The West African lungfish provides insights into the evolution of tetrapod tail regeneration

Kellen Matos Verissimo[1], Louise Neiva Perez[1], Aline Cutrim Dragalzew[1], Gayani Senevirathne[2], Sylvain Darnet[1], Wainna Renata Barroso Mendes[1], Ciro Ariel dos Santos Neves[1], Erika Monteiro dos Santos[1], Cassia Nazare de Sousa Moraes[1], Neil Shubin[2], Nadia Belinda Frobisch[3], Josane de Freitas Sousa[1] and Igor Schneider[1*].

[1] Instituto de Ciências Biológicas, Universidade Federal do Pará, 66075-900, Belém, Brazil.
[2] Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA.
[3] Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, 10115 Berlin, Germany.

*Correspondence to: Igor Schneider, Instituto de Ciências Biológicas, Universidade Federal do Pará, Rua Augusto Corrêa, 01, Belém, 66075-900, Brazil. Phone: 55 91 3201-7009, e-mail: ischneider@ufpa.br

Keywords: lungfish, tail, evolution, regeneration, transcriptome

Abstract

Salamanders, frog tadpoles, and lizards possess the remarkable ability to regenerate tails. The fossil record suggests that this capacity is an ancestral tetrapod trait, yet its evolutionary history remains unclear. Here we examine tail regeneration in a living representative of the sister group of tetrapods, the West African lungfish *Protopterus annectens*. We show that, as seen in salamanders, lungfish tail regeneration occurs via formation of a proliferative blastema and restores original structures including muscle, skeleton and spinal cord. Contrary to lizards and similar to salamanders, lungfish regenerate spinal cord neurons and reconstitute dorsoventral patterning of the tail. Similar to salamander and frog tadpoles, we show that Shh is required for lungfish tail regeneration. Through RNA-seq analysis of uninjured and regenerating tail blastema we show that lungfish deploy a genetic program comparable to that of tetrapods, showing upregulation of genes and signaling pathways previously implicated in amphibian and lizard tail regeneration. Furthermore, the tail blastema showed marked upregulation of genes encoding post-transcriptional RNA processing components and transposon-derived genes. Collectively, our study establishes the lungfish as a valuable research system for regenerative biology and provides insights into the evolution of cellular and molecular processes underlying vertebrate tail regeneration.

Keywords: lungfish, tail, regeneration, evolution, transcriptome

1. Introduction

Salamanders and *Xenopus* tadpoles have the remarkable capacity to regenerate various body parts, including their tails and spinal cord. In both species, tail regeneration initiates through a comparable sequence of events: soon after tail amputation a wound epithelium forms and covers the injured site. Wound healing is followed by the thickening of the wound epidermis, giving rise to the apical epithelial cap (AEC). After initial wound healing, tail regeneration in salamanders and *Xenopus* tadpoles shows significant differences. In salamanders, undifferentiated cells accumulate at the amputation site, giving rise to the blastema [1]. Conversely, in tadpoles, a mass of undifferentiated cells accumulates around the neural ampulla and

notochord tip forming a blastema-like cell population also known as the regeneration bud [2]. In addition, in salamanders, dedifferentiated cells play a large role in restoring muscle tissue [3], whereas in tadpoles this is achieved chiefly via proliferation of Pax7+ satellite cells [4, 5]. Nevertheless, tadpoles and salamanders restore all tail associated tissues such as spinal cord, muscle, vasculature and other tissues.

Lizards are the only amniotes capable of tail regeneration, which often occurs in association to autotomy, when the animal sheds off its own tail. After loss of the tail, a wound epidermis forms and numerous mesenchymal cells of undetermined origin accumulate to produce the regenerative blastema. However, instead of regenerating the spinal cord and associated skeleton as seen in uninjured tails, a simple ependymal tube forms enclosed by an unsegmented cartilage tube [6-7]. In lizards, this simpler version of the original spinal cord and tail skeleton results from failure to recapitulate the embryonic dorsoventral expression pattern of genes responsible for establishing roof plate, floor plate, and lateral domains of the neural tube. Specifically, contrary to salamanders, floor plate markers *Shh* and *FoxA2* in lizards are expressed along the entire regenerating ependymal tube, which consequently acquires a floor plate identity [8].

Great advances have been made in the past 15 years in our understanding of the molecular program of tail regeneration. Studies in *Xenopus* tadpoles have identified a sequence of major molecular events leading to successful tail regeneration. In summary, upon injury, TGF-β signaling mediated by smad2/3 is required for formation of the wound epidermis [9]. Extracellular matrix (ECM) remodeling, reactive oxygen species (ROS) signaling and inflammation are also among the first responses to tail injury [10]. Transcriptomic analysis of genes upregulated in the tail AEC identified pan-AEC factors that act both in regenerating limbs and tails, as well as two tail-specific AEC factors, *Angptl2* and *Egfl6* [11]. Recently, single-cell RNA-seq identified an epidermal cell type termed regeneration-organizing cell (ROC) that relocates to the wound epithelium at amputation plane and is required for tail regeneration [12]. ROCs express genes from pathways known to be necessary for regeneration, such as the Wnt and Fgf pathways. As the regeneration bud forms and tail outgrowth occurs, additional signaling pathways are deployed, such as Bmp, Egf, Shh, Tgf-β and Notch pathways, among others [3-4,10,13]. RNA-seq analysis of isolated proliferating blastema cells revealed blastema-enriched genes such as *Il11*, *Cse1l* and *L1td1-like* [14].

Pharmacological approaches have established roles for histone deacetylases [15] and for Hyaluronan, an ECM component [16]. Finally, functional studies have identified specific genes required for regeneration, as CRISPR/Cas9-mediated knockdown of *Il11* blocks tail regeneration in *Xenopus* [17].

Salamanders are the only tetrapods that can fully regenerate tails as adults. Various pathways required for tail regeneration in *Xenopus* tadpoles are also necessary for salamander tail regeneration. Pharmacological inhibition of Wnt, Fgf and Tgf-β pathways [13,18], ion channels [19], or Shh signaling [18,20] block salamander tail regeneration. Pharmacological inhibition of Wnt signaling alters the expression profiles of genes associated with multiples pathways implicated in tail regeneration in *Xenopus*, such as Egf, Notch, and others [13]. Axolotl regenerating spinal cords also display upregulation of genes linked to immune and inflammatory response, ECM remodeling, and genes encoding morphogens such as Shh, Bmps, Wnts and Fgfs [21]. Finally, specific genes have been associated with the early steps of tail regeneration, as morpholino-mediated knockdown of the axolotl Marcks-like protein (*AxMlp*), an extracellular protein, blocks tail regeneration [22].

The molecular profile of lizard tail regeneration has also been examined in recent years. RNA-seq analysis comparing tail *versus* limb blastema in the common wall lizard (*Podaris muralis*) revealed genes exclusively upregulated in the regenerating tail, which include those coding for growth factors (Wnts, Shh, Bmps, Fgfs) as well as genes within broad categories of ECM, inflammation and immunity, metabolism and cytoskeleton [23]. A similar molecular profile was observed in the regenerating tail proteome of the northern house gecko, *Hemidactylus flaviviridis*, which showed enrichment for peptides involved in immune response, ECM remodeling, Fgfs and Bmps [24]. In the green anole lizard (*Anolis carolinensis*), the distal tip of regenerating tail at 25 days post amputation (dpa) is enriched for genes that fall in gene ontology (GO) categories of wound response, immune response, hormonal regulation and embryonic morphogenesis [25]. Furthermore, studies on tail regeneration in geckos have uncovered roles for ROS signaling [26], Tgf-β signaling [27] and Fgf signaling [28].

Gene expression data from tail regeneration in various tetrapod species is beginning to shed light into a general molecular profile of successful tail regeneration. Whether tetrapods share an ancestral tail regeneration program is yet unknown. Recent fossil evidence of caudal autotomy has been described for ancient reptiles dating from

the early Permian [29] and evidence of tail regeneration has been uncovered in lepospondyl microsaurs on the amniote stem group [30]. These findings suggest that tail regeneration may be an ancestral feature of tetrapods. Lungfish are the extant sister group of tetrapods and, like salamanders, are capable of regenerating complete tails as adults (figure 1). This remarkable capacity was first reported nearly 150 years ago [31] and documented in the laboratory half a century ago [32]. Since then, despite its potential as a model research system for regenerative medicine and important position as a phylogenetic outgroup to tetrapods, no reports on lungfish tail regeneration have followed.

Here we show that lungfish tail regeneration involves the formation of a proliferative blastemal cell population and the restoration of original tail structures including muscle, spinal cord and tail fin skeleton. As seen in salamanders and frogs, lungfish tail skeletal and spinal cord tissues regenerate neurons and display dorsoventral patterning similar to that observed in the original tail. Further, we show that Shh signaling, fundamental for amphibian tail dorsoventral patterning and regeneration, is also required for lungfish tail regeneration. RNA-seq analysis of lungfish uninjured and regenerating tail blastema revealed that lungfish deploy a blastema genetic program similar to that reported in tetrapods, showing upregulation of genes involved in Fgf, Wnt, Tgf-β, Shh, Bmp, Egf, signaling pathways as well as in inflammatory response, ECM remodeling and stem cell maintenance. Interestingly, lungfish tail blastema showed marked upregulation of transposon-derived genes and components of post-transcriptional RNA processing. Our findings suggest that tetrapods and lungfish share a core genetic program for tail regeneration and establish lungfishes as a novel research system for tail regeneration.

2. Methods

(a) Animals and surgical procedures

A total of 30 juvenile *Protopterus annectens* specimens, ranging from 15 to 35 cm in length were maintained in individual tanks in a recirculating freshwater system at 24-28°C with aeration. Lungfish were anesthetized in 0.1% MS-222 (Sigma). The point of amputation was determined as follows: the distance from the snout to the attachment

point of the pelvic fin was measured in centimeters and divided by 1,75. The resulting number was then used as the distance from the attachment point of the pelvic fin to the point of tail amputation. Samples were labeled as uninjured and blastema and stored in RNAlater (Sigma-Aldrich) for RNA extraction or embedded in Tissue Tek O.C.T compound (Sakura Finetek), frozen in dry ice for histology and *in situ* hybridization (stored at -80 ^{o}C) or fixed in 4% paraformaldehyde (PFA) overnight at 4 ^{o}C for cell proliferation and section-immunohistochemistry.

(b) External morphology and histology of tail fin regeneration

Tails were photographed at 1 dpa and weekly to document the changes on external morphology during regeneration. For histology, frozen tissues were allowed to equilibrate to cryostat temperature (-20 ^{o}C) for 30 min and then 20 μm longitudinal sections were obtained on ColorFrost Plus microscope slides (Thermo Fisher Scientific). Sections were fixed in 3% paraformaldehyde for 5 min, rinsed twice in 0.01M PBS, and dehydrated in graded ethanol series (70, 95 and 100%) for 2 min each. Slides were stored in -80 ^{o}C. Sections were stained with haematoxylin (Sigma-Aldrich) and eosin (Sigma-Aldrich) and imaged on a SMZ1000 stereoscope (Nikon). Tails were cleared and stained as described previously [33].

(c) Cell proliferation assay and immunohistochemistry

5-bromo-2-deoxyuridine (BrdU) was injected intraperitoneally into anesthetized lungfish (80 mg per kg of body weight), 24 h before tail tissue collection to observe cell proliferation. Overnight-fixed tissues were transferred to 30% sucrose, flash frozen in OCT blocks; longitudinal and transverse sections (20uM thickness) were obtained as described in the preceding section. Sections were permeabilized in 2N HCl solution at 37 ^{o}C for 15 minutes, followed by washes in 0.1M borate buffer and in PBS tween (0.1% tween in 0.01M PBS). A treatment with 0.1% trypsin at 37 ^{o}C for 15 minutes was performed and followed by a wash in PBS tween. Unspecific labelling was blocked with 5% normal goat serum diluted in 0.01 M PBS with 0.5% Tween and 1% bovine serum albumin for 1 h at room temperature. Sections were then incubated with mouse anti-BrdU primary antibody (1:200, Sigma-Aldrich, cat. number B8434) in 0.01M PBS with

1% bovine serum albumin and 0.5% Tween overnight at 4 °C. On the next day, sections were incubated with the Alexa 488 conjugated goat anti-mouse secondary antibody (1:400, ThermoFisher Scientific, cat. number A-11001) for 2 h at room temperature and slides were mounted and counterstained with Fluoromount with DAPI (ThermoFisher Scientific). For immunohistochemistry, βIII-Tubulin immunostaining (mouse monoclonal, 1:500, Sigma-Aldrich, cat. number ab78078) was performed using the same procedure, followed by incubation with the Alexa Fluor 594 goat anti-mouse secondary antibody (1:1000, ThermoFisher Scientific, cat. number A-11005).

(d) Library preparation and Illumina sequencing

For transcriptome sequencing, total RNA from tail tissues was extracted using TRIzol Reagent (Life Technologies). A two-step protocol, with the RNeasy Mini Kit (Qiagen) and DNaseI treatment (Qiagen), was used to purify the RNAs and remove residual DNA. mRNA sequencing libraries were constructed using NEXTflex® Rapid Directional qRNA-Seq™ Kit (Illumina). Lungfish reference transcriptomes and transcript abundance estimation were obtained from the sequencing of three biological replicates of blastemas at 14 dpa and three biological replicates of uninjured tail tissue, performed on an Illumina 2500 HiSeq platform with 150 bp paired-end reads. Reads from 6 additional runs of other regenerating tail stages (1 dpa and 21 dpa) showed high variation in read counts among replicas and were used only to help produce a comprehensive *de novo* lungfish reference transcriptome assembly.

(e) Bioinformatic analysis

The West African lungfish reference transcriptome was assembled *de novo* using Trinity with default parameters [34]. For each run, all read datasets were mapped to reference transcriptomes using CLC genomic workbench with default parameters (Qiagen). Expression data per transcript were summed by human homolog gene cluster using a bash script (HHGC). As previously described [35], the HHGCs were defined by grouping transcripts with an e-value of 10^{-3} when compared by BLASTx against Human NCBI RefSeq database (11/2016). For each HHGC, the expression was calculated in TPM, and the comparison was based on t-test considering two conditions (Uninjured tail

and tail blastema) with three independent biological replicates. Similarity matrix between runs was calculated using TPM values for each HHGC, using Spearman rank correlation in Morpheus software (https://software.broadinstitute.org/morpheus). A list of enriched GO terms was produced using WebGestalt 2019 [36]. Differentially expressed genes with false discovery rate (FDR) adjusted *P* values smaller than 0.05 were ranked from highest to lowest fold change values, and the corresponding ranked list of gene symbols was used for GO enrichment analysis. GO enriched categories were significant when *P* values were 0.05 or less. Venn diagrams were generated using BioVenn [37].

(f) Statistical analysis

For each transcript and HHGC, mean TPM value between uninjured tails and tail blastemas were compared with a two-tailed t-test. A transcript or HHGC was deemed differentially expressed when its fold change is superior to 2 or inferior to –2 and FDR adjusted *P* value is inferior to 0.05. GO enrichment and pathway over-representation analyses were performed using WebGestalt 2019 web-based tool.

(g) *In situ* hybridization

Regenerating tails at 14 dpa (n = 3) were sampled and embedded and frozen in OCT (TissueTek). Frozen sections (20 µm) were obtained on a Leica CM1850 UV cryostat and positioned on Color Frost Plus microscope slides (Thermo Fisher Scientific). Sections were fixed as previously described [35] and stored at -80°C for haematoxylin and eosin staining, or *in situ* hybridization. Riboprobe templates for *in situ* hybridization were produced by a two-round PCR strategy: first-round PCR produced specific fragments (400-500 bp) of selected genes and in a second PCR a T7 promoter sequence was included at either 5'or 3'end of the fragments for generation of templates for sense or anti-sense probes. The primers used were: *Col12a1* forward: 5'-GGCCGCGGTTGATGCTCCCATTTGGTTAG-3' and reverse: 5'-CCCGGGGCGAAACCCAGGAACAAGAGGTC-3', *Hmcn2* forward 5'-GGCCGCGGTTGAGCAGAACCAGCTTCATT-3' and reverse 5'-CCCGGGGCTTAGTGGGGCAGACAATCAAC-3', *Inhbb* forward 5'-GGCCGCGGCCGTGCTTGAACCACTAAAAA and reverse 5'-CCCGGGGCTTTGCAGAGACAGATGACGTG-3',

3’-T7 universal 5’-AGGGATCCTAATACGACTCACTATAGGGCCCGGGGC-3’, 5’-T7 universal 5’-GAGAATTCTAATACGACTCACTATAGGGCCGCGG-3’ (linker sequences for annealing of the 3' or 5’_T7 universal primer are underlined). The riboprobes were synthesized using T7 RNA polymerase (Roche) and DIG-labelling mix (Roche). *In situ* hybridization was performed as previously described [35], using 375 ng of DIG-labelled riboprobe per slide. Slides were photographed on Nikon Eclipse 80i microscope and the images were processed on the NIS-Element D4.10.1 program.

3. Results

(a) Establishment of a proliferative blastemal cell population during lungfish tail regeneration

We evaluated regeneration in juvenile West African upon tail amputation. At 7 dpa, a wound epithelium covers the amputation site and tail outgrowth is negligible. At 14 dpa, tail outgrowth is visible at the level of the midline. At 21 dpa, the regenerating tail reaches 1 cm in length and the skin is highly pigmented. In the following weeks the tail continues to extend and by 56 dpa it has nearly reached its length prior to amputation (figure 2a). Histological sections showed that at 1 dpa, a 1 to 2 cell layer wound epithelium covers the amputation site. At 7 dpa, the wound epithelium thickens, and a mass of mesenchymal cells accumulate subjacent to it, posterior to the severed postcaudal cartilage. At 21 dpa, the regenerated postcaudal cartilage bar and ependymal tube are visible (figure 2b). At our latest experimental end point (60 dpa), the postcaudal cartilage was undergoing segmentation and cartilaginous neural and haemal arches and spines were visible (figure 2c).

Next, we assessed cell proliferation during the first 3 weeks of tail regeneration. BrdU staining revealed that at 1 dpa, proliferating cells are mostly found in the wound epithelium. At 14 dpa, proliferating cells are found distal to the amputation plane in the region of the presumptive tail blastema. At 21 dpa, cell proliferation is observed posterior to the amputation site across the entire regenerated tail (figure 2d). Our results indicate that lungfish tail regeneration proceeds via morphological events comparable to those involved in salamander tail regeneration, with the establishment of a wound epithelium, which thickens to form an AEC, the formation of a mass of proliferating blastemal cells, and restoration of original tail tissue organization.

(b) Lungfish regeneration restores spinal cord neurons original dorsoventral patterning of original tail

A fundamental difference between tail regeneration in salamanders and lizards is the loss of dorsoventral patterning and lack of spinal cord neurons in the regenerated tail in lizards [8], resulting in a radially symmetric structure consisting of an ependymal tube enveloped in a cartilage bar. Conversely, salamanders restore the dorsoventral pattern of an ependymal tube that gives rise to the spinal cord, located dorsally relative to the notochord. Our transversal sections of lungfish uninjured and regenerating tails show that at 21 dpa, the regenerating ependymal tube is dorsally positioned relative to the notochord (figure 3a). Newly formed blood vessel ventral to the notochord and regenerating muscle laterally are also visible. In addition, immunostaining for β-tubulin revealed new spinal cord neurons forming as early as 28 dpa (figure 3b).

Shh is a key signaling molecule expressed in the floor plate of the ependymal tube in salamanders and in the regenerating notochord in frog tadpoles, which is necessary for tail regeneration in both species [20,38]. Likewise, pharmacological inhibition of Shh signaling via administration of the cyclopamine blocked lungfish tail regeneration (figure 3c).

Together, these results show the reestablishment of the dorsoventral tail patterning and spinal cord neurons, as well as the requirement of Shh during lungfish tail regeneration.

(c) Differential gene expression analysis of tail blastema versus uninjured tail

To identify genes differentially expressed in the tail blastema relative to uninjured tail tissue, we produced RNA-seq libraries from uninjured tail tissues and regenerating tails at 14 dpa, a stage when proliferative blastemal cells were identified. Principal component analysis showed two distinct clusters representing uninjured and blastemal tail samples, and Spearman correlation coefficients among biological replicas were greater than 0.78, corroborating the reproducibility of RNA-seq runs (electronic supplementary material, figure S1). DGE analysis of the lungfish uninjured *versus* 14 dpa tail revealed 1072 upregulated genes (FC > 2, FDR < 0.05). Among the

upregulated gene dataset, we identified components of the various pathways previously associated to tail regeneration in tetrapods, including the Wnt pathway (*Wnt5a*, *Wnt5b*, *Axin2*, *Gsk3b*, *Ctnnb1*), Fgf (*Fgfr1*), Bmp (*Bmp1*, *Bmp4*, *Smad2*), Shh (*Ptch2*, *Gli2*), Notch (*Notch2*), Tgf-β (*Inhbb*, *Tgfbi*), Egf (*Vegfa*, *Megf6*, *Megf10*), ECM components and remodelers (*C1qtnf3*, *Col11a1*, *Fbn2*, *Mmp11*, *Adamts14*), Hyaluron pathway (*Hyal2*, *Has2*), immune and inflammatory response (*Il11*, *Mdk*, *Nfkbiz*) and stem cell maintenance (*Sox4*, *Sall4*, *Chd2*) (figure 4a). Some components of pathways involved in tetrapod tail regeneration were moderately upregulated (FC > 1.5, FDR < 0.05), including *Bmp2*, *Hif1a*, *Tgfb1*, *Tgfb2*, *Myc* and *Hdac7* (electronic supplementary material, table S1).

GO enrichment analysis identified categories such as mitotic cell cycle phase transition, extracellular structure organization and regulation of RNA metabolic processing (electronic supplementary material, figure S3). Likewise, pathway enrichment analysis revealed that the lungfish tail blastema shows high overrepresentation of genes in collagen biosynthesis and modifying enzymes, pathways related to mitotic cell division and ECM organization, consistent with the major events occurring at the blastema stage of tail regeneration (figure 4b). *In situ* hybridization in 14 dpa lungfish tails of 2 genes encoding ECM components (*Col12a1* and *Hmcn2*) and a gene encoding a member of the TGF-β family of cytokines (*Inhbb*) showed similar expression pattern, with anti-sense probe signal predominantly detected in the AEC (figure 4c), and no specific signal observed in sense-control probes (electronic supplementary material, Fig S1)

Il11, a gene highly upregulated and required for tail regeneration in *Xenopus* tadpole, was also among the most highly upregulated in our dataset (FC = 34.95) (electronic supplementary material, table S1). The lungfish ortholog of *AxMlp*, required for axolotl tail regeneration, showed moderate upregulation in lungfish tail and FDR value just above our cutoff (FC = 1.41, FDR = 0.06). The lungfish ortholog of *Vwde*, a gene highly expressed and required for axolotl limb regeneration and associated with successful *Xenopus* tadpole tail regeneration, was highly upregulated in lungfish, however with an FDR value above our cutoff (FC > 57.45, FDR = 0.08) [39]. Furthermore, *Angptl2* and *Egfl6*, identified recently as tail-specific AEC factors, are both upregulated in our lungfish tail blastema dataset (FC = 3.67 and 7.62, respectively) (electronic supplementary material, table S1). Finally, we examined a set of 10 genes

recently reported to be expressed preferentially in the tail blastema of *Xenopus* tadpoles relative to embryonic tail bud [14]. We found that of 3 out of 10 genes were *Xenopus*-specific genes, and 1 gene was not contained in our annotated lungfish reference transcriptome. Of the 6 remaining, 3 were upregulated in our lungfish dataset, namely *Il11*, *Cse1l* (FC = 2.77), *L1td1* (FC = 10.52), and 1 gene, *cd200*, was upregulated with an FDR value above our 0.05 cutoff (FC = 5.94, FDR = 0.09).

Taken together, the transcriptomic profile of lungfish tail blastema relative to uninjured tail tissue showed remarkable resemblance to the expression profiles reported for tetrapods, particularly those of salamander and *Xenopus*.

(d) Genes enriched in lungfish tail blastema versus pectoral fin blastema

Next we sought to compare our tail blastema dataset to previously published data on South American lungfish pectoral fin regeneration [33]. Comparison of 1072 upregulated genes in tail blastema to the 843 genes upregulated in pectoral fin blastema revealed an overlap of 225 genes. Pathway enrichment analysis showed that this overlapping dataset included genes involved in collagen metabolism, ECM organization and mitotic cell cycle, all of which represent categories commonly found in regenerating tissues (figure 4d). Interestingly, genes exclusively enriched in tail blastema relative to pectoral fin blastema were involved in pathways related to post-transcriptional RNA processing, including transport of mature transcript to cytoplasm, processing of capped intron-containing pre-mRNA, mRNA splicing and metabolism of RNA (figure 4d). These results suggest that post-transcriptional RNA processing may play a more significant role in tail versus pectoral fin regeneration.

(e) Upregulated transposon-derived genes in the lungfish tail blastema

Transposable elements and genes derived from retrotransposon have been previously implicated in regeneration in species as distantly related as sea cucumbers [40] and axolotls [41]. Two genes derived from transposons were found enriched in the tail blastema of *Xenopus* tadpoles [14]: *oax*, which is a *Xenopus*-specific repetitive element [42], and *l1td1-like*, which is similar to *l1td1*, a mammal-specific L1 transposon-derived gene previously identified as a marker of human embryonic stem cells [43]. We searched our lungfish dataset for transposon-derived genes and found 16 genes,

including an ortholog of *l1td1-like*, the harbinger transposon-derived gene *Harbi1*, and the piggyBac transposon-derived gene *Pgbd4* (figure 4e). Our findings indicate that transposon-derived genes may consist in important components of the regenerative response induced upon tail injury in the lungfish.

4. Discussion

Our results reveal extensive morphological and molecular similarities between lungfish and salamander tail regeneration. Lungfish tail amputation is followed by the formation of a highly proliferative wound epithelium at 1 dpa, which thickened to form an AEC. As in salamanders, a blastemal cell population accumulates distally and medially relative to the amputation site. Regenerated lungfish tails restore the original pattern of skeleton and spinal cord along the dorsoventral axis, forming new spinal cord neurons at 28 dpa. While this condition is similar to that of salamanders and frogs, tail regeneration in lizards represents a derived character state, where tail structures fail to recapitulate the tissue organization and composition seen in the intact, original tail.

At 14 dpa, blastemal cells are proliferating, and RNA-seq analysis revealed marked upregulation of signaling pathways previously linked to tail regeneration in salamanders and frogs, such as Wnt, Fgf, Shh, Notch, Tgf-β and Egf, as well as ECM and inflammatory response. Comparison of regenerating tail transcriptomes to those of lungfish pectoral fin blastemata suggests that RNA processing plays a more significant role in tail regeneration. Post-transcriptional RNA processing provides means of generating RNA variants and may be linked to the greater tissue complexity of the regenerating tail compared to the pectoral fin.

Transposable elements and genes derived from retrotransposon are widely upregulated in the lungfish tail blastema. Previous study on the West African lungfish transcriptome identified large diversity of active transposable elements, which may have played a role in the expansion of the lungfish genome [44]. This was also shown to be the case in the Iberian ribbed newt *Pleurodeles waltl* [45], the axolotl [46] and the coelacanth [47, 48, 49]. In *P. waltl*, the transposable element family that expanded the most was the Harbinger transposon family, which has given rise to two vertebrate protein-coding genes, *Harbi1* [50] and *Naif1* [51]. In the Coelacanth, Harbinger elements account to 4% of the genome and were shown to possess transcriptional and

enhancer activities in vivo [48]. In our dataset, the lungfish *Harbi1* orthologue was among the highest differentially expressed transposon-derived transcripts. Interestingly, *Harbi1* is not upregulated during regeneration of the lungfish pectoral fin [33] or *P. waltl* limb [44], suggesting that it may play a role in cell populations specific to the tail. Future studies aimed at functionally evaluating the roles of transposon-derived genes such as *Harbi1* may help uncover context-specific roles in development and in regeneration.

Lungfish are an emerging model system for studies in tail and spinal cord regeneration that can inform both the history and mechanisms of regeneration. Our results support the hypothesis that tail regeneration in salamanders, frogs, and lungfish share a common evolutionary history. A more detailed investigation of initial stages of regeneration leading to the blastemal formation in both lungfishes and salamanders may uncover a shared molecular and cellular program of regeneration.

Ethics

All experimental procedures and animal care were conducted in accordance to the Ethics Committee for Animal Research at the Universidade Federal do Pará, under the approved protocol number 037-2015.

Authors' contributions

KMV, JFS, NHS, NBF and IS designed the research; KMV, JFS, ACD, WRBM, CASN, EMS, CNSM and IS performed regeneration assays; KMV, WBM, CNSM, GS and EMS, performed cell proliferation assay, immunohistochemistry and pharmacological experiments; LNP, KMV, JFS, ACD, SD, NBF and IS analyzed transcriptome data; IS supervised this work and wrote the manuscript with input from all authors. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Data accessibility

Sequence data that support the findings of this study have been deposited in GenBank with the following BioProject accession numbers: PRJNA491932, with accession number for uninjunred (SRR7880018, SRR7880019 e SRR7880016), 14 dpa blastema

(SRR7880020, SRR7880021 e SRR7880024) and additional libraries of tail blastemas at 1 dpa (SRR7880017, SRR7880022 e SRR7880023) and 21 dpa (SRR7880025, SRR7880026 e SRR7880027). The authors declare that all other relevant data supporting the findings of this study are available on request.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by funding from CNPq Universal Program [grant number 403248/2016-7], CAPES/Alexander von Humboldt Foundation fellowship, CAPES/DAAD PROBRAL [grant number 88881.198758/2018-01] to I.S. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Acknowledgments
We would like to thank Patricia Schneider for insightful comments on the manuscript.

References

1. Tanaka EM. 2003 Cell differentiation and cell fate during urodele tail and limb regeneration. *Current Opinion in Genetics & Development* **13**, 497–501. (doi:10.1016/j.gde.2003.08.003)
2. Gargioli C. 2004 Cell lineage tracing during Xenopus tail regeneration. *Development* **131**, 2669–2679. (doi:10.1242/dev.01155)
3. Tseng A-S, Levin M. 2008 Tail Regeneration in *Xenopus laevis* as a Model for Understanding Tissue Repair. *J Dent Res* **87**, 806–816. (doi:10.1177/154405910808700909)
4. Mochii M, Taniguchi Y, Shikata I. 2007 Tail regeneration in the Xenopus tadpole: Tail regeneration in Xenopus. *Development, Growth & Differentiation* **49**, 155–161. (doi:10.1111/j.1440-169X.2007.00912.x)

5. Beck CW, Izpisúa Belmonte JC, Christen B. 2009 Beyond early development: *Xenopus* as an emerging model for the study of regenerative mechanisms. *Dev. Dyn.* **238**, 1226–1248. (doi:10.1002/dvdy.21890)
6. Gilbert EAB, Delorme SL, Vickaryous MK. 2015 The regeneration blastema of lizards: an amniote model for the study of appendage replacement: The Lizard Blastema. *Regeneration* **2**, 45–53. (doi:10.1002/reg2.31)
7. McLean KE, Vickaryous MK. 2011 A novel amniote model of epimorphic regeneration: the leopard gecko, Eublepharis macularius. *BMC Dev Biol* **11**, 50. (doi:10.1186/1471-213X-11-50)
8. Sun AX, Londono R, Hudnall ML, Tuan RS, Lozito TP. 2018 Differences in neural stem cell identity and differentiation capacity drive divergent regenerative outcomes in lizards and salamanders. *Proc Natl Acad Sci USA* **115**, E8256–E8265. (doi:10.1073/pnas.1803780115)
9. Ho DM, Whitman M. 2008 TGF-β signaling is required for multiple processes during Xenopus tail regeneration. *Developmental Biology* **315**, 203–216. (doi:10.1016/j.ydbio.2007.12.031)
10. Li J, Zhang S, Amaya E. 2016 The cellular and molecular mechanisms of tissue repair and regeneration as revealed by studies in *Xenopus*. *Regeneration* **3**, 198–208. (doi:10.1002/reg2.69)
11. Okumura A, Hayashi T, Ebisawa M, Yoshimura M, Sasagawa Y, Nikaido I, Umesono Y, Mochii M. 2019 Cell type-specific transcriptome analysis unveils secreted signaling molecule genes expressed in apical epithelial cap during appendage regeneration. *Develop. Growth Differ.* **61**, 447–456. (doi:10.1111/dgd.12635)
12. Aztekin C, Hiscock TW, Marioni JC, Gurdon JB, Simons BD, Jullien J. 2019 Identification of a regeneration-organizing cell in the *Xenopus* tail. *Science* **364**, 653–658. (doi:10.1126/science.aav9996)
13. Ponomareva LV, Athippozhy A, Thorson JS, Voss SR. 2015 Using Ambystoma mexicanum (Mexican axolotl) embryos, chemical genetics, and microarray analysis to identify signaling pathways associated with tissue regeneration. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **178**, 128–135. (doi:10.1016/j.cbpc.2015.06.004)
14. Tsujioka H, Kunieda T, Katou Y, Shirahige K, Kubo T. 2015 Unique Gene Expression Profile of the Proliferating Xenopus Tadpole Tail Blastema Cells Deciphered

by RNA-Sequencing Analysis. *PLoS ONE* **10**, e0111655. (doi:10.1371/journal.pone.0111655)
15. Tseng A-S, Carneiro K, Lemire JM, Levin M. 2011 HDAC Activity Is Required during Xenopus Tail Regeneration. *PLoS ONE* **6**, e26382. (doi:10.1371/journal.pone.0026382)
16. Contreras EG, Gaete M, Sanchez N, Carrasco H, Larrain J. 2009 Early requirement of Hyaluronan for tail regeneration in Xenopus tadpoles. *Development* **136**, 2987–2996. (doi:10.1242/dev.035501)
17. Tsujioka H, Kunieda T, Katou Y, Shirahige K, Fukazawa T, Kubo T. 2017 interleukin-11 induces and maintains progenitors of different cell lineages during Xenopus tadpole tail regeneration. *Nat Commun* **8**, 495. (doi:10.1038/s41467-017-00594-5)
18. Singh BN, Weaver CV, Garry MG, Garry DJ. 2018 Hedgehog and Wnt Signaling Pathways Regulate Tail Regeneration. *Stem Cells and Development* **27**, 1426–1437. (doi:10.1089/scd.2018.0049)
19. Franklin BM, Voss SR, Osborn JL. 2017 Ion channel signaling influences cellular proliferation and phagocyte activity during axolotl tail regeneration. *Mechanisms of Development* **146**, 42–54. (doi:10.1016/j.mod.2017.06.001)
20. Schnapp E. 2005 Hedgehog signaling controls dorsoventral patterning, blastema cell proliferation and cartilage induction during axolotl tail regeneration. *Development* **132**, 3243–3253. (doi:10.1242/dev.01906)
21. Monaghan JR, Walker JA, Page RB, Putta S, Beachy CK, Voss SR. 2006 Early gene expression during natural spinal cord regeneration in the salamander Ambystoma mexicanum: Gene expression and spinal cord regeneration. *Journal of Neurochemistry* **101**, 27–40. (doi:10.1111/j.1471-4159.2006.04344.x)
22. Sugiura T, Wang H, Barsacchi R, Simon A, Tanaka EM. 2016 MARCKS-like protein is an initiating molecule in axolotl appendage regeneration. *Nature* **531**, 237–240. (doi:10.1038/nature16974)
23. Vitulo N, Dalla Valle L, Skobo T, Valle G, Alibardi L. 2017 Transcriptome analysis of the regenerating tail vs. the scarring limb in lizard reveals pathways leading to successful vs. unsuccessful organ regeneration in amniotes: Tail and Limb Transcriptome in Regenerating Lizard. *Dev. Dyn.* **246**, 116–134. (doi:10.1002/dvdy.24474)

24. Murawala H, Ranadive I, Patel S, Desai I, Balakrishnan S. 2018 Protein expression pattern and analysis of differentially expressed peptides during various stages of tail regeneration in Hemidactylus flaviviridis. *Mechanisms of Development* **150**, 1–9. (doi:10.1016/j.mod.2018.02.001)
25. Hutchins ED *et al.* 2014 Transcriptomic Analysis of Tail Regeneration in the Lizard Anolis carolinensis Reveals Activation of Conserved Vertebrate Developmental and Repair Mechanisms. *PLoS ONE* **9**, e105004. (doi:10.1371/journal.pone.0105004)
26. Zhang Q *et al.* 2016 Reactive oxygen species generated from skeletal muscles are required for gecko tail regeneration. *Sci Rep* **6**, 20752. (doi:10.1038/srep20752)
27. Gilbert RWD, Vickaryous MK, Viloria-Petit AM. 2013 Characterization of TGFβ signaling during tail regeneration in the leopard Gecko (*Eublepharis macularius*): Tgfβ Expression During Tissue Regeneration. *Dev. Dyn.* **242**, 886–896. (doi:10.1002/dvdy.23977)
28. Pillai A, Patel S, Ranadive I, Desai I, Balakrishnan S. 2019 Fibroblast growth factor-2 signaling modulates matrix reorganization and cell cycle turnover rate in the regenerating tail of Hemidactylus flaviviridis. *Acta Histochemica* , 151464. (doi:10.1016/j.acthis.2019.151464)
29. LeBlanc ARH, MacDougall MJ, Haridy Y, Scott D, Reisz RR. 2018 Caudal autotomy as anti-predatory behaviour in Palaeozoic reptiles. *Sci Rep* **8**, 3328. (doi:10.1038/s41598-018-21526-3)
30. Fröbisch NB, Bickelmann C, Olori JC, Witzmann F. 2015 Deep-time evolution of regeneration and preaxial polarity in tetrapod limb development. *Nature* **527**, 231–234. (doi:10.1038/nature15397)
31. Traquair RH. 1871 On the restoration of the tail in Protopterus annectens. *Rep. Brit. Ass. Adv. Sci.* , 143.
32. Conant EB. 1970 Regeneration in the African lungfish,Protopterus. I. Gross aspects. *J. Exp. Zool.* **174**, 15–31. (doi:10.1002/jez.1401740103)
33. Nogueira AF *et al.* 2016 Tetrapod limb and sarcopterygian fin regeneration share a core genetic programme. *Nat Commun* **7**, 13364. (doi:10.1038/ncomms13364)
34. Grabherr MG *et al.* 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644–652. (doi:10.1038/nbt.1883)
35. Darnet S *et al.* 2019 Deep evolutionary origin of limb and fin regeneration. *Proc Natl Acad Sci USA* **116**, 15106–15115. (doi:10.1073/pnas.1900475116)

36. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019 WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research* **47**, W199–W205. (doi:10.1093/nar/gkz401)
37. Hulsen T, de Vlieg J, Alkema W. 2008 BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488. (doi:10.1186/1471-2164-9-488)
38. Taniguchi Y, Watanabe K, Mochii M. 2014 Notochord-derived hedgehog is essential for tail regeneration in Xenopus tadpole. *BMC Dev Biol* **14**, 27. (doi:10.1186/1471-213X-14-27)
39. Leigh ND *et al.* 2019 von Willebrand Factor D and EGF Domains is an evolutionarily conserved and required feature of blastemas capable of multi-tissue appendage regeneration. (doi:10.1101/842948)
40. Mashanov VS, Zueva OR, García-Arrarás JE. 2012 Posttraumatic regeneration involves differential expression of long terminal repeat (LTR) retrotransposons. *Dev. Dyn.* **241**, 1625–1636. (doi:10.1002/dvdy.23844)
41. Zhu W *et al.* 2012 Retrotransposon long interspersed nucleotide element-1 (LINE-1) is activated during salamander limb regeneration. *Develop. Growth Differ.* **54**, 673–685. (doi:10.1111/j.1440-169X.2012.01368.x)
42. Whitford KL, Oakes JA, Scholnick J, Saha MS. 2000 Tissue-specific developmental expression of OAX, a Xenopus repetitive element. *Mechanisms of Development* **94**, 209–212. (doi:10.1016/S0925-4773(00)00307-5)
43. Wong RC-B, Ibrahim A, Fong H, Thompson N, Lock LF, Donovan PJ. 2011 L1TD1 Is a Marker for Undifferentiated Human Embryonic Stem Cells. *PLoS ONE* **6**, e19355. (doi:10.1371/journal.pone.0019355)
44. Biscotti MA, Gerdol M, Canapa A, Forconi M, Olmo E, Pallavicini A, Barucca M, Schartl M. 2016 The Lungfish Transcriptome: A Glimpse into Molecular Evolution Events at the Transition from Water to Land. *Sci Rep* **6**, 21571. (doi:10.1038/srep21571)
45. Elewa A *et al* 2017 Reading and editing the Pleurodeles waltl genome reveals novel features of tetrapod regeneration. *Nat Commun* **8**, 2286. (doi:10.1038/s41467-017-01964-9)
46. Nowoshilow S *et al.* 2018 The axolotl genome and the evolution of key tissue formation regulators. *Nature* **554**, 50–55. (doi:10.1038/nature25458)

47. Nikaido M *et al.* 2013 Coelacanth genomes reveal signatures for evolutionary transition from water to land. *Genome Research* **23**, 1740–1748. (doi:10.1101/gr.158105.113)
48. Smith JJ, Sumiyama K, Amemiya CT. 2012 A Living Fossil in the Genome of a Living Fossil: Harbinger Transposons in the Coelacanth Genome. *Molecular Biology and Evolution* **29**, 985–993. (doi:10.1093/molbev/msr267)
49. Amemiya CT *et al.* 2013 The African coelacanth genome provides insights into tetrapod evolution. *Nature* **496**, 311–316. (doi:10.1038/nature12027)
50. Kapitonov VV, Jurka J. 2004 *Harbinger* Transposons and an Ancient HARBI1 Gene Derived from a Transposase. *DNA and Cell Biology* **23**, 311–324. (doi:10.1089/104454904323090949)
51. Sinzelle L, Kapitonov VV, Grzela DP, Jursch T, Jurka J, Izsvak Z, Ivics Z. 2008 Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. *Proceedings of the National Academy of Sciences* **105**, 4715–4720. (doi:10.1073/pnas.0707746105)

Figure legends

Figure 1. Phylogenetic distribution of complete tail regeneration in sarcopterygians. Complete tail regeneration is denoted in green, intermediate capacity in yellow and black denotes absence.

Figure 2. Characterization of lungfish tail regeneration. (a) shows progression of lungfish tail regeneration and extent of growth up to 56 dpa. (b) histological sections of regenerating lungfish tail. (c) regeneration of skeletal elements of the tail at 60 dpa. (d) BrdU staining of proliferative cells during tail regeneration. we, wound epithelium; aec, apical epithelial cap; bl, blastema; et, ependymal tube; ptc.c, postcaudal cartilage; ns, neural spine; na, neural arch; hs, haemal spine; ha, haemal arch. Scale bars of 1 cm (a), 1 mm (b,d), 0.5 cm (c).

Figure 3. Establishment of dorsoventral organization and requirement for Shh signaling during lungfish tail regeneration. (a) histological transversal sections of uninjured and 21 dpa regenerating tail. (b) Effect of DMSO and cyclopamine treatment in tail

regeneration. m, muscle; ptc.c, postcaudal cartilage; et, ependymal tube; bv, blood vessel. Scale bars of 1 mm (a - panoramic views - and b), 0.5 mm (a, enlarged view).

Figure 4. Upregulated genes and overrepresented pathways in lungfish tail blastema relative to uninjured tail. (a) Volcano plot showing differentially expressed genes in lungfish uninjured tail tissue and 14 dpa tail blastema (FDR < 0.05, FC > 2), select lungfish orthologs up or downregulated in the blastema are noted in red. (b) Pathways overrepresented in the tail blastema. (c) *in situ* hybridization of genes upregulated in the blastema (d) Area-proportional Venn diagram showing commonly upregulated genes in lungfish tail and fin datasets, enriched pathways in the shared tail and fin dataset, and pathways enriched exclusively on tail blastema. (e) Transposon-derived genes upregulated in the tail blastema. Scale bars of 1 mm (panoramic views) and 0.25 mm (enlarged view).

Figure S1. Statistics of the lungfish reference transcriptome and similarity among RNA-seq libraries. (a) Statistics or the lungfish reference transcriptome. (b) Busco assessment of transcriptome completeness. (c) Principal component analysis (PCA) of uninjured and 14 dpa lungfish tail libraries. (d) Spearman correlation matrix of uninjured and 14 dpa lungfish tail libraries.

Figure S2. Gene ontology analysis of genes upregulated in the lungfish 14 dpa tail blastema. (a) Biological process. (b) Cellular component. (c) Molecular function.

Figure S3. In situ hybridization of sense-control probes for select genes upregulated in the 14 dpa lungfish tail blastema relative to uninjured tail tissue. Scale bars of 1 mm (panoramic views) and 0.25 mm (enlarged view).

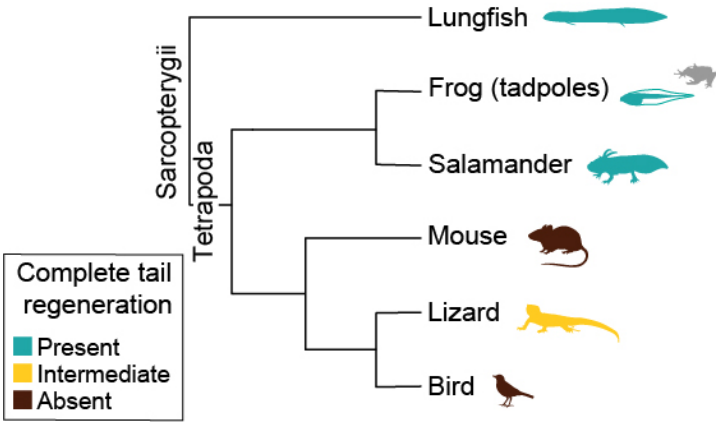


Figure 1. Phylogenetic distribution of complete tail regeneration in sarcopterygians. Complete tail regeneration is denoted in green, intermediate capacity in yellow and black denotes absence.

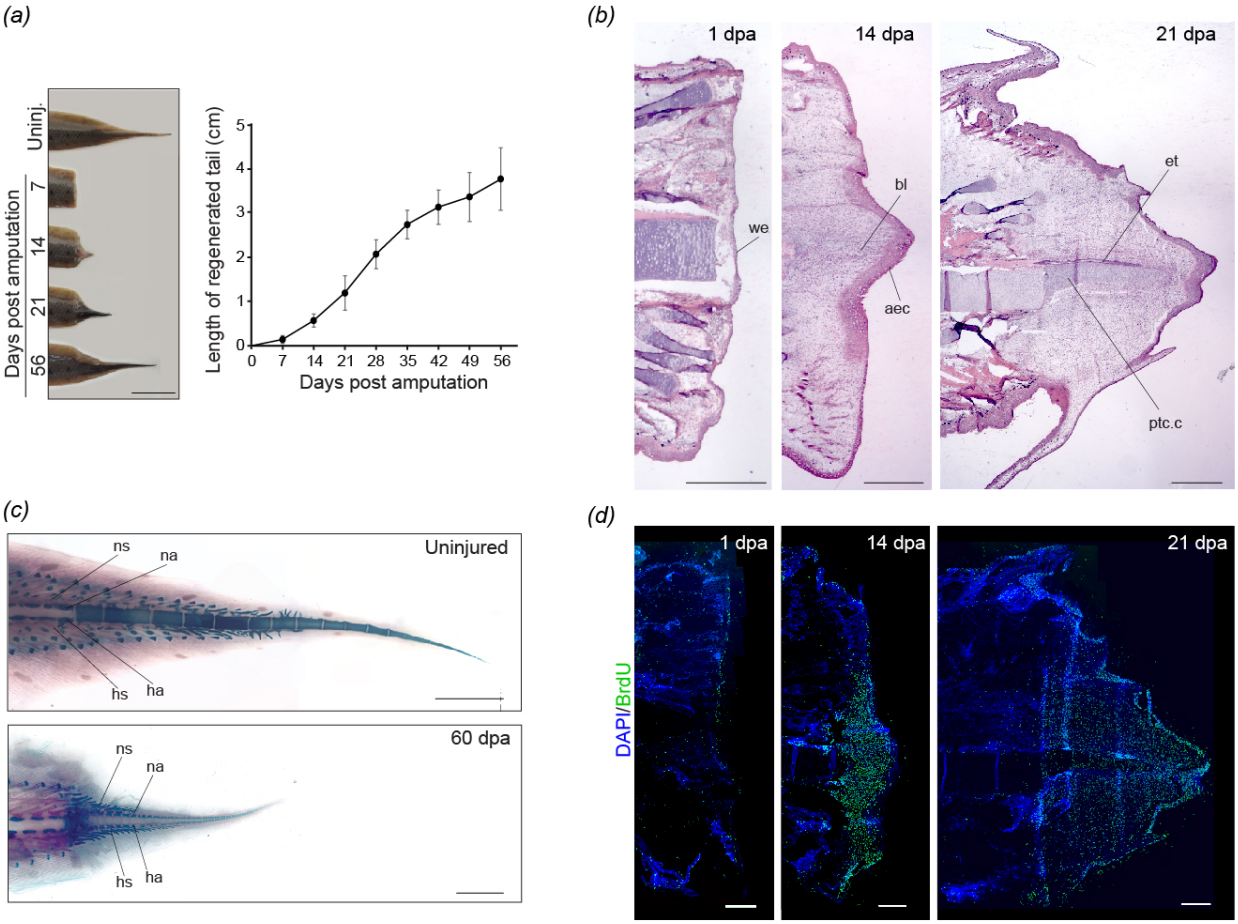


Figure 2. Characterization of lungfish tail regeneration. (a) shows progression of lungfish tail regeneration and extent of growth up to 56 dpa. (b) histological sections of regenerating lungfish tail. (c) regeneration of skeletal elements of the tail at 60 dpa. (d) BrdU staining of proliferative cells during tail regeneration. we, wound epithelium; aec, apical epithelial cap; bl, blastema; et, ependymal tube; ptc.c, postcaudal cartilage; ns, neural spine; na, neural arch; hs, haemal spine; ha, haemal arch. Scale bars of 1 cm (a), 1 mm (b,d), 0.5 cm (c).

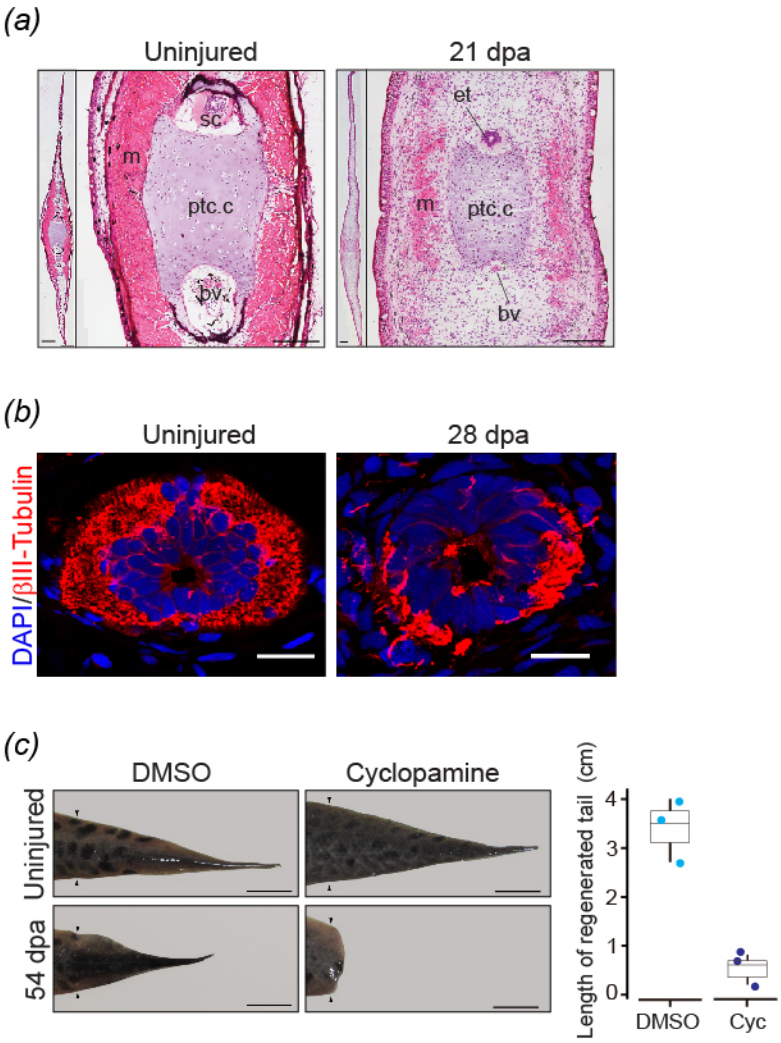
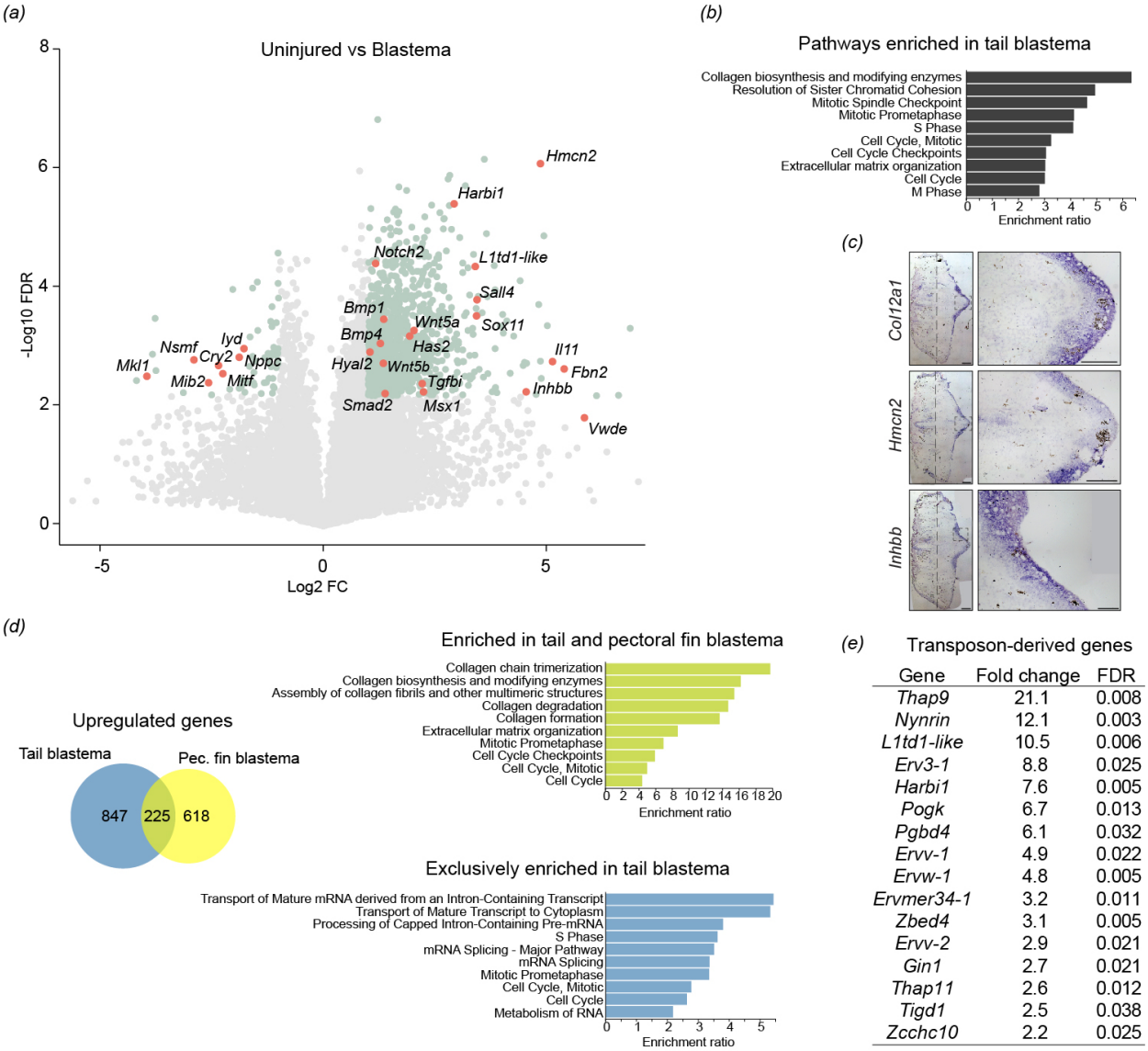


Figure 3. Establishment of dorsoventral organization and requirement for Shh signaling during lungfish tail regeneration. (a) histological transversal sections of uninjured and 21 dpa regenerating tail. (b) immunostaining of DAPI and βIII-tubulin in uninjured and 28 dpa regenerating spinal cord (c) Effect of DMSO and cyclopamine treatment in tail regeneration. m, muscle; ptc.c, postcaudal cartilage; et, ependymal tube; bv, blood vessel. Scale bars of 1 mm (a - panoramic views - and b), 0.5 mm (a, enlarged view).



Gene	Fold change	FDR
Thap9	21.1	0.008
Nynrin	12.1	0.003
L1td1-like	10.5	0.006
Erv3-1	8.8	0.025
Harbi1	7.6	0.005
Pogk	6.7	0.013
Pgbd4	6.1	0.032
Ervv-1	4.9	0.022
Ervw-1	4.8	0.005
Ervmer34-1	3.2	0.011
Zbed4	3.1	0.005
Ervv-2	2.9	0.021
Gin1	2.7	0.021
Thap11	2.6	0.012
Tigd1	2.5	0.038
Zcchc10	2.2	0.025

Figure 4. Upregulated genes and overrepresented pathways in lungfish tail blastema relative to uninjured tail. (a) Volcano plot showing differentially expressed genes in lungfish uninjured tail tissue and 14 dpa tail blastema (FDR < 0.05, FC > 2), select lungfish orthologs up or downregulated in the blastema are noted in red. (b) Pathways overrepresented in the tail blastema. (c) *in situ* hybridization of genes upregulated in the blastema (d) Area-proportional Venn diagram showing commonly upregulated genes in lungfish tail and pectoral fin datasets, enriched pathways in the shared tail and pectoral fin dataset, and pathways enriched exclusively on tail blastema. (e) Transposon-derived genes upregulated in the tail blastema. Scale bars of 1 mm (panoramic views) and 0.25 mm (enlarged view).

(a)

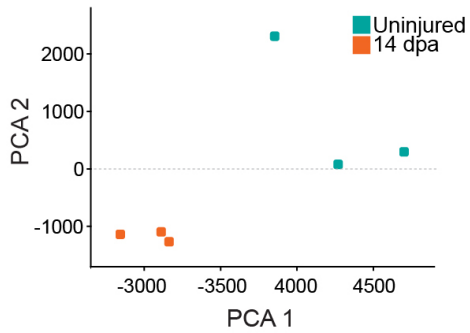
Statistics of *de novo* transcriptome assembly	
Number of reads	279,685,470
Number of mapped reads	221,312,168
Number of contigs:	75,447
Number of contigs >=1kb:	21,637
Number of contigs in N50:	10,769
Min contig length:	201
Max contig length:	17,116
Mean contig length:	997.51
Standard deviation of contig length:	1250.86
Median contig length:	442
N50 contig length:	1,987
Number of bases in all contigs:	75,258,961
Number of bases in contigs >=1kb:	53,191,743
%GC	39.95

(b)

BUSCO quality data from the lungfish tail transcriptome data before and after Evidential gene.

	Before Evidential gene	After Evidential gene
Complete BUSCOs	951 (97.2%)	949 (97.0%)
Complete and single-copy BUSCOs	669 (68.4%)	937 (95.8%)
Complete and duplicated BUSCOs	282 (28.8%)	12 (1.2%)
Fragmented BUSCOs	19 (1.9%)	18 (1.8%)
Missing BUSCOs	8 (0.9%)	11 (1.2%)
Total BUSCO groups searched	978	978

(c)



(d)

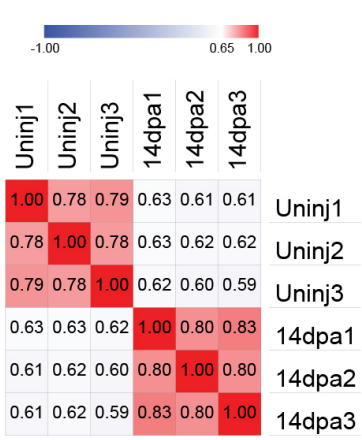


Figure S1. Statistics of the lungfish reference transcriptome and similarity among RNA-seq libraries. (a) Statistics or the lungfish reference transcriptome. (b) Busco assessment of transcriptome completeness. (c) Principal component analysis (PCA) of uninjured and 14 dpa lungfish tail libraries. (d) Spearman correlation matrix of uninjured and 14 dpa lungfish tail libraries.

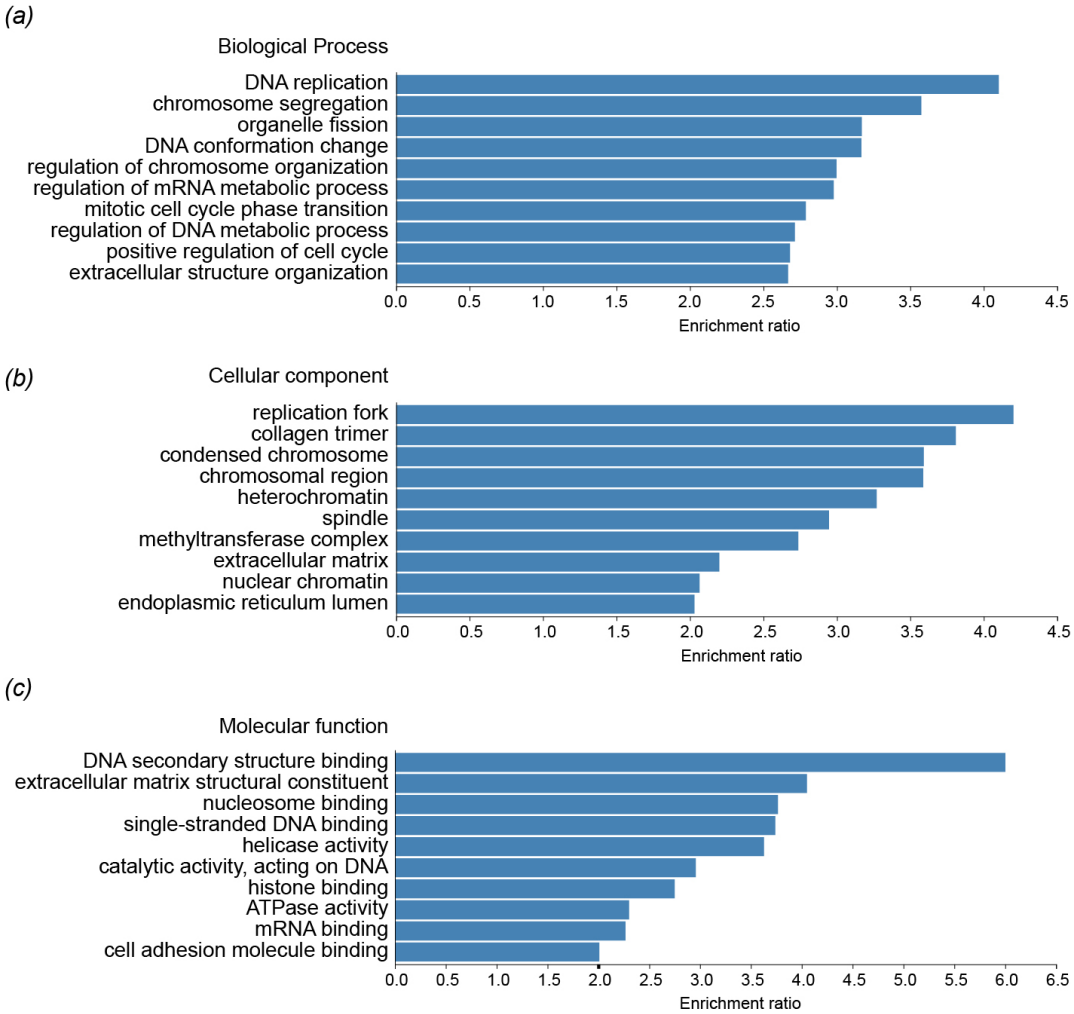


Figure S2. Gene ontology analysis of genes upregulated in the lungfish 14 dpa tail blastema. (a) Biological process. (b) Cellular component. (c) Molecular function.

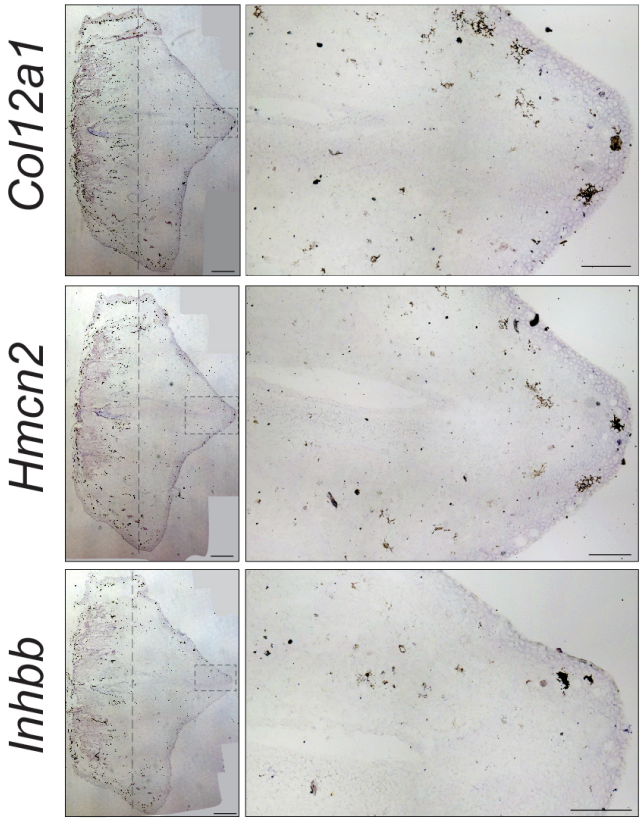


Figure S3. In situ hybridization of sense-control probes for select genes upregulated in the 14 dpa lungfish tail blastema relative to uninjured tail tissue. Scale bars of 1 mm (panoramic views) and 0.25 mm (enlarged view).