


RESEARCH

Open Access



# Characterization of regeneration initiating cells during *Xenopus laevis* tail regeneration

Radek Sindelka<sup>1\*†</sup> , Ravindra Naraine<sup>1†</sup>, Pavel Abaffy<sup>1</sup>, Daniel Zucha<sup>1</sup>, Daniel Kraus<sup>1</sup>, Jiri Netusil<sup>1</sup>, Karel Smetana Jr.<sup>2</sup>, Lukas Lacina<sup>2,3</sup>, Berwini Beduya Endaya<sup>4</sup>, Jiri Neuzil<sup>4,5,6,7</sup>, Martin Psenicka<sup>8</sup> and Mikael Kubista<sup>1</sup>

<sup>†</sup>Radek Sindelka and Ravindra Naraine are first coauthors.

\*Correspondence: sindelka@ibt.cas.cz

<sup>†</sup>Laboratory of Gene Expression, Institute of Biotechnology of the Czech Academy of Sciences, Vestec 252 50, Czech Republic  
Full list of author information is available at the end of the article

## Abstract

**Background:** Embryos are regeneration and wound healing masters. They rapidly close wounds and scarlessly remodel and regenerate injured tissue. Regeneration has been extensively studied in many animal models using new tools such as single-cell analysis. However, until now, they have been based primarily on experiments assessing from 1 day post injury.

**Results:** In this paper, we reveal that critical steps initiating regeneration occur within hours after injury. We discovered the regeneration initiating cells (RICs) using single-cell and spatial transcriptomics of the regenerating *Xenopus laevis* tail. RICs are formed transiently from the basal epidermal cells, and their expression signature suggests they are important for modifying the surrounding extracellular matrix thus regulating development. The absence or deregulation of RICs leads to excessive extracellular matrix deposition and defective regeneration.

**Conclusion:** RICs represent a newly discovered transient cell state involved in the initiation of the regeneration process.

**Keywords:** Regeneration, *Xenopus laevis*, RICs, ROCs

## Background

Regeneration is the complete restoration of “missing” tissue with a fully functional and essentially identical replica. It is different from the process of repair, which is associated with scar formation and impaired function [1, 2]. Fishes and amphibians have nearly perfect regenerative capacity during early development, and some show partial regeneration of specific organs like the heart, retina, liver, limb, and kidney even in adulthood [3]. In addition, partial tissue regeneration can also be observed in birds [4] and reptiles [5, 6]. Mammals can regenerate specific tissues such as amputated digit tips and heal wounds in youthhood. This capacity decreases during maturation and is lost in adults except for a few instances, such as liver regeneration [7], regrowth of skin of the spiny



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

mouse [8], and deer antler [9]. Healing capacity also declines with age, leading to age-related disorders such as leg ulcers, diabetic wounds, heart attacks, and cerebral infarction [10].

Regeneration is a complex multi-step process involving many cell types and pathways. Healing is sometimes considered as the initial step of regeneration [11]. Immediately after injury, the cell sheets around the wound are activated to prevent further loss of biological material. In fishes and amphibians, the wound epithelium, sometimes referred to as the apical epithelial cap, covers the injured site and acts as the signaling center [12, 13]. This is followed by the transformation of the wound area into the blastema, which comprises specialized progenitor cells [14]. The blastema stimulates local migration of additional cell types, proliferation, and differentiation [15]. This early structure determines the extent of the functional recovery observed in the later phase of tissue regrowth. Many factors are well conserved among vertebrates during the initial step, such as the production of small molecules, including reactive oxygen species [16–18], and activation of early response genes followed by the activation of remodeling enzymes [19].

Inflammation burst in the early phases of the regenerative process activates the immune system to protect against infections and stimulates the removal of tissue debris. The role of the immune system during regeneration is multifaceted and is characterized by the types of immune cells involved and the duration of the immune response [20, 21], but generally it is required for the regeneration process [22]. The less aggressive immune response in the embryo, in contrast to the strong immune reaction in adults and phylogenetically advanced organisms like mammals, is suggested as a critical factor promoting regeneration [23]. The age of the mammal, including humans, is also associated with the modulation of the inflammatory response affecting healing [24].

Regeneration continues with cell proliferation and matrix remodeling to replace the missing structures. Repair results in functionally suboptimal scar formation. Typically, specialized cell types enter the wound, the scar is organized, and the extracellular matrix (ECM), composed mainly of collagen, is established. ECM serves as a provisional scaffold for the remodeling, facilitating cell migration and differentiation during the healing process. ECM remodeling enzymes are essential for blastema formation [25] and regeneration [26]. The collagen matrix is re-organized, and inflammation is reduced by mechanisms that are not yet fully understood [27]. In wounds that do not regenerate, ECM transforms into a fibrous scar that limits tissue remodeling. Many differences in ECM composition between regenerative and repair models have been found and are considered vital factors determining regeneration efficiency [28, 29].

Our understanding of the regeneration mechanisms and cellular interactions leaped forward with data from high-throughput methods such as single-cell sequencing [30]. Their application to traditional regenerative models such as the limb of the axolotl (*Ambystoma mexicanum*) [31], the fin of zebrafish (*Danio rerio*) [32, 33], and the digit tip of the mouse (*Mus musculus*) [31] has revealed new cell subpopulations required for their regeneration. A breakthrough in understanding embryonic regeneration came from studies of tail regeneration at the level of individual cells in *Xenopus laevis* [32]. Regeneration organizing cells (ROCs) were discovered, and their migration to the site of injury, where they promoted developmentally related signaling pathways, was essential

to stimulate the regrowth of the tail [32]. However, the mechanism of their migration and their attraction to the injury site remained unclear.

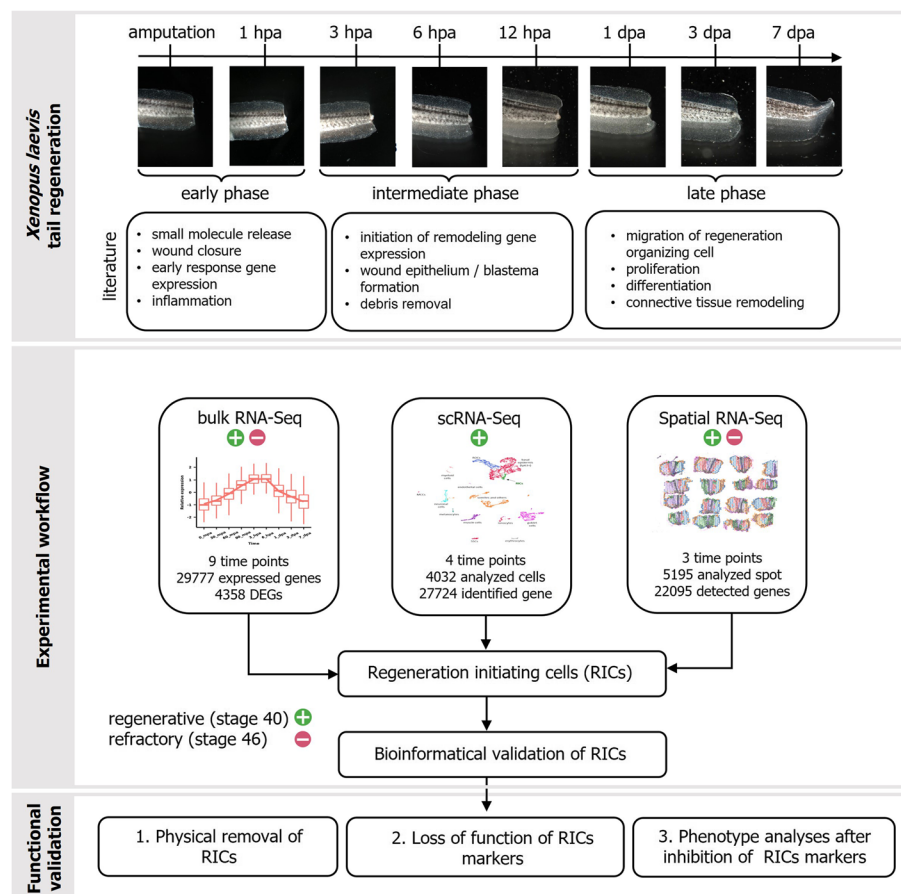
A unique advantage of the *Xenopus* model is its transient refractory stage, where the embryo loses its regenerative capability [21, 33–35], making it ideal for comparative studies. Comparison of myeloid cells in regenerative and refractory developmental stages divided them into inflammatory and regenerative subpopulations that have an opposing effect on regeneration. This resembles the difference between mammalian M1 (inflammatory) and M2 (regenerative) macrophages [20].

Standard single-cell profiling studies lack information about cells' positions and their surroundings. This can be addressed with spatially resolved transcriptomics [36]. So far, spatial transcriptomic studies of regeneration have been limited to the fibroblast fate during tissue repair [37] and mouse digit regeneration [38]. Here, we address the healing of the *Xenopus* tadpole tail, combining results from three high throughput methods: bulk RNA-Seq, single-cell RNA-Seq, and spatial transcriptomics, studying regeneration from the very early initiation hours post amputation (hpa) throughout days post amputation (dpa). We discovered a new cell population that appears transiently in the very early critical phase of regeneration, which we call the regeneration initiating cells (RICs). With functional studies, we demonstrated that RICs are critical for regeneration.

## Results

### Gene expression in the early phase after injury initiates regeneration via a conserved mechanism

With bulk RNA-Seq, we studied the temporal regulation of gene expression by comparing regenerative and refractory stage embryos (Figs. 1 and 2). In total, 4358 differentially expressed genes (DEGs) associated with regeneration were identified (adjusted  $p$ -value < 0.01; minimum of 20 transcripts in at least one sample). Based on the temporal expression profiles, the DEGs were divided based on activity into early, intermediate, and late phases (Additional file 1: Table S1). Early DEGs (1637) were divided into two subgroups based on their profiles. Eight hundred and one DEGs active during the early phase were characterized by an expression burst 0.5 to 1 hpa followed by a rapid decline (Fig. 2B). Many of these (e.g., *fos*, *jun*, *egr1*) have previously been identified in injury models including *Xenopus* embryonic wound healing [19]. Other early phase DEGs with a peak at 1 hpa, code for ATPases regulate the function of muscle cells (*myh4*, 8, 11 and 13, *myl1* and 4, *myod1*), oxidative phosphorylation (*ndufa* genes, *cytochrome c oxidase* genes, *sdhc*, *sdhd* and *ubiquinol-cytochrome c reductase* genes), reactive oxygen species production (*sod1*, *cox* genes), and other metabolic changes such as the catabolic processes (*aldh1l2*, *glud1*, *got1*, *got2*). Eight hundred and thirty-six early DEGs were downregulated immediately after the injury (Fig. 2B). These are affiliated with GOs associated with muscle activity and ATP metabolic processes (Additional file 1: Table S2). Seven hundred and seventy-four DEGs showed increased expression within 1.5–6 hpa (Fig. 2B) and are referred to as the intermediate phase. Their functions are predicted to be associated with tissue remodeling (*mmp1*, 8 and 9, *timp1* and 3), cell migration (*epcam*, *integrin subunits*, *muc1*, *vim*), and control of the developmental processes (*notch1*, *shh*, *wnt10a*, *sox11*, *bmp4*). One thousand one hundred and seventy-nine DEGs with increased expression after 1 dpa (Fig. 2B) were classified as the late phase genes.

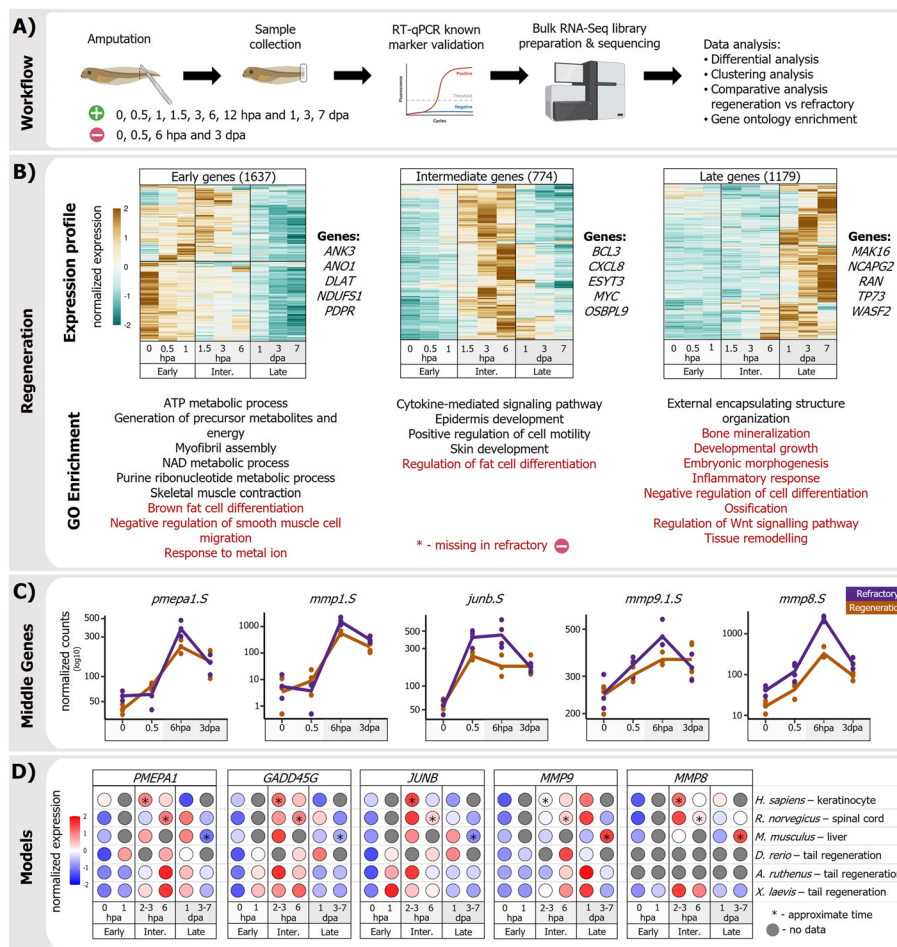


**Fig. 1** Study design of regeneration initiation using *Xenopus laevis* tail amputation model. Green positive sign represents the regeneration experiment. Red negative sign represents the refractory experiment

Their functions are linked to developmental regulation and signaling pathways such as the Wnt (*axin2*, *dkk1* and 3, *lrp1* and 4, *ctnnb1*, *frzb*, *wnt5b* and 11) and the TGF $\beta$  pathway (*bmp1*, 5 and 7, *smad1*, 4 and 9, *tgfb1* and 2). Late response was also observed for many *hox* genes, keratins, and genes required for proliferation (rRNA metabolism). The list of all the DEGs active in each phase and the associated enriched GO terms are in Additional file 1.

Comparison of the DEGs in regenerative and refractory stages (Additional file 1: Table S3) revealed a clear difference in the late phase, with many developmental processes reduced or missing (Fig. 2B, Additional file 1: Table S4). This suggests that the intermediate DEGs determine the ultimate regenerative properties of the injured tissue.

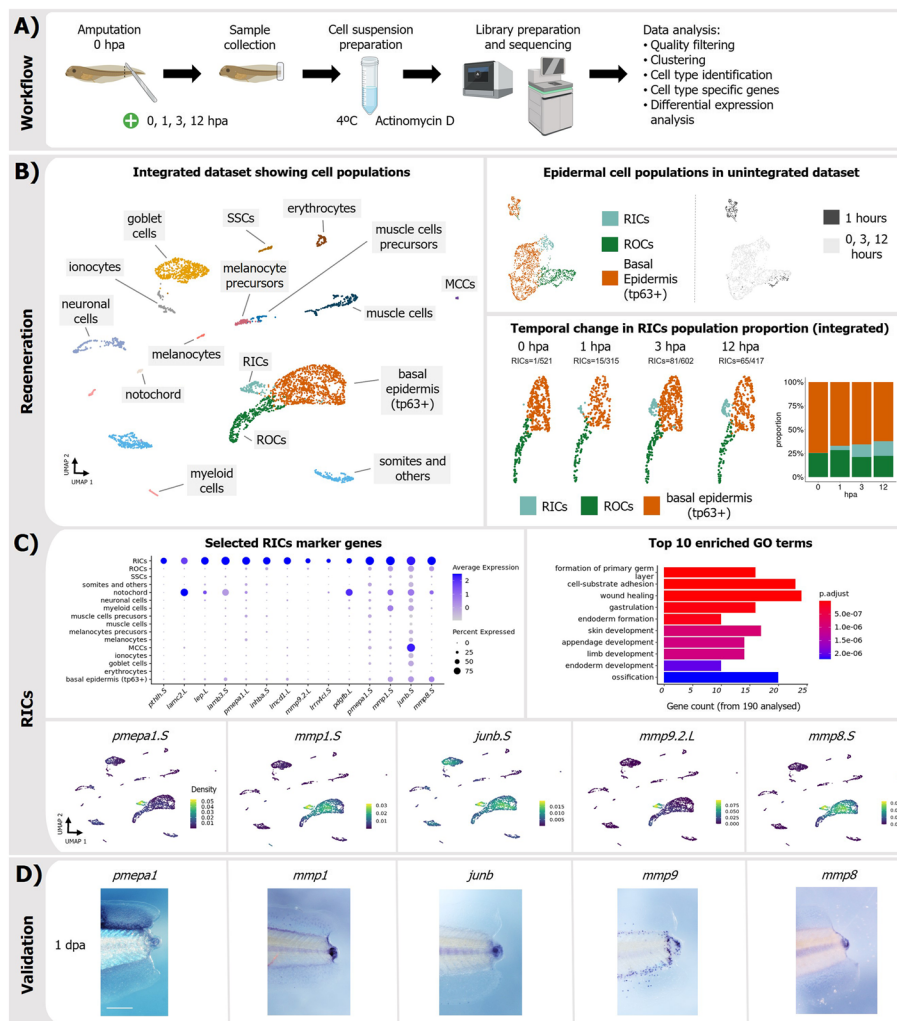
To test if the patterning of gene expression after injury is also relevant to other species, we studied the healing model using a scratch assay of the human fibroblast layer, regeneration of the rat spinal cord, mouse liver, and fish tails. They all showed similar changes in the expression of selected intermediate genes 3–6 hpa, suggesting a conserved mechanism for regeneration initiation (Fig. 2D).



**Fig. 2** Temporal bulk RNA-Seq comparison of regenerative and refractory embryos. **A** Scheme of experiment. Green positive sign indicates regeneration. Red negative sign refractory experiment. **B** Regeneration is divided into three phases: early, intermediate, and late. Heatmaps showing the z-score for the mean normalized expression of selected genes together with enriched GO terms (red: not significantly enriched in refractory samples). **C** Bulk expression of five genes from the intermediate group (later identified as RICs markers) in regenerative (orange) and refractory (purple) samples. **D** Expression (z-score) of selected intermediate genes measured with RT-qPCR in other regenerative models

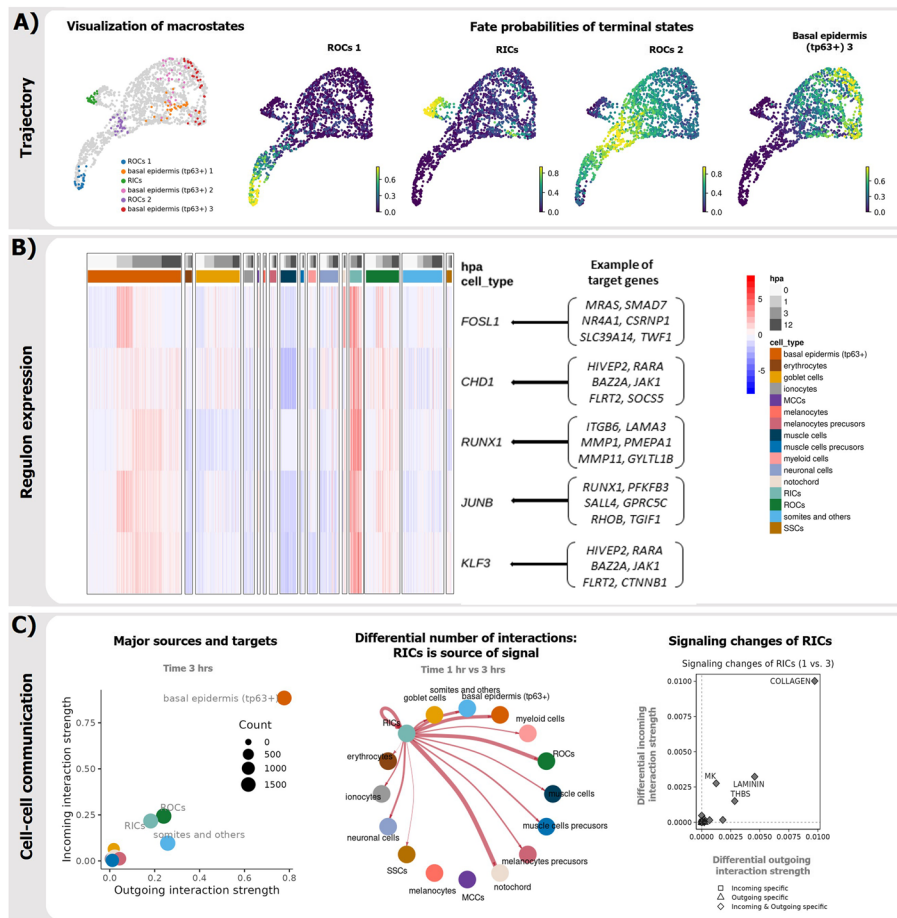
### Regeneration initiation cells (RICs) appear when regeneration is initiated

We profiled 4032 single cells during the early and intermediate phases of regeneration with an average coverage of ~6000 genes per cell (Figs. 1 and 3A). The dissociation protocol was validated by RT-qPCR assessment of cell type specific genes, comparing samples extracted from the whole tail with those from cell suspension, and showed negligible differences (Additional file 2: Fig. S1). Quality controls such as the levels of mitochondrial genes' and ribosomal genes' expressions and unique molecular identifiers can be found in Additional file 2: Fig. S2. After clustering of the data and cell type annotation (Fig. 3B, Additional file 2: Fig. S3, S4, Additional file 3: Table. S1), the basal epidermal cells showed the largest temporal changes (Fig. 3B, Additional file 2: Fig. S3). It is among these we identified the new regeneration initiating cell (RIC, Fig. 3B) subpopulation. Both RICs, and the previously reported ROCs, expressed the top 20 marker genes (fold



**Fig. 3** Single-cell analysis of regeneration at early phase. **A** Scheme of the scRNA-Seq experiment. **B** UMAP visualization of the integrated data sets, identifying the regeneration initiating cells (RICs), regeneration organizing cells (ROCs), small secretory cells (SSCs), and multiciliated cells (MCCs). Temporal changes in the epidermal cell populations (*tp63+*) for time separated and integrated data. **C** Expression profiles of selected RIC markers within the different cell populations. Top ten enriched GO terms for the RIC markers. **D** In situ hybridization of selected RIC markers at 1 dpa. Representative bright field images from at least seven biological replicates (scale size 300 μm). Green positive sign indicates the regeneration experiment

change  $> 2x$ ,  $P_{adj} < 0.05$ ) for basal epidermal cells, but RNA velocity and CellRank analysis revealed they are different subpopulations (Fig. 4A, Additional file 2: Fig. S5, S6). RICs were not present at time 0, which indicates that some basal epidermal cells undergo a transition forming the RICs following the amputation. RIC population increases in time, reaching a maximum of 10% of the total basal epidermal cells at 12 hpa (Fig. 3B). CellRank identified 5 macrostates, whereby the initial population is derived from a subpopulation within the basal epidermis (*tp63+*) and leads to the differentiation to the terminal states of the basal epidermal cells, the "ROCs 2" subpopulation comprising primarily ROCs and lying in-between the divergence to the terminal cells of the RICs and "ROCs 1" sub-populations (Fig. 4A, Additional file 2: Fig. S6). We identified several



**Fig. 4** Trajectory, regulon and cell–cell communication pathway analysis associated with RICs. **A** CellRank analysis showing the UMAP projections of the clustered locations for the selected cell states. "Basal epidermal (*tp63+*) 1" represents the initiating population. UMAP showing the cells that are identified as terminal states and the fate probability of each cell towards a given cell state. **B** Gene regulatory network of the RIC's specific transcription factors that are differentially (fold change  $> 2x$ ,  $P_{adj} < 0.05$ ) expressed. Shown are examples of the correlated associated regulons. Heatmap is based on z-score of the AUC (area under the curve) scores for each cell. Columns are ordered by cell type. **C** Cell–cell communication network comparison between 1 and 3 hpa showing the major sources and targets for the signals, all signal changes associated with the RICs, and the differential pathways involving RICs

potential driver genes specific to each of the terminal population (Additional file 2: Fig. S6).

We looked for the association of correlated transcription factors with their target genes (regulons) in the different cell populations. We observed *FOSL1*, *CHD1*, *RUNX*, *JUNB*, and *KLF3* regulons within the RIC population (Fig. 4B). Cell–cell communication indicated that the RICs are most active (receiving/sending signals) at 3 hpa and 12 hpa (Fig. 4C, Additional file 2: Fig. S7A). Therefore, we focused on the differences between time 1 and 3 hpa to identify the main communication pathways that are being enhanced for the RIC population. We observed that at 3 hpa many signals are being received/sent by the RIC population from/to other cell types relative to 1 hpa. Most of these signals are sent to the notochord, basal epidermis (*tp63+*), and RICs (Additional file 2: Fig. S7D). The receiving cells also contribute to signals sent to the RICs (Fig. 4C, Additional file 2:

Fig. S7D). Collagen, laminin, MK (midkine), and THBS (thrombospondin) represent the main signaling pathways associated with the RICs and were enhanced at 3 hpa relative to 1 hpa (Fig. 4C, Additional file 2: Fig. S7D, E). We looked for specifically upregulated signaling ligand-receptor pairs between the RICs (sender) and ROCs (receiver) and found upregulations of the collagen, laminin, CD99, and NOTCH pathways (Fig. 4C, Additional file 2: Fig. S7E, G). Pathways WNT, ACTIVIN, PDGF, and LIFR had the strongest outgoing signaling strength coming from RICs relative to other cells at 3 hpa, while ADGRE, EPHB, and LIFR were the strongest incoming signals to the RICs (Additional file 2: Fig. S7B, G).

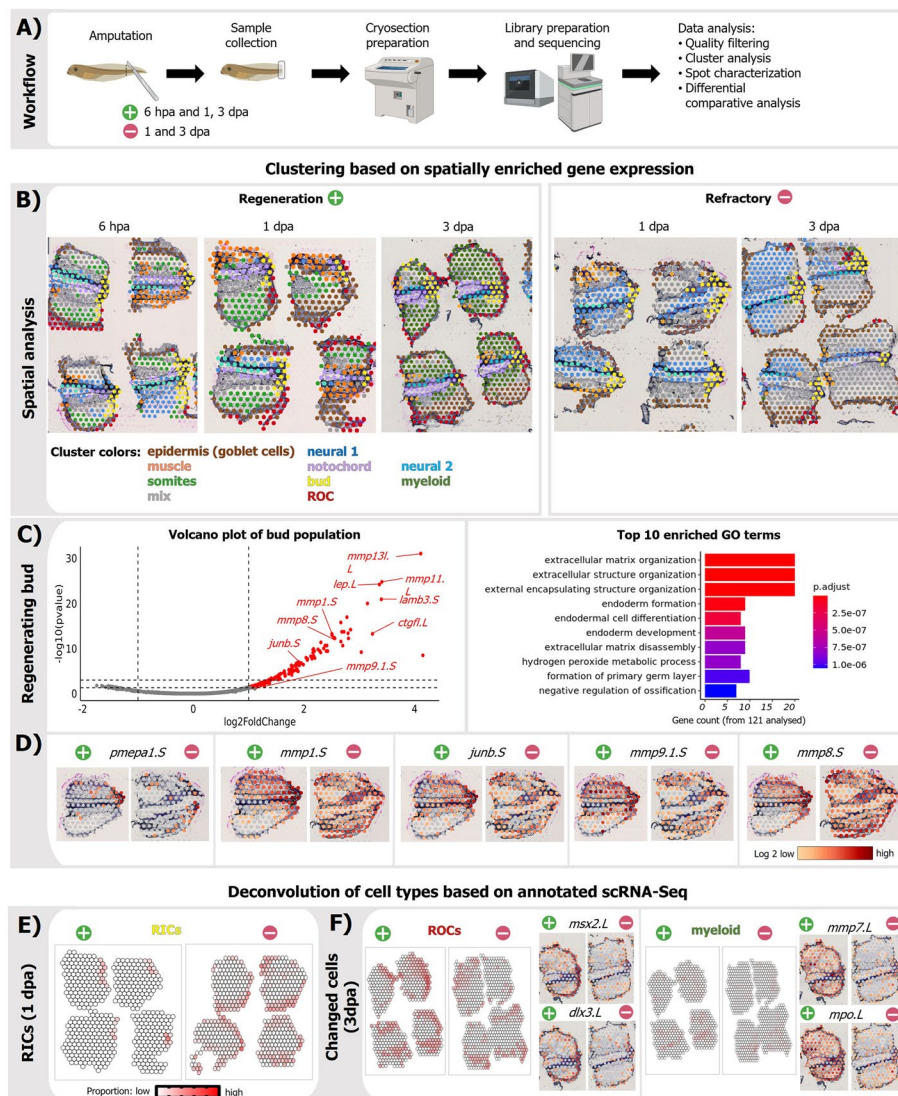
A previously published single-cell RNA-Seq (scRNA-Seq) study that assessed only later timepoints of *Xenopus* regeneration reported an expression signature characteristic of the RICs population [32], but the RICs were referred to as laminin-rich cells and their function was not characterized. Reanalyzing the published data, we found >50% overlap between markers for the laminin-rich cells and the RICs. In our scRNA-Seq dataset, we identified 272 genes as RIC markers (fold change >2x,  $P_{adj} < 0.05$ ) with the highest enrichment of *palld.L/S*, *lep.L*, *pmepa1.S*, *inhba.L*, *pthlh.S*, *lamb3.L*, *mmp1.S*, *mmp8.L/S*, and *lamc2.L* transcripts (Additional file 3: Table. S1). Further analysis revealed GO enrichment for wound healing, cell adhesion, cell migration, and ECM remodeling (Fig. 3C, Additional file 3: Table. S2). The RIC markers overlapped (hypergeometric test:  $p < 0.001$ ) with the DEGs from the intermediate phase of the bulk RNA-Seq (43%), indicating the important contribution from RICs to the intermediate regeneration phase. Five RIC markers of biological relevance were localized in the wound 1 dpa using in situ hybridization (Fig. 3D; Additional file 2: Fig. S8). These markers are predominantly present in the regeneration bud, indicating that the RICs form an organizing center after the injury, onsetting the regeneration.

### **RICs accumulate in the regeneration bud during the regenerative stage of the embryo**

To better understand the spatial arrangement of the RIC population within the regenerating tissue, we performed spatial transcriptomics on embryos in regenerative and refractory stages (Figs. 1 and 5A), assessing times within the intermediate and late phases. After data quality control (Additional file 2: Fig. S9) and clustering, the majority of cell types identified in the scRNA-Seq dataset were recovered in the spatial data through canonical markers. These include clusters of epidermal, muscle, somite, neural, notochord, myeloid, ROCs, and regeneration bud (blastema) markers (Fig. 5B, Additional file 4: Table. S1, Additional file 2: Fig. S10, S11).

Spatial composition at 6 hpa and 1 dpa of regenerative and refractory stages were compared. In the regenerative bud “spots,” we identified 165 upregulated genes (fold change >2x,  $P_{adj} < 0.05$ ) relative to the rest of the tissue. Forty-six percent of these were RIC markers in our scRNA-Seq analysis, suggesting high RIC abundance in the regeneration bud (Fig. 5C, Fig. 6A). Functions of the regenerative bud genes include ECM organization, collagen degradation, cell differentiation, and regulation of developmental processes (Fig. 5C, Additional file 4: Table. S2). A distinct bud cluster was also identified in the refractory samples. This cluster was characterized by overexpression of 161 genes with 40% overlap with markers of the regenerative bud spots (Fig. 6B, Additional file 5: Table. S1). Comparing the regenerative and refractory samples, we observed

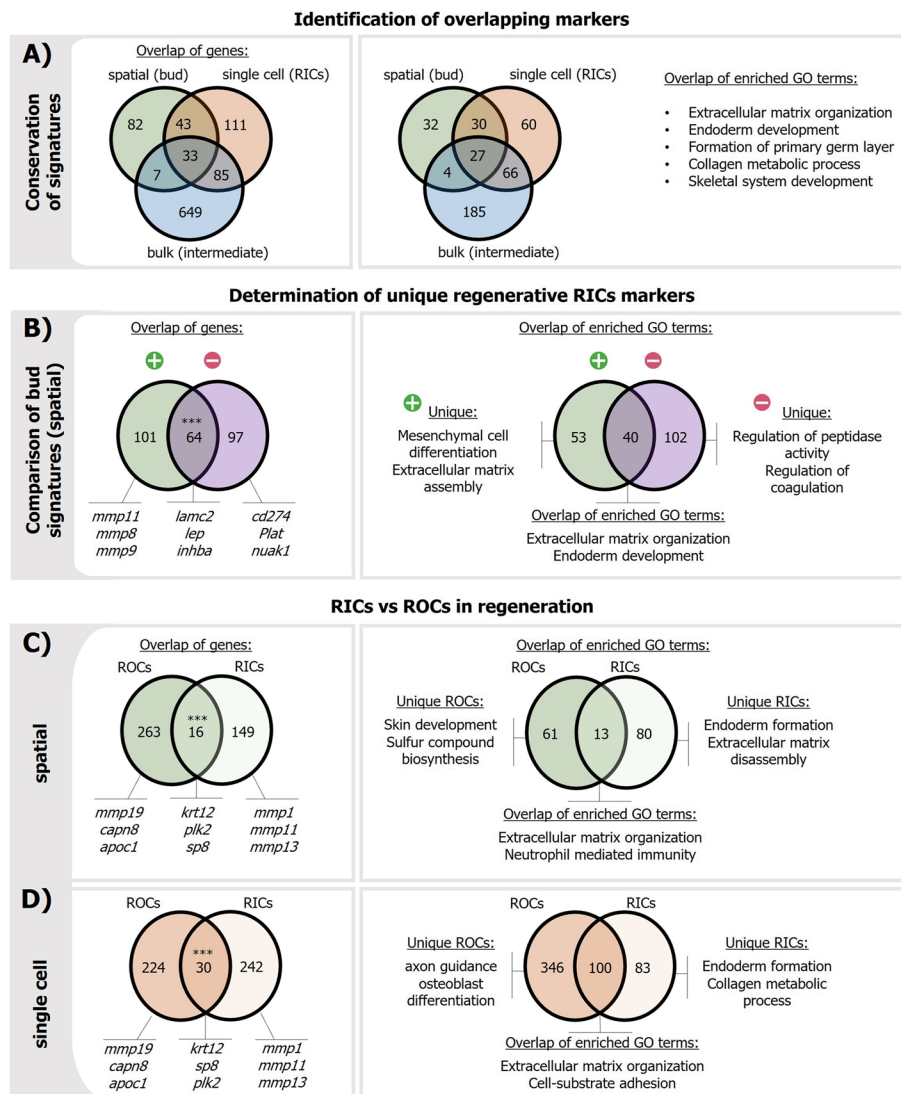




**Fig. 5** Spatial transcriptomics of tail regeneration. **A** Scheme of the spatial transcriptomics experiment. **B** Clusters formed based on enriched expression in regenerative and refractory samples. Ten clusters in the Loupe software were manually annotated based on top cluster markers and comparison with scRNA-Seq data. **C** Volcano plot showing differential expression between regenerative bud and the remaining spots. Bar plot shows the top 10 enriched GO terms associated with the overexpressed genes. **D** Comparison of spatial expression of selected RIC markers in regenerative and refractory samples at 1 dpa. **E** Distribution of RICs in regenerative and refractory samples at 1 dpa based on deconvolution using scRNA-Seq data. **F** ROC and myeloid markers (based on scRNA-Seq) in regenerative and refractory samples at 3 dpa. Green positive and red negative signs indicate regenerative and refractory sample, respectively. Sections are shown from anterior (left) to posterior (right) of the tail bud region

overexpression of *mmp8.L/S*, *mmp9.1.S*, *pmepa1.S*, *mmp1.S*, and *junb.S* in the bud of the regenerative stage, while dispersed throughout the tissue during the refractory period (Fig. 5D).

The cellular composition of bud spots in the spatial dataset was assessed by deconvolution using our annotated scRNA-Seq profiles (Fig. 5E, Additional file 2: Fig. S10, S12, S13). In the regeneration stage, RICs were primarily localized in the regeneration bud,



**Fig. 6** Comparative analysis of descriptive datasets. Overlap of significantly expressed transcripts and their associated enriched GO terms in: **A** intermediate bulk RNA-Seq, RICs in single cell, and regenerative bud in spatial datasets; **B** specific regenerating genes identified by comparison of bud clusters from spatial transcriptomics data in regenerative and refractory samples; **C** RIC and ROC marker genes in spatial transcriptomics data, and **D** in single-cell data sets. Significance was assessed using a hypergeometric test (1-tailed) (ns not significant, \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ ). Green positive and red negative sign represents regenerative and refractory sample, respectively

while in the refractory stage, they were dispersed on the surface of the embryo at 1 dpa (Fig. 5E). The ROCs, which are necessary for continued regeneration [32], localized to the amputation site of the regenerative samples, primarily during the later phase at 3 dpa based on deconvoluted spatial plots using canonical ROC markers (*msx2.L* and *dlx3.L*). In the refractory samples at 1 dpa and 3 dpa, ROCs were dispersed (Fig. 5F, Additional file 2: Fig. S12, S13). Marker gene expression of the remaining cell types (epidermal, muscle, neural and notochord tissues) in the spatial dataset was consistent with their expected locations (Additional file 2: Fig. S12). Reference-free cell type deconvolution based on STdeconvolve confirmed the reference-based deconvolution localizing the

RICs to the tail bud region as early as 6 hpa and the ROCs appearing around 1 dpa for the regenerative samples (Additional file 2: Fig. S12, S14). For the refractory sample, the spatial data showed dispersed RIC populations and negligible ROC populations at 1 and 3 dpa (Additional file 2: Fig. S12, S14). The reference-based and reference-free methods used to analyze the spatial transcriptomics data showed high correlation in annotations (Additional file 2: Fig. S15), confirming the difference in response to injury between the regenerative and refractory embryos, with the accumulation of RICs in the regenerative bud being a plausible important step that initiates regeneration. Additional file 2: Fig. S16 shows the spatial distribution of some RIC markers as derived by scRNA-Seq.

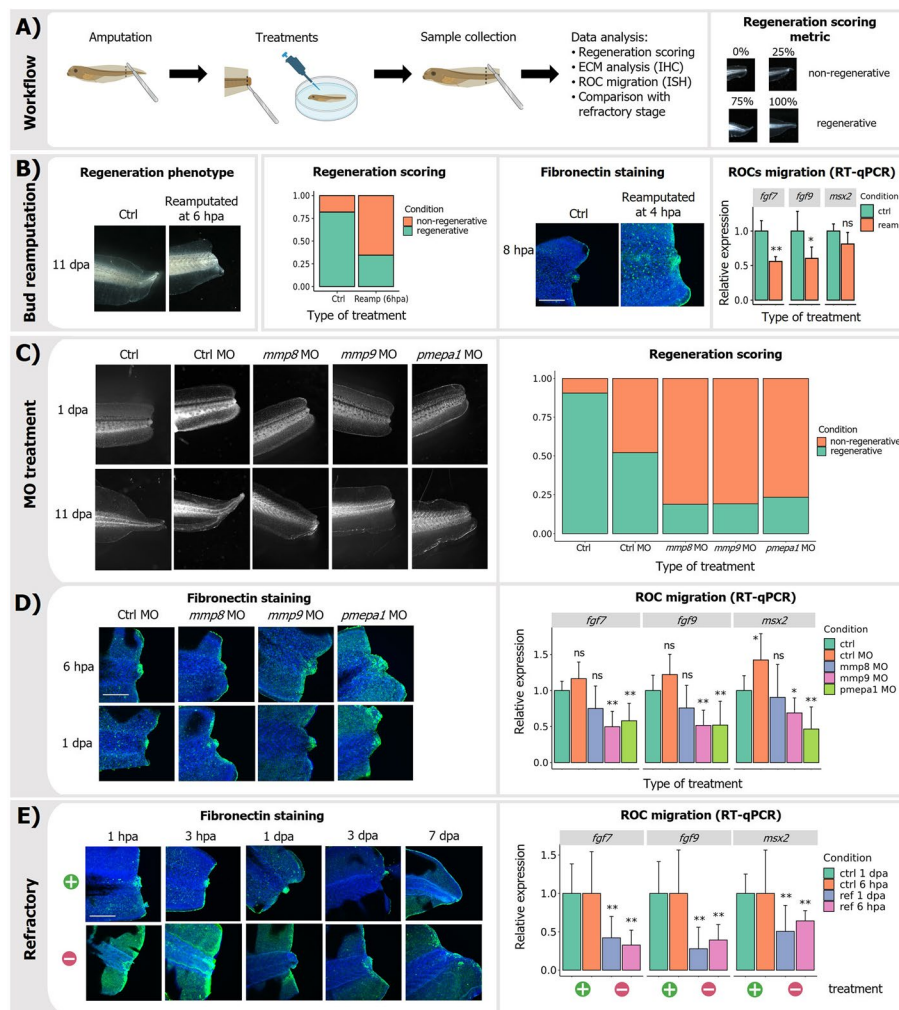
### **RIC marker genes expressed at intermediate time in the bud are essential to initiate regeneration**

Above, we reported studies of regeneration initiation using the high-throughput methods bulk RNA-Seq, scRNA-Seq, and spatial transcriptomics. The data were analyzed to (1) identify gene signatures in time and space, (2) identify regeneration-linked genes by comparing bud expression signatures in regenerative and refractory samples, and (3) identify relation between the RICs and ROCs in regenerative samples.

Genes upregulated between 1.5 and 6 hpa in bulk RNA-Seq were compared to RIC markers identified by scRNA-Seq and with genes upregulated in bud spots of regenerative embryos using spatial transcriptomics. In the three datasets, 33 regeneration-initiating genes (e.g., *mmp1*, *lep.L*, *mmp8*, *sox11.S*) were conserved, and 27 GOs of biological processes, such as ECM organization and regulation of development, were shared (Fig. 6A, Additional file 5).

The bud gene expression signatures in the spatial datasets were compared between regenerative and refractory embryos. Out of 165 markers of the regeneration bud, 64 were shared (e.g., *lamc2.L*, *inhba.L*, *fn1.S*, *mmp13.L*, *lep.L*, *gadd45g.L*), while 101 were unique to the regeneration bud (Fig. 6B, Additional file 5: Table. S1). Among the 33 regeneration-initiating RIC markers, 20 were specific only for regenerative buds (missing in refractory) based on spatial analysis (Fig. 6A). This included ECM remodeling markers (*mmp8*, 9, and 11) and regulation of cell differentiation (*rgcc.L*, *sox11.S*, *fibin.S*), which was also confirmed on the level of GOs of biological processes (Fig. 6B, Additional file 5).

In the third analysis, the relationship between RICs and ROCs was studied. Both populations are formed from the basal epidermal cells (*tp63+*) but differ in their formation in time, behavior, and gene expression signatures. To better discern the RIC-ROC relationship, we compared their signatures in the spatial and single-cell datasets (Fig. 6C, 6D, Additional file 5). As revealed by GO analysis, RICs are involved in extracellular space modifications and regulation of development, while ROCs are important regulators of epidermis development, limb and gland morphogenesis, and metabolic processes (Fig. 6C). Both ROCs and RICs seem important for ECM remodeling, but the ECM-related genes they express are different. A complete list of the ECM and remodeling genes is presented in Additional file 2: Fig. S17 and selected spatial localization in Additional file 2: Fig. S18.



**Fig. 7** Functional validation of regenerative RICs. **A** Scheme of functional experiments. **B** Removal of the regenerative bud at 6 hpa and its phenotype (11 dpa), scoring, fibrosis (fibronectin staining, 8 hpa), and defective ROC migration (RT-qPCR of ROC markers 1 dpa). **C** Loss of function of three RIC markers: *pmepa1*, *mmp8*, and *mmp9* using Vivo MO. **D** Phenotypes and scoring with extensive fibrosis (fibronectin IHC) and defective ROC migration (RT-qPCR for ROC markers) in refractory samples. Brightfield and confocal images scale size is 300  $\mu$ m. RT-qPCR was prepared from at least biological triplicates, each containing at least five dissected regenerating tails. Significance tested with Student's *t*-test (ns not significant, \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ ). Fibronectin staining is shown using one representative sample from at least seven biological replicates. Green positive sign indicates regenerative while red negative sign indicates refractory sample

### Silencing of RIC markers blocks regeneration by fibrosis

Above, we have identified several genes required for regeneration initiation. To assess their relevance, we performed functional experiments removing either the bud or the RICs. We also transiently inhibited three important RIC marker genes (Fig. 7A). These phenotypes were compared to those of the refractory samples.

The regenerative bud, which primarily contains RICs (Fig. 5B), was manually removed at 4–6 hpa, and the embryos were cultivated in parallel with controls. The removal of the bud resulted in the loss of regeneration capability (Fig. 7B) due to extensive fibrosis produced through the accumulation of Fibronectin, which resembles scarring in the

regenerating area. Also, migration of ROCs was perturbed, as indicated by the reduced expression of ROC markers (Fig. 7B). A similar phenotype was observed in embryos that had *mmp8*, *mmp9*, or *pmepa1* individually inhibited using Vivo-morpholino oligonucleotides (MOs). The concentration range of MOs and application periods was optimized in pilot experiments. The optimal concentration of MOs was applied on the first day of regeneration when RICs appeared in the regenerative bud. This reduced regeneration efficacy is accompanied with fibrosis and defective migration of ROCs (Fig. 7C, D). Refractory samples were prepared for comparison (Fig. 7E). Migration defect of refractory ROCs was assessed using bulk RNA-Seq and spatial transcriptomics and was characterized by decreased levels of *msx2* and *dlx3* at 3 dpa in refractory compared to regenerative samples.

## Discussion

We propose a mechanism for regeneration initiation that depends on the activity of the new cell population, which we have characterized and refer to as regeneration initiating cells (RICs). The RIC population originates from the basal epidermal population and appears at the edge of the tail tissue within a few hours after amputation and remodels the surrounding ECM and seems to promote the migration of cells required for regeneration. *Xenopus* tail regeneration depends on the migration of ROCs, but there may also be other cell types involved as indicated from other models, such as the connective fibroblasts in axolotl [39, 40], interaction of fibroblasts and keratinocytes in mammalian wound healing [41], and the fibroblast subtype in reindeer antler regeneration [42]. The RICs have not been described before, as previous single-cell studies in the regeneration field have focused on later time points after their appearance. Our targeted expression analysis reveals a similar increase in RIC specific gene expression in the early phases post injury in many models of regeneration, including mammals, which suggests it is an evolutionary conserved mechanism. An interesting aspect is whether to refer to the RICs as a “cell state” or “cell type” or perhaps even “cell subtype.” During *Xenopus* regeneration, RICs display cell state characteristics because of their transient temporal appearance. However, in other models, it shows cell type behavior with persistent and stable presence (Fig. 2D).

Applying descriptive analyses combined with functional validation, we consistently identify a group of genes conserved in time and space that is expressed in the specific RIC subpopulation. Comparison of regenerative and refractory samples indicates the importance of combining all three approaches to reliably determine the mechanism of regeneration initiation, especially in the spatial context. The temporal expression profile of RIC markers based on bulk RNA-Seq data is characterized by rapid increase during 3–6 hpa, followed by declined expression after 1 dpa. An expression burst of early response genes just before the appearance of the RICs suggests a potential link. Several studies have already confirmed a dependence of the expression of remodeling enzymes on the early response genes [43, 44]. As we show here, the RICs are not present in the uninjured embryos (time 0). They form upon injury, which contrasts the ROCs that are present at the tip of the regrowing tail [32]. RICs and ROCs originate from the basal epidermal cells, but the RICs are newly formed at the amputated plane and later accumulate in the regeneration bud. The basal position of the RICs

and the ROCs in the epithelium seems to be a general feature of vertebrates. Proliferating keratinocytes capable of wound closure are located in the basal layer of the epidermis in mammals and may be employed in wound therapy in the clinic [45]. No sign of active migration of RICs was observed in our spatial analysis, but cell migration is required for other cell types (ROCs [32] and myeloid cells [20, 46]) in order to regulate regeneration. We hypothesize that the role of RICs during regeneration is to stimulate migration of the ROCs, as indicated by their expression markers, which include ECM remodeling enzymes. Other migratory cells, such as primitive myeloid cells, presumably use alternative routes as they were not affected by the silencing of RICs (Additional file 2: Fig. S19).

Based on our single-cell data, RICs express more than a hundred specific genes with various functions required for regeneration. For example, *junb* was identified in *Xenopus* tail regeneration and has a role in the regulation of cell proliferation induced by TGF $\beta$  signaling [47]. The PMEPA1/TMEPA1 protein encoded by *pmepa1* is a negative regulator of TGF $\beta$  signaling and reduces cell migration [48]. RICs also express *tgfbi* and *inhba*, which are members of the TGF $\beta$  family regulating cell adhesion [49] and modulating TGF $\beta$  pathway activity [50]. Gene regulatory pathway analysis performed using scRNA-Seq data suggests a potentially important regulatory cascade where the early expressed *junb* activates the expression of *runx1* factor, which later activates the expression of many RIC markers, including *pmepa1*, *mmps*, and ECM components. Our cell-cell communication analysis also found that the ACTIVIN pathway is enriched in the RICs presumably serving as a major signal source with ROCs being one of the receiver cells. However, whether the expressed TGF $\beta$  members in *Xenopus* RICs act positively or negatively to TGF $\beta$  signaling cannot be concluded from our results. There are also contradictory results showing the promotion of malignant cell migration by PMEPA1/TGF $\beta$  [51], supporting the well-documented context dependency of TGF $\beta$  signaling in cancer. The homeolog of another RIC marker (*gadd45g*) Gadd45 is involved in wound healing in *Drosophila* and protects cells from DNA damage at high ROS levels [52]. Gadd45b is also required for liver regeneration [53].

Among the most dominant RIC markers are genes coding for enzymes remodeling the extracellular matrix, especially *mmps*. There is extensive literature about the function of *mmps*/MMPs, their substrates, and activities [54, 55]. MMPs are enzymes with multiple functions. They are expressed by a variety of cell types and can be active in normal as well as pathological situations [56–59]. The many members of the MMP family make detailed experimental characterization challenging [60]. The most important function of MMPs is the cleavage of ECM components that leads to changes in the behavior of neighboring cells. However, cleavage of ECM can also stimulate signaling. In the context of the cleavage of laminin 111, it was suggested to regulate the epithelial-mesenchymal transition [61], laminin 5 gamma 2 chain to induce cell migration [62], and E-cadherin to facilitate migration of epithelial cells [63]. The importance of MMPs during regeneration and healing is evident in various animal models. Loss of *mmp9* function in zebrafish results in defective wound healing and regeneration via excessive collagen formation [64]. The *mmps* and their inhibitors were also identified in zebrafish fin regeneration [65], and several *mmps* including *mmp9* were also induced during axolotl limb regeneration [66]. MMPs 1, 2, 3, 9, 10, 11, and 13 are highly expressed during the early phase of

wound healing in humans after laser treatment [67]. *Mmps 1, 8, and 13* are active anti-fibrotic and proliferative factors during mouse liver regeneration [68]. Increased expression of several *mmps* (8, 11, 13) was also observed in axolotl and lungfish late phase limb and fin regeneration [69]. Cooperative action of individual *mmps* has also been suggested during wound healing, which includes the RIC markers *MMP8* and *MMP9* [70].

ECM composition is one of the key factors determining cell migration and behavior, and both RICs and ROCs express many ECM components. Interestingly, the ECM expression pattern is different between the RICs and the ROCs, suggesting ECM types produced lead to different cell phenotypes.

### Proof of RICs biological function

In this study, the RIC-dependent phase of regeneration (3 hpa–1dpa) was studied, and two complementary approaches were used to prove the biological function of the RICs: (1) physically removing the bud tip containing the RICs; (2) transiently silencing the RIC markers. Both experiments produced phenotypes lacking the migration of ROCs with concomitant excessive fibrosis that blocked regeneration. These inhibited systems resembled the post-amputation phenotype observed during the refractory stage embryos. Based on these observations, we hypothesize that the accumulation of RICs in the bud is a prerequisite for ECM remodeling following amputation. The RICs are probably also important to initiate the migration of other cells like the ROCs. Possibly, the regeneration bud acts as a signaling center, producing chemoattractants for the migratory cells. The nature of these putative attractants remains, however, unclear. Additional functional experiments will be valuable to elucidate RICs alternative functions and fate. However, their transitory state makes the preparation of stable cell lines very challenging. This became apparent to us through our attempts to prepare a stable cell line where we encountered several issues. Several attempts to transplant RICs into refractory embryos resulted in no rescue of the regenerative potential. We believe that the RIC's state requires a specific environment initiated by injury, and it is quickly lost after their dissection. However, in our other preliminary experiments, we have obtained partial rescue in refractory embryos using chemicals mimicking RICs function.

Several upregulated RIC genes, like *itga2*, *lamb3*, and *lamc2*, were recently found to be predictors of an aggressive course in pancreatic cancer [71]. Possibly illegitimate activation of RICs in pancreatic cancer positions them as regulators in the tumor micro-environment. Notably, pancreatic cancer is often associated with extensively developed stromal components and the extent of fibroplasia resembles scar tissue [72].

Some studies link the initial steps of regeneration that occur within hours after injury to the regeneration capability. Limb regeneration in adult *Xenopus* following amputation is minimal. However, it can be dramatically improved by treating the stump with a pro-regenerative multidrug mixture the first day after injury resulting in regeneration, although it does take 18 months. Application of the two compounds, 1,4-DPCA (1,4-dihydrophenanthrolin-4-one-3-carboxylic acid), which reduces excessive collagen deposition, and resolving D5, which has anti-inflammatory activity, is critical [73]. A recent pre-print demonstrates the role of an *mmp-1* homolog in the regeneration of the planarian flatworm *Schmidtea mediterranea* [74], which indicates similarities with the regulation of regeneration in invertebrates.

In summary, we have described a novel cell population, the RICs, involved in regeneration initiation, and we have identified several putative targets to improve poor healing and regeneration incapacities. Additional cellular analysis based on lineage tracing of RICs in a suitable zebrafish embryonic model may provide further valuable information to understand the mechanisms behind the RICs. There are similarities in the expression profiles we present here to those observed during prenatal development, wound repair, and cancer [75]. This includes producing ECM and its remodeling, which is essential for wound healing and cancer dissemination [76, 77]. Based on the complex mechanism of healing, an effective therapy will hardly be based on a single factor, but rather on a combination of several factors applied during different phases of regeneration.

## Conclusions

Our research provides evidence that there is a unique cell population, the regeneration initiation cells (RICs), which is transiently formed soon after the amputation of the *X. laevis* tail. We hypothesized it is an important hub that modifies the ECM and is likely to stimulate the migration of other cell types that promote regeneration. The absence or deregulation of the RICs leads to excessive extracellular matrix deposition and regeneration defects.

## Methods

### Embryo preparation and amputation

The handling, preparation, and experimentation of the animal models were done following the protocols approved by the animal committee of the Czech Academy of Sciences and EU legislation. *Xenopus laevis* females were stimulated with 500 U of human chorionic gonadotropin (Sigma-Aldrich). Eggs were collected the following day and fertilized by testes suspension which were surgically obtained from the male. After removing the jelly coats using a 2% cysteine treatment, embryos were incubated in  $0.1 \times$  MBS until the experimental procedure. Embryonic stages were determined based on the Nieuwkoop and Faber table [78]. Amputation was performed manually (removal of ~30% of tail tissue) using tricaine anesthetized tadpoles at stage 40 (regenerative) and stage 46 (refractory). Embryos were immediately transferred to the  $0.1 \times$  MMR solution with gentamycin. The solution was changed every day. Embryos for RNA work were again anesthetized at defined time points, and the regenerating tissue was dissected, collected, and (1) stored at  $-80^\circ\text{C}$  (bulk RNA-Seq, RT-qPCR), (2) dissociated into cell suspension (scRNA-Seq), or (3) embedded to cutting medium (spatial transcriptomics). Whole embryos were fixed in 4% paraformaldehyde for immunohistochemistry or in situ hybridization.

### RT-qPCR

Regeneration and refractory tissues were pooled from five embryos in at least biological triplicates and stored at  $-80^\circ\text{C}$ . Regenerating tissues were also prepared from the tail of sturgeon embryos, spinal cord tissue in rats, fibroblast culture using scratch assay from humans, the liver in mice, and the tail in the zebrafish embryos. Collaborating laboratories performed injuries, collected control and regenerating tissues (at least biological triplicates for all conditions) in defined periods, and stored them at  $-80^\circ\text{C}$ . Total RNA



was extracted using TriReagent extraction and LiCl precipitation (Sigma) according to the manufacturer's instructions. The concentration of total RNA was determined using a spectrophotometer (Nanodrop 2000; Thermo Fisher Scientific), and the quality of RNA was assessed using a Fragment Analyzer (Agilent, Standard Sensitivity RNA analysis kit, DNF-471).

cDNA was prepared using 100 ng of total RNA, 0.5  $\mu$ l of oligo dT and random hexamers (50  $\mu$ M each), 0.5  $\mu$ l of dNTPs (10 mM each), 0.5  $\mu$ l of Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific), 0.5  $\mu$ l of recombinant ribonuclease inhibitor (RNaseOUT, Invitrogen), and 2  $\mu$ l of 5  $\times$  Maxima RTA buffer (Thermo Scientific), which were mixed with Ultrapure water (Invitrogen) to a final volume 10  $\mu$ l. Samples were incubated for 5 min at 65  $^{\circ}$ C, followed by 10 min at 4  $^{\circ}$ C, 10 min at 25  $^{\circ}$ C, 30 min at 50  $^{\circ}$ C, 5 min at 85  $^{\circ}$ C and cooling to 4  $^{\circ}$ C. Obtained cDNA samples were diluted to a final volume of 60  $\mu$ l and stored at  $-20^{\circ}$ C.

qPCR reaction contained 5  $\mu$ l of TATAA SYBR Grand Master Mix, 0.5  $\mu$ l of forward and reverse primers mix (mixture 1:1, 10  $\mu$ l each), and 2  $\mu$ l of cDNA and 2.5  $\mu$ l of RNase-free water. qPCR was performed using the CFX384 Real-Time system (BioRad) with conditions: initial denaturation at 95  $^{\circ}$ C for 3 min, 40 repeats of denaturation at 95  $^{\circ}$ C for 10 s, annealing at 60  $^{\circ}$ C for 20 s, and elongation at 72  $^{\circ}$ C for 20 s. Melting curve analysis was performed after to test reaction specificity, and only one product was detected for all assays.

### **Bulk RNA-Seq**

Bulk RNA-Seq was done to analyze the temporal changes in gene expression during the regeneration and refractory periods. The first regeneration experiment assessed time intervals (0, 0.5, 1.5, 3, 6 hpa and 1, 3, 7 dpa), while the second, which was run in parallel with the refractory, analyzed time intervals (0, 0.5, 6 hpa and 3 dpa). The first experiment was done in triplicates, while the second used four replicates. All samples were stored at  $-80^{\circ}$ C. Total RNA was extracted from 20 embryos per sample at regenerative and refractory stages. Total RNA was extracted and validated using the same protocol as for RT-qPCR. One hundred nanograms of total RNA was used for library preparation (Lexogen SENSE Total RNA-Seq Library Prep Kit). Library quality was tested by capillary electrophoresis on Fragment Analyzer (Agilent, NGS High Sensitivity kit DNF-474). The libraries were pooled and sequenced using Illumina NextSeq 500, 2  $\times$  76 bp. On average, approximately 4.6 M reads per sample were obtained after quality control. Reads were filtered for low quality reads and adaptor sequences using TrimmomaticPE (v. 0.36) [79], while ribosomal RNA reads were filtered out using Sortmerna (v. 2.1b) [80] (default parameters). The cleaned reads were then aligned using STAR (v. 2.7.9a) [81] to the *Xenopus laevis* genome version 10.1 and annotation version XENLA\_10.1\_GCF [82]. A count table was then generated using the python script htseq-count (v. 0.6.1p1) [83] with the parameter “-m union”. The counts were normalized and analyzed using DESeq2 (v. 1.32.0) [84] under R (v. 4.1.0) [85]. Outlier samples were assessed by analyzing the PCA of the 500 most variable genes. The differentially expressed genes (DEGs) were assessed across the time point for each separate experiment, using the Design: ~ Replicate + Condition, Test: LRT and Reduced model: ~ Replicate. A DEG was defined by a *p*-adjusted value < 0.01 and with a transcript count of > 20 in at least one sample. Clustering to

identify DEGs with similar temporal profiles was done using the `degPatterns` function from `DEGreport` (v. 1.28.0) [86]. Parameters for the clustering were restricted to DEGs with a reproducible minimum of 1.5-fold (short time period) or 2-fold (long time period) difference between any given time point. The resulting clusters were then manually curated to create three final clusters representing early genes (highest expression found between 0 and 1.5 hpa, intermediate genes with highest expression between 3 and 6 hpa, and late genes with highest expression between 1 and 7 dpa. Gene Ontology (GO) enrichment for each cluster was analyzed using `EnrichGO/compareCluster` from `ClusterProfiler` (v. 4.0.5) [87], with the background genes set as the annotatable genes from *X. laevis* (v. 10.1) as reference where applicable, but the GO terms from the reference human database `org.Hs.eg.db` (v. 3.13.0) [88]. GO terms were deemed as significantly enriched using a *p*-adjusted value of 0.01 after multiple hypothesis correction using Benjamini–Hochberg method. The DEGs and GO terms were then compared between the similar clusters of the different conditions to identify similarities and differences.

### Single-cell analysis (scRNA-Seq)

Single-cell experiment was performed at 0, 1, 3, and 12 hpa of *Xenopus* tail collected at stage 40. To avoid artificial activation of gene expression and keep cells at a low temperature, the process of the whole tissue dissociation was done in a refrigerated room at 4 °C. Embryos were anesthetized and the regenerating part was dissected and collected into 0.5 ml of 2/3 PBS (Sigma D8537, diluted by RNase-free water) with actinomycin D (ActD, 50 µg/ml; Sigma A1410, storage solution 5 mg/ml in DMSO). Tail pieces from 50 embryos were collected per one sample. Using a test experiment, the artificial expression or loss of cell types within the cell suspension relative to the collected regenerating tissue was assessed using RT-qPCR of known early response genes and cell population markers.

Tubes were spun down for 5 s using a table centrifuge, the medium removed, and tissue resuspended in 200 µl of dissociation solution I containing papain (SERANA, RPL-001-100 ml), BSA (40 µg/ml, Thermo Fisher AM2616), ActD (50 µg/ml), protease (0.5 mg/ml; Sigma P5380, storage solution 250 mg/ml in PBS), and DNase I (50 µg/ml; Roche 11,284,932,001, storage solution 10 mg/ml in RNase-free water). Dissociation was gently resuspended using P200 wide-bore tip for 1 min and gently rotated for 2 min. After three repeated resuspending/rotating steps, samples were shortly spun down, and supernatant with cells were collected into tubes with 1 ml of fetal bovine serum (FBS, Gibco) prechilled on ice. Undissociated pieces were resuspended in another 200 µl of dissociation solution I and dissociated for another three steps of resuspending and rotating. The supernatant with cells was collected in a new tube with 1 ml of FBS. Next, the tissue was resuspended in 200 µl of dissociation solution II (papain with CaCl<sub>2</sub> (5 mM), BSA (40 µg/ml), ActD (50 µg/ml), protease (5 mg/ml), and DNase I (50 µg/ml), and dissociation was performed for three rounds of 10 min with dissociation the same as described above. In total, five tubes with single-cell suspension in FBS were collected per sample. This allowed for both the preservation of the quality of sensitive cells (released first) and also the collection of the inner cells (collected last).

All tubes were centrifuged (300 g, 6 min, 2 °C, minimal acceleration, no break), medium was removed (minimal volume was retained to avoid air contact with cells), and

cell pellets from one sample were resuspended using P1000 wide-bore tip and pooled together into one 1.5 ml centrifugation tube using 500  $\mu$ l of 2/3 PBS with BSA (40  $\mu$ g/ml) and ActD (5  $\mu$ g/ml). Cells were washed twice using centrifugation (100 g, 7 min, 2 °C, minimal acceleration, no break) and resuspended in 1 ml of 2/3 PBS with BSA and ActD. After the third centrifugation (100 g, 7 min, 2 °C, minimal acceleration, no break), the cell pellet was resuspended in 200  $\mu$ l of 2/3 PBS with BSA and without ActD. Cell suspensions were filtered using a 50- $\mu$ m filter (CellTricks), and cell concentrations were measured using TC20 Automated Cell Counter (BioRad). Cells larger than 7  $\mu$ m were counted, and samples with higher than 80% viability were used in the next step.

Sequencing libraries were prepared according to the manufacturer's manual "Chromium Single Cell 3' Reagent Kits User Guide (v 3.1)." In brief, a total of 2400 cells per sample were loaded into the Chromium chip. Library quality was tested by capillary electrophoresis on Fragment Analyzer (Agilent, NGS High Sensitivity kit, DNF-474). The sample libraries were then pooled and sequenced on Illumina NovaSeq 2000, targeting 100,000 read pairs per cell.

Data were processed using STAR v2.7.9a [81]. Reads were aligned and counted against *Xenopus laevis* genome (v 10.1) with annotation XENLA\_10.1\_GCF obtained from Xenbase [82]. STAR parameters were set "--soloType CB\_UMI\_Simple --soloCB-whitelist 3 M-february-2018.txt --soloCBstart 1 --soloCBlen 16 --soloUMIstart 17 --soloUMIlen 12 --soloBarcodeReadLength 0 --soloFeatures Gene --soloCBmatchWltype 1MM\_multi\_Nbase\_pseudocounts --soloUMIdedup 1MM\_Directional\_UMItools --soloUMIfiltering MultiGeneUMI --clipAdapterType CellRanger4 --sjdbGTFfeatureExon exon --sjdbGTFtagExonParentTranscript transcript\_id --sjdbGTFtagExonParentGene gene\_name --sjdbOverhang 119." Raw unfiltered data were processed using R packages. Firstly, droplets with cells were selected using DropletUtils v1.14.1 [89, 90] using command emptyDrops with parameter "lower=1000" and  $FDR \leq 0.001$ . The filtered matrix was then processed using Seurat v4.1.0 [91]. The normalization and integration of the data followed the standard Seurat protocol. Cells were kept that had less than 15% reads from mitochondrial genes and number of UMIs and counts greater than 2500. The counts were normalized using SCTransform, after which the normalized data from all time points were integrated using integration anchor points selected from the most 10,000 variable genes. Identification of nearest neighbors and clusters were done using the default parameters from the Seurat package (k.parameter of 20, annoy nearest neighbor method using the Euclidean distance metric) and a resolution of 0.5 on the 1:25 PCAs. Uniform manifold approximation and projection (UMAP) was then used to visualize these clusters in a two-dimensional space. Clusters that showed mixed annotation were removed from the final UMAP visualization. The quality and accuracy of the clusters were assessed using SCCAF (v. 0.0.10) under the default condition [92]. Before identifying marker genes, PrepSCTFindMarkers was run to standardize the sequencing depth. FindAllMarkers function was then used to identify marker genes for each cluster using the default parameters. Identified genes were used for manually curated annotation using provided database (<https://marionilab.cruk.cam.ac.uk/XenopusRegeneration/>, [32]). Annotations were verified using known markers (Additional file 2: Fig. S4A). The expression of the genes was visualized using the corrected SCT assay produced after the PrepSCTFindMarkers.

The cellular trajectory of the basal epidermal (*tp63* +), RICs, and ROCs cells was assessed using *scvelo* (v. 0.3.1) using their dynamic model [93]. *CellRank* (v. 0.2.0.2) was used to infer the initial cell population and the resulting terminal states based on the RNA velocity and gene similarity matrix using the default parameters [94]. Genes were first filtered to identify the top 2000 variable genes with a minimum of 20 counts shared across the cells. The counts were then normalized and log transformed. Principle components and moments for velocity estimation were calculated using the default parameters. The putative number of macrostates was selected by analyzing the inflection point of the plotted top 20 eigenvalues. The course-grained transition matrix was used to visually check the validity of the states. The terminal and initiating states from the selected macrostates were then automatically determined using the inbuilt functions in *CellRank*. Genes with greater than at least  $r=0.6$  (and  $p$  value  $<0.01$ ) correlation between its gene expression and the cell fate probabilities for a terminal state were deemed as putative driver genes.

Cell–cell communication between the cell populations were inferred using *CellChat* (v. 2.1.1), using the ligand-receptor database *CellChatDB.human* [95]. The scRNA-Seq datasets were split into separate assays for each time point. Each RNA assay was log normalized, and the *X. laevis* gene symbol was converted to its most probable human ortholog. 24,413 genes had a human ortholog, of which 9894 had paralogs and had to be given a unique identifier. A *CellChat* object was created for each time point, requiring a minimum of 10 cells for a given cluster. The merged *CellChat* objects were then compared to identify the major cell population sources of pathway alterations (maximum senders and receivers of signal). Data at 1 versus 3 hpa were analyzed in-depth to determine the signals that changed between the RIC populations at the two time points and also relative to the other cell populations. Additionally, from functional similarities between ligand-receptor pairs at 1 versus 3 hpa, the most distant pathways were identified including specific upregulated ligand-receptor pairs.

*pySCENIC* (v. 0.12.1) was used to infer the cell states through the association of known human transcription factors and their target genes [96]. The normalized counts were first filtered to remove genes with counts less than 40 across the cells. Genes were further filtered to keep those with a human gene symbol. *Homo sapiens* (hg38, refseq\_r80, mc\_v10\_clust) feather database with the ranking and scores for the motifs and genes, the motif2tf annotations v10, and the list of transcription factors were downloaded from the *pySCENIC* resources (resources.aertslab.org). The software was run using the default parameters (mode: “custom\_multiprocessing”). An initial run of *pySCENIC* was done to identify and remove paralog genes that were not correlated with a transcription factor. The regulons were then filtered against the RICs marker genes, and the AUC values were visualized using a heatmap.

### Spatial transcriptomics

Given the small size of *X. laevis* tails, biological replicates (6 for refractory and 8 for regenerative embryos) for each condition (Fig. 4A) were prepared and analyzed using the Visium platform from 10 × Genomics for fresh frozen tissue (poly-A fraction), and the produced data initially queried using the 10 × Genomic Loupe software (<https://>

[support.10xgenomics.com/spatial-gene-expression/software](https://support.10xgenomics.com/spatial-gene-expression/software)). Together we analyzed the following: (1) regenerative 6 hpa and 1 dpa, (2) refractory 6 hpa and 1 dpa, (3) regenerative and refractory 3 dpa. Dissected tissues were embedded and oriented in 50% optimal cutting temperature (OCT) medium, rapidly frozen using dry ice, and then transferred to  $-80^{\circ}\text{C}$  for a maximum of 6 weeks of storage. Samples were sectioned sagittally ( $20\ \mu\text{m}$  thickness) using Leica CM1950 cryostat (Leica Microsystems). Sections collected for the  $10\times$  Visium Spatial Gene Expression processing were then stored at  $-80^{\circ}\text{C}$ .

Fixation, staining, and imaging were performed strictly according to manufacturer's manuals as in "Methanol fixation + H&E Staining Demonstrated protocol" (CG000160) and "Imaging Guidelines Technical Note" (CG000241). In brief, the sections were shortly incubated to thaw them, after which they were methanol-fixed, isopropanol-incubated, and H&E stained. Stained sections were imaged using Carl Zeiss AxioZoomV16 upright microscope equipped with Plan-Neofluar Z objective ( $2.3\times$  magnification, 0.57 NA, 10.6 WD) at a total of  $63.0\times$  zoom. Samples were imaged with 5% overlap and stitched using the ZEN blue pro 2012 software.

Sequencing libraries were prepared according to the manufacturer's manual "Visium Spatial Gene Expression Reagent Kits User Guide" (CG000239). Sections were permeabilized for 9 min, as was previously determined by performing an optimization experiment. Reverse transcription was performed on the slide, followed by 2nd cDNA strand synthesis. Double-stranded cDNA was transferred, PCR-amplified, enzymatically fragmented, size selected, and tagged by Illumina sequencing adapters. Library quality was tested by capillary electrophoresis on Fragment Analyzer (Agilent, NGS High Sensitivity kit DNF-474). The sample libraries were then pooled and sequenced on Illumina NovaSeq 2000 targeting 50,000 read pairs per spot. We selected the clearest four neighboring biological replicates for visualization in Fig. 5. In total, 5195 spots were analyzed, and on average, 3800 genes per spot were identified (in total, 22,095 genes were detected). Ten spot clusters were created (K-means), annotated based on known predominant marker genes, and confirmed using overlap with the single-cell data.

For deconvolution of spatial data using single-cell results, the raw sequencing data were processed using the recommended set of Space ranger function (v. 1.2.2,  $10\times$  Genomics) for processing of fresh frozen samples. Binary base call files were demultiplexed using mkfastq function with default parameters. The resulting fastq files were mapped separately to Space ranger reference (XENLA\_10.1\_GCF) using the count function which takes a microscope slide image and fastq files, performs alignment, tissue detection, fiducial detection, and barcode/UMI counting.

Spatial regeneration and refractory datasets were analyzed using Seurat (v. 5) [97]. Spots with greater than 20% reads for the mitochondrial genes were removed. Each dataset was normalized using SCTransform using vst.flavor v2. Reference dependent deconvolution based on robust cell type decomposition (RCTD) was performed using the R package spacexr (v. 2.2.1) using the multi doublet mode [98]. We used our annotated regeneration scRNA-Seq dataset as the reference. However, the MCCs, melanocyte precursors, muscle precursors, and "somites and others" were removed from the reference dataset due to either low cell counts, precursor cell status, or mixed cell status. The subweights for the possible cell types within each spot were extracted and visualized as a pie chart using the vizAllTopics function from STdeconvolve (v. 1.4.0) [99].

Reference independent deconvolution was done using STdeconvolve [99]. The spatial datasets were first filtered to remove genes with less than 10 reads across the spots and spots with less than 100 detected genes. The top 3000 over dispersed genes present between 2.5 and 95% of the pixels were selected to model the topics. The topics were predicted using a range of potential unique cell types from 2 to 20 in increments of 1. The optimal model was selected based on the one with the lowest perplexity where the cell proportions were below 5%. The topics were then annotated based on the enrichment of the top 20 discriminatory gene markers for the scRNA-Seq cell populations as selected using scMAGS (Additional file 2: Fig. S20) [100]. Topics that showed the same cell type were merged into one singular topic and visualized as a pie chart using the vizAllTopics function. Correlation between the cell type proportions found in each spot of the four replicates between the RCTD and STdeconvolve methods was calculated, and the top matches were plotted as a correlation plot.

### Comparative analysis

The marker genes (*PMEPA*, *GADD45G*, *JUNB*, *MMP9*, *MMP8*) for the RICs population were assessed in other healing/regenerating tissues from other models to determine if there was also a peak expression during the 2–6 hpa period. The following temporal RT-qPCR experiments were assessed in the regenerating tail in sturgeon, the spinal cord in rats, fibroblasts from humans, the liver in mice, and the tail in zebrafish. Values were normalized against the time 0 and presented as relative quantities.

A comparison of the shared genes between the markers for RICs And ROCs population in the single-cell, spatial, and intermediate genes from the bulk RNA-Seq was done. ROCs and RICs markers were filtered to select those from the scRNA-Seq and spatial experiment with *p*-adjusted values <0.05 and fold change >2x. The enriched GOs for these populations were compared between the experiments and also different cell populations using the same method from EnrichGO/compareCluster as described earlier. Comparisons were also made of the expression profile of genes associated with collagen, extracellular matrix, integrins, keratins, laminins, metalloproteinase, and metalloproteinases. The expression profiles were observed for the regeneration and refractory bulk RNA-Seq data and the averaged cellular expression (using Seurat's AverageExpression function) from the RICs, ROCs, basal epidermis (*tp63+*), and all other cell types grouped into the category "other" from the 3 hpa regeneration scRNA-Seq dataset. Analysis was also done on the spatial tail regeneration at 6 hpa. Approximately 10 spots were selected from the central part of the tail bud region (posterior) and the central part of the region furthest from it (anterior) for each replicate. Seurat's AverageExpression function was used to calculate the average spot expression of the queried gene in the areas of interest across the replicates.

*Xenopus laevis* genome (v 10.1) with annotation XENLA\_10.1\_GCF [82] was used for all alignments and annotations. The unannotated protein coding transcripts were analyzed for their closest ortholog relative to the *H. sapiens* proteome (GRCh38.p14) using the reciprocal best alignment heuristic tool Proteinortho (v. 6.0.31) along with DIAMOND (v. 2.0.11) [101, 102].

### Functional experiments

Reamputation was performed using manual dissection of regeneration bud containing RICs by scalpel at 6 hpa, and embryos were incubated in  $0.1 \times$  MMR with gentamycin at 16 °C. Scoring was performed at 7–10 dpa. It was based on three main criteria: notochord length, notochord shape, and epidermis attachment to the notochord and scaled in the range 0 (no regeneration), to 0.25 (poor regeneration), to 0.5 (partial regeneration), to 0.75 (minor defect in regeneration), to 1 (perfect regeneration). To simplify visualization, values  $>0.5$  were called regenerative and  $<0.5$  non-regenerative. Samples for fibronectin staining and RT-qPCR of ROCs markers were collected at 1 dpa.

Three Vivo morpholino antisense oligos (MOs) were designed and ordered from Gene Tools, LLC (Philomath, Oregon USA). MOs targeted exon–intron junction near the start of the coding sequence of *mmp8*, *mmp9*, and *pmepa1* genes (*mmp8* MO: 5′-ATG AAAACCAATCTACTTACCTCAG-3′; *mmp9* MO: 5′-CATGATCAATAATCCCCTCAC-3′; *pmepa1* MO: 5′-CCTTTATAATTGCTACTTACAGATC-3′). Standard Control Vivo MO was used in parallel to test for concentration toxicity and specificity. Stock solution with a concentration of 0.5 mM was prepared in UltraPure DNase/RNase free water (Invitrogen) and stored at room temperature. A range of concentrations (1000–100 nM) and inhibitory time (a few hours–5 days treatment) were tested to reach an inhibitory effect while achieving low mortality. Five testing experiments were performed and revealed the optimal conditions: working solution in a final 250 nM concentration (below the suggested toxic level for cell cultures, which is 10 μM). Embryos were transferred to MO solutions immediately after tail amputation and incubated until 1 dpa, followed by additional incubation in fresh  $0.1 \times$  MMR, which was changed daily. Higher concentrations and longer incubations resulted in high mortality in contrast to variable effects when using lower and shorter incubations. The mortality rate in optimal concentration was ~5% for untreated and control MO and 20–30% for *mmp8*, *mmp9*, and *pmepa1* MOs. RT-qPCR was used for inhibition quantification, and we detected ~50% decrease of *pmepa1* and *mmp9* mRNA in their respective MO treated embryos' expression, with no change of expression in control MO embryos compared to untreated embryos (Additional file 2: Fig. S21). Scoring and phenotype experiments were performed in 10 independent experiments, and in total,  $>100$  embryos were analyzed per condition. Embryos were collected and fixed at 1 dpa (RT-qPCR, Fibronectin staining) or 7–10 dpa (regeneration scoring). Refractory embryos were prepared the same way as MOs.

Embryos for in situ hybridization and immunohistochemistry were fixed in 4% paraformaldehyde overnight at 4 °C and then stored in 100% methanol (Penta chemicals) at –20 °C. Whole mount in situ hybridization protocol was adopted from Sive et al. [103]. At least seven embryos per condition were prepared. Pictures were taken using Nikon SMZ 1500 microscope. Fibronectin immunohistochemistry was performed using at least five embryos per condition. After three washes in PBT for 15 min, the primary antibody (anti-Fibronectin, Sigma F3648) was added in concentration 1:150 for overnight incubation at 4 °C. The next day, a secondary antibody was added to the concentration at 1:500 and incubated overnight at 4 °C. DAPI 1:1000 was used as a reference nuclei label.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03396-3>.

Additional file 1: Supplementary Tables. Tables S1 to S4 (file contains analysis from bulk RNA-Seq analysis of regeneration and refractory conditions in *Xenopus laevis*).

Additional file 2: Supplementary Figures. Figures S1 to S21 (file contains information on quality control analysis, gene markers used to annotate clusters, spatial deconvolution, trajectory analysis, cell–cell communication analysis, and functional analysis) analysis from scRNA-Seq analysis of regeneration condition in *Xenopus laevis*.

Additional file 3: Supplementary Tables. Tables S1 to S2 (file contains analysis from scRNA-Seq analysis of regeneration condition in *Xenopus laevis*).

Additional file 4: Supplementary Tables. Tables S1 to S2 (file contains analysis from spatial regeneration and refractory conditions in *Xenopus laevis*).

Additional file 5: Supplementary Tables. Tables S1 to S2 (file contains comparative analysis between the results from the bulk RNA-Seq, scRNA-Seq and spatial regeneration or refractory conditions in *Xenopus laevis*).

Additional file 6: Review history.

## Acknowledgements

We thank the other members of the Laboratory of Gene Expression for their support. Image schematics for the experimental procedures were created with BioRender.com.

## Review history

The review history is available as Additional file 6.

## Peer review information

Stefania Giacomello and Veronique van den Berghe were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

## Authors' contributions

Conceptualization: RS. Data curation: RS, PA, RN, DK. Formal Analysis: RS, PA, RN, DK. Funding acquisition: RS, MK. Investigation: RS, PA, DZ, RN, DK, JN. Methodology: RS, PA, DZ, RN, DK, JN. Project administration: RS. Resources: RS, KS, LL, BBE, JN, MP. Supervision: RS. Validation: RS, DK, JN. Visualization: RS, PA, RN, DK, JN. Writing—original draft: RS. Writing—review and editing: RS, PA, DZ, RN, DK, JN, KS, LL, BBE, JN, MP, MK.

## Authors' X handles

X handles: @labgenex (Mikael Kubista).

## Funding

86652036 from RVO

The Grant Agency of the Czech Republic GA24-12027S

Operational Programme Research, Development, and Education within the projects: Centre for Tumour Ecology—Research of the Cancer Microenvironment Supporting Cancer Growth and Spread (reg. No. CZ.02.1.01/0.0/0.0/16\_019/000785)

Project National Institute for Cancer Research (Programme EXCELES, ID Project No. LX22NPO5102)—funded by the European Union—Next Generation EU

Charles University project Cooperation ONCO

NW24-03-00459 from Ministry of Health of the Czech Republic

## Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) [104], accession number GSE245320 [105] (GSE245312, GSE245313, GSE245317, GSE245318). The scRNA-Seq analysis (UMAP, cluster annotations) has been deposited online on Nygen's ScarfWeb (<https://scarfweb.nygen.io/eu-central-1/public/jhexdxjk>).

## Declarations

### Ethics approval and consent to participate

All experimental procedures involving *X. laevis* were carried out in accordance with the Czech Law 246/1992 on animal welfare. Protocols involving *X. laevis* were approved by the animal committee of the Czech Academy of Sciences (137/2016 and 65–2024-P).

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.



**Author details**

<sup>1</sup>Laboratory of Gene Expression, Institute of Biotechnology of the Czech Academy of Sciences, Vestec 252 50, Czech Republic. <sup>2</sup>First Faculty of Medicine, Institute of Anatomy, Charles University, Prague 2 128 00, Czech Republic. <sup>3</sup>Department Dermatovenereology, First Faculty of Medicine and General University Hospital, Charles University, Prague, Czech Republic. <sup>4</sup>Laboratory of Molecular Therapy, Institute of Biotechnology of the Czech Academy of Sciences, Vestec 252 50, Czech Republic. <sup>5</sup>School of Pharmacy and Medical Science, Griffith University, Southport, QLD, Australia. <sup>6</sup>Faculty of Science, Charles University, Prague 2, Czech Republic. <sup>7</sup>First Faculty of Medicine, Charles University, Prague 2, Czech Republic. <sup>8</sup>Faculty of Fisheries and Protection of Waters, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, University of South Bohemia in Ceske Budejovice, Vodnany 389 25, Czech Republic.

Received: 5 October 2023 Accepted: 19 September 2024

Published online: 01 October 2024

**References**

- Krafts KP. Tissue repair: the hidden drama. *Organogenesis*. 2010;6:225–33.
- Murawala P, Tanaka EM, Currie JD. Regeneration: the ultimate example of wound healing. *Semin Cell Dev Biol*. 2012;23:954–62.
- Dall'Agnese A, Puri PL. Could we also be regenerative superheroes, like salamanders? *Bioessays*. 2016;38:917–26.
- Coleman CM. Chicken embryo as a model for regenerative medicine. *Birth Defects Res C Embryo Today*. 2008;84:245–56.
- Jacyniak K, McDonald RP, Vickaryous MK. Tail regeneration and other phenomena of wound healing and tissue restoration in lizards. *J Exp Biol*. 2017;220:2858–69.
- Vonk AC, Zhao X, Pan Z, Hudnall ML, Oakes CG, Lopez GA, Hasel-Kolossa SC, Kuncz AWC, Sengelmann SB, Gamble DJ, Lozito TP. Single-cell analysis of lizard blastema fibroblasts reveals phagocyte-dependent activation of Hedgehog-responsive chondrogenesis. *Nat Commun*. 2023;14:4489.
- So J, Kim A, Lee SH, Shin D. Liver progenitor cell-driven liver regeneration. *Exp Mol Med*. 2020;52:1230–8.
- Seifert AW, Kiama SG, Seifert MG, Goheen JR, Palmer TM, Maden M. Skin shedding and tissue regeneration in African spiny mice (*Acomys*). *Nature*. 2012;489:561–5.
- Li C, Zhao H, Liu Z, McMahon C. Deer antler—a novel model for studying organ regeneration in mammals. *Int J Biochem Cell Biol*. 2014;56:111–22.
- Smetana K Jr, Dvorankova B, Lacina L. Phylogeny, regeneration, ageing and cancer: role of microenvironment and possibility of its therapeutic manipulation. *Folia Biol (Praha)*. 2013;59:207–16.
- Galliot B, Tanaka E, Simon A. Regeneration and tissue repair: themes and variations. *Cell Mol Life Sci*. 2008;65:3–7.
- Beck CW, Izpisua Belmonte JC, Christen B. Beyond early development: *Xenopus* as an emerging model for the study of regenerative mechanisms. *Dev Dyn*. 2009;238:1226–48.
- Christensen RN, Tassava RA. Apical epithelial cap morphology and fibronectin gene expression in regenerating axolotl limbs. *Dev Dyn*. 2000;217:216–24.
- Kragl M, Knapp D, Nacu E, Khattak S, Maden M, Epperlein HH, Tanaka EM. Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature*. 2009;460:60–5.
- Tanaka EM, Reddien PW. The cellular basis for animal regeneration. *Dev Cell*. 2011;21:172–85.
- Love NR, Chen Y, Ishibashi S, Kritsiligkou P, Lea R, Koh Y, Gallop JL, Dorey K, Amaya E. Amputation-induced reactive oxygen species are required for successful *Xenopus* tadpole tail regeneration. *Nat Cell Biol*. 2013;15:222–8.
- Helston O, Amaya E. Reactive oxygen species during heart regeneration in zebrafish: lessons for future clinical therapies. *Wound Repair Regen*. 2021;29:211–24.
- Carbonell MB, Zapata Cardona J, Delgado JP. Post-amputation reactive oxygen species production is necessary for axolotl limb regeneration. *Front Cell Dev Biol*. 2022;10:921520.
- Abaffy P, Tomankova S, Naraine R, Kubista M, Sindelka R. The role of nitric oxide during embryonic wound healing. *BMC Genomics*. 2019;20:815.
- Aztekin C, Hiscock TW, Butler R, De Andino FJ, Robert J, Gurdon JB, Jullien J. The myeloid lineage is required for the emergence of a regeneration-permissive environment following *Xenopus* tail amputation. *Development*. 2020;147:dev185496.
- Aztekin C, Storer MA. To regenerate or not to regenerate: vertebrate model organisms of regeneration-competency and -incompetency. *Wound Repair Regen*. 2022;30:623–35.
- Godwin JW, Brookes JP. Regeneration, tissue injury and the immune response. *J Anat*. 2006;209:423–32.
- Godwin JW, Rosenthal N. Scar-free wound healing and regeneration in amphibians: immunological influences on regenerative success. *Differentiation*. 2014;87:66–75.
- Zivicova V, Lacina L, Mateu R, Smetana K Jr, Kavkova R, Drobna Krejci E, Grim M, Kvasilova A, Borsky J, Strnad H, et al. Analysis of dermal fibroblasts isolated from neonatal and child cleft lip and adult skin: developmental implications on reconstructive surgery. *Int J Mol Med*. 2017;40:1323–34.
- Santosh N, Windsor LJ, Mahmoudi BS, Li B, Zhang W, Chernoff EA, Rao N, Stocum DL, Song F. Matrix metalloproteinase expression during blastema formation in regeneration-competent versus regeneration-deficient amphibian limbs. *Dev Dyn*. 2011;240:1127–41.
- Vinarsky V, Atkinson DL, Stevenson TJ, Keating MT, Odelberg SJ. Normal newt limb regeneration requires matrix metalloproteinase function. *Dev Biol*. 2005;279:86–98.
- Arenas Gomez CM, Sabin KZ, Echeverri K. Wound healing across the animal kingdom: crosstalk between the immune system and the extracellular matrix. *Dev Dyn*. 2020;249:834–46.
- Godwin J, Kuraitis D, Rosenthal N. Extracellular matrix considerations for scar-free repair and regeneration: insights from regenerative diversity among vertebrates. *Int J Biochem Cell Biol*. 2014;56:47–55.

29. Lu P, Takai K, Weaver VM, Werb Z. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harb Perspect Biol.* 2011;3:a005058.
30. 2013 MotY: Method of the year 2013. *Nat Methods* 2014;11:1. <https://www.nature.com/articles/nmeth.2801>.
31. Johnson GL, Masias EJ, Lehoczy JA. Cellular heterogeneity and lineage restriction during mouse digit tip regeneration at single-cell resolution. *Dev Cell.* 2020;52(5):525–540. e525.
32. Aztekin C, Hiscock TW, Marioni JC, Gurdon JB, Simons BD, Jullien J. Identification of a regeneration-organizing cell in the *Xenopus* tail. *Science.* 2019;364:653–8.
33. Slack JM, Beck CW, Gargioli C, Christen B. Cellular and molecular mechanisms of regeneration in *Xenopus*. *Philos Trans R Soc Lond B Biol Sci.* 2004;359:745–51.
34. Phipps LS, Marshall L, Dorey K, Amaya E. Model systems for regeneration: *Xenopus*. *Development.* 2020;147:dev180844.
35. Beck CW, Christen B, Slack JM. Molecular pathways needed for regeneration of spinal cord and muscle in a vertebrate. *Dev Cell.* 2003;5:429–39.
36. MotY: Method of the year spatially resolved transcriptomics. *Nat Methods.* 2020;2021(18):1. <https://www.nature.com/articles/s41592-020-01042-x>.
37. Foster DS, Januszky M, Yost KE, Chinta MS, Gulati GS, Nguyen AT, Burcham AR, Salhotra A, Ransom RC, Henn D, et al. Integrated spatial multiomics reveals fibroblast fate during tissue repair. *Proc Natl Acad Sci U S A.* 2021;118:e2110025118.
38. Tower RJ, Busse E, Jaramillo J, Lacey M, Hoffseth K, Guntur AR, Simkin J, Sammarco MC. Spatial transcriptomics reveals metabolic changes underlying age-dependent declines in digit regeneration. *Elife.* 2022;11:e71542.
39. Gerber T, Murawala P, Knapp D, Masselink W, Schuez M, Hermann S, Gac-Santel M, Nowoshilow S, Kageyama J, Khattak S, et al. Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science.* 2018;362:eaag0681.
40. Lin TY, Gerber T, Taniguchi-Sugiura Y, Murawala P, Hermann S, Grosser L, Shibata E, Treutlein B, Tanaka EM. Fibroblast dedifferentiation as a determinant of successful regeneration. *Dev Cell.* 2021;56(15):1541–1551. e1546.
41. Gal P, Brabek J, Holub M, Jakubek M, Sedo A, Lacina L, Strnadova K, Dubovy P, Hornychova H, Ryska A, Smetana K Jr. Autoimmunity, cancer and COVID-19 abnormally activate wound healing pathways: critical role of inflammation. *Histochem Cell Biol.* 2022;158:415–34.
42. Sinha S, Sparks HD, Labit E, Robbins HN, Gowing K, Jaffer A, Kutluerk E, Arora R, Raredon MSB, Cao L, et al. Fibroblast inflammatory priming determines regenerative versus fibrotic skin repair in reindeer. *Cell.* 2022;185(4):717–4736. e4725.
43. Benbow U, Brinckerhoff CE. The AP-1 site and MMP gene regulation: what is all the fuss about? *Matrix Biol.* 1997;15:519–26.
44. Vincenti MP, Brinckerhoff CE. Transcriptional regulation of collagenase (MMP-1, MMP-13) genes in arthritis: integration of complex signaling pathways for the recruitment of gene-specific transcription factors. *Arthritis Res.* 2002;4:157–64.
45. Dvorankova B, Holikova Z, Vacik J, Konigova R, Kapounkova Z, Michalek J, Pradn M, Smetana K Jr. Reconstruction of epidermis by grafting of keratinocytes cultured on polymer support—clinical study. *Int J Dermatol.* 2003;42:219–23.
46. Costa RM, Soto X, Chen Y, Zorn AM, Amaya E. spib is required for primitive myeloid development in *Xenopus*. *Blood.* 2008;112:2287–96.
47. Nakamura M, Yoshida H, Takahashi E, Wlizla M, Takebayashi-Suzuki K, Horb ME, Suzuki A. The AP-1 transcription factor JunB functions in *Xenopus* tail regeneration by positively regulating cell proliferation. *Biochem Biophys Res Commun.* 2020;522:990–5.
48. Fournier PG, Juarez P, Jiang G, Clines GA, Niewolna M, Kim HS, Walton HW, Peng XH, Liu Y, Mohammad KS, et al. The TGF-beta signaling regulator PMEPA1 suppresses prostate cancer metastases to bone. *Cancer Cell.* 2015;27:809–21.
49. Corona A, Blobel GC. The role of the extracellular matrix protein TGFBI in cancer. *Cell Signal.* 2021;84: 110028.
50. He Z, Liang J, Wang B. Inhibin, beta A regulates the transforming growth factor-beta pathway to promote malignant biological behaviour in colorectal cancer. *Cell Biochem Funct.* 2021;39:258–66.
51. Nie Z, Wang C, Zhou Z, Chen C, Liu R, Wang D. Transforming growth factor-beta increases breast cancer stem cell population partially through upregulating PMEPA1 expression. *Acta Biochim Biophys Sin (Shanghai).* 2016;48:194–201.
52. Weavers H, Wood W, Martin P. Injury activates a dynamic cytoprotective network to confer stress resilience and drive repair. *Curr Biol.* 2019;29(38):3851–3862. e3854.
53. Columbano A, Ledda-Columbano GM, Pibiri M, Cossu C, Menegazzi M, Moore DD, Huang W, Tian J, Locker J. Gadd45beta is induced through a CAR-dependent, TNF-independent pathway in murine liver hyperplasia. *Hepatology.* 2005;42:1118–26.
54. Page-McCaw A, Ewald AJ, Werb Z. Matrix metalloproteinases and the regulation of tissue remodeling. *Nat Rev Mol Cell Biol.* 2007;8:221–33.
55. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol.* 2014;15:786–801.
56. Bassiouni W, Ali MAM, Schulz R. Multifunctional intracellular matrix metalloproteinases: implications in disease. *FEBS J.* 2021;288:7162–82.
57. Nagase H, Visse R, Murphy G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res.* 2006;69:562–73.
58. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell.* 2010;141:52–67.
59. Kessenbrock K, Wang CY, Werb Z. Matrix metalloproteinases in stem cell regulation and cancer. *Matrix Biol.* 2015;44–46:184–90.
60. Martins VL, Caley M, O'Toole EA. Matrix metalloproteinases and epidermal wound repair. *Cell Tissue Res.* 2013;351:255–68.

61. Horejs CM, Serio A, Purvis A, Gormley AJ, Bertazzo S, Poliniewicz A, Wang AJ, DiMaggio P, Hohenester E, Stevens MM. Biologically-active laminin-111 fragment that modulates the epithelial-to-mesenchymal transition in embryonic stem cells. *Proc Natl Acad Sci U S A*. 2014;111:5908–13.
62. Sadowski T, Dietrich S, Koschinsky F, Ludwig A, Proksch E, Titz B, Sedlacek R. Matrix metalloproteinase 19 processes the laminin 5 gamma 2 chain and induces epithelial cell migration. *Cell Mol Life Sci*. 2005;62:870–80.
63. McGuire JK, Li Q, Parks WC. Matrilysin (matrix metalloproteinase-7) mediates E-cadherin ectodomain shedding in injured lung epithelium. *Am J Pathol*. 2003;162:1831–43.
64. LeBert DC, Squirrel JM, Rindy J, Broadbridge E, Lui Y, Zakrzewska A, Eliceiri KW, Meijer AH, Huttenlocher A. Matrix metalloproteinase 9 modulates collagen matrices and wound repair. *Development*. 2015;142:2136–46.
65. Bai S, Thummel R, Godwin AR, Nagase H, Itoh Y, Li L, Evans R, McDermott J, Seiki M, Sarras MP Jr. Matrix metalloproteinase expression and function during fin regeneration in zebrafish: analysis of MT1-MMP, MMP2 and TIMP2. *Matrix Biol*. 2005;24:247–60.
66. Yang EV, Gardiner DM, Carlson MR, Nugas CA, Bryant SV. Expression of Mmp-9 and related matrix metalloproteinase genes during axolotl limb regeneration. *Dev Dyn*. 1999;216:2–9.
67. Sherrill JD, Finlay D, Binder RL, Robinson MK, Wei X, Tiesman JP, Flagler MJ, Zhao W, Miller C, Loftus JM, et al. Transcriptomic analysis of human skin wound healing and rejuvenation following ablative fractional laser treatment. *PLoS ONE*. 2021;16: e0260095.
68. Duarte S, Baber J, Fujii T, Coito AJ. Matrix metalloproteinases in liver injury, repair and fibrosis. *Matrix Biol*. 2015;44–46:147–56.
69. Darnet S, Dragalzew AC, Amaral DB, Sousa JF, Thompson AW, Cass AN, Lorena J, Pires ES, Costa CM, Sousa MP, et al. Deep evolutionary origin of limb and fin regeneration. *Proc Natl Acad Sci U S A*. 2019;116:15106–15.
70. Gutierrez-Fernandez A, Inada M, Balbin M, Fueyo A, Pitiot AS, Astudillo A, Hirose K, Hirata M, Shapiro SD, Noel A, et al. Increased inflammation delays wound healing in mice deficient in collagenase-2 (MMP-8). *FASEB J*. 2007;21:2580–91.
71. Islam S, Kitagawa T, Baron B, Abiko Y, Chiba I, Kuramitsu Y. ITGA2, LAMB3, and LAMC2 may be the potential therapeutic targets in pancreatic ductal adenocarcinoma: an integrated bioinformatics analysis. *Sci Rep*. 2021;11:10563.
72. Novak S, Kolar M, Szabo A, Vernerova Z, Lacina L, Strnad H, Sachova J, Hradilova M, Havranek J, Spanko M, et al. Desmoplastic crosstalk in pancreatic ductal adenocarcinoma is reflected by different responses of Panc-1, MIA-PaCa-2, PaTu-8902, and CAPAN-2 cell lines to cancer-associated/normal fibroblasts. *Cancer Genomics Proteomics*. 2021;18:221–43.
73. Murugan NJ, Vigran HJ, Miller KA, Golding A, Pham QL, Sperry MM, Rasmussen-Ivey C, Kane AW, Kaplan DL, Levin M. Acute multidrug delivery via a wearable bioreactor facilitates long-term limb regeneration and functional recovery in adult *Xenopus laevis*. *Sci Adv*. 2022;8:eabj2164.
74. Benham-Pyle BW, Mann FG, Brewster CE, Dewars ER, Nowotarski SH, Guerrero-Hernández C, Malloy S, Hall KE, Maddera LE, Chen S, et al. Stem cells partner with matrix remodeling cells during regeneration. *bioRxiv*. 2022.03.20.485025.
75. Plzak J, Boucek J, Bandurova V, Kolar M, Hradilova M, Szabo P, Lacina L, Chovanec M, Smetana K Jr. The head and neck squamous cell carcinoma microenvironment as a potential target for cancer therapy. *Cancers (Basel)*. 2019;11:440.
76. Mustafa S, Koran S, AlOmair L. Insights into the role of matrix metalloproteinases in cancer and its various therapeutic aspects: a review. *Front Mol Biosci*. 2022;9: 896099.
77. Perrin L, Gligorijevic B. Proteolytic and mechanical remodeling of the extracellular matrix by invadopodia in cancer. *Phys Biol*. 2022;20:015001.
78. Nieuwkoop PD, Faber J. Normal table of *Xenopus laevis* (Daudin): a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis. New York: Garland Pub; 1994.
79. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
80. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*. 2012;28:3211–7.
81. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
82. Fortriede JD, Pells TJ, Chu S, Chaturvedi P, Wang D, Fisher ME, James-Zorn C, Wang Y, Nenni MJ, Burns KA, et al. Xenbase: deep integration of GEO & SRA RNA-seq and ChIP-seq data in a model organism database. *Nucleic Acids Res*. 2020;48:D776–82.
83. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9.
84. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550–550.
85. Team RC. R: a language and environment for statistical computing. Vienna: Austria; 2021.
86. Pantano L. DESeq2: report of DEG analysis. 2021.
87. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284–7.
88. Carlson M. org.Hs.eg.db: Genome wide annotation for Human. 2021.
89. Griffiths JA, Richard AC, Bach K, Lun ATL, Marioni JC. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat Commun*. 2018;9:2667.
90. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T. participants in the 1st Human cell atlas J, Marioni JC. empty-Drops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol*. 2019;20:63.
91. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411–20.

92. Miao Z, Moreno P, Huang N, Papatheodorou I, Brazma A, Teichmann SA. Putative cell type discovery from single-cell gene expression data. *Nat Methods*. 2020;17:621–8.
93. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*. 2020;38:1408–14.
94. Lange M, Bergen V, Klein M, Setty M, Reuter B, Bakhti M, Lickert H, Ansari M, Schniering J, Schiller HB, et al. Cell Rank for directed single-cell fate mapping. *Nat Methods*. 2022;19:159–70.
95. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using cell chat. *Nat Commun*. 2021;12:1088.
96. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, Seurinck R, Saelens W, Cannoodt R, Rouchon Q, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc*. 2020;15:2247–76.
97. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(3573–3587):e3529.
98. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, Irizarry RA. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol*. 2022;40:517–26.
99. Miller BF, Huang F, Atta L, Sahoo A, Fan J. Reference-free cell type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nat Commun*. 2022;13:2339.
100. Baran Y, Dogan B. scMAGS: Marker gene selection from scRNA-seq data for spatial transcriptomics studies. *Comput Biol Med*. 2023;155: 106634.
101. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2014;12(1):59–60.
102. Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*. 2011;12:1–9.
103. Sive HL, Grainger RM, Harland RM. Early development of *Xenopus laevis*: a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 2000.
104. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002;30:207–10.
105. Sindelka R, Naraine R, Abaffy P, Zucha D, Kraus D, Netusil J, Smetana KJ, Lacina L, Endaya BB, Neuzil J, et al. Characterization of regeneration initiating cells during *Xenopus laevis* tail regeneration. Datasets; Gene Expression Omnibus 2023. <https://identifiers.org/geo:GSE245320>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.