered in this analysis.

scRNA-seq: Gene-ontology and gene set enrichment analysis

Putative marker genes for ROCs were identified from the HVGs using *findMarkers* from *scran*; the top 50 genes were selected. Gene-ontology analysis was then performed on *Homo sapiens* orthologs using http://geneontology.org/page/go-enrichment-analysis (*36*). Single cell gene set enrichment analysis (scGSEA) was performed using the *AUCell* R package (*37*), using manually curated gene lists (*38*), provided in Table S3, and all genes as the background gene set. Cell cycle phase was inferred from the transcriptome using *CellCycleScoring* (R package *Seurat*) (*39*), with the gene set defined previously (*40*) (see also Table S3). To compute the change in proliferation during regeneration (Fig. 3B), we computed the fraction of each cell type in either G2/M, or S phases, and compared regeneration-competent 2 dpa and incompetent intact tail samples.

These samples were chosen to have the same developmental stage. Estimates for proliferating fraction are noisy for cell types with low abundance (< 10 in either sample); these are not plotted.

Microsurgery

Graft material was collected from the posterior tail bud of donor embryos; prepared in one of two ways which we denote as "large" or "small". Large grafts are taken directly from the posterior tail bud at stage NF 24-25; these grafts likely contain multipotent tailbud progenitors, in addition to the overlying skin in which ROCs are present. Small grafts are produced by either peeling off the skin region directly from NF stage 24-25 tails, or by isolating skin from the "large" grafts by cleaning them with a needle. For the trunk-to-trunk grafts, donor material is sourced by peeling different dorsal or lateral regions of trunk. Then, different regions of trunk (dorsal vs. lateral vs. ventral) of an acceptor embryo NF Stage 24-28 were prepared for grafting by peeling the skin. Anterior edges of the donor tissues were then put onto the peeled trunk area. These manipulations were carried out in high-salt solution (0.7 mM $CaCl_2$, 20 mM NaCl in 1X MBS) on clay, and embryos remained on clay for 1 hour before transfer to 0.1X MMR. Stereoscope images were taken 2 days after grafting. For experiments using GFP labelled host or donor embryos, Xla.Tg(CMV:GFP)Brown transgenic testes were used to generate tadpoles, and tadpoles were sorted for GFP positivity prior to grafting. Images of tadpoles were taken on a Leica stereomicroscope with microscope camera DFC7000 T or M80.

Immunofluorescence and Imaging

Tadpoles were fixed with 4% formaldehyde in 1X PBS for 40 min, permeabilized by three 20 min incubations in PBS-T (1X PBS + 0.1 Tween-100), blocked with 100% CAS-Block (Invitrogen, 008120) for 1 hour, and then stained with primary antibodies in 100% CAS-Block overnight. Samples were then washed with PBS-T for 10 min 3 times, re-blocked in 100% CAS-Block for 30 min, and stained with secondary antibodies for 2 hours. Samples were then washed 3 times 10 min in PBS-T followed by 3 times 20 min in PBS. Samples were incubated in DAPI solution (Invitrogen, P36962) for 10 minutes. The entire procedure is carried out at room temperature, except for incubation of the primary antibody at 4 degrees. Samples were mounted in 0.8% ultra-low gelling temperature agar (Sigma, A5030) in 0.1X MMR, and imaged on a Leica SP8 confocal microscope. A 10x/0.4 HC PL Apo CS2 Air objective was used for all images except: Fig. 3B, which is taken with 20x/0.7 HC PL Apo CS2 Air, and Fig S3B and Fig S5B, which are taken with a 63x/1.2 HC PL Apo CS2 Water objective. The following laser lines were used: 405 nm (DAPI), 470-670 nm White Laser (EGFP, Alexa Fluor 594). Fig S6D and S7E immunofluorescence images were taken via a Leica stereomicroscope with a microscope camera DFC7000 T. Fiji was used for maximum projection of z-stacks and to adjust contrast to highlight biological relevance. If needed, images were cropped, flipped, and/or rotated to highlight biological relevance. Primary antibodies used: TP63 (clone 4A4) (Abcam, ab735, 1:100), EGFP (Invitrogen, A-21311, 1:500), PCNA (Acris, SM1421P, 1:100), PHH3 (Abcam, ab10543, 1:100). The following primary antibodies were purchased from Developmental Studies Hybridoma Bank and used with 1:5 dilution: 12/101, NCAM1 (4D), COL2A1 (CIIC1). Secondary antibodies used were: Alexa Fluor 488 secondary antibody (Invitrogen, A11001, 1:500), Alexa Fluor 594 secondary antibody (Invitrogen, R37121, 1:500).

Live imaging

pbin7Lef:GFP transgenic testes were used to generate tadpoles. After amputations, regeneration-competent tadpoles were cultured in 0.8% ultra-low gelling temperature agar (Sigma, A5030) in 0.1X MMR, and imaged once every 2 hours using a Leica stereomicroscope with a microscope camera DFC7000 T. Fiji was used to adjust contrast to highlight biological relevance. If needed, images were cropped, flipped, and/or rotated to highlight biological relevance.

Chemical perturbations to assess ROCs mobilization

pbin7Lef:GFP transgenic testes was used to generate tadpoles. Lef1 positive tadpoles were sorted prior to experiments. Amputated tadpoles were placed in 0.1X MMR containing 0.1% DMSO, Diphenyleneiodonium chloride (DPI) (Sigma, D2926), SU5402 (Sigma, 0443), SB-505124 (Sigma, S4696) with indicated concentrations. Tadpoles were fixed after 16 hours post amputation and stereoscope images were taken. The degree of ROCs mobilization was quantified by measuring the distance of leading Lef1 positive cells (either dorsal or ventral) from the edges of the amputation area towards notochord.

Genetic ablation experiments

The 1853 bp region upstream of the Krt.L coding-sequence was amplified from Xla genomic DNA and cloned into p2 pTransgenesis vector (*41*) (XLA KRT 1-8 F: caaaaaagcaggctCgCCaGTCGACACTAATGTTTTTTTTAACATATTGTGTTGGACCCA, R: TTGTATTCTATAGTGTCACCTAAATTCTGCGCAACAACACTGTTTAGACAA). CFP-Nitroreductase was amplified from pCS2.CFPNTR (*42*) vector via standard PCR protocols and cloned into p3 pTransgenesis vector (*41*) using In-Fusion cloning (CFPNTR F: TTCTTGTACAAAGTGGGGATCCACCATGGTGAGCAAGGG,
R: CTCACTATAGTTCTAGAGGCTCGAGCACTTCGGTTAAGGTGATGTTTTGCG). Krt.L:CFP-NTR, Krt.L:VenusGFP, CMV:VenusGFP, and CMV:dTomato pTransgenesis vectors containing I-SceI cut site were generated by Gateway cloning using LR Clonase II Plus (Invitrogen, 12538120) and by cloning into pDEST Transgenesis 4-2 (Tol2/ISce-1-CH4-SAR/Tol2). F_0 transgenic lines were generated by using I-SceI mediated transgenesis as previously described (*43*). Briefly, 1-cell stage embryos were injected with a 3:1 mix of p4_Krt.L:CFPNTR, and p4_CMV:dTomato or p4_CMV:VenusGFP constructs with a I-SceI enzyme (NEB, R0694L). After incubation at 14 degrees overnight or until neurula stages, embryos were switched to 23 degrees, then grown to the indicated stages. Transgenesis efficiency was assessed by evaluating on an epifluorescence stereomicroscope the extent of tomato or venusGFP fluorescence in injected tadpoles. Broad signal emitting tadpoles were considered as high, and mosaic signal emitting tadpoles were considered as low transgenesis. For regeneration assays, tadpoles with high or low dTomato or VenusGFP expression were incubated in 10 mM MTZ (Sigma, M1547). MTZ was added one day before amputation, after amputation, and 2 days post amputation, then removed on 6 dpa. For the assessment of ablation experiments, Krt.L:NTR pTransgenesis was performed on embryos generated using pbin7LEF:GFP testes. Tadpoles were first sorted for high *Lef1* reporter activity and then put through the ablation protocol. During the regeneration assay, ablation of ROCs upon MTZ treatment was checked in these tadpoles at 3 DPA using an epifluorescence stereomicroscope. As MTZ is light-sensitive, all samples were maintained in the dark in foiled dishes.

Genetic ablation combined with grafting experiments

To ablate ROCs from donor grafts, Krt.L:NTR transgenic embryos (together with CMV:VenusGFP, or CMV:dTomato) were generated and broad-high signal emitting embryos were sorted at early NF Stage 20 as described above; wild type embryos acted as controls. Donor embryos at stages 23-25 were used to generate "small grafts", as described above, which were then implanted onto different trunk regions of wild-type acceptor embryos. After grafting, embryos were randomly allocated in wells with or without 10 Mm MTZ for 2 days (MTZ changed daily). The extent of the ectopic outgrowths was quantified as the length between the visible proximal and distal ends of the outgrowths (Fiji). As MTZ is light-sensitive, all samples were maintained in the dark in foiled dishes.

Statistical tests

For Fig. 3D, and fig. S11D-E, we fit a linear model to the data (function lm() in R) to test for significant differences, accounting for experimental batch effects. For fig. S7C and S7E, we first binarized the data: "Excellent" and "Good" into one group; "Partial" and "None" into another. We then perform logistic regression (function glm() in R, with family = binomial(link = "logit")), again accounting for batch effects. For each comparison, we report the p-value as the likelihood ratio test between the data and a null model where the variables that are being compared are merged (function anova() in R, with method = "LRT"). For Fig. S8C, we performed a two tailed t-test assuming unequal variances.

Data availability

We are currently uploading our data to ArrayExpress (E-MTAB-7716).

Code availability

Code will be deposited on github prior to publication.

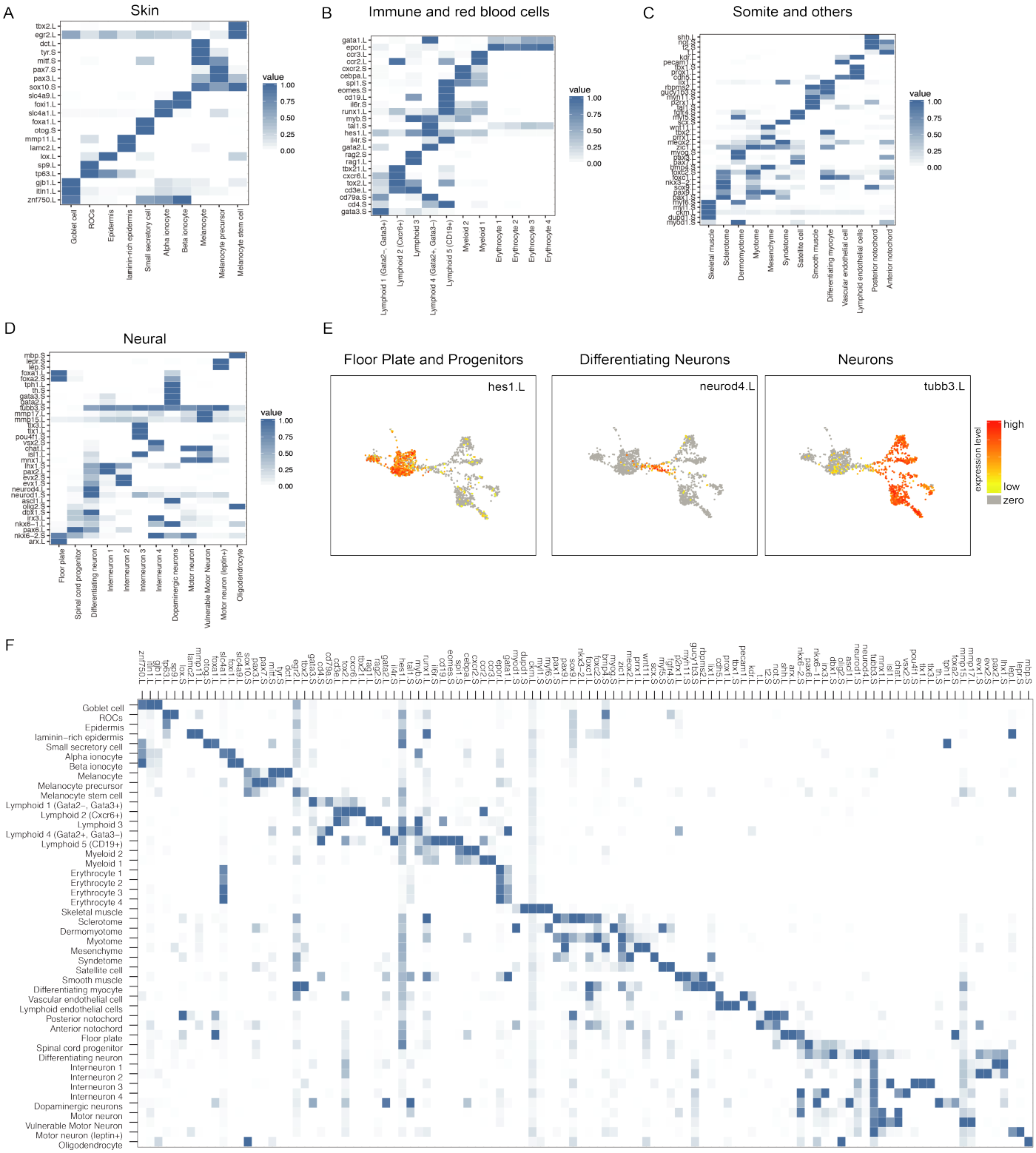

Fig. S1. Identification of cell states corresponding to 46 scRNAseq clusters based on established markers of cell identity.

Heatmaps showing selected marker genes for **(A)** skin; **(B)** immune system and red blood cells (RBCs); **(C)** somite and other tissue types; and **(D)** nervous system. Detected mRNA values are normalized to the highest expressing cell type. **(E)** Expression of selected genes to visualize differentiation of neurons from progenitors: *Hes1* is detected specifically in the floor plate and neural progenitor cells. *Neurod4* expression is seen only in actively differentiating neurons, and not in neural progenitors, nor in terminally differentiated neurons. *Tubb3* is only present in terminally differentiated neurons. **(F)** Heatmap combining all marker genes and cell types from (A-D).

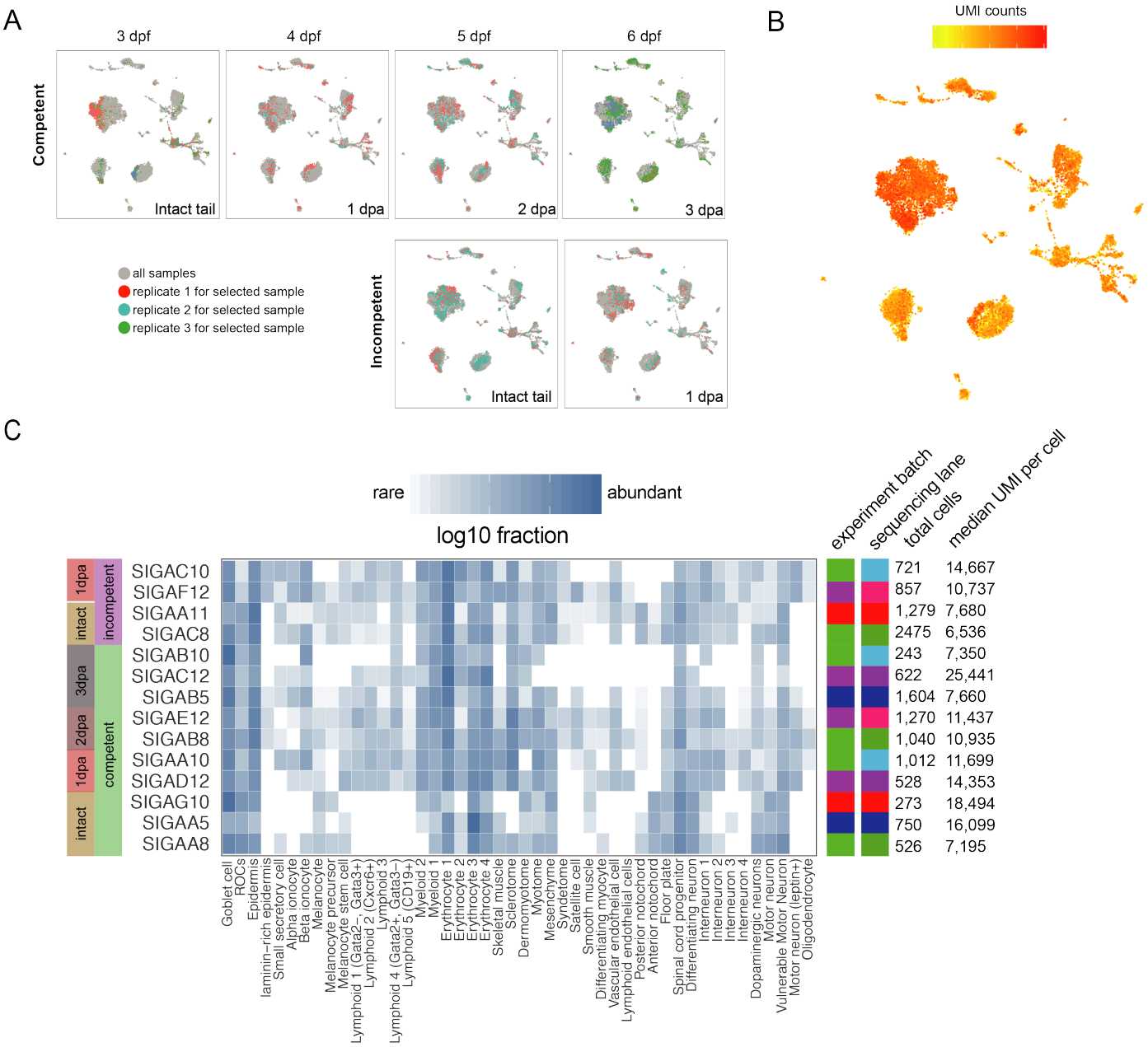

Fig. S2. Contribution of biological replicates to the UMAP projection of tail scRNA-seq data.
(A) UMAP visualization of cells from all biological replicates with cells derived from each replicate colour-coded. Grey dots: cells from all samples; red, blue, green dots: cells from different biological replicates for selected sample. Most clusters show mingling of cells from replicates. Regeneration-competent intact tail and 3 dpa samples have three replicates, all other samples are from two replicates. **(B)** UMI counts (log10) plotted on the UMAP visualization. **(C)** Heatmap of cell type abundances for each biological replicate. The color scale corresponds to the fraction of cells allocated to each cell type per sample (plotted on a log10 scale). Also shown are the experimental batch and sequencing lane.

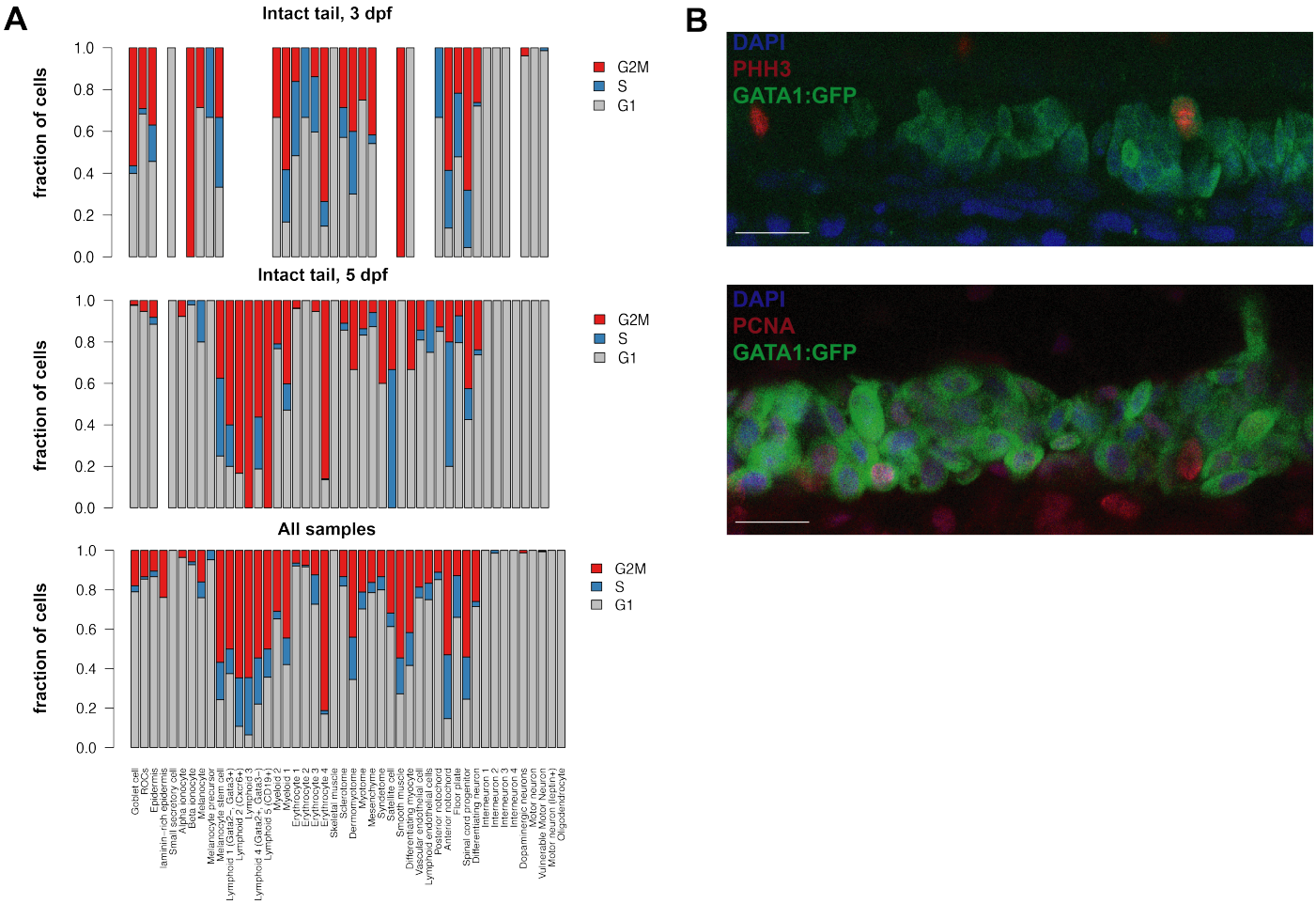

Fig. S3. Computational inference of cell cycle phase reveals differences between progenitors, terminally differentiated cells and erythrocytes.

(A) Barplot showing the cell cycle status of each cell type as a fraction of cells in G1, G2/M and S phases for (Top) Intact tail regeneration-competent, (Mid) Intact tail regeneration–incompetent, and (Bottom) all samples. **(B)** Erythrocytes are positive for cycling protein markers in the scRNA-Seq data. GATA1:GFP transgenic line is used to visualize erythrocytes and stained with cycling markers, PCNA (n=7 from 3 biological replicates) and PHH3 (n=5 from 3 biological replicates) at Stage 40. Green, GATA1:GFP, Red: PCNA or PHH3, Blue: DAPI. Scale, 25 µm. Double positive cells (GATA1/cell cycling markers) confirm the presence of cycling erythrocytes at this developmental stage.

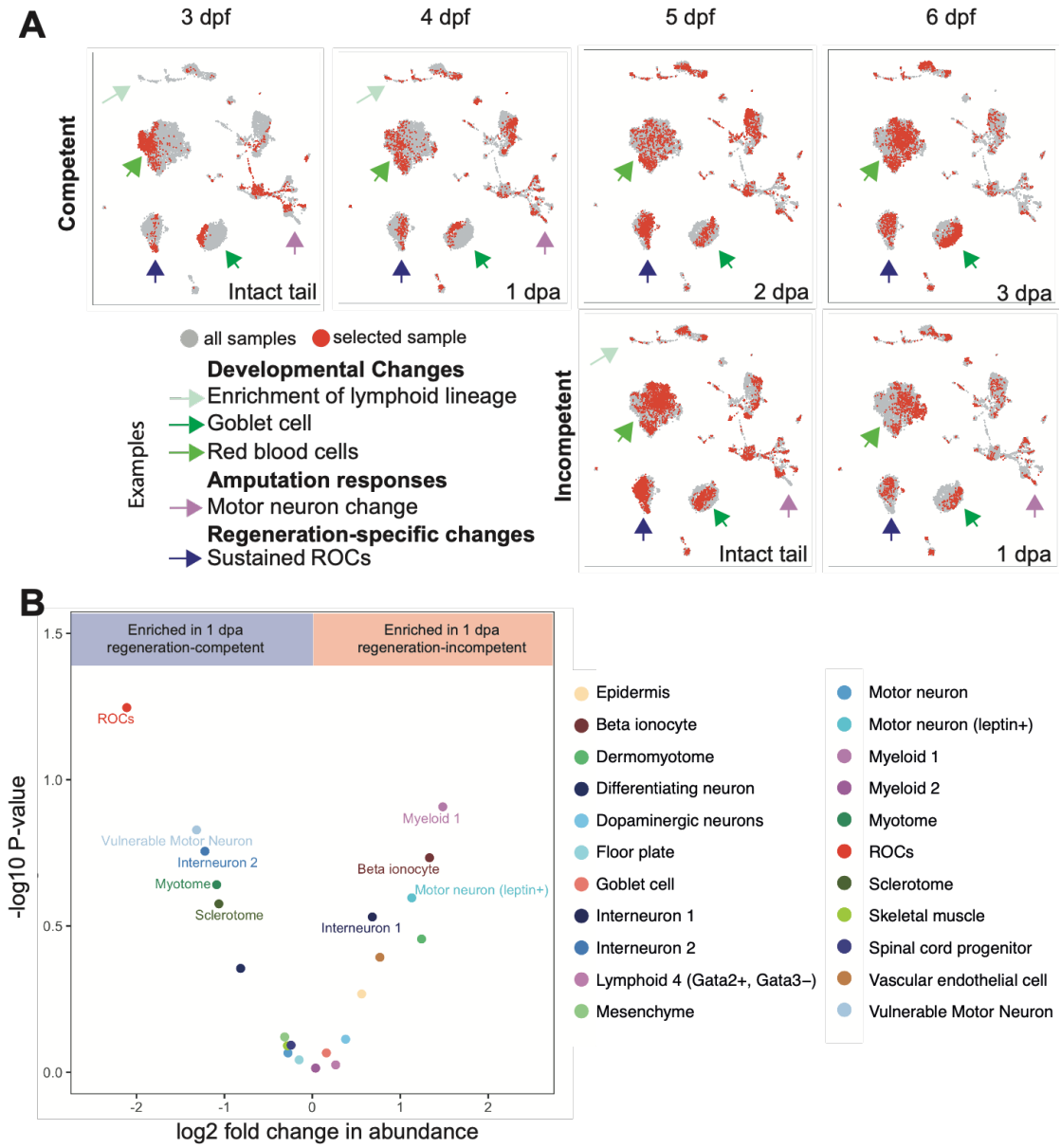

Fig. S4. Developmental, amputation and regeneration related changes are revealed by pairwise comparison of samples.

(A) UMAP visualization for each sample overlaid onto all pooled samples. Red dots represent cells specific to a given sample whereas grey dots represent cells from all pooled samples. Days post-fertilization (dpf) and days post-amputation (dpa) used for comparisons are indicated for each sample. Transcriptional and cell type changes occurring as development proceeds (increasing dpf), regardless of regeneration potential, are classified as developmental (examples shown as green arrows). Cell type changes observed between intact tail and 1 dpa, regardless of regeneration potential, are classified as amputation response (examples shown as purple arrows). Cell type changes that are specific to regeneration-competent samples are classified as regeneration-specific (example shown as blue arrow). **(B)** Regeneration specific changes at 1 dpa were assessed using differential abundance analysis between regeneration-competent and incompetent tadpoles. Colored dots represent different putative cell types shown in Fig 1B. The ranking of the test statistics by cell type highlights the change in ROCs abundance as the most significant change, consistent with the conclusions of fig. S4A. The absolute values of the test statistics are large, but this is to be expected given the low replicate number (n = 2).

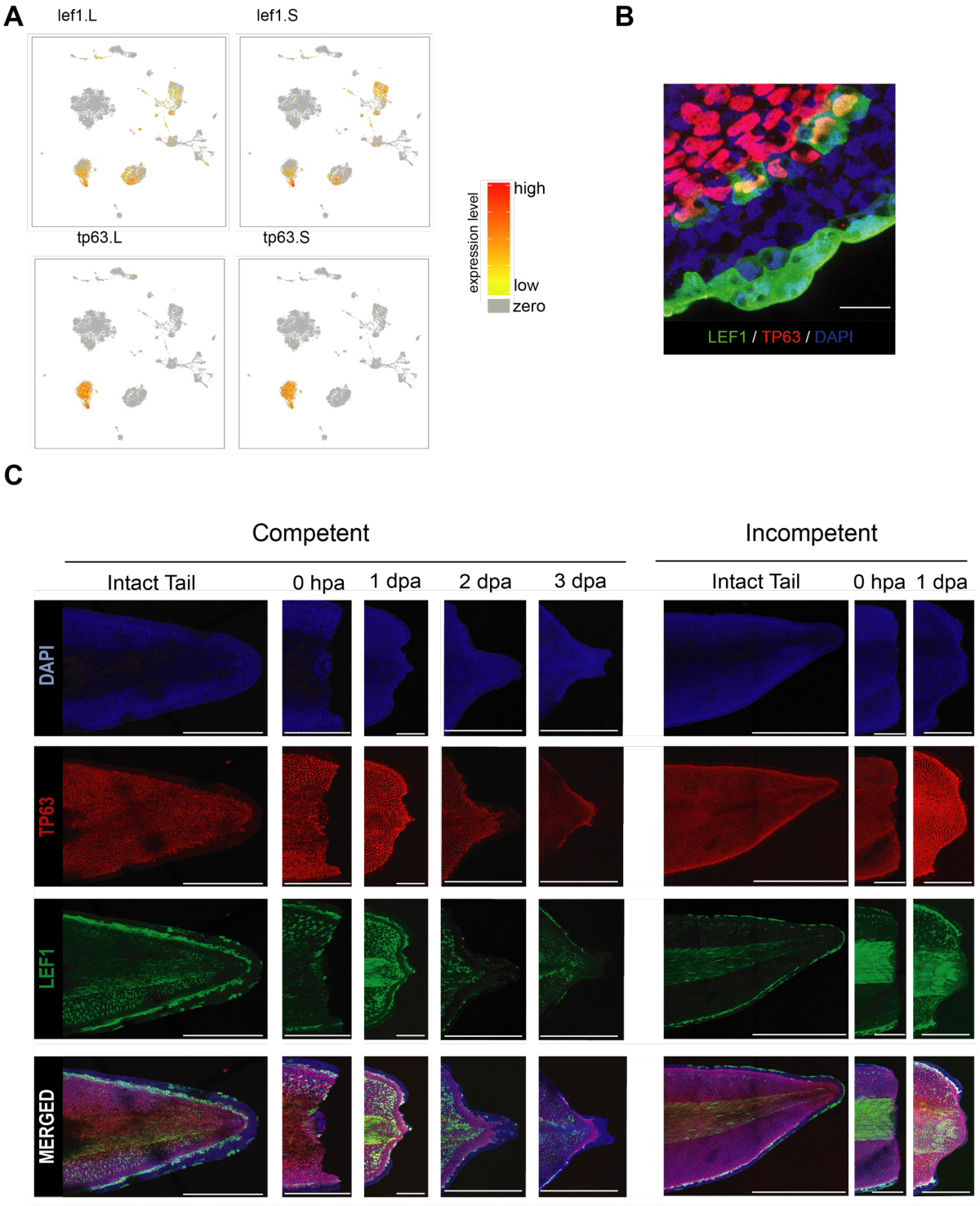

Fig. S5. Identification of ROCs as a regeneration-specific epidermal cell type localized to the amputation plane.

(A) UMAP visualization for *Tp63.L*, *Tp63.S*, *Lef1.L*, and *Lef1.S* expressing cells. **(B)** High magnification image of ROCs (TP63+/LEF1+). Green, pbin7Lef; Red, TP63; Blue, DAPI. Scale bar: 25 µm. **(C)** ROCs (TP63+/LEF1+) are present in regeneration–competent and -incompetent intact tails; are removed following amputation; and reappear at the wound epidermis specifically in regeneration competent tadpoles. Green, pbin7LEF:GFP; Red, TP63; Blue, DAPI. Scale bars: intact tails 500 µm, otherwise 250 µm; (A total ≥3 tadpoles per conditions were imaged from 2 biological replicates). Note, the 0 hpa and 1 dpa Lef1 images are identical to Fig. 3C.

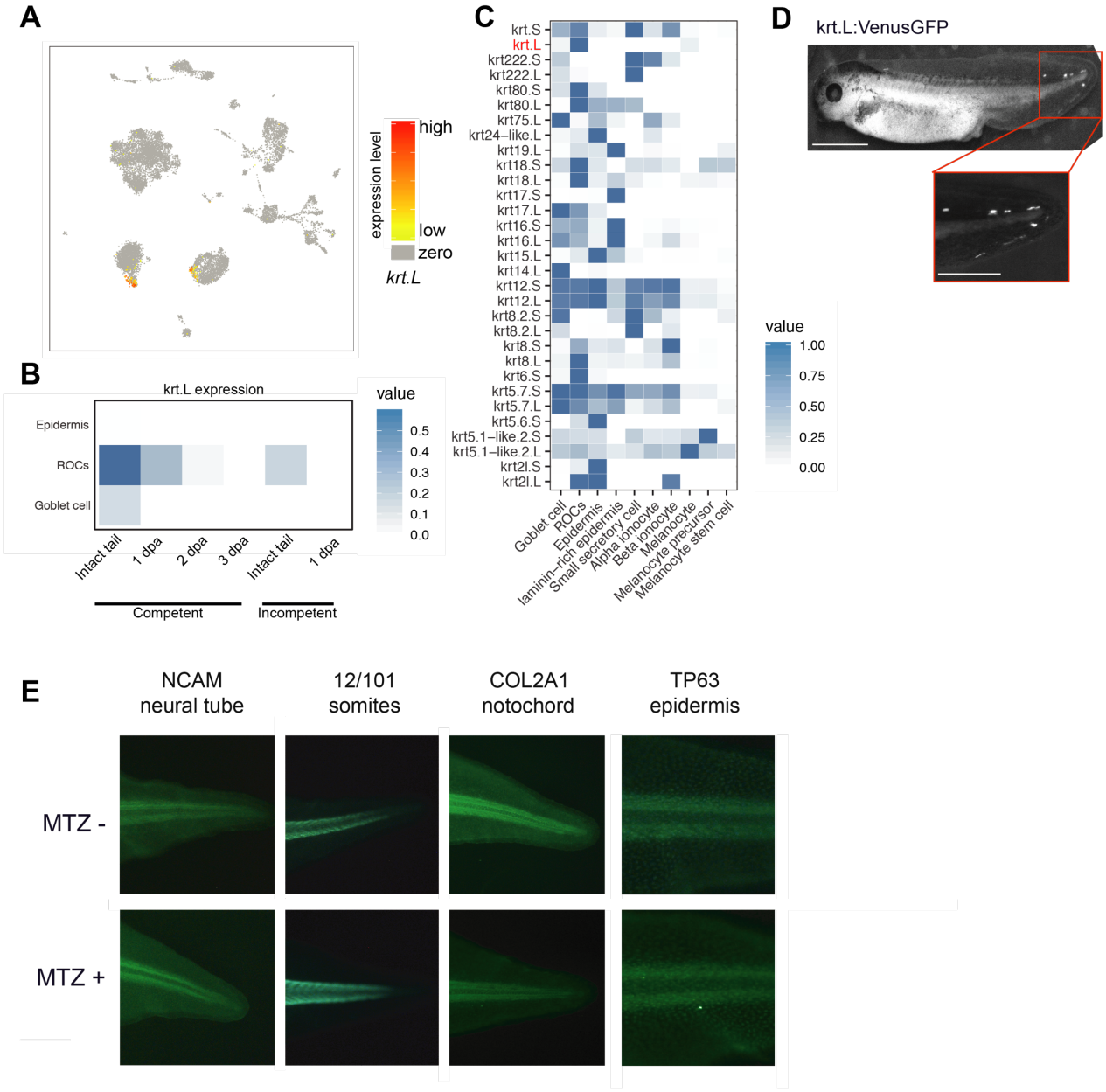

Fig. S6. Ablation of ROCs using the *Krt.L:NTR* / MTZ system
(A) UMAP visualization for *Krt.L* expressing cells. **(B)** Heatmap showing *Krt.L* expression in ROCs, epidermis and goblet cells, across the different biological conditions used in this study. **(C)** Heatmap of keratin-family gene expression (from both large and short chromosomes) within the skin, confirming the high specificity of *Krt.L* to ROCs. **(D)** Mosaic staining of ROCs after GFP immunolabelling in Krt.L:venusGFP F_0-tadpoles. Scale bar: 1 mm. **(E)** Antibody staining of different tissue types following ablation with MTZ confirms the absence of gross off-target effects.

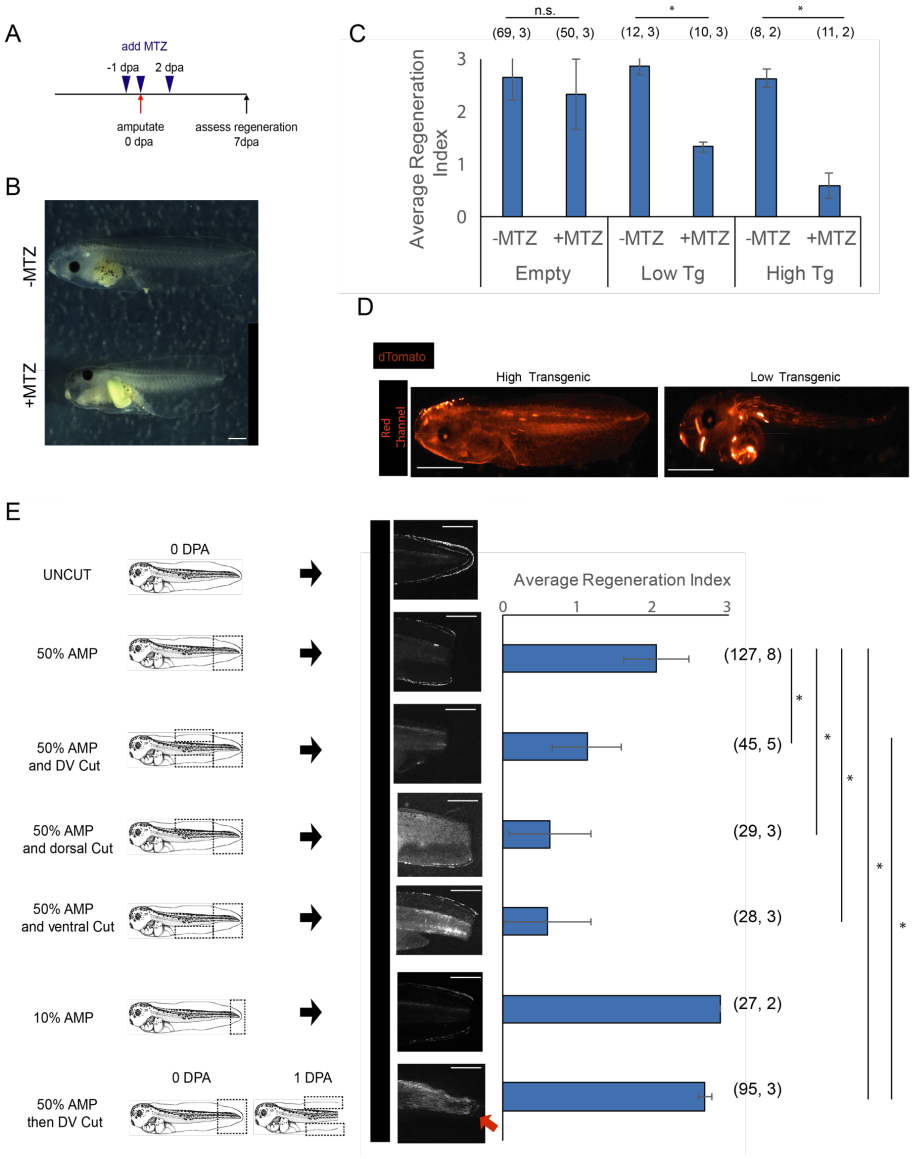

Fig. S7. Removal of ROCs by genetic ablation or manual extirpation reduces regeneration.
(A) Schematic outline of MTZ treatment timing. **(B)** Brightfield images of regeneration-competent tadpoles at 3 dpa, with or without MTZ-based ROCs ablation. Scale bar: 1mm. Lef1:GFP images of the same tadpoles are provided in Fig. 4A. **(C)** Average regeneration index in Krt.L:NTR ablation experiments. Scorings are shown for samples with or without MTZ treatment in high or low transgenic tadpoles as well as in non-transgenic tadpoles. In non-transgenic tadpoles, MTZ treatment does not significantly affect regeneration outcomes compared to no MTZ treatment ($p > 0.05$). In transgenic tadpoles, MTZ treatment significantly reduces regeneration score. All data are mean±std across replicates, *, $p < 0.001$;. n.s, not significant. Replicates are indicated in figures inside brackets as (total tadpole number, biological replicate number). **(D)** Transgenesis levels were assessed based on dTomato fluorescence intensity from a CMV:dTomato vector co-injected with a Krt.L:NTR vector. Examples of high- (left) and low-transgenesis (right) are shown. Scale, 1 mm. (**E**) Extent of ROC removal during amputation correlates with regeneration outcome. Schematics of posterior trunk ROCs removal protocols are shown on the left. Amputations were carried out to remove either 50% or 10% of the tail (50% and 10% AMP, respectively). In the case of 50% tail amputation, additional posterior trunk removals were also performed in the dorsal, ventral, or dorsal and ventral regions of the posterior trunk (50% AMP and dorsal cut, 50% AMP and ventral cut, and 50% AMP and DV cut, respectively). Lastly dorsal and ventral posterior trunk removal was also carried out 1 day after amputation (AMP then DV cut). The extent of ROCs removal is evaluated by imaging GFP fluorescence in pbin7LEF:GFP tadpoles (middle panel) at 1 dpa. Regeneration scores were assessed at 7 dpa and shown for each condition (left). Replicates are indicated in figures inside brackets as (total tadpole number, biological replicate number). *, $p < 0.01$. All data are mean±std across replicates.

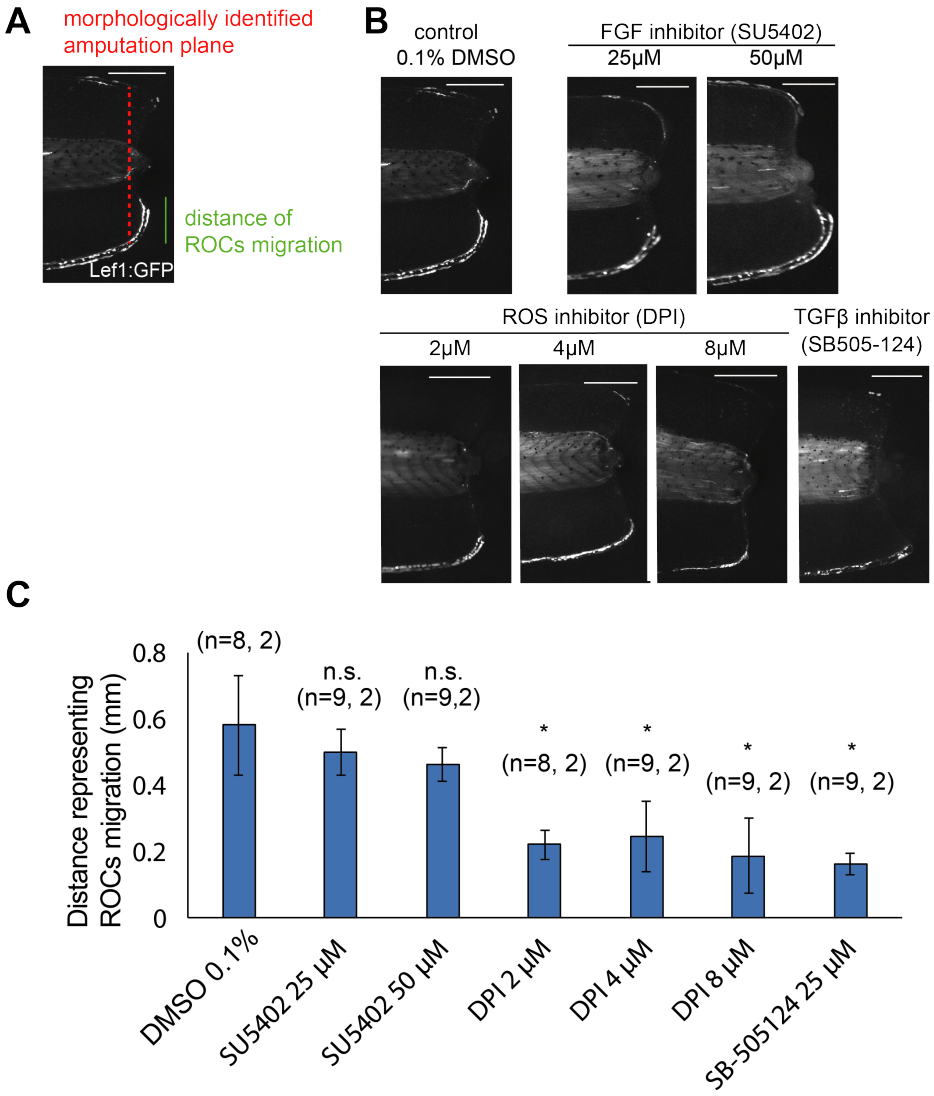

Fig S8. Interfering with early signalling pathways inhibits ROCs mobilization.
(A) Method of quantifying cell migration on the amputation plane. **(B)** Representative images of Lef1 positive tadpoles treated with chemicals that interfere with regeneration and wound epidermis formation. Scale bars = 0.5 mm. Note, the control 0.1% DMSO image is identical to Fig. S8A. **(C)** Chemicals targeting early regeneration signalling pathways (DPI, SB-505124) block ROCs migration, meanwhile inhibition of FGF pathway via SU5402 does not inhibit ROCs migration at 16 hours post amputation. Replicates are indicated in figures inside brackets as (total tadpole number, biological replicate number); *, *p*-value < 0.001. All data are ±std, two tailed t-test.

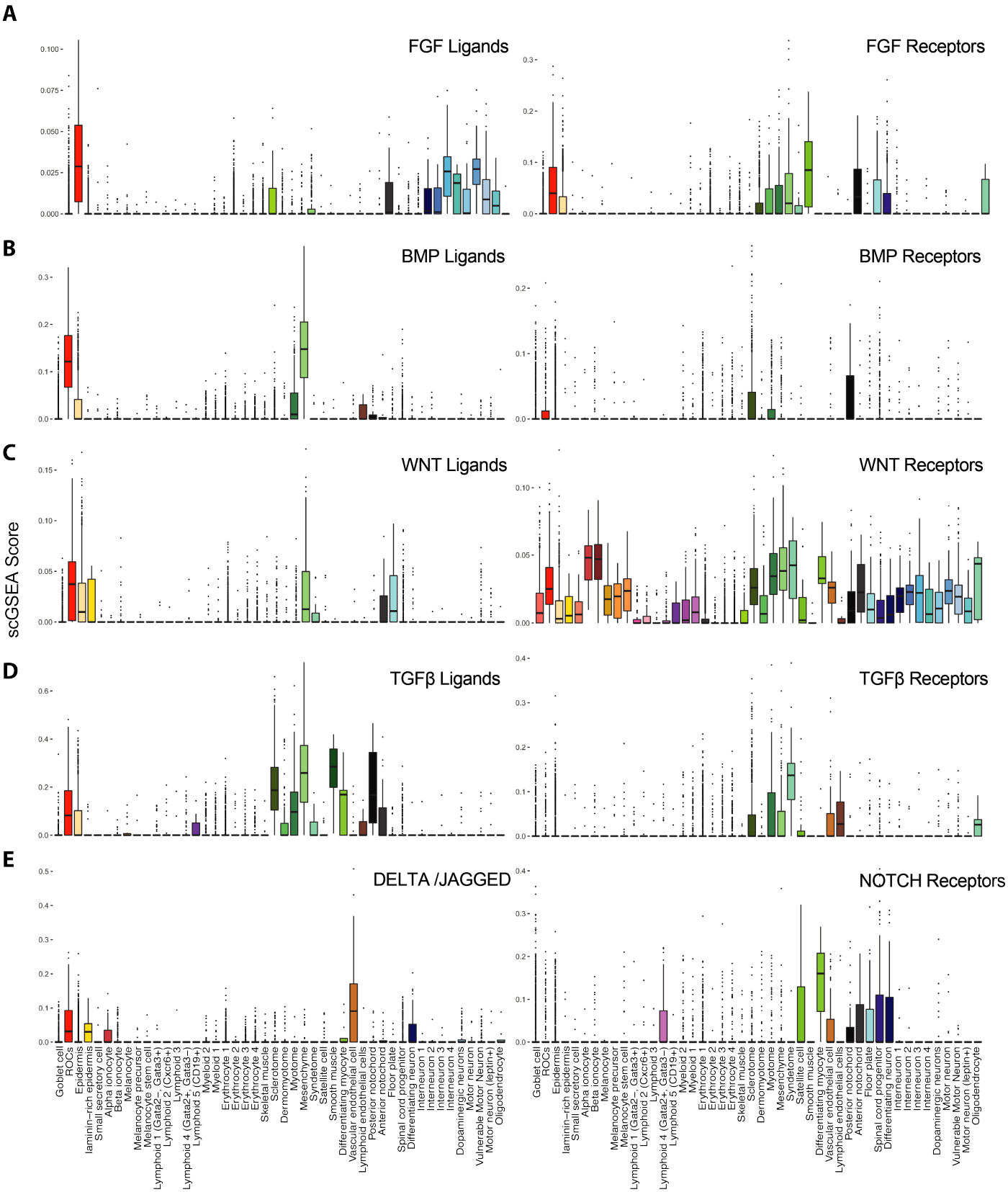

Fig. S9. ROCs and progenitors express complementary receptor/ligand pairs.
Expression of ligands (Left) and receptors (Right) for different signalling pathway were evaluated by scGSEA for all cell types, **(A)** FGF; **(B)** BMP; **(C)** WNT; **(D)** TGFβ; and **(E)** Notch-Delta. ROCs express high levels of ligands, whereas progenitor cell types express high levels of receptors.

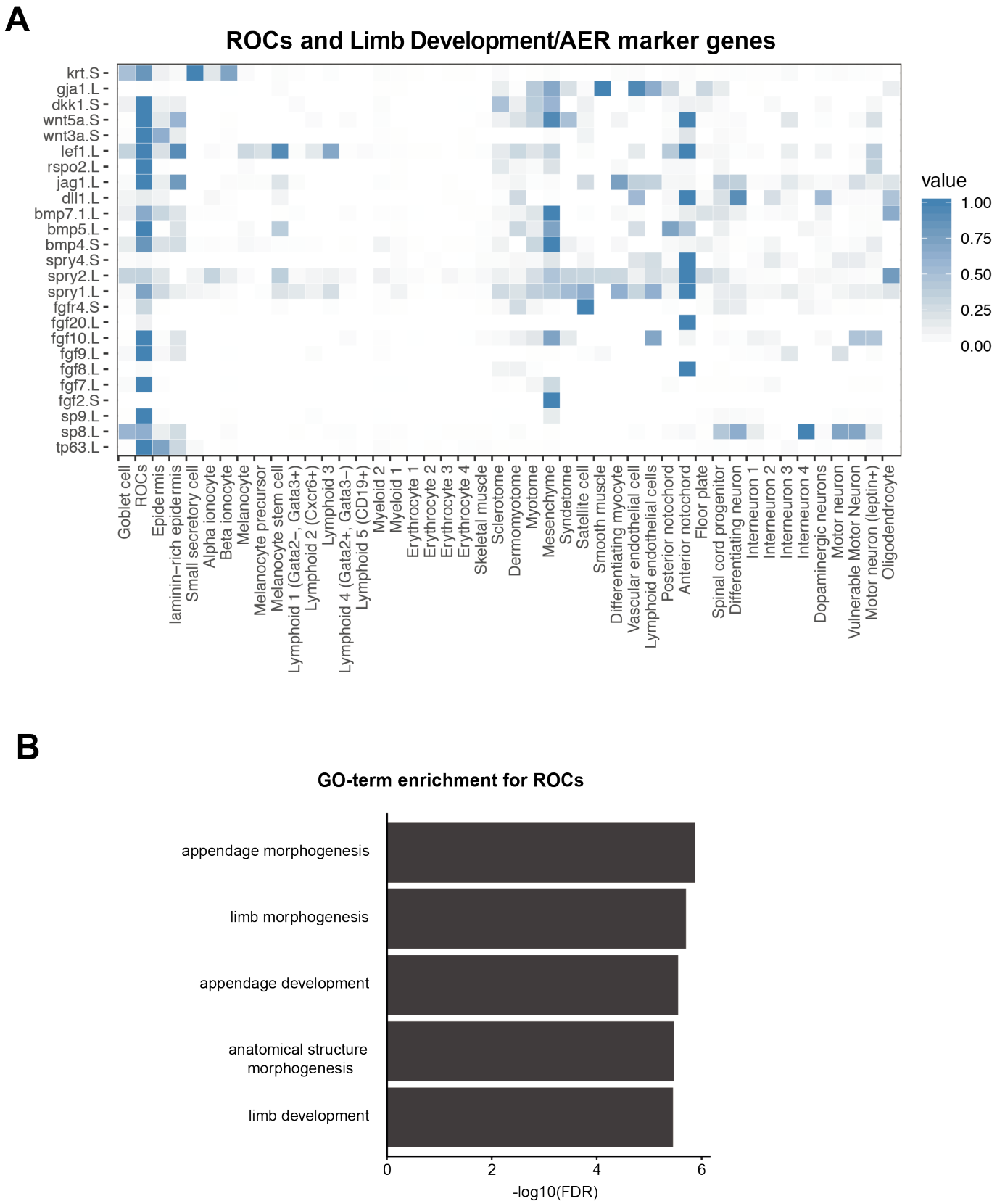

Fig S10. ROCs express genes related to limb development.
(A) Heatmap indicating the relative expression of selected ROCs and limb development genes. **(B)** Transcriptome of ROCs is enriched for limb development-related GO-terms.

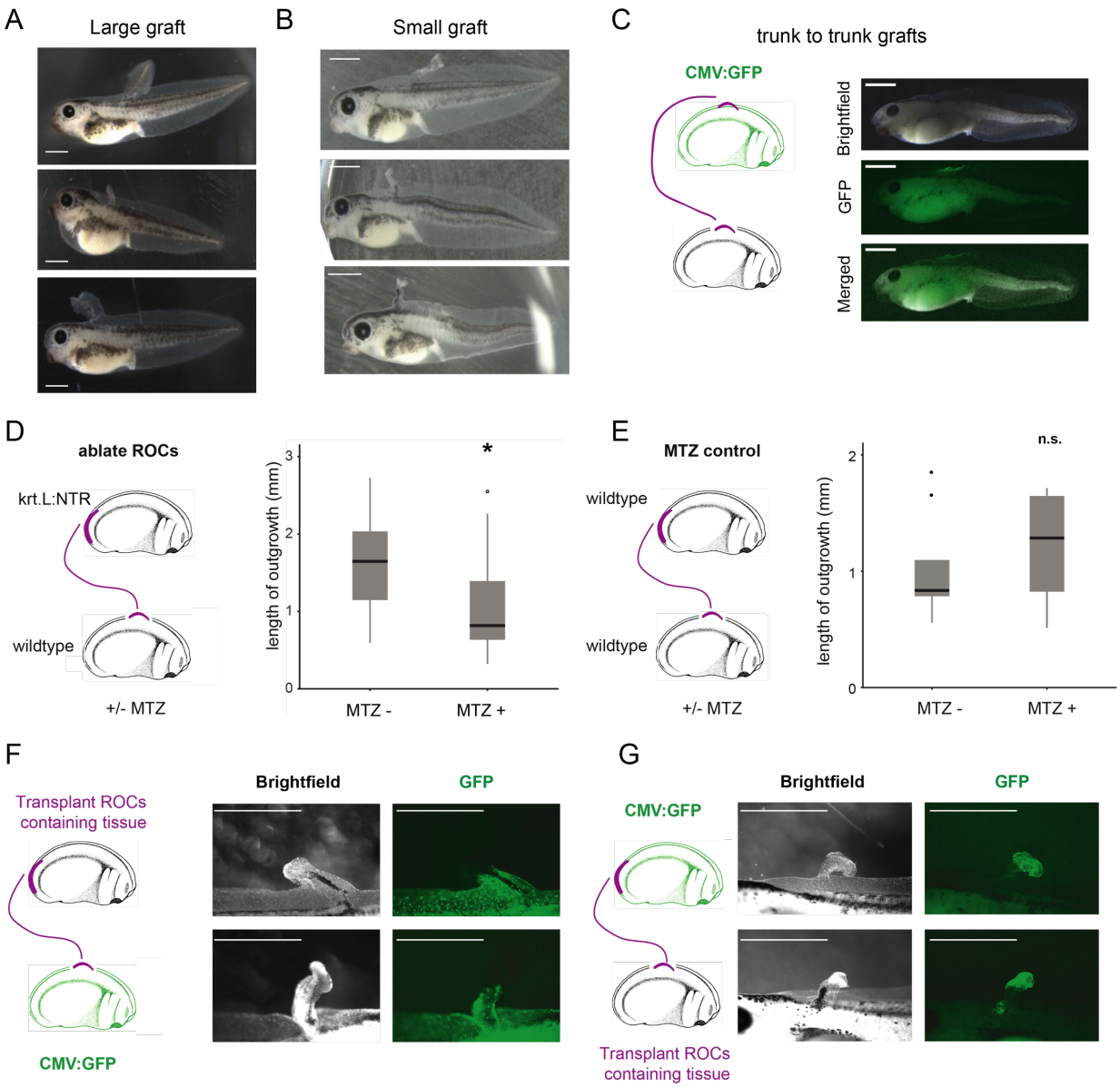

Fig. S11. Transplants of ROC-containing regions induce ectopic structures.
(A) "Large" grafts: Additional images of tail-like ectopic structures obtained by grafting large ROCs containing posterior tail bud regions to the trunk region of host embryos. Scale 1 mm. **(B)** "Small" grafts: Additional images of fin-enriched ectopic structures are obtained following the "small" graft procedure. Scale bars: 1 mm. **(C)** Control trunk-to-trunk grafts display no overt outgrowth phenotype. CMV:GFP embryos were used as donor to ensure successful grafting (n=11, from 3 biological replicates). **(D)** "Small" grafts of ROCs-containing tissues result in significantly reduced outgrowth length when treated with the *Krt.L:NTR* /MTZ ROCs ablation protocol (n=27, and n=28 for with and without MTZ added samples, respectively. Both are from 6 biological replicates, *: $p < 0.001$). **(E)** MTZ treatment alone does not significantly impact length of the ectopic outgrowths (n=9, and n=7 for with and without MTZ added samples, respectively. Both are from 3 biological replicates, n.s.: not significant, $p > 0.05$). **(F-G)** Grafting non-labelled ROCs containing tissues to CMV:GFP embryos, or *vice-versa*, induce ectopic outgrowths with varying contribution from acceptor embryos. All ROCs-containing donor tissues remain at the tip of the ectopic structures. Scale bars: 500 µm.

Table S1: Summary and quality control of single-cell RNA-seq data sets of >13000 *Xenopus* cells.

This Excel file contains a summary of each single-cell RNA-sequencing experiment. Samples are labelled depending on their developmental stage reflective of their regenerative ability and days post amputation status. Sequencing batch, numbers of sequenced cells, mean reads per cell, median genes per cell, median UMI counts per cell, total genes detected, and fraction of reads in cells are reported.

Table S2: Published ROCs marker gene expression labelling the midline edge of epidermis.

This Excel file lists ROCs marker genes that have reported mRNA in situ hybridization results in the literature. These reported findings are provided with their PMID, or PMCID, or DOI, or Xenbase link.

Table S3: Gene lists for downstream analysis used in this study and ROCs marker genes.

This Excel file contains the gene lists that are used for cell cycle analysis, and scGSEA analysis. It also includes ROCs marker genes.