A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors

Bryant, Donald M; Johnson, Kimberly; DiTommaso, Tia; Tickle, Timothy; Couger, Matthew Brian; Payzin-Dogru, Duygu; Lee, Tae J; Leigh, Nicholas D; Kuo, Tzu-Hsing; Davis, Francis G; Bateman, Joel; Bryant, Sevara; Guzikowski, Anna R; Tsai, Stephanie L; Coyne, Steven; Ye, William W; Freeman, Robert M; Peshkin, Leonid; Tabin, Clifford J; Regev, Aviv; Haas, Brian J; Whited, Jessica L

Published in:
Cell Reports

DOI:
10.1016/j.celrep.2016.12.063

2017

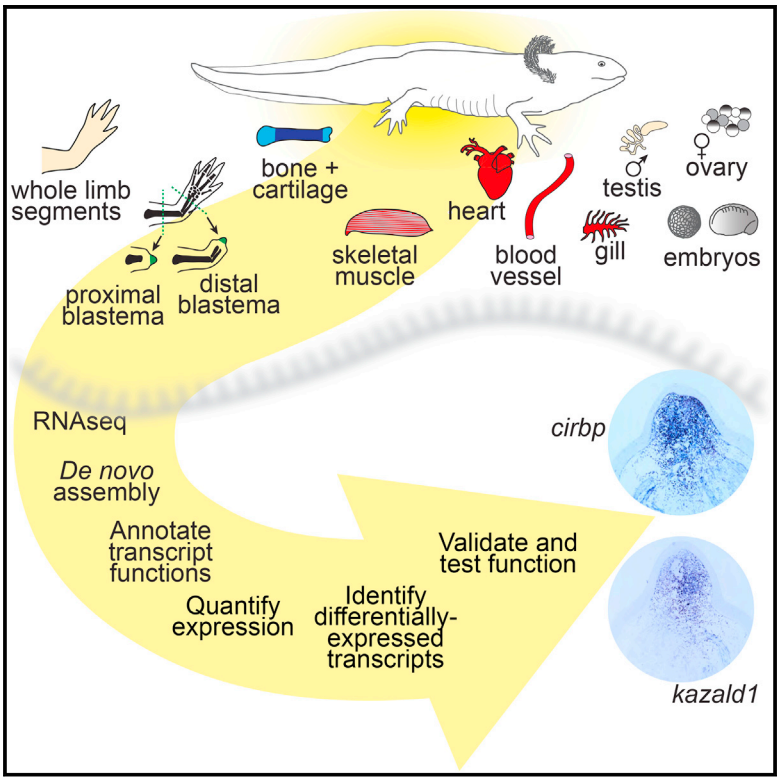
Document Version:
Publisher's PDF, also known as Version of record

Resource

Cell Reports

A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors

Graphical Abstract



Authors

Donald M. Bryant, Kimberly Johnson, Tia DiTommaso, ..., Aviv Regev, Brian J. Haas, Jessica L. Whited

Correspondence

bhaas@broadinstitute.org (B.J.H.), jwhited@bwh.harvard.edu (J.L.W.)

In Brief

Discovery of genes driving axolotl limb regeneration has been challenging, due to limited genomic resources. Bryant et al. have created a transcriptome with near-complete sequence information for most axolotl genes, identified transcriptional profiles that distinguish blastemas from differentiated limb tissues, and uncovered functional roles for *cirbp* and *kazald1* in limb regeneration.

Highlights

- Creation of a transcriptome with near-complete sequence data for 88% of axolotl genes
- Expression analyses identify tissue-enriched transcripts for key tissues
- The RNA-binding protein *cirbp* plays a cytoprotective role in limb regeneration
- Knockdown and overexpression of *kazald1* in blastema cells impair limb regeneration

Accession Numbers

GSE92429

Bryant et al., 2017, Cell Reports *18*, 762–776
January 17, 2017
http://dx.doi.org/10.1016/j.celrep.2016.12.063

OPEN ACCESS
CellPress

Cell Reports
Resource

A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors

Donald M. Bryant,[1,6] Kimberly Johnson,[1,6] Tia DiTommaso,[1] Timothy Tickle,[2] Matthew Brian Couger,[3] Duygu Payzin-Dogru,[1] Tae J. Lee,[1] Nicholas D. Leigh,[1] Tzu-Hsing Kuo,[1] Francis G. Davis,[1] Joel Bateman,[1] Sevara Bryant,[1] Anna R. Guzikowski,[1] Stephanie L. Tsai,[4] Steven Coyne,[1] William W. Ye,[1] Robert M. Freeman, Jr.,[5] Leonid Peshkin,[5] Clifford J. Tabin,[4] Aviv Regev,[2] Brian J. Haas,[2,*] and Jessica L. Whited[1,7,*]

[1]Harvard Medical School, Harvard Stem Cell Institute, and Department of Orthopedic Surgery, Brigham & Women's Hospital, 65 Landsdowne St., Cambridge, MA 02139, USA
[2]Broad Institute of MIT and Harvard and Klarman Cell Observatory, 7 Cambridge Center, Cambridge, MA 02142, USA
[3]Department of Microbiology and Molecular Genetics, Oklahoma State University, 307 Life Sciences East, Stillwater, OK 74078, USA
[4]Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
[5]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA
[6]Co-first author
[7]Lead Contact
*Correspondence: bhaas@broadinstitute.org (B.J.H.), jwhited@bwh.harvard.edu (J.L.W.)
http://dx.doi.org/10.1016/j.celrep.2016.12.063

SUMMARY

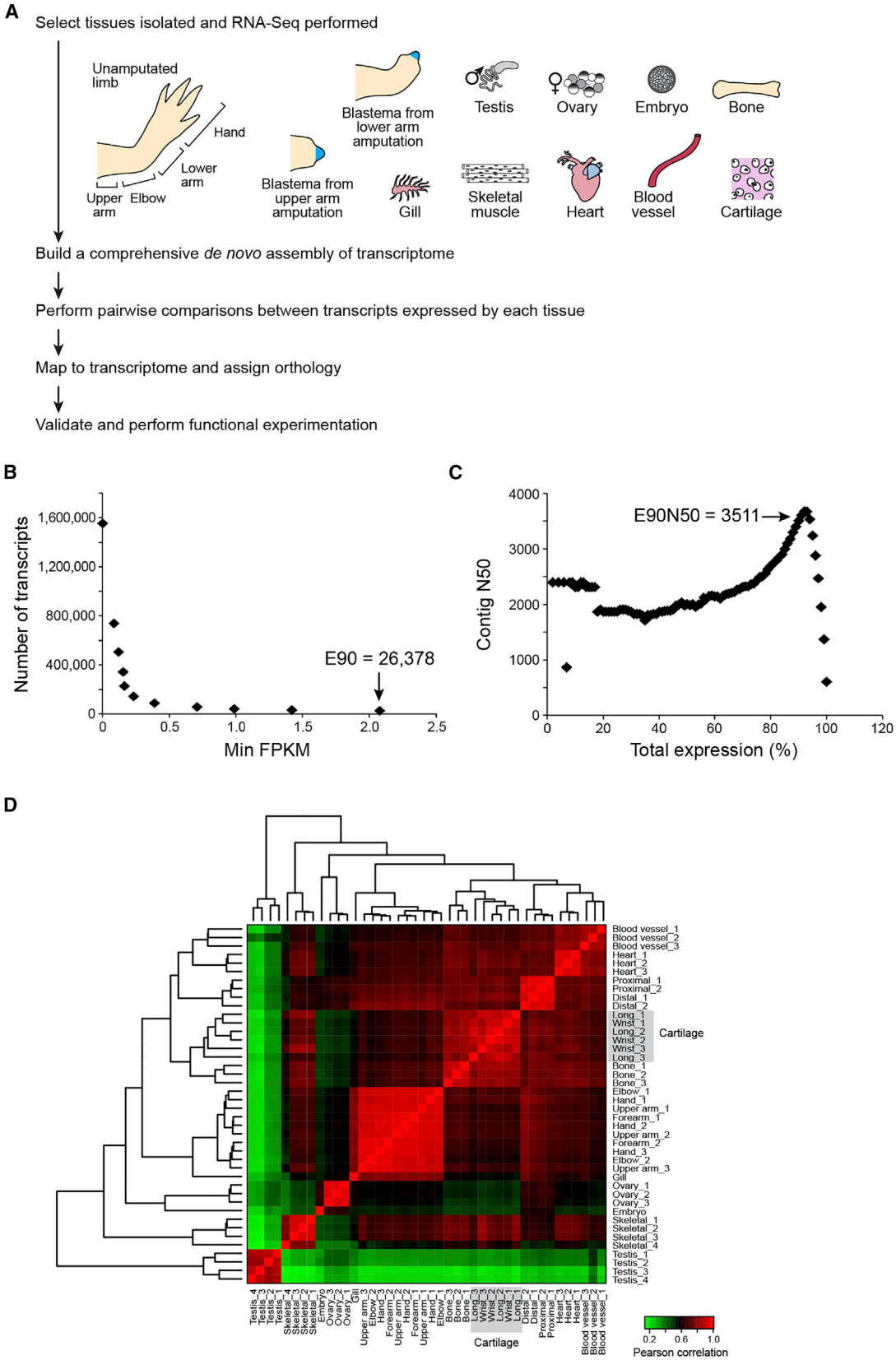
Mammals have extremely limited regenerative capabilities; however, axolotls are profoundly regenerative and can replace entire limbs. The mechanisms underlying limb regeneration remain poorly understood, partly because the enormous and incompletely sequenced genomes of axolotls have hindered the study of genes facilitating regeneration. We assembled and annotated a de novo transcriptome using RNA-sequencing profiles for a broad spectrum of tissues that is estimated to have near-complete sequence information for 88% of axolotl genes. We devised expression analyses that identified the axolotl orthologs of *cirbp* and *kazald1* as highly expressed and enriched in blastemas. Using morpholino anti-sense oligonucleotides, we find evidence that *cirbp* plays a cytoprotective role during limb regeneration whereas manipulation of *kazald1* expression disrupts regeneration. Our transcriptome and annotation resources greatly complement previous transcriptomic studies and will be a valuable resource for future research in regenerative biology.

INTRODUCTION

The limited capacity of humans to regenerate many tissues, organs, and appendages is a formidable clinical hurdle (Ziegler-Graham et al., 2008). Conversely, some animals, including invertebrates such as planaria and vertebrates such as amphibians, have remarkable regenerative capacity. Among those, many salamanders, including axolotls, can regenerate entire limbs throughout life (reviewed in Whited and Tabin, 2009). Elucidating the molecular mechanisms that enable such profound regenerative capacity may provide key insights relevant to human regenerative medicine.

The axolotl community has made significant strides in advancing our understanding of limb regeneration, but our knowledge of the molecular mechanisms that underlie axolotl regeneration is still very limited. Unbiased genomics and transcriptomics can often unlock the molecular components of systems that have not been genetically tractable. Unfortunately, the axolotl genome remains mostly unsequenced and poses major challenges at ~32 Gb in size (Keinath et al., 2015; Smith et al., 2009; Straus, 1971). RNA sequencing followed by de novo transcriptome assembly (Haas et al., 2013; Robertson et al., 2010; Schulz et al., 2012) has offered investigators an alternative for identifying near-full-length transcripts and performing differential gene expression analyses without genome mapping. Recent axolotl transcriptome studies (Knapp et al., 2013; Li et al., 2014; McCusker et al., 2015; Monaghan et al., 2009; Stewart et al., 2013; Voss et al., 2015; Wu et al., 2013) have focused on and significantly advanced our understanding of the changes in transcription over time in the regenerating portion of the limb. However, an important missing component of the existing datasets is deep transcriptional information about each of the presumed parent tissue types within the limb, which contribute progenitors and serve as the template for the future regenerate limb. Thus, examining the transcripts that define them in the differentiated, homeostatic state will be critical for future comparisons with progenitor cells along the temporal path of regeneration.

Here we combined RNA sequencing (RNA-seq) of diverse tissues with de novo transcriptome assembly, computational analysis, and experimental validation to develop a systematic map of the axolotl transcriptome. This assembly facilitated identification of specific transcripts and classes of genes whose expression is associated with successful limb regeneration. We experimentally validated our transcriptome's accuracy by analyzing mRNA expression of identified transcripts using in situ hybridization. Furthermore, we experimentally modified the expression of *cirbp*



(legend on next page)

and *kazald1* (two blastema-enriched transcripts) and uncovered functional roles for these genes in axolotl limb regeneration. The transcriptome and analyses that we provide significantly complement prior research and will be an important resource for future studies of limb regeneration, as well as for inquiries using the axolotl that extend beyond those of limb regeneration.

RESULTS

An RNA-Seq Catalog for Limb Tissues and Regeneration

To build a reference map of axolotl limb regeneration, we profiled 42 samples across 16 different tissues (Figure 1A). First, we profiled intact, unamputated limbs to reflect the starting—and end—point for limb regeneration. We sampled four positions along the proximal (shoulder) to distal (fingertip) axis to identify any location-specific transcriptional differences. Second, we sequenced the blastema, a bud-like outgrowth at the tip of the regenerating limb that contains activated progenitor cells that regenerate the internal structures of the limb. To identify transcripts whose regulation distinguishes blastema cells, we removed the regenerate epithelium of blastemas at the medium-bud stage before it began to differentiate into the various tissues of the regenerating limb. Third, it is thought that axolotls are able to redeploy transcriptional programs that guide embryogenesis during limb regeneration, and a recent study demonstrated that genes with roles in germline cell renewal are required for axolotl limb regeneration (Zhu et al., 2012). To facilitate the systematic identification of germline and embryonic transcriptional programs that are reactivated during regeneration, we generated transcriptional profiles for testes, ovaries, and embryos (one-cell to pre-hatch stage). Fourth, following amputation, individual tissues within the limb, including skeletal muscle, cartilage, bone, and blood vessels, may contribute activated progenitor cells to the regenerating limb (Kragl et al., 2009; Muneoka et al., 1986; Sandoval-Guzmán et al., 2014); reviewed in Knapp and Tanaka, 2012). We therefore sampled and analyzed each of these tissues to define tissue-specific expression and marker transcripts for the differentiated tissue types in limbs. We also included transcripts from heart and gill filaments in our assembly to increase the comprehensiveness of the transcriptome and to provide resources for researchers interested in pursuing questions outside of limb regeneration. Fifth, the blastema is an autonomous unit, programmed from the onset with spatial coordinates that instruct the regrowth of precisely the portion of the limb that has been lost (Crawford and Stocum, 1988a, 1988b; Echeverri and Tanaka, 2005; McCusker and Gardiner, 2013; Mercader et al., 2005; Stocum and Melton, 1977), such that amputation of the hand (a "distal" amputation) results in regeneration of a hand, but amputation of an entire arm (a "proximal" amputation) results in regeneration of an arm. To assess how information is differentially encoded in proximal versus distal blastemas, we sequenced medium-bud-stage blastemas derived from both proximal and distal amputations.

A De Novo Assembled Axolotl Transcriptome

We assembled 42 RNA-seq samples totaling ~1.3 billion 100-base paired-end reads (Table S1) using Trinity (Grabherr et al., 2011; Haas et al., 2013). We first combined reads from all samples and performed in silico normalization, retaining 6.6% of reads for assembly (86 million paired-end reads). We then used Trinity (Table S1) to generate an assembly of 1,554,055 transcript contigs clustered into 1,388,798 "gene" groupings (Data S1 and S2; median transcript length: 288 bases, N50 of 606 bases). Although our in silico normalization filtered out the vast majority of the reads prior to assembly, 80% of the original reads mapped back to the assembly, with most mapping as properly paired reads (Figure S1A). The majority of the transcriptome corresponds to lowly expressed transcript contigs (Figure 1B), as 90% of the total transcription is represented by an "E90 transcript set" of 26,378 transcripts (20,506 genes, E90-N50 of 3,511) (Table S1; Figure 1C).

Functional Annotation with Trinotate and Assessment of Transcriptome Completeness

Annotating de novo assembled transcripts in the absence of an assembled genome, as in the case of the axolotl, remains a bioinformatics challenge with few available tools. We therefore developed Trinotate (http://trinotate.github.io), an annotation protocol and toolkit for de novo assembled transcriptomes. Trinotate extracts predicted coding regions using TransDecoder (http://transdecoder.github.io), searches both the entire transcripts and the predicted coding regions separately for protein homology by BLAST search against Swiss-Prot, and annotates the coding regions for domain content, signal peptides, and transmembrane domains. Trinotate reports its annotations in a convenient tab-delimited file (Data S3).

We applied Trinotate to the 1.6 million transcripts in our assembly, finding a large number of likely full-length transcripts enriched in the E90 set. Trinotate reported 109,180 transcripts matching 29,529 unique Swiss-Prot proteins, with 27,056 Trinity transcripts matching 13,501 proteins across at least 80% of the matching protein's length. These transcripts likely represent fully or near-fully reconstructed transcripts with detectable homologs in other species, of which over a third (9,950) are significantly enriched (22-fold; p value $< 2.2 \times 10^{-16}$, Fisher's exact test) among the E90 transcript set.

By subjecting our transcriptome to BUSCO analysis (Simão et al., 2015), we identified by this measure that our transcriptome

Figure 1. Pipeline and Assessment of Read Representation for De Novo Axolotl Transcriptome

(A) Strategy for deriving a tissue-coded de novo transcriptome for axolotl.

(B) The count of the most highly expressed transcripts is plotted as a function of minimum expression value. 90% of the total expression (E90) is accounted for by the 26,378 most highly expressed transcripts. FPKM, fragments per kilobase of transcript per million mapped reads.

(C) The contig N50 value is computed for cumulative sets of the most highly expressed transcripts.

(D) Expression values for all tissue types were compared and Pearson correlation values were computed. Samples were clustered according to Pearson correlation values, indicating high similarity among sample replicates and between similar tissue types.

See also Figure S1, Table S1, and Data S1, S2, S3, S4, and S5.

Table 1. BUSCO Analysis of Transcriptome Completeness

Reference	Species	Summary in BUSCO Annotation
This manuscript	*A. mexicanum*	C:88%[D:53%],F:4.5%,M:7.3%,n:3023
Jiang et al., 2016	*A. mexicanum*	C:89%[D:42%],F:2.7%,M:7.8%,n:3023
http://www.ambystoma.org/genome-resources/5-gene-and-est-database	*A. mexicanum*	C:42%[D:0.9%],F:17%,M:40%,n:3023
Li et al., 2014	*A. mexicanum*	C:48%[D:21%],F:10%,M:41%,n:3023
Stewart et al., 2013	*A. mexicanum*	C:17%[D:1.6%],F:15%,M:67%,n:3023
Knapp et al., 2013	*A. mexicanum*	C:9.7%[D:0.7%],F:8.5%,M:81%,n:3023
Abdullayev et al., 2013	*N. viridescens*	C:84%[D:41%],F:4.6%,M:11%,n:3023
http://sandberg.cmb.ki.se/redspottednewt	*N. viridescens*	C:87%[D:56%],F:4.1%,M:8.3%,n:3023
Looso et al., 2013	*N. viridescens*	C:30%[D:7.0%],F:10%,M:58%,n:3023
Nakamura et al., 2014	*C. pyrrhogaster*	C:82%[D:34%],F:4.6%,M:13%,n:3023

C, complete; D, duplicated; F, fragmented; M, missing; n, number of BUSCO genes. For comparative purposes, we have included several recent newt transcriptome assemblies. See also Figure S1, Table S1, and Data S1, S2, and S3.

contains near-complete gene sequence information for 88% of the genes in the axolotl genome (Table 1). We finally applied the Core Eukaryotic Genes Mapping Approach (CEGMA) (Parra et al., 2007), which further indicated that our transcriptome is more than 98% complete (Table S1).

Differential Expression Analysis Identified Tissue-Specific Gene Expression and Splicing

We estimated the expression profiles for all transcripts across all samples, and found excellent correlation between biological replicates (average Pearson correlation between replicates: 0.90; Figure 1D). As expected, samples derived from similar tissues (e.g., sections of the arm segments; bone and cartilage; proximal and distal blastemas) had more highly correlated expression than more distinct tissues such as the testes, which exhibited the most divergent transcriptional profile.

We next identified differentially expressed transcripts, conservatively focusing on transcripts identified as differentially expressed by multiple methods (Figure S1B) and restricted to samples with biological replicates (this excludes gill filament and embryos, which were used only for assembly). We identified 60,355 transcripts corresponding to 41,697 genes as at least 2-fold significantly differentially expressed (FDR $\leq$ 0.05) between at least two tissue-type comparisons (Data S4 and S5). Finally, we used a graph-based analysis to identify genes that are differentially expressed across each set of tissues (Figure 2A). In this approach, a graph is constructed for each differentially expressed transcript such that each tissue type is represented by a node, and an edge is drawn from a significantly upregulated tissue node to the downregulated tissue node. The graph is then partitioned to identify the maximal set of upregulated tissues for that transcript. Visualization of the graphical output for a given transcript intuitively presents the tissue(s) with the highest expression level for the transcript of interest (Figure 2A).

Of the differentially expressed transcripts, 14,594 transcripts (11,523 genes) were most tissue enriched given all pairwise tissue comparisons (Figure 2B; Table S2). Testes and ovaries had the most tissue-enriched transcripts (7,795 and 2,247, respectively, together 69%), followed by skeletal muscle tissue (961 transcripts) and the combination of all arm-segmented tissues (elbow, forearm, upper arm, and hand; 892) compared to each of the remaining tissues, including those tissues that compose the arm (e.g., bone). The tissue-specific transcripts were enriched for physiologically relevant functional categories based on enrichment analysis of Gene Ontology (GO) terms assigned from Swiss-Prot annotations of homologous proteins using Trinotate (e.g., spermatogenesis was enriched in testis-specific genes [FDR $< 5.5e^{-44}$], muscle contraction for skeletal muscle [FDR $< 5.1e^{-20}$], etc.; Table S2).

We experimentally validated by in situ hybridization and RT-PCR some of the tissue-enriched expression predictions to demonstrate our ability to correctly predict diverse tissue-specific gene expression patterns (Figures 2C–2J). For example, as predicted by our analysis, *speriolin* (*spermatogenesis and centriole associated 1*; *spatc1*) is enriched in germ cells within the testes (Figure 2C). This is consistent with limited studies on Speriolin that demonstrate this protein associates with centrosomes in spermatocytes and exhibits testis-specific expression in mice (Goto and Eddy, 2004; Goto et al., 2010).

Finally, we identified differentially expressed transcripts that are also associated with differential isoform usage across tissues. Of the 60,355 differentially expressed transcripts (41,697 genes), there are 29,750 transcripts (11,092 genes) with putative alternatively spliced isoforms. For 7,300 genes, the alternative isoforms have different patterns of tissue enrichment, suggestive of alternate functional roles (Table S3). Most of these genes (4,097 of 7,300) have at least one isoform significantly enriched in either testes or ovaries, with far fewer genes with evidence for tissue-specific isoforms in other tissues (e.g., 96 genes involving alternate isoform enrichment between heart and skeletal muscle). Notably, only 346 genes of the 11,523 genes with strong tissue-specific expression (Figure 2B) show evidence for tissue-specific isoform usage, the vast majority (315 genes) in either testes or ovaries. These data suggest axolotls may be less reliant upon alternative splicing in directing tissue-specific

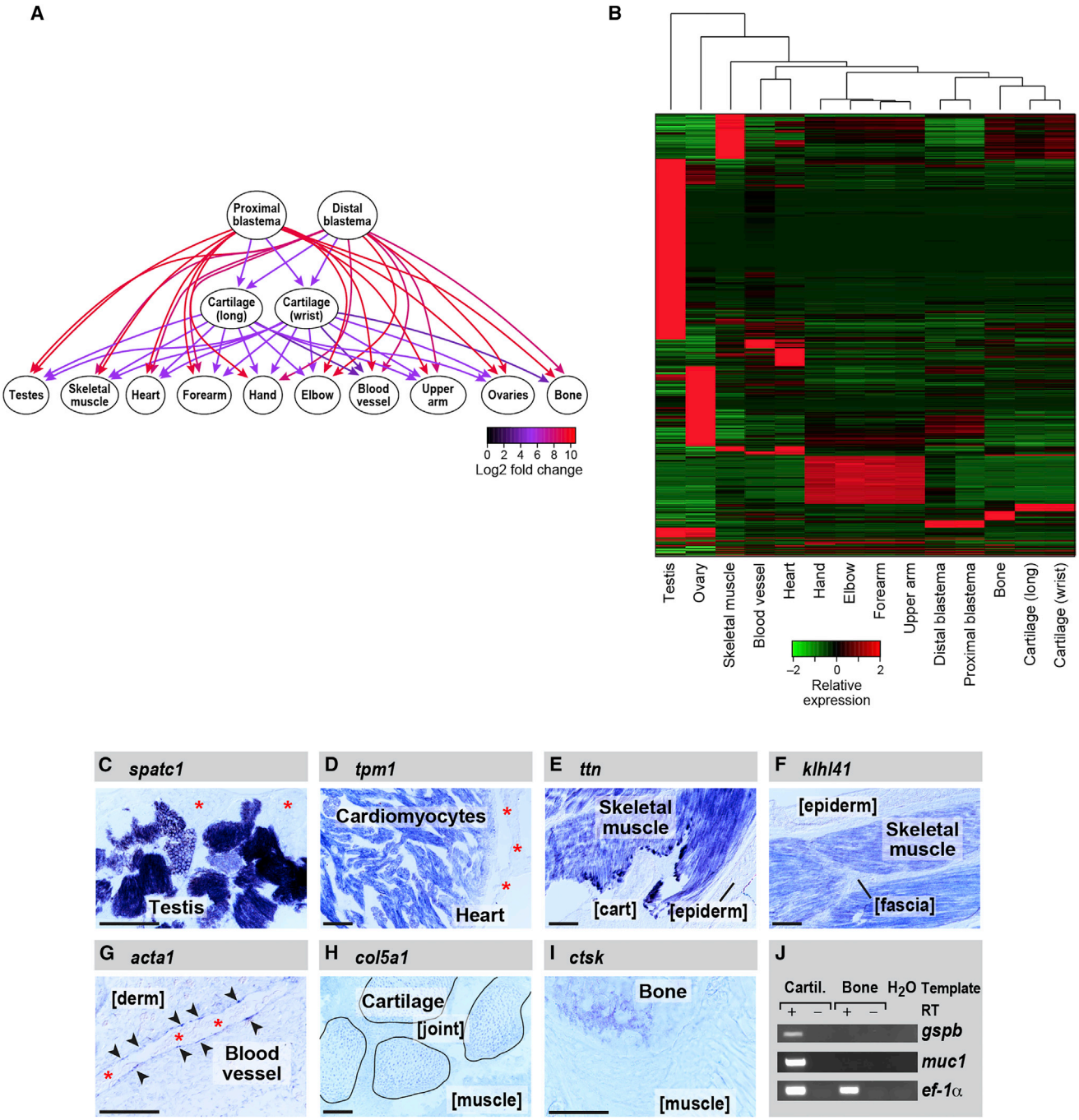


Figure 2. Differential Gene Expression Analysis across Each Set of Tissues Identifies Transcripts Most Enriched in Specific Tissue Types

(A) Graph illustrating the methodology for the identification of genes that are tissue enriched in the context of all tissue pairwise comparisons using *kazal-type serine peptidase inhibitor domain 1* (*kazald1*) as an example. Directed edges are drawn from upregulated to downregulated tissues, and fold changes in expression are indicated by edge colors.

(B) Heatmap showing all transcripts that are enriched in specific tissue types.

(C–J) RNA in situ hybridization performed on tissue sections.

(C) *speriolin* (*spatc1*) is enriched in the germ cells in the testes but is not detectable in adjacent support cells (asterisks).

(D) *tropomyosin 1* (*tpm1*) is enriched in cardiomyocytes within the heart and not detectably expressed by other heart cell types such as epicardium (asterisks).

(E) *titin* (*ttn*) is enriched in limb skeletal muscle but is not detectable in adjacent cartilage (cart) and epidermis (epiderm).

(F) *kelch repeat and BTB domain-containing protein 10* (*klhl41*) is highly enriched in skeletal muscle and not detectable in adjacent tissues such as epidermis (epiderm) and fascia.

(legend continued on next page)

gene expression than has been documented in mammals (for example, Wang et al., 2008).

A Transcriptional Program for Blastemas Implicates the RNA Life Cycle in Limb Regeneration

We next analyzed the 159 transcripts (151 genes) differentially induced in all blastema samples (proximal and distal) as compared to all other sampled tissues (Figure 3A; Table S4). Notably, many of the blastema-specific genes (64/151) do not have a similar sequence identified in another organism in the UniRef90 database ($E \leq 10^{-3}$; BLASTX). Nonetheless, several blastema-specific genes have known blastema relevance (e.g., *twist family bhlh transcription factor 1* [*twist1*] [Satoh et al., 2008]) or are predicted to encode proteins in pathways implicated in salamander limb regeneration (*retinol-binding protein 2a* [*rbp2a*] [Maden, 1983]; *matrix metallopeptidase 11* [*mmp11*] [Yang et al., 1999]). In particular, we identified the T box transcription factor *tbx5*, previously reported as enriched in forelimb blastemas (Khan et al., 2002), discriminating its tissue-specific expression among 13 total *tbx* genes detected in this axolotl transcriptome. Furthermore, when analyzing forelimb blastemas, our differential expression analysis did not erroneously uncover *tbx4*, which has been documented to only be upregulated in hindlimb blastemas (Khan et al., 2002).

Several blastema-specific genes are predicted to encode RNA-binding proteins, highlighting a potential role for the regulation of the RNA life cycle in regeneration (Figure 3A, highlighted). Among the blastema-induced transcripts are three predicted *heterogeneous nuclear ribonucleoprotein* (*hnrnp*) transcripts, an *RNA-binding motif* (*rbm3*) ortholog, a *serine/arginine-rich RNA splicing factor* (*sfrs1*) ortholog, and the axolotl ortholog of *cold-inducible RNA-binding protein* (*cirbp*). Interestingly, a newt ortholog of *cirbp* was recently identified in a subtractive hybridization screen as upregulated in limb blastemas compared to non-amputated limbs (Jiang et al., 2014), raising the possibility that *cirbp* and its homologs may be important regulators in regeneration across species. We also validated several of the blastema-enriched genes by in situ hybridization (Figure 3B), and confirmed they are not appreciably expressed in unamputated limbs (Figure 3C).

We further explored the expression of *cirbp* during limb regeneration and found that *cirbp* expression was apparent as early as 3 days post-amputation (Figure 4A), persisted as the blastema consolidates, and grows through palette formation (Figures 4B–4D). Past studies have demonstrated that CIRBP proteins regulate cell fate by inhibiting pro-apoptotic pathways during newt spermatogenesis (Eto et al., 2009) and in the mammalian nervous system (Zhang et al., 2015). To further explore an anti-apoptotic role for axolotl *cirbp* in limb regeneration, we delivered a translation-blocking morpholino to early blastemas (Figure S2) and assayed cell death using TUNEL staining 4 days later. Administration of *cirbp*-targeting morpholino caused a significant increase in the percentage of TUNEL-positive blastema cells relative to PBS and standard morpholino controls (*cirbp*-targeting morpholino: n = 19 limbs; standard control morpholino: n = 11 limbs; PBS control: n = 5 limbs) (Figures 4E–4G). These results suggest a conserved role for Cirbp-related proteins in regulating cell death across vertebrates and highlight a cytoprotective role for RNA-binding proteins within blastema cells.

We also examined the 60 transcripts specifically downregulated in blastemas (Figure S3). These included transcripts that encode proteins that likely regulate fatty acid metabolism (*prkab2*, *cyp2a13*, and *faah*), glucose metabolism (*pgm1* and *gp1*), the cell cycle (*ccng1*), and mitochondrial respiration (*cox7a2* and *cox7a2l*). The expression of most of these blastema-repressed transcripts is higher in skeletal muscle than in bone and cartilage (e.g., *casq1*), suggesting that blastema cells may be more actively repressing a muscle program than a skeletal program.

Functional Studies Indicate a Critical Role for *kazald1* in Limb Regeneration

The most blastema-enriched transcript identified by our analyses was *kazald1*, whose transcription was >10-fold upregulated in blastemas versus all other tissues examined (Figure 5A). Kazd1 is predicted to have a Kazal-type serine protease inhibitor domain, a follistatin-like domain, and an immunoglobulin domain. We found that *kazald1* expression is not detectable in the intact limb pre-amputation, increases shortly after amputation, is maintained during the blastemal stages of regeneration, and is dramatically downregulated near the end of regeneration (Figures 5B and 5C). Notably, *kazald1* is not detectable in the developing limb bud within limb bud progenitor cells (Figure 5C). Therefore, its function in blastemas may be regeneration specific, in contrast to other genes whose expression is similar between limb buds and blastemas (for example, *sonic hedgehog* [Torok et al., 1999]; *hoxd-8*, *-10*, and *-11* [Torok et al., 1998]).

We hypothesized that diminishing *kazald1* expression during regeneration would have deleterious consequences and found that *kazald1*-targeting morpholinos delayed limb regeneration (Figures 5D–5F; Figure S4) and resulted in significantly smaller blastemas as compared to controls (Figure 5F: n = 48 limbs per condition; Figures S4B–S4D: n = 48 control MO2-treated limbs, n = 44 *kazald1*-MO2-treated limbs). Further, the *kazald1*-morpholino-treated animals had delayed chondrification and differentiation of the digits at the palette-to-early-digit stage, with less than 50% total Alcian blue-staining area relative to controls (Figures 5G–5K: n = 46 control MO1-treated limbs, n = 45

(G) *actin, alpha 1, skeletal muscle* (*acta1*) mRNA is enriched in the very thin layer of vascular endothelial cells lining the blood vessels (arrowheads) and absent from adjacent dermis (derm). Asterisks mark red blood cell clumps in the vessel lumen.
(H) *collagen type V alpha 1* (*col5a1*) expression is highly enriched in cartilage; shown are four carpals (outlined) within the wrist. Expression in joint (between carpals) and adjacent muscle is diminished.
(I) A bone-enriched marker, *cathepsin k* (*ctsk*), is highly expressed in ossified portions of the humerus and low in adjacent muscle.
(J) *platelet-binding protein GspB* (*gspb*) and *mucin 1* (*muc1*) are detected in cartilage but not in bone by RT-PCR. *eukaryotic translation elongation factor 1 alpha 1* (*ef-1a*) serves as the loading control.
Scale bars, 100 μm. See also Figure S7, Tables S2 and S3, and Data S1, S2, S3, S4, and S5.

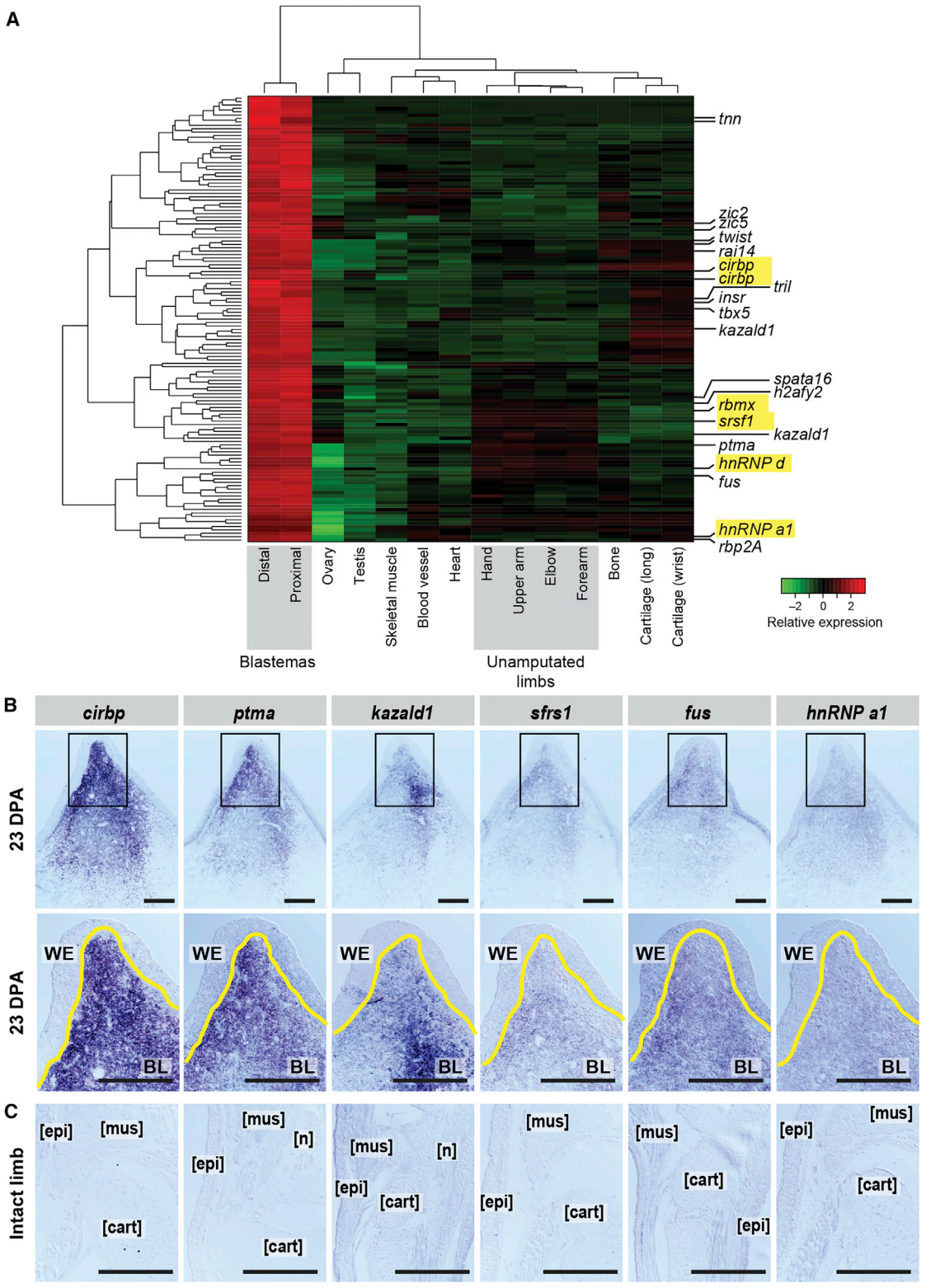


Figure 3. Identification and Validation of Blastema-Enriched Transcripts

(A) We identified 159 transcripts (151 genes) enriched in the blastema (the combination of proximal and distal blastema tissue) as compared to all other tissues. Those predicted to encode proteins with RNA-binding/regulation properties are highlighted in yellow.

(legend continued on next page)

kazald1-MO1-treated limbs; Figures S4E–S4I: n = 48 control MO2-treated limbs, n = 44 *kazald1*-MO2-treated limbs). These data strongly suggest that *kazald1* is critical for the proper progression of limb regeneration.

Motivated by the finding that *kazald1* is essential for normal regeneration, we leveraged our transcriptome to create *kazald1*-targeted deletions in both embryos and unamputated limbs via electroporation (Supplemental Information; Figure S5) to further complement our studies with morpholino knockdown. Although we were able to successfully edit the *kazald1* gene, we did not achieve a sufficient level of editing efficiency to study the effect of genetic loss of function of *kazald1* on limb regeneration with a high degree of confidence (17%–46% efficiency; Figure S5); we did not observe regenerative defects in mosaic animals or electroporated limbs, most likely due to low editing efficiency (data not shown). Nevertheless, future studies with complete genetic loss of function of *kazald1* are warranted and will prove highly valuable to further examining this gene's mechanistic role in axolotl limb regeneration.

kazald1's intense expression within the transient blastema structure suggests that it has a time-delineated function during regeneration. To test this, we used a replication-incompetent retrovirus (Whited et al., 2013) to constitutively overexpress *kazald1* in regenerating limbs (Figure S6A). A majority (15/22) of *kazald1*-infected limbs regenerated abnormally compared to control EGFP infections (0/18) (Figure S6B). The most commonly observed defect following *kazald1* overexpression was syndactyly (8/15 abnormal limbs), and other defects included missing digits, truncated digits, severe clinodactyly, and ectopic tissue growth on the palm and digits (Figure S6C).

Our morpholino and retroviral experiments highlight a functional role for *kazald1* during limb regeneration, as perturbing its expression during this process adversely affected the progression and outcome of limb regeneration. Another report also discovered enriched expression of *kazald1* in a regenerative context in vitro (Athippozhy et al., 2014). However, in this study, the expression of the transcript was not profiled in intact or regenerating limbs, and the bulk of the expression was attributed to nerves. Here we show that the blastema cells themselves produce large amounts of *kazald1* transcript. More studies will be needed to determine whether blastema cells are stimulated by the ingrowing regenerating nerve fibers to upregulate this transcript.

Comparison of Tissues at Different Anatomical Locations Identifies Candidate Transcripts of Positional Memory along Intact and Limb Regeneration Sites

Axolotl limb tissues maintain a "positional memory" throughout life allowing for the precise regeneration of only lost limb elements following amputation, suggesting that specific genes may be differentially expressed along the length of the limb and control positional memory. To uncover such genes, we assayed samples of similar tissue type but of diverse positional/anatomical context. First, we profiled cartilage from the wrist (carpals) and the ends of long bones (humerus, radius, and ulna). Second, we collected samples along the length of the limb, and blastemas arising from either proximal (mid-humerus) or distal (distal radius/ulna) amputations. We then analyzed their expression profiles to identify differentially expressed genes for each context.

Arm Segments

We performed two separate analyses to address the transcriptional differences between different locations along the proximal-distal axis of the arm (Figure 6). We first performed pairwise comparisons to identify transcripts significantly differentially expressed between any pair of arm segments. In the second approach, we performed analyses to uncover factors that show graded enrichment across the proximal-distal axis.

Using the pairwise approach, we uncovered a relatively modest number of transcriptional differences between different segments of the forelimb ("arm"): 636 differentially expressed transcripts identified between at least one pair of arm segments and 83 transcripts involving all arm segments (Figures 6A and 6B; Table S5), especially in the hand (48/83). Many of these likely reflect differences in the cell-type composition of the different arm segments. For example, transcripts enriched in the forearm, elbow, and upper arm are also more highly enriched in skeletal muscle and bone tissue; transcripts enriched in upper arm are also enriched in bone and cartilage (Figure S7A). Interestingly, 27 of the 48 transcripts enriched in the hand relative to other arm segments are also hand specific relative to other non-arm tissues. These include keratins and other filamentous proteins (6/27) with known structural roles in skin, consistent with the increased surface area of the hand versus the rest of the arm. Notably, when comparing highly similar tissue types, we found little signal of position-specific expression. In particular, comparing cartilage between wrist and long bone, we found only minor distinctions (Figure S7B), such that the samples were very highly correlated ($r = 0.98$), with only 22 differentially expressed transcripts, mostly in lowly expressed or non-cartilage-specific genes.

The arm segment analysis highlighted key transcripts with known function in limb polarity and candidates with uncharacterized functions in limb polarity (Figure 6B). First, we found *hoxa13*, which is known to mark distal elements in developing appendages (Haack and Gruss, 1993). Second, we found and validated a specific keratin, *krt17*, as exclusively expressed in the hand, restricted to the tubercles on the ventral/palmar surface of the hand (Figures 6B and 6C). Additional hand-enriched transcripts include a homolog of *alpha tectorin* (*tecta*), which is required for hearing in mammals (reviewed in Legan et al., 2014). Finally, one of the few transcripts that distinguishes upper-arm segments from other segments of intact limbs encodes a predicted ortholog of Chordin-like 1 (*crdl1*) (Figure 6B), a diffusible morphogen (reviewed in Zakin and De Robertis, 2010).

(B and C) In situ hybridization for six highly blastema-enriched transcripts at 23 days post-amputation (DPA (B) and on intact limbs (C). (B) Lower: higher magnifications of the boxed areas (upper). Yellow lines mark the WE/BL boundary. WE, wound epidermis; BL, blastema; mus, muscle; epi, epidermis; cart, cartilage; n, nerve. Scale bars, 500 μm.
See also Figure S3, Table S4, and Data S1, S2, S3, S4, and S5.

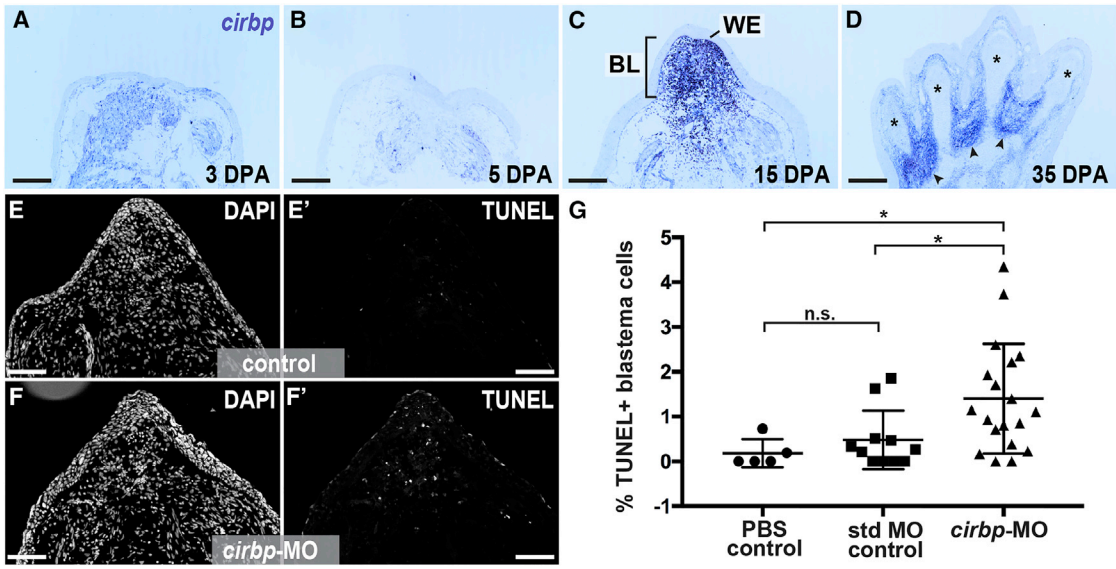


Figure 4. Axolotl Cold-Inducible RNA-Binding Protein Is a Cytoprotective Factor for Blastema Cells

(A–D) In situ hybridization for *cirbp* over the course of regeneration in regenerating limb tissue sections. (A) Three days post-amputation. (B) Five days post-amputation. (C) 15 days post amputation. (D) 35 days post-amputation. Asterisks denote newly differentiated cartilage in digits; arrowheads denote interdigital regions. WE, wound epidermis; BL, blastema; mus, muscle; epi, epidermis; cart, cartilage; n, nerve. Scale bars, 500 μm.

(E–F′) Tissue sections from regenerating limbs treated with standard morpholinos (MO) control (E and E′) and *cirbp*-targeting morpholino (F and F′) stained for nuclei (DAPI; E and F) and TUNEL (E′ and F′). Scale bars, 500 μm.

(G) Quantification of the percentage of TUNEL$^+$ blastema cell nuclei in control and *cirbp*-MO-treated limbs. *p < 0.05; n.s., not significant.

See also Figure S2 and Data S1, S2, and S3.

We uncovered 875 transcripts demonstrating a significant (FDR ≤ 0.05) gradient in expression intensity along the proximal-distal axis of the unamputated limb (Figure 6D; Table S5). Among these, we uncovered two transcripts orthologous to *emx2* (*empty spiracles homeobox 2*) that were most highly expressed in the hand and had diminishing expression toward the proximal sample sites. The newt ortholog of *emx2* has previously been shown to exhibit distal-enriched expression in unamputated limbs (Beauchemin et al., 1998). Conversely, we identified *apcdd1l* (orthologous to Apcdd1-like) as the most significantly graded transcript, with high levels in the proximal segments. Apcdd1 is an extracellular membrane-bound signaling protein that can inhibit Wnt signaling (Shimomura et al., 2010).

Proximal versus Distal Blastema

There are 305 transcripts (275 genes) differentially expressed between proximal and distal blastema tissue (Data S4 and S5), of which 81 transcripts (70 genes; Table S5) are also enriched in blastema tissue relative to other tissues (Figures 6E and 6F). We validated *cd38* and *shox*, whose expression is significantly enriched in proximal blastemas (Figure 6G). *cd38* is predicted to encode the axolotl ortholog of mammalian CD38, a cell-surface glycoprotein and ADP-ribosyl cyclase (reviewed in Wei et al., 2014). CD38 expression at both the mRNA and protein levels can be induced by exposure of cells to retinoic acid (Drach et al., 1993). Because retinoic acid is a proximalizing agent in limb regeneration (Maden, 1983), enrichment of *cd38* transcripts in proximal blastemas may reflect activation of retinoic acid signaling. *shox* (*short-stature homeobox*) encodes a predicted homeobox-containing transcription factor. Experimentally removing the most closely related mouse gene, *Shox2*, in limb buds results in the near-complete loss of the humerus, a proximal skeletal element, whereas remaining structures are mostly unaffected (Cobb et al., 2006). Future studies may reveal a functional role for these genes in controlling proximal-distal identity during limb regeneration.

DISCUSSION

Here we report the assembly and validation of an axolotl transcriptome centered on tissues relevant for limb regeneration that provides near-complete sequence information for exonic regions of ~88% of axolotl genes with a high degree of confidence (Table 1), generating a valuable annotated resource for researchers using axolotl. We used differential expression analysis within the context of a biologically informed design to identify tissue-specific and anatomically specific gene expression patterns that we were able to validate experimentally, demonstrating the validity and the breadth of experimental design that can be conducted with our catalog.

Within the blastema, the nexus of limb regeneration, we uncovered an enriched profile of transcripts for members of the heterogeneous nuclear ribonucleoprotein complex (hnRNP) such as *cirbp*, *fus*, *roa1*, and *hnrnpd*. In other biological contexts, RNA-binding proteins are capable of modulating large repertoires of mRNA targets, coordinating cellular responses, and maintaining cellular phenotypes (Kafasla et al., 2014; Zhang et al., 2013). During limb regeneration, modulation of

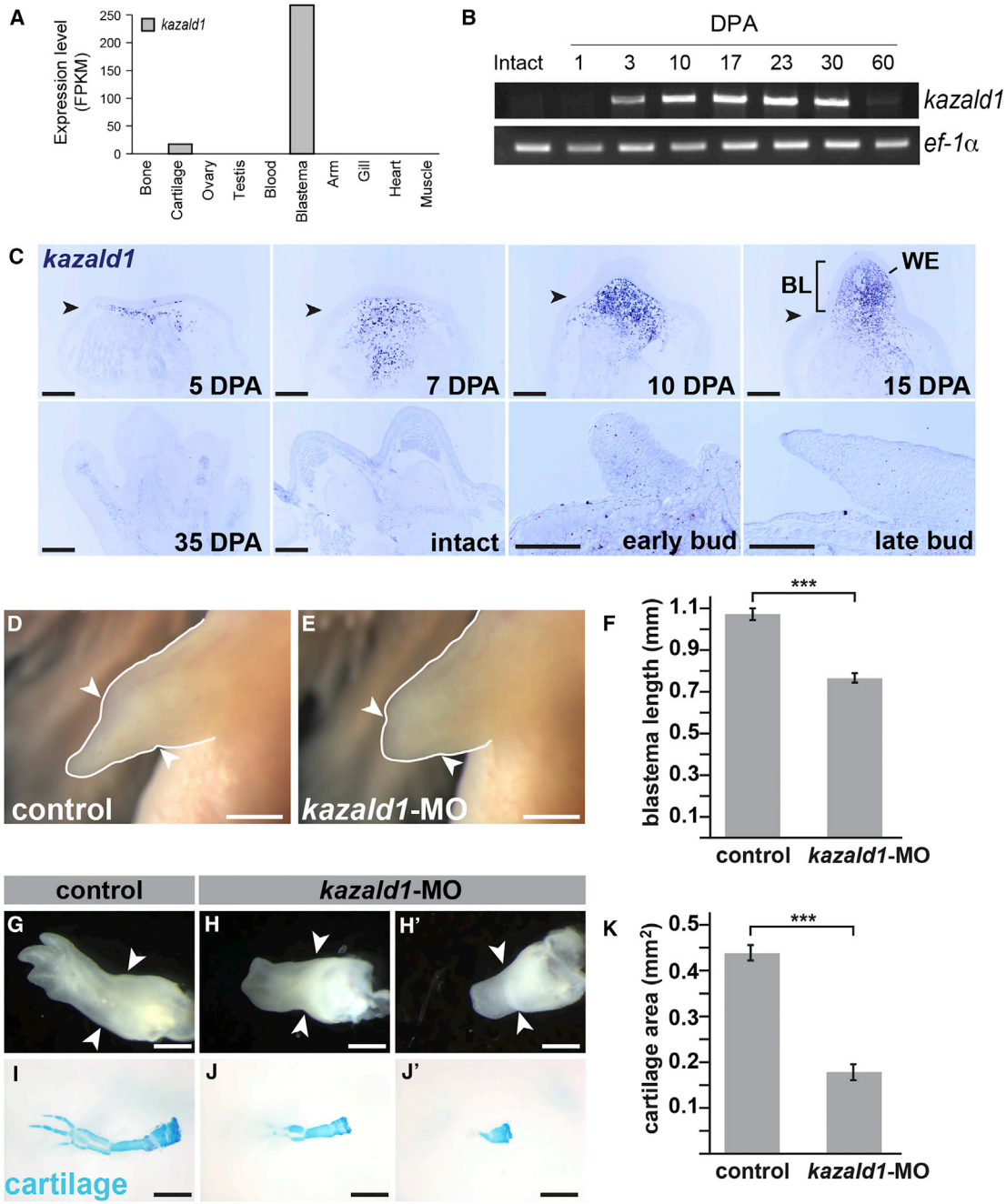


Figure 5. *kazald1*, the Most Robust Blastema Marker, Is Required for Limb Regeneration

(A) Differential tissue expression analysis identifies *kazald1* as the most blastema-enriched transcript compared to all other tissues sequenced.

(B) RT-PCR performed on blastema cDNA samples throughout the course of regeneration for *kazald1* expression. *kazald1* was not detected in intact limbs and at 1 DPA, and has dramatically diminished by 60 DPA.

(C) In situ hybridization for *kazald1* in the blastema over the course of regeneration (top panels). *Kazald1* is not detectable in regenerated limbs (35 DPA), intact limbs, or developing limb buds (lower panels).

(D–F) Regenerating limbs at 19 DPA treated with control (D) or *kazald1*-targeting morpholino (E); quantified in (F).

(G–J′) Regenerating limbs at 28 DPA treated with control (G) or *kazald1*-targeting morpholino (H and H′).

(I–K) Same specimens stained with Alcian blue to visualize cartilage; (I) is the skeletal preparation of the limb shown in (G); (J) is the skeletal preparation of the limb shown in (H); (J′) is the skeletal preparation of the limb shown in (H′). Not that (G) and (I) are a control, while (H) and (J) and (H′) and (J′) are specimens treated with *kazald1*-targeting morpholino. Results are quantified in (K).

Scale bars, 500 μm (C) and 1 mm (D, E, and H–J′). ***$p < 0.001$. Error bars are SEM. Arrowheads mark the amputation plane in each image.

See also Figures S4–S6 and Data S1, S2, S3, S4, and S5.

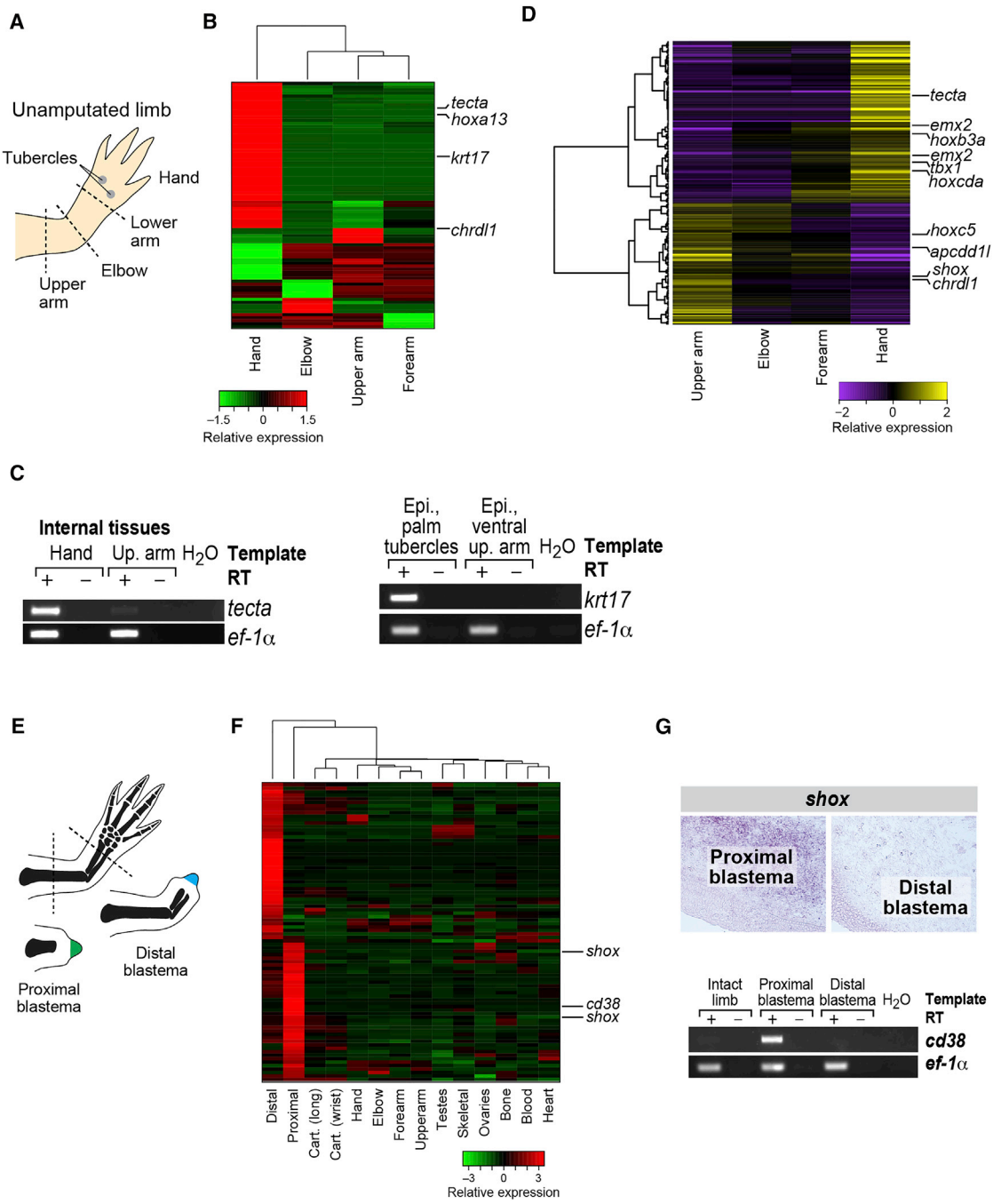


Figure 6. Transcripts Differentially Expressed in Proximal versus Distal Elements

(A) Schematic illustrating specific elements of the hand and arm.

(B) Differential gene expression analysis identifies transcripts that are enriched in distinct sections of the intact limb.

(C) RT-PCR validation of select transcripts identified by differential expression analysis.

(D) Gradient gene expression analysis identifies transcripts enriched in a gradient from proximal to distal or distal to proximal.

(E) Schematic illustrating amputation planes for sampling of proximal and distal blastemas.

(F) Differential expression analysis identifies transcripts that are enriched in proximal versus distal blastemas.

(G) In situ and RT-PCR validation of computational predictions of differentially expressed transcripts in proximal and distal blastemas.

See also Figure S7, Table S5, and Data S1, S2, S3, S4, and S5.

RNA-binding proteins and post-translational regulation of mRNA may constitute major mechanisms for executing the vast cellular changes necessary to regenerate a complex structure. Specifically, we highlighted the axolotl ortholog of *cirbp* as highly blastema enriched and uncovered a cytoprotective function for this gene during early blastema development. *cirbp* encodes a putative RNA-binding protein orthologous to mammalian cold-inducible RNA-binding protein (CIRBP) and highly similar to RBM3 (orthologous axolotl *rbmx* was also identified as blastema enriched). Both *cirbp* and *rbm3* have been demonstrated to be transcriptionally upregulated in response to specific cellular stressors including cold shock, hypoxia, and UV irradiation in mammalian cells (Danno et al., 1997; Nishiyama et al., 1997; Wellmann et al., 2004). Our experiments suggest that one function of *cirbp* in regeneration is to protect progenitor cells from cell death, possibly to protect blastema progenitors from extensive cellular changes occurring following amputation. We also identified *cirbp* transcripts at the early-to-late-digit stage (35 days post-amputation) within the interdigital regions, which are presumed to be relatively undifferentiated and will be eliminated as the regenerating digits are sculpted (reviewed in Montero and Hurlé, 2010). Further, *fus* mRNA has been identified as a major target of mammalian CIRBP (Morf et al., 2012), both of which are significantly enriched in the blastema. Thus, understanding regulatory loops between RNA-binding proteins to modulate cell-fate pathways within blastema cells may be emerging as a theme in limb regeneration, supporting a need for future studies.

We also experimentally tested the consequence of manipulating expression of *kazald1*, the gene whose transcript best distinguishes blastema cells from all other samples sequenced in our data. We found that reducing expression of *kazald1* via morpholinos resulted in an outward lag in the progression of limb regeneration compared to controls whereas constitutive expression of *kazald1* within the blastema resulted in regeneration of deformed limbs, underscoring the importance of temporal and/or spatial regulation of this transcript. Interestingly, a similar gene has been implicated in regeneration in hydra, an invertebrate with extensive regenerative abilities (Chera et al., 2006), raising the possibility that an ancient and conserved role for *kazald1* in regeneration exists. Other Kazal-type serine protease inhibitors in mammals serve as biomarkers for various forms of cancer, but their exact functions are only now emerging (reviewed in Fradet, 2009). Because aspects of limb regeneration share some similarity to cancer (for instance, activation of stem cells and dedifferentiation, cell proliferation, reactivation of developmental pathways, and the wound response) (reviewed in Pomerantz and Blau, 2013), links between the two processes warrant further investigation. Future work will be necessary to determine whether connections exist between the roles of Kazal-type inhibitors in other organisms and the role of *kazald1* in axolotl limb regeneration.

Our assembly and experimentation will serve as a valuable resource for researchers using axolotls in a variety of fields, as well as for those studying regeneration in other models. Our differential gene expression analysis can provide starting points for studying genes that may drive the biological function of particular tissues. It can also be used to identify reliable markers of particular tissue types and, with further experimentation, possibly specific cell types within particular tissues. In many cases, our assembly will provide the full-length coding sequence for future functional studies, including the requisite sequence information to design genome-editing strategies and provide data to minimize potential off-targets. The axolotl transcriptome we report can serve as a basis for mining tissue-enriched or -repressed transcripts, and will also serve as a powerful reference for more directed future studies. All reconstructed axolotl transcripts, expression data, and feature annotations are available via our web portal at http://portals.broadinstitute.org/axolotlomics.

EXPERIMENTAL PROCEDURES

Trinity De Novo RNA-Seq Assembly

Trinity version trinityrnaseq_r2013-02-25 was used for de novo transcriptome assembly and analysis as follows. The combined set of 1.3 billion RNA-seq reads spanning all sampled tissues was combined into a single pair of fastq files, quality trimmed using Trimmomatic (Bolger et al., 2014) (parameters LEADING:5 TRAILING:5 MINLEN:36), and subsequently normalized using the in silico normalization step incorporated into Trinity ($TRINTY_HOME/util/normalize_by_kmer_coverage.pl–seqType fq–JM 100G–left ALL_AXOLOTL_READS.left.fq.P.qtrim.fq–right ALL_AXOLOTL_READS.right.fq.P.qtrim.fq–pairs_together–JELLY_CPU 10–PARALLEL_STATS–max_cov 50). The resulting normalized reads were then assembled using Trinity (Trinity.pl–left left.fq–right right.fq–seqType fq–JM 100G–CPU 10).

Transcript Annotation Using Trinotate

We developed Trinotate as a bioinformatics protocol and software system for annotating putative functional characteristics of transcripts. Similarities to known proteins were detected by a BLASTX search (Camacho et al., 2009) ($E \leq 1e^{-5}$) of a comprehensive protein database formed by merging Swiss-Prot (Boeckmann et al., 2005) with UniRef90 (UniProt Consortium, 2015) protein databases downloaded from UniProt (versions available on July 5, 2014). Likely coding regions within transcripts were predicted using TransDecoder (http://transdecoder.github.io), and resulting protein products were searched for sequence similarities against the comprehensive protein database described above (BLASTP; $E \leq 1e^{-5}$) and for conserved protein domains using Hmmer (http://hmmer.org/) and PFam (Finn et al., 2014). Signal peptides were predicted using SignalP (Petersen et al., 2011), and transmembrane region predictions were predicted using TmHMM (Krogh et al., 2001). All results were parsed by Trinotate, stored in an SQLite relational database, and then reported as a tab-delimited transcript annotation summary file (Data S3). Gene Ontology identifiers were transitively assigned to transcripts based on available GO annotations of best-matching Swiss-Prot entries (Data S4 and S5).

Transcript Abundance Estimation and Differential Expression Analysis

RSEM (Li and Dewey, 2011) and MMSEQ (Turro et al., 2011) software were used to estimate expression values for transcripts, and edgeR (Robinson et al., 2010), EBSeq (Leng et al., 2013), and mmdiff (Turro et al., 2014) were used to identify significantly differentially expressed transcripts. Further details can be found in Supplemental Experimental Procedures.

Animal Experimentation

All animal experiments were performed in accordance with Harvard Medical School's Institutional Animal Care and Use Committee regulations and in accordance with Animal Experimentation Protocol 04160. For all survival surgeries, animals were anesthetized in 0.1% tricaine and allowed to recover overnight in 0.5% sulfamerazine.

Library Preparation and Sequencing

Total RNA was purified using either TRIzol or QIAGEN RNeasy columns. The Illumina TruSeq v2 protocol was used throughout to generate bar-coded sequencing libraries. Paired-end, 100-bp sequencing was performed on the Illumina HiSeq 2500 at Harvard Medical School's Biopolymers Facility.

In Situ Hybridization

Sequences were amplified from cDNA and cloned into the pGEM-T Easy vector and sequenced. Depending upon orientation, T7 or Sp6 polymerase was used to transcribe the probe. In situ hybridization was performed as in Whited et al. (2011).

Statistical Analyses

All results from morpholino knockdown and retroviral experiments were quantified by at least one blinded observer. Blastema length and chondrification area were measured using ImageJ (NIH). A two-tailed Student's t test was used to assess statistical significance. The statistical significance of the phenotypic outcomes from retroviral overexpression of *kazald1* was assessed by performing a two-tailed Fisher's exact test on a 2 × 2 contingency table.

For detailed methods, see Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the raw sequencing reads reported in this paper is SRA: PRJNA300706 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA300706). The accession number for the transcriptome data reported in this paper is Transcriptome Shotgun Assembly Database: GFBM000000000. The accession number for all data reported in this paper is GEO: GSE92429.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, five tables, and five data files and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.12.063.

AUTHOR CONTRIBUTIONS

D.M.B., K.J., T.D., B.J.H., and J.L.W. designed the experiments; D.M.B., K.J., T.D., C.J.T., A.R., B.J.H., and J.L.W. wrote the paper; B.J.H., T.T., and M.B.C. performed the computations and developed related resources; R.M.F. and L.P. performed the embryo sequencing; and D.M.B., K.J., T.D., D.P.-D., T.J.L., N.D.L., T.-H.K., F.G.D., J.B., S.B., A.R.G., S.L.T., S.C., W.W.Y., and J.L.W. performed the sample preparations, validations, and functional experimentation. All authors contributed to manuscript editing.

ACKNOWLEDGMENTS

This work was supported by the Harvard Stem Cell Institute (C.J.T.), Brigham & Women's Hospital (J.L.W.), Richard and Susan Smith Family Foundation (J.L.W.), NIH/NICHD (1R03HD083434 and 1DP2HD087953, to J.L.W.), NIH/NICHD (R01 HD073104, to L.P.), and Howard Hughes Medical Institute (A.R.). This research was supported by the National Cancer Institute of the NIH under award 1U24CA180922-01 (to B.J.H. and A.R.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. D.M.B. was supported by an HHMI Gilliam Fellowship. We thank Borja Sese Ballesteros and Yick Fong for their helpful discussions on using CRISPR-Cas9; Jana Hersch for germ cell identification; Josh Gorham and Jim Pancoast for assistance in RNA extraction; Harvard Biopolymers for sequencing; Esther Pearl and Marko Horb for sequencing the embryo samples; Rich Lee for use of shared equipment; Leslie Gaffney for help with figures; and Benjamin Lewis for discussions. A subset of computations in this paper was run on the Orchestra cluster supported by the Research Information Technology Group at Harvard Medical School. We thank the Ambystoma Genetic Stock Center for providing some of the animals (Lexington, KY; NIH grant P40-OD019794).

Received: May 25, 2016
Revised: October 26, 2016
Accepted: December 20, 2016
Published: January 17, 2017

REFERENCES

Abdullayev, I., Kirkham, M., Björklund, A.K., Simon, A., and Sandberg, R. (2013). A reference transcriptome and inferred proteome for the salamander *Notophthalmus viridescens*. Exp. Cell Res. *319*, 1187–1197.

Athippozhy, A., Lehrberg, J., Monaghan, J.R., Gardiner, D.M., and Voss, S.R. (2014). Characterization of in vitro transcriptional responses of dorsal root ganglia cultured in the presence and absence of blastema cells from regenerating salamander limbs. Regeneration (Oxf) *1*, 1–10.

Beauchemin, M., Del Rio-Tsonis, K., Tsonis, P.A., Tremblay, M., and Savard, P. (1998). Graded expression of Emx-2 in the adult newt limb and its corresponding regeneration blastema. J. Mol. Biol. *279*, 501–511.

Boeckmann, B., Blatter, M.C., Famiglietti, L., Hinz, U., Lane, L., Roechert, B., and Bairoch, A. (2005). Protein variety and functional diversity: Swiss-Prot annotation in its biological context. C. R. Biol. *328*, 882–899.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.

Chera, S., de Rosa, R., Miljkovic-Licina, M., Dobretz, K., Ghila, L., Kaloulis, K., and Galliot, B. (2006). Silencing of the hydra serine protease inhibitor Kazal1 gene mimics the human SPINK1 pancreatic phenotype. J. Cell Sci. *119*, 846–857.

Cobb, J., Dierich, A., Huss-Garcia, Y., and Duboule, D. (2006). A mouse model for human short-stature syndromes identifies Shox2 as an upstream regulator of Runx2 during long-bone development. Proc. Natl. Acad. Sci. USA *103*, 4511–4515.

Crawford, K., and Stocum, D.L. (1988a). Retinoic acid coordinately proximalizes regenerate pattern and blastema differential affinity in axolotl limbs. Development *102*, 687–698.

Crawford, K., and Stocum, D.L. (1988b). Retinoic acid proximalizes level-specific properties responsible for intercalary regeneration in axolotl limbs. Development *104*, 703–712.

Danno, S., Nishiyama, H., Higashitsuji, H., Yokoi, H., Xue, J.H., Itoh, K., Matsuda, T., and Fujita, J. (1997). Increased transcript level of RBM3, a member of the glycine-rich RNA-binding protein family, in human cells in response to cold stress. Biochem. Biophys. Res. Commun. *236*, 804–807.

Drach, J., Zhao, S., Malavasi, F., and Mehta, K. (1993). Rapid induction of CD38 antigen on myeloid leukemia cells by all *trans*-retinoic acid. Biochem. Biophys. Res. Commun. *195*, 545–550.

Echeverri, K., and Tanaka, E.M. (2005). Proximodistal patterning during limb regeneration. Dev. Biol. *279*, 391–401.

Eto, K., Eda, K., Hayano, M., Goto, S., Nagao, K., Kawasaki, T., Kashimura, H., Tarui, H., Nishimura, O., Agata, K., and Abe, S. (2009). Reduced expression of an RNA-binding protein by prolactin leads to translational silencing of programmed cell death protein 4 and apoptosis in newt spermatogonia. J. Biol. Chem. *284*, 23260–23271.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., et al. (2014). Pfam: the protein families database. Nucleic Acids Res. *42*, D222–D230.

Fradet, Y. (2009). Biomarkers in prostate cancer diagnosis and prognosis: beyond prostate-specific antigen. Curr. Opin. Urol. *19*, 243–246.

Goto, M., and Eddy, E.M. (2004). Speriolin is a novel spermatogenic cell-specific centrosomal protein associated with the seventh WD motif of Cdc20. J. Biol. Chem. *279*, 42128–42138.

Goto, M., O'Brien, D.A., and Eddy, E.M. (2010). Speriolin is a novel human and mouse sperm centrosome protein. Hum. Reprod. *25*, 1884–1894.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Haack, H., and Gruss, P. (1993). The establishment of murine Hox-1 expression domains during patterning of the limb. Dev. Biol. *157*, 410–422.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. *8*, 1494–1512.

Jiang, D., Zhu, X.L., Zhao, J.F., Zhou, Y.K., Zhong, C., Zhang, J., and Huang, X. (2014). Subtractive screen of potential limb regeneration related genes from *Pachytriton brevipes*. Mol. Biol. Rep. *41*, 1015–1026.

Jiang, P., Nelson, J.D., Leng, N., Collins, M., Swanson, S., Dewey, C.N., Thomson, J.A., and Stewart, R. (2016). Analysis of embryonic development in the unsequenced axolotl: waves of transcriptomic upheaval and stability. Dev. Biol. Published online July 27, 2016. http://dx.doi.org/10.1016/j.ydbio.2016.05.024.

Kafasla, P., Skliris, A., and Kontoyiannis, D.L. (2014). Post-transcriptional coordination of immunological responses by RNA-binding proteins. Nat. Immunol. *15*, 492–502.

Keinath, M.C., Timoshevskiy, V.A., Timoshevskaya, N.Y., Tsonis, P.A., Voss, S.R., and Smith, J.J. (2015). Initial characterization of the large genome of the salamander *Ambystoma mexicanum* using shotgun and laser capture chromosome sequencing. Sci. Rep. *5*, 16413.

Khan, P., Linkhart, B., and Simon, H.G. (2002). Different regulation of T-box genes Tbx4 and Tbx5 during limb development and limb regeneration. Dev. Biol. *250*, 383–392.

Knapp, D., and Tanaka, E.M. (2012). Regeneration and reprogramming. Curr. Opin. Genet. Dev. *22*, 485–493.

Knapp, D., Schulz, H., Rascon, C.A., Volkmer, M., Scholz, J., Nacu, E., Le, M., Novozhilov, S., Tazaki, A., Protze, S., et al. (2013). Comparative transcriptional profiling of the axolotl limb identifies a tripartite regeneration-specific gene program. PLoS ONE *8*, e61352.

Kragl, M., Knapp, D., Nacu, E., Khattak, S., Maden, M., Epperlein, H.H., and Tanaka, E.M. (2009). Cells keep a memory of their tissue origin during axolotl limb regeneration. Nature *460*, 60–65.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. *305*, 567–580.

Legan, P.K., Goodyear, R.J., Morín, M., Mencia, A., Pollard, H., Olavarrieta, L., Korchagina, J., Modamio-Hoybjor, S., Mayo, F., Moreno, F., et al. (2014). Three deaf mice: mouse models for TECTA-based human hereditary deafness reveal domain-specific structural phenotypes in the tectorial membrane. Hum. Mol. Genet. *23*, 2551–2568.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035–1043.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Li, B., Fillmore, N., Bai, Y., Collins, M., Thomson, J.A., Stewart, R., and Dewey, C.N. (2014). Evaluation of de novo transcriptome assemblies from RNA-seq data. Genome Biol. *15*, 553.

Looso, M., Preussner, J., Sousounis, K., Bruckskotten, M., Michel, C.S., Lignelli, E., Reinhardt, R., Höffner, S., Krüger, M., Tsonis, P.A., et al. (2013). A de novo assembly of the newt transcriptome combined with proteomic validation identifies new protein families expressed during tissue regeneration. Genome Biol. *14*, R16.

Maden, M. (1983). The effect of vitamin A on the regenerating axolotl limb. J. Embryol. Exp. Morphol. *77*, 273–295.

McCusker, C.D., and Gardiner, D.M. (2013). Positional information is reprogrammed in blastema cells of the regenerating limb of the axolotl (*Ambystoma mexicanum*). PLoS ONE *8*, e77064.

McCusker, C.D., Athippozhy, A., Diaz-Castillo, C., Fowlkes, C., Gardiner, D.M., and Voss, S.R. (2015). Positional plasticity in regenerating *Amybstoma mexicanum* limbs is associated with cell proliferation and pathways of cellular differentiation. BMC Dev. Biol. *15*, 45.

Mercader, N., Tanaka, E.M., and Torres, M. (2005). Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. Development *132*, 4131–4142.

Monaghan, J.R., Epp, L.G., Putta, S., Page, R.B., Walker, J.A., Beachy, C.K., Zhu, W., Pao, G.M., Verma, I.M., Hunter, T., et al. (2009). Microarray and cDNA sequence analysis of transcription during nerve-dependent limb regeneration. BMC Biol. *7*, 1.

Montero, J.A., and Hurlé, J.M. (2010). Sculpturing digit shape by cell death. Apoptosis *15*, 365–375.

Morf, J., Rey, G., Schneider, K., Stratmann, M., Fujita, J., Naef, F., and Schibler, U. (2012). Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally. Science *338*, 379–383.

Muneoka, K., Fox, W.F., and Bryant, S.V. (1986). Cellular contribution from dermis and cartilage to the regenerating limb blastema in axolotls. Dev. Biol. *116*, 256–260.

Nakamura, K., Islam, M.R., Takayanagi, M., Yasumuro, H., Inami, W., Kunahong, A., Casco-Robles, R.M., Toyama, F., and Chiba, C. (2014). A transcriptome for the study of early processes of retinal regeneration in the adult newt, *Cynops pyrrhogaster*. PLoS ONE *9*, e109831.

Nishiyama, H., Higashitsuji, H., Yokoi, H., Itoh, K., Danno, S., Matsuda, T., and Fujita, J. (1997). Cloning and characterization of human CIRP (cold-inducible RNA-binding protein) cDNA and chromosomal assignment of the gene. Gene *204*, 115–120.

Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics *23*, 1061–1067.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods *8*, 785–786.

Pomerantz, J.H., and Blau, H.M. (2013). Tumor suppressors: enhancers or suppressors of regeneration? Development *140*, 2502–2512.

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S.D., Mungall, K., Lee, S., Okada, H.M., Qian, J.Q., et al. (2010). De novo assembly and analysis of RNA-seq data. Nat. Methods *7*, 909–912.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sandoval-Guzmán, T., Wang, H., Khattak, S., Schuez, M., Roensch, K., Nacu, E., Tazaki, A., Joven, A., Tanaka, E.M., and Simon, A. (2014). Fundamental differences in dedifferentiation and stem cell recruitment during skeletal muscle regeneration in two salamander species. Cell Stem Cell *14*, 174–187.

Satoh, A., Bryant, S.V., and Gardiner, D.M. (2008). Regulation of dermal fibroblast dedifferentiation and redifferentiation during wound healing and limb regeneration in the axolotl. Dev. Growth Differ. *50*, 743–754.

Schulz, M.H., Zerbino, D.R., Vingron, M., and Birney, E. (2012). Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics *28*, 1086–1092.

Shimomura, Y., Agalliu, D., Vonica, A., Luria, V., Wajid, M., Baumer, A., Belli, S., Petukhova, L., Schinzel, A., Brivanlou, A.H., et al. (2010). APCDD1 is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. Nature *464*, 1043–1047.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics *31*, 3210–3212.

Smith, J.J., Putta, S., Zhu, W., Pao, G.M., Verma, I.M., Hunter, T., Bryant, S.V., Gardiner, D.M., Harkins, T.T., and Voss, S.R. (2009). Genic regions of a large salamander genome contain long introns and novel genes. BMC Genomics *10*, 19.

Stewart, R., Rascón, C.A., Tian, S., Nie, J., Barry, C., Chu, L.F., Ardalani, H., Wagner, R.J., Probasco, M.D., Bolin, J.M., et al. (2013). Comparative RNA-seq analysis in the unsequenced axolotl: the oncogene burst highlights early gene expression in the blastema. PLoS Comput. Biol. *9*, e1002936.

Stocum, D.L., and Melton, D.A. (1977). Self-organizational capacity of distally transplanted limb regeneration blastemas in larval salamanders. J. Exp. Zool. *201*, 451–461.

Straus, N.A. (1971). Comparative DNA renaturation kinetics in amphibians. Proc. Natl. Acad. Sci. USA *68*, 799–802.

Torok, M.A., Gardiner, D.M., Shubin, N.H., and Bryant, S.V. (1998). Expression of HoxD genes in developing and regenerating axolotl limbs. Dev. Biol. *200*, 225–233.

Torok, M.A., Gardiner, D.M., Izpisúa-Belmonte, J.C., and Bryant, S.V. (1999). Sonic hedgehog (shh) expression in developing and regenerating axolotl limbs. J. Exp. Zool. *284*, 197–206.

Turro, E., Su, S.Y., Gonçalves, Â., Coin, L.J., Richardson, S., and Lewin, A. (2011). Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. Genome Biol. *12*, R13.

Turro, E., Astle, W.J., and Tavaré, S. (2014). Flexible analysis of RNA-seq data using mixed effects models. Bioinformatics *30*, 180–188.

UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204–D212.

Voss, S.R., Palumbo, A., Nagarajan, R., Gardiner, D.M., Muneoka, K., Stromberg, A.J., and Athippozhy, A.T. (2015). Gene expression during the first 28 days of axolotl limb regeneration I: experimental design and global analysis of gene expression. Regeneration (Oxf) *2*, 120–136.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wei, W., Graeff, R., and Yue, J. (2014). Roles and mechanisms of the CD38/cyclic adenosine diphosphate ribose/Ca(2+) signaling pathway. World J. Biol. Chem. *5*, 58–67.

Wellmann, S., Bührer, C., Moderegger, E., Zelmer, A., Kirschner, R., Koehne, P., Fujita, J., and Seeger, K. (2004). Oxygen-regulated expression of the RNA-binding proteins RBM3 and CIRP by a HIF-1-independent mechanism. J. Cell Sci. *117*, 1785–1794.

Whited, J.L., and Tabin, C.J. (2009). Limb regeneration revisited. J. Biol. *8*, 5.

Whited, J.L., Lehoczky, J.A., Austin, C.A., and Tabin, C.J. (2011). Dynamic expression of two thrombospondins during axolotl limb regeneration. Dev. Dyn. *240*, 1249–1258.

Whited, J.L., Tsai, S.L., Beier, K.T., White, J.N., Piekarski, N., Hanken, J., Cepko, C.L., and Tabin, C.J. (2013). Pseudotyped retroviruses for infecting axolotl in vivo and in vitro. Development *140*, 1137–1146.

Wu, C.H., Tsai, M.H., Ho, C.C., Chen, C.Y., and Lee, H.S. (2013). De novo transcriptome sequencing of axolotl blastema for identification of differentially expressed genes during limb regeneration. BMC Genomics *14*, 434.

Yang, E.V., Gardiner, D.M., Carlson, M.R., Nugas, C.A., and Bryant, S.V. (1999). Expression of Mmp-9 and related matrix metalloproteinase genes during axolotl limb regeneration. Dev. Dyn. *216*, 2–9.

Zakin, L., and De Robertis, E.M. (2010). Extracellular regulation of BMP signaling. Curr Biol. *20*, R89–92.

Zhang, L., Prak, L., Rayon-Estrada, V., Thiru, P., Flygare, J., Lim, B., and Lodish, H.F. (2013). ZFP36L2 is required for self-renewal of early burst-forming unit erythroid progenitors. Nature *499*, 92–96.

Zhang, H.T., Xue, J.H., Zhang, Z.W., Kong, H.B., Liu, A.J., Li, S.C., and Xu, D.G. (2015). Cold-inducible RNA-binding protein inhibits neuron apoptosis through the suppression of mitochondrial apoptosis. Brain Res. *1622*, 474–483.

Zhu, W., Pao, G.M., Satoh, A., Cummings, G., Monaghan, J.R., Harkins, T.T., Bryant, S.V., Voss, S.R., Gardiner, D.M., and Hunter, T. (2012). Activation of germline-specific genes is required for limb regeneration in the Mexican axolotl. Dev. Biol. *370*, 42–51.

Ziegler-Graham, K., MacKenzie, E.J., Ephraim, P.L., Travison, T.G., and Brookmeyer, R. (2008). Estimating the prevalence of limb loss in the United States: 2005 to 2050. Arch. Phys. Med. Rehabil. *89*, 422–429.

Cell Reports, Volume *18*

Supplemental Information

A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors

Donald M. Bryant, Kimberly Johnson, Tia DiTommaso, Timothy Tickle, Matthew Brian Couger, Duygu Payzin-Dogru, Tae J. Lee, Nicholas D. Leigh, Tzu-Hsing Kuo, Francis G. Davis, Joel Bateman, Sevara Bryant, Anna R. Guzikowski, Stephanie L. Tsai, Steven Coyne, William W. Ye, Robert M. Freeman, Jr., Leonid Peshkin, Clifford J. Tabin, Aviv Regev, Brian J. Haas, and Jessica L. Whited

Supplemental Information

Supplemental Figures

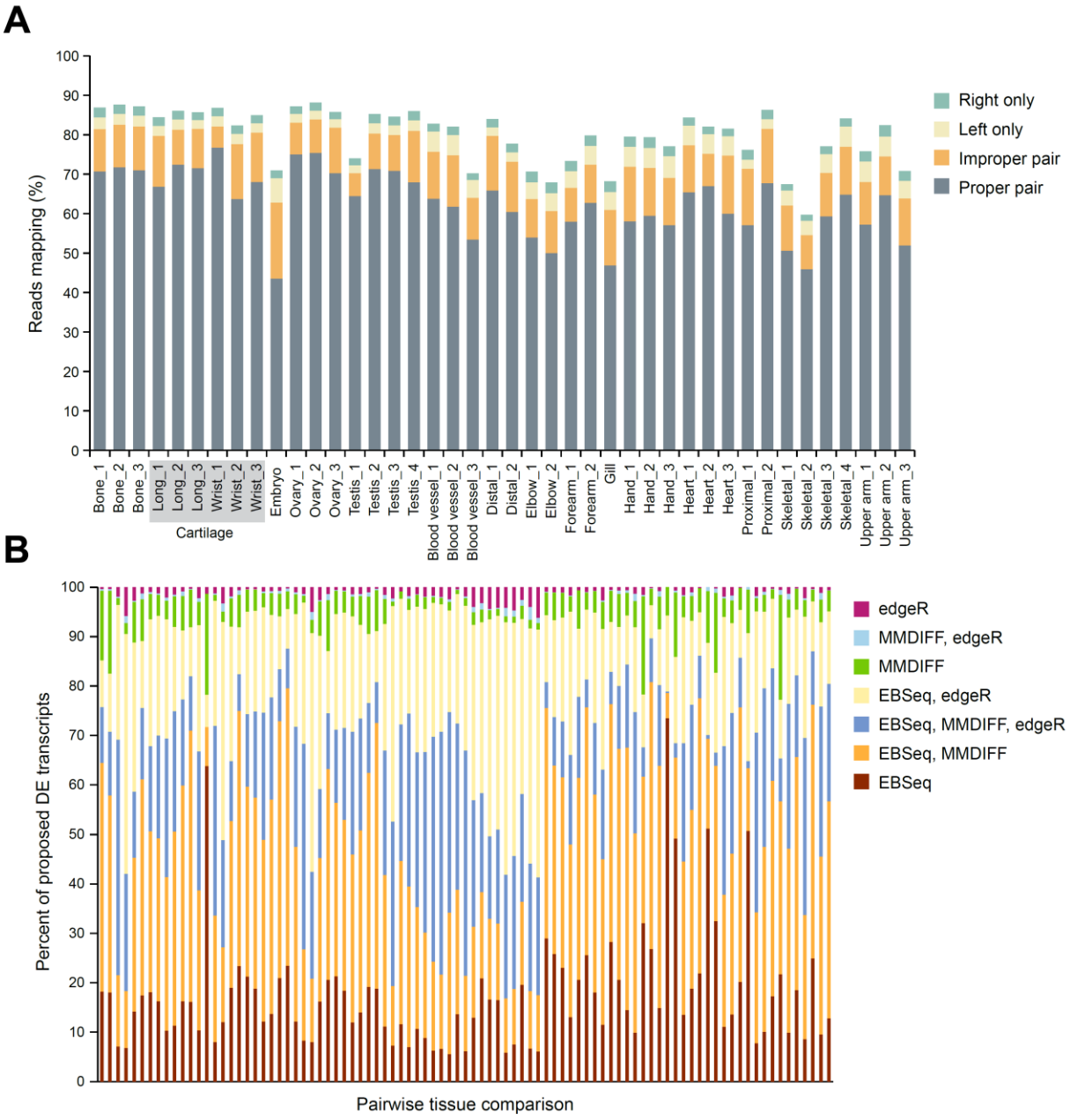


Figure S1. Transcriptome read representation and comparison of methods for differential gene expression analysis. Related to Figure 1 and Table 1. (A) Approximately 80% of RNA-Seq reads from each sample and replicate are represented by the transcriptome assembly. (B) Most transcripts identified as significantly differentially expressed were agreed upon by at least two different computational methods.

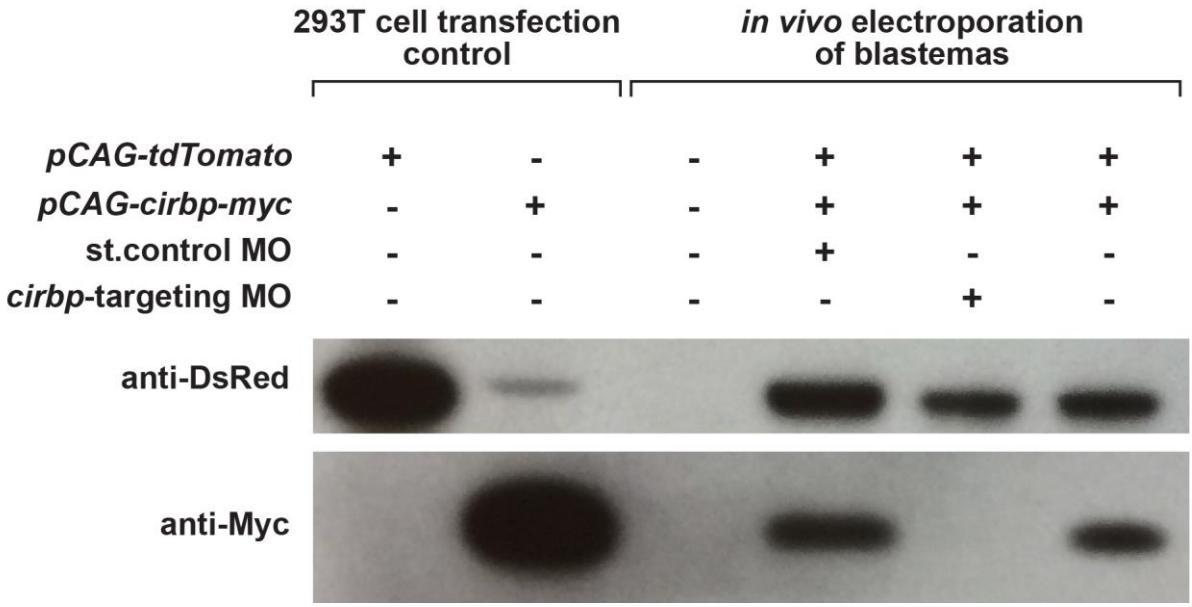


Figure S2. Administration of *cirbp*-targeting morpholino diminishes CIRBP protein level in axolotl tissue. Related to Figure 4. Shown is a western blot labeled with anti-DsRed (top) and anti-myc (bottom) antibodies. On the left are protein samples from HEK293T cells transfected with DNA constructs as shown for validation of protein size and as positive controls. On the right are protein samples from blastemas electroporated with DNA constructs and either the standard control morpholino (st. control MO) or the *cirbp*-targeting morpholino (*cirbp*-targeting MO). Each lane was loaded with 2 μg of total protein.

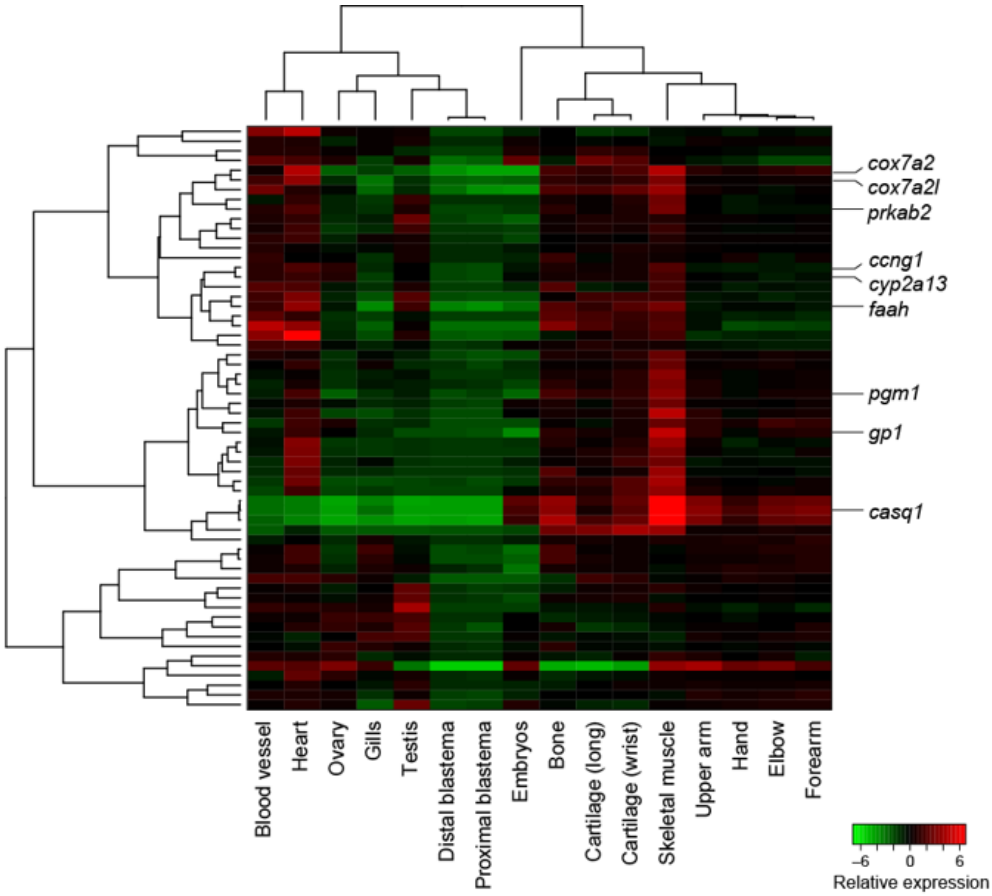


Figure S3. Transcripts found globally repressed in blastemal tissue. Related to Figure 3. Many of the transcripts repressed in blastemal tissue are found highly expressed in skeletal muscle tissue.

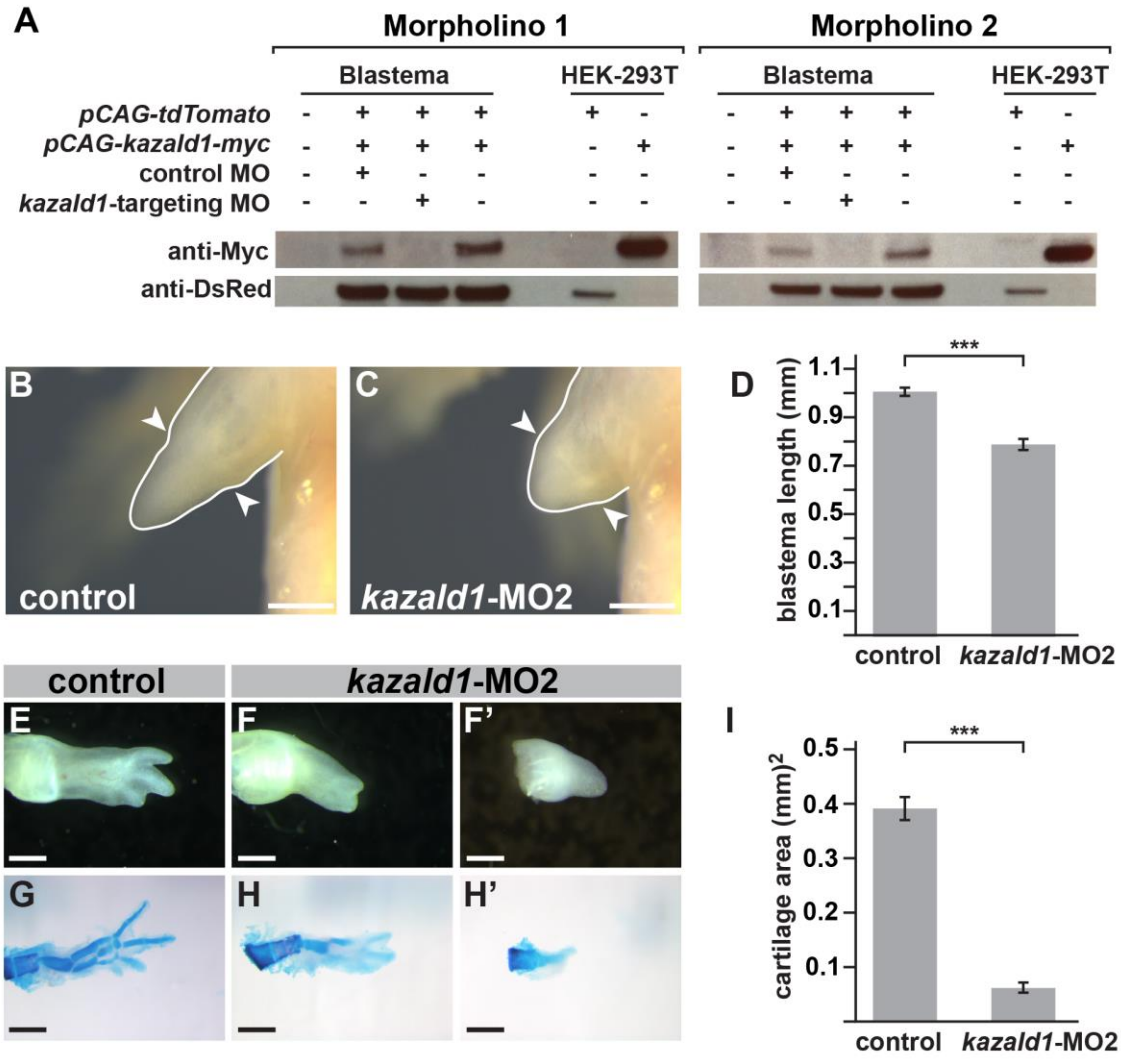


Figure S4. Administration of *kazald1*-targeting morpholinos diminish Kazd1 protein level in blastema cells and a second *kazald1*-targeting morpholino also disrupts regeneration. Related to Figure 5. (A) Shown are western blots probed with anti-myc (top) and anti-DsRed (bottom). Left blot is control morpholino (MO) and *kazald1*-targeting MO for targeting sequence 1, right blot is control MO2 and *kazald1*-targeting MO2; left four lanes are protein samples from limb blastemas electroporated with DNA constructs shown and with the *kazald1*-targeting MO or the related control MO (inverted sequence). Empty lane separates blastema samples from 293T cell samples. 293T cell protein samples transfected with individual DNA constructs were used as the positive controls and to verify size. (B-D) MO2 administration impairs blastema growth. (B-C) Regenerating limbs at 19 days post-amputation treated with control (B, inverted MO2 sequence) or a second *kazald1*-targeting morpholino (*kazald1*-MO2) (C); quantified in (D). (D-H) MO2 administration causes a delay in chondrification. (D-G') Whole-mount brightfield images (top row, E-F') of regenerating limbs with Alcian blue-stained skeleton pictured below (G-H'). Limbs were harvested at 28 days post-amputation. (I) Quantification of cartilage area in control versus *kazald1*-MO2-treated. Control refers to inverted MO sequence. Scale bar is 1 mm. *** indicates $p<0.001$ and error bars are SEM. Arrowheads mark amputation plane in each image.

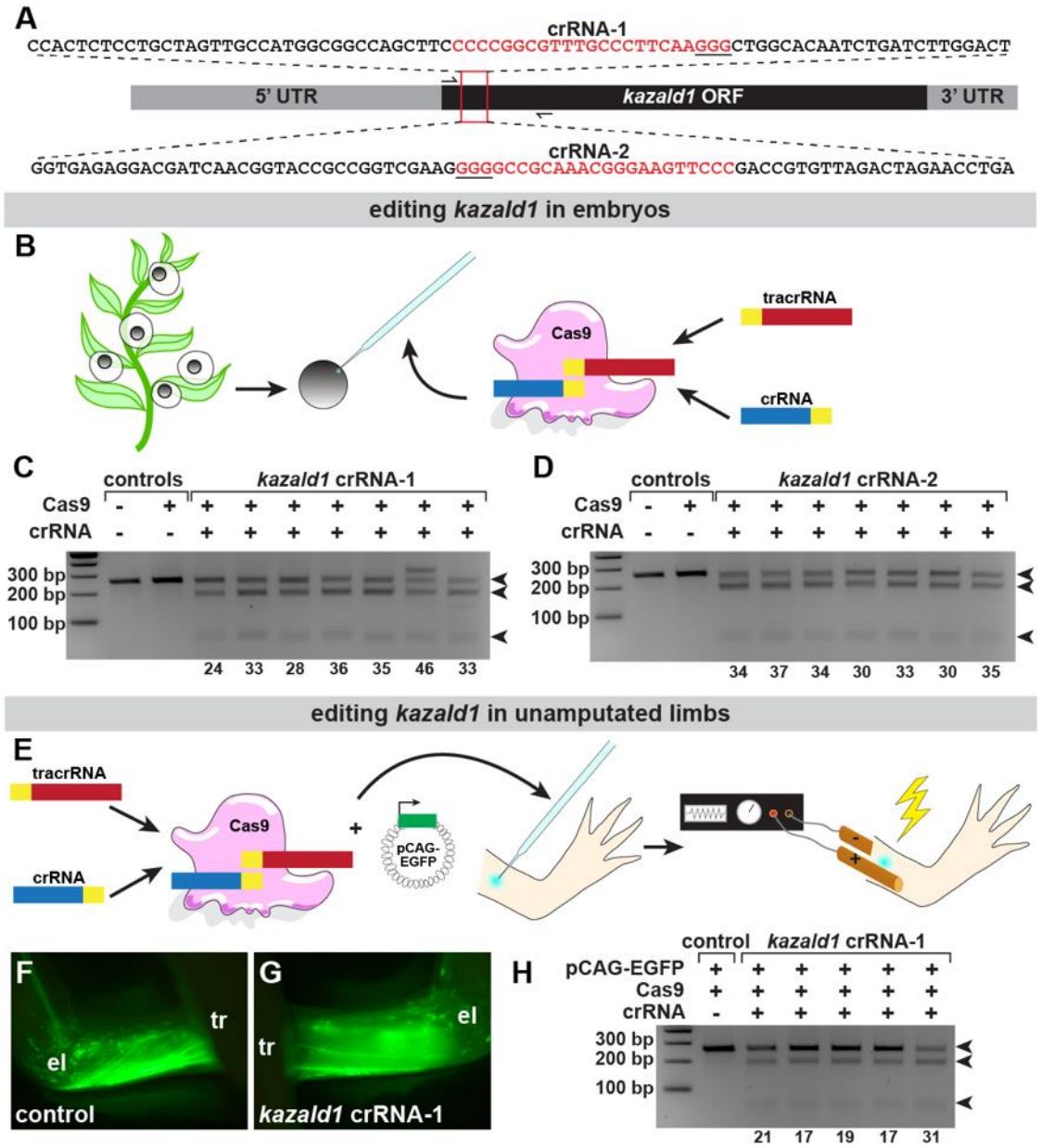


Figure S5. CRISPR-mediated deletions in *kazald1* locus in embryos and unamputated limbs. Related to Figure 5. (A) *Kazald1* genomic locus and targeting scheme. Primers for PCR amplification for genotyping are indicated. (B) Experimental schematic for targeting in embryos via injection. (C-D) Embryo genotyping via PCR followed by T7 endonuclease assay for embryos injected with crRNA-1 (C) and crRNA-2 (D). (E) Experimental schematic for targeting in unamputated limbs via electroporation. (F-G) EGFP expression in control (F, left limb) and crRNA-1-treated (G, right limb) limbs at 25 days post-electroporation. Landmarks: el (elbow), tr (trunk). (H) Limb genotyping via PCR followed by T7 endonuclease assay. For all genotyping, each lane indicates an individual specimen. Editing efficiency is estimated by the indel percentage beneath each lane.

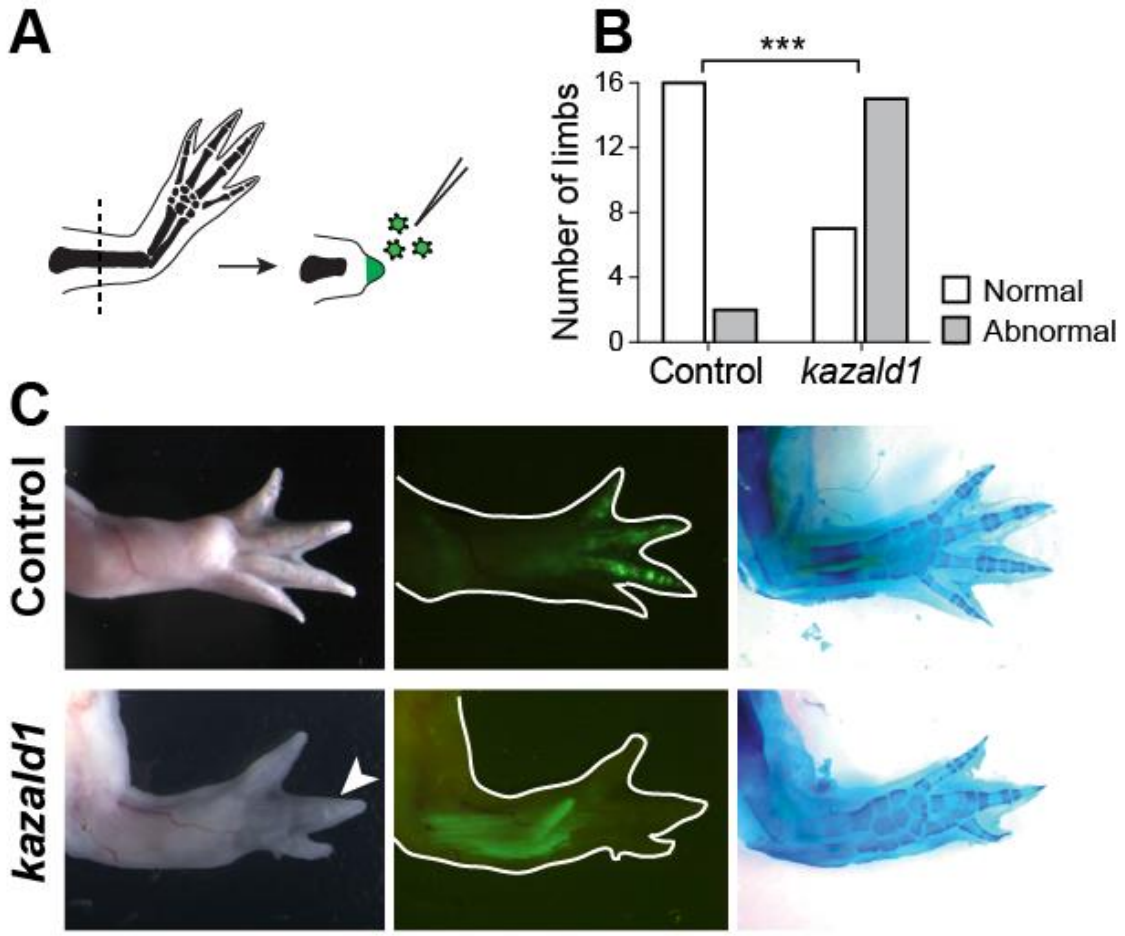


Figure S6. Expression of the most robust blastema marker must be temporally or spatially restricted to achieve perfect regeneration. Related to Figure 5. (A) Experimental setup for virus-driven constitutive misexpression. (B) Prolonged misexpression of *kazald1* causes profound regenerative defects as compared to controls (n=22 experimental, 18 control; p=0.004 with Fisher's exact test). (C) The most common defect observed was syndactyly (indicated by arrow). Scale bar is 500 microns.

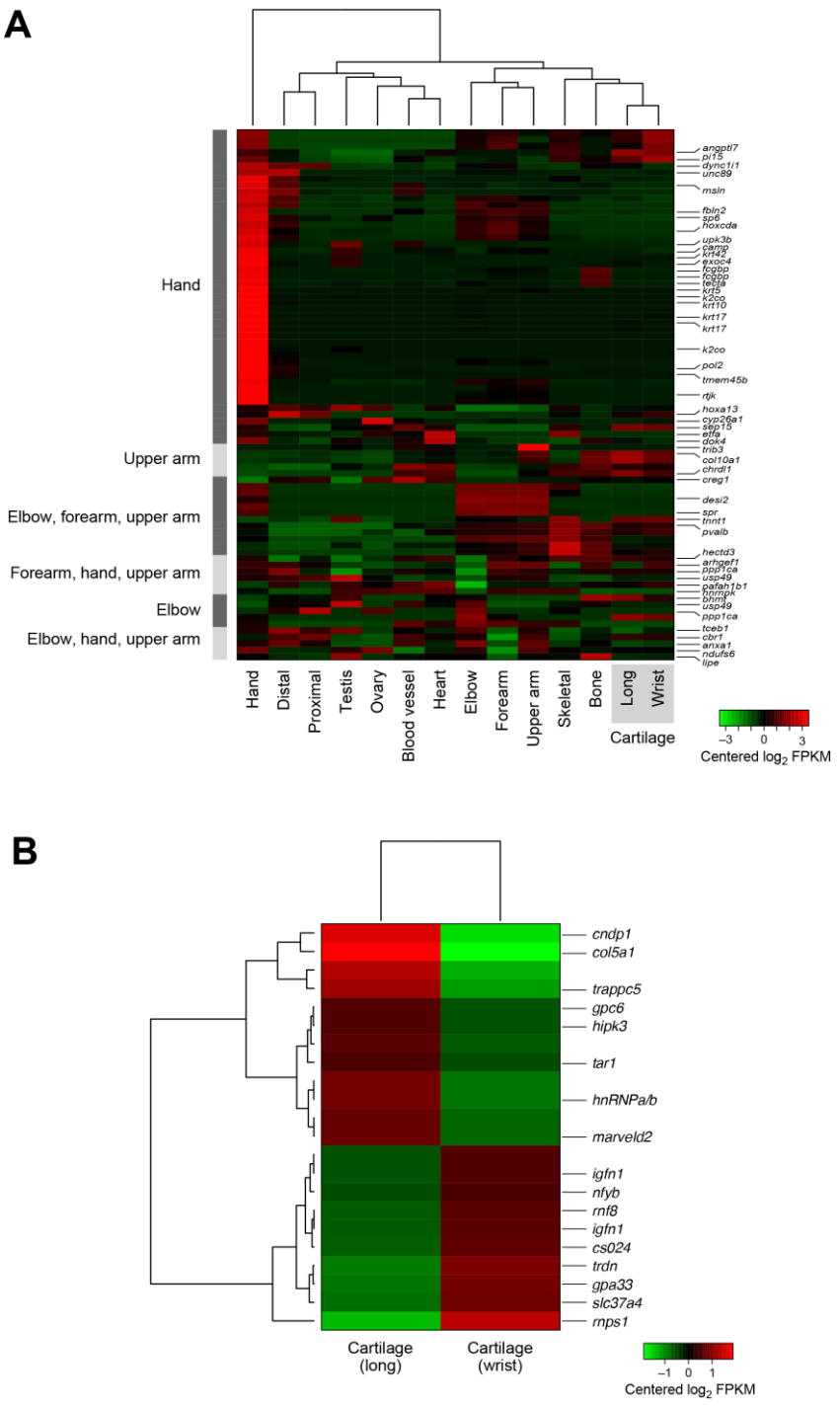


Figure S7. (A) Expression of arm-gradient enriched transcripts across all tissue types. Related to Figure 2 and Figure 6. Many arm segment enriched transcripts have transcriptional enrichment that largely reflects the tissue composition, such as relative skeletal muscle or cartilage content. Hand stands out as having a unique transcriptional program with many hand-specific transcripts being expressed. (B) A paucity of differentially expressed transcripts identified between wrist and long bone cartilage. Differences in expression between these two tissue types are subtle, and this bulk analysis of these two very similar tissue types likely simply lacks the resolution required to reveal key molecular differences between them.

Supplemental Tables

Table S1. Axolotl fastq read counts and Trinity assembly statistics. Related to Table 1 and Figure 1.

Table S2. Tissue-enriched transcripts and Gene Ontology enrichment or depletion. Related to Figure 2.

Table S3. Differential isoform usage. Related to Figure 2.

Table S4. Blastema-enriched transcripts. Related to Figure 3.

Table S5. Arm-segment enrichment patterns, arm gradient correlations, and transcripts exhibiting enrichment in distal vs. proximal blastemas and vice versa. Related to Figure 6.

Supplemental Data Files

Data S1. Axolotl Trinity *de novo* assembly Part 1. Related to Figures 1-6 and Table 1. Data S1 includes the first half of our transcriptome and can be concatenated with Data S2 to reconstruct the full transcriptome. File name: File_S1A_Axolotl.TrinityAssembly.Part1.fasta.gz.

Data S2. Axolotl Trinity *de novo* assembly Part 2. Related to Figures 1-6 and Table 1. Data S2 includes the second half of our transcriptome and can be concatenated with Data S1 to reconstruct the full transcriptome. File name: File_S1A_Axolotl.TrinityAssembly.Part2.fasta.gz.

Data S3. Trinotate annotations. Related to Figures 1-6 and Table 1.
File name: File_S1B_Axo.Mar2014.Trinotate.xls.gz.

Data S4. Differential expression results, gene ontology full results, raw transcript abundance counts produced with RSEM, and normalized transcript abundance expressed as FPKM. Related to Figures 1-3, Figure 5, and Figure 6. Data S4 consists of a zipped archive that contains the full list of differentially expressed genes identified by multiple differential expression analysis programs (included as "File_S2A_.Axolotl.TrinityAssembly.DiffExpressionMin2ProgsAgree.dat.gz"), full gene ontology results (included as "File_S2B_.Axolotl.TrinityAssembly.GeneOntology.dat.gz"), raw transcript counts produced by RSEM (included as "File_S2C_.Axolotl.TrinityAssembly.RSEM_counts.matrix.gz"), and normalized transcript abundances expressed as FPKM (included as "File_S2D_.Axolotl.TrinityAssembly.TMM_normalized_FPKM.matrix.gz"). In order to be included in this list of differential expression results, a transcript needed to be classified as differentially expressed by at least two differential expression algorithms.

Data S5. Transcripts identified as differentially expressed by edgeR, EBSeq, and MMDiff. Related to Figures 1-3, Figure 5, and Figure 6. Data S5 consists of a zipped archive that contains the full list of differentially expressed transcripts identified by edgeR (included as "File_S2E_.edgeR.trans.pairwise_summary.FDR0.05.dat.gz"), EBSeq (included as "File_S2F_.EBSeq.pairwise.trans.post0.95.dat.gz"), and MMDiff (included as "File_S2G_.MMDIFF.trans.pairwise_summary.PostProb0.2.dat.gz").

Supplemental Experimental Procedures

Transcript abundance estimation and differential expression analysis

Counts of RNA-Seq fragments mapping to transcripts were determined for each tissue type using the RSEM software (Li and Dewey, 2011) (estimated RNA-Seq fragment counts per transcript and TMM-normalized FPKM values are provided in **Data S4**). The RSEM abundance estimates were input into edgeR (Robinson et al., 2010) and EBSeq (Leng et al., 2013) tools to identify differentially expressed transcripts in each of the pairwise tissue comparisons, employing a false discovery threshold of 0.05 (**Data S5**). To further complement our DE analyses, we separately estimated expression values and identified differentially expressed transcripts using mmseq (Turro et al., 2011) and mmdiff (Turro et al., 2014), respectively. Transcripts defined by mmdiff as having a posterior probability of differential isoform usage >= 0.2 were captured as differentially expressed (**Data S5**). Transcripts found to be significantly differentially expressed by at least two of the DE analysis methods and to have

at least a 2-fold change in expression were retained as candidate differentially expressed transcripts (**Data S4**) and subject to further analysis.

Computing the transcript contig E-statistic

Transcripts are ordered by expression values, and for each top-most *X* number of transcripts, the expression values are summed and the percentage of total expression represented by that set of transcripts is stated as the *E* percentage count of transcripts. For example, if the top *X* most highly expressed transcripts corresponds to 90% of the total expression data (by summing FPKM values), then E90 = *X*. The contig N50 statistic (Miller et al., 2010) is then computed for that corresponding set of transcripts. The contig N50 for the E90 set of transcripts is referred to as the E90N50 value. A script in the Trinity software suite is available for computing these transcriptome statistics and can be found as ${TRINITY_HOME}/util/misc/contig_ExN50_statistic.pl .

Identification of tissue-enriched transcripts

For each differentially expressed gene, a directed graph data structure was constructed with tissue types as nodes and directions of significant differential expression encoded by edges; an edge being drawn from up-regulated tissue *A* to down-regulated tissue *B* (for example, see **Figure 2A**). Those directed graphs that represented all tissue types were selected as best candidates for further study of tissue-enriched expression. Patterns of tissue-enriched expression were defined as collections of up- or down-regulated tissues, and genes were partitioned accordingly (**Table S2**). Genes exhibiting patterns of differential isoform usage were identified as those encoding at least two transcript isoforms, each significantly differentially expressed between at least two tissues, and demonstrating different patterns of tissue enrichment (**Table S3**). Tissue types for gill filament and embryo were excluded here due to not having biological replicates required for reliable detection of differentially expressed transcripts. The software we developed for identifying tissue-enriched transcripts is included in our Trinity software tool suite (https://github.com/trinityrnaseq/trinityrnaseq/tree/master/Analysis/DifferentialExpression/TissueEnrichment). Statistical enrichment of GO terms for groups of differentially expressed transcripts was performed using GO-Seq (Young et al., 2010).

Identification of transcripts expressed along a positional gradient in the limb

Transcript expression (TMM-normalized FPKM) values for arm segments were examined for significant Pearson correlation with ordered position as follows. Arm segments (upper arm, elbow, forearm, and hand) were assigned ordinal positions (1, 2, 3, 4). Those transcripts having at least a 2-fold difference in expression between the most distal positions (upper arm and hand) and having an expression of at least 1 TMM-normalized FPKM in any arm segment RNA-Seq replicate were tested for significant Pearson correlation with their ordinal positions using the *cor.test()* function in R (**Table S5)**. Multiple testing corrections were performed via the Q-value method (Storey and Tibshirani, 2003)).

RNA Isolation and RT-PCR

Tissues were harvested while animals were deeply anesthetized (for limb specimens) or after they had been sacrificed (for all other specimens). RNA (from tissues described below) was isolated either by Trizol extraction or by Qiagen RNeasy column for either library preparation or RT-PCR. For bone, ossified central areas of the humerus, radius, ulna, and metacarpals were isolated and immediately snap-frozen in liquid nitrogen. A similar approach was used for cartilage. Mortar and pestle were used to pulverize the tissue, which was immediately transferred to lysis buffer or Trizol. Soft tissues and embryos were placed directly in lysis buffer or Trizol. Blastemas were harvested at 23 days post-amputation (medium bud stage), and wound epidermis was manually removed prior to homogenization. DNase digestion (Qiagen RNase-free DNase) of genomic DNA, if required, was performed either on-column (for RNA purified by column) or following pellet resuspension (for RNA purified by Trizol). Following DNase treatment in suspension, samples were cleaned up by RNeasy MinElute column. RNA concentration was measured by spectrophotometer, and cDNA construction was performed using High Capacity cDNA Reverse Transcription Kit (Life Technologies) with template RNA seeded at 300 ng/20 µL reaction. Water was used in place of reverse transcriptase in the no-RT control reactions. Resulting cDNA was diluted 1:2 and used to seed RT-PCR reactions (1 µL in a 25 µL reaction). Standard NEB Taq was used for amplification, and RT-PCR reactions were conducted with 35 cycles. The entire reaction was loaded onto the gel. All RT-PCR results represent at least three individual biological replicates.

Morpholino design and delivery

Morpholinos were designed and synthesized by GeneTools. The *cirbp* targeting morpholino (5'-TTAGTTCTCTCGGATCAGGAGCACT-3') is predicted to bind upstream of the translational start site at -29 bp through -4 bp with respect to ATG. The standard control morpholino sequence is 5'-CCTCTTACCTCAGTTACAATTTATA-3'. The *kazald1* targeting morpholino (5'-TGGCAGCTCACAGTGACAGCCATTG-3') is predicted to bind 2 bp upstream through 23 bp coding sequence, while the control morpholino is the inverted sequence (5'-GTTACCGACAGTGACACTCGACGGT-3'). The sequence of the second *kazald1* targeting morpholino is 5'- ACAGTGCCACAAGTTTAGGTGACCA – 3', and the sequence for the corresponding alternative control morpholino is 5' – ACCAGTGGATTTGAACACCGTGACA – 3'. The second morpholino is predicted to bind upstream of the translational start site at -52 bp through -28 bp with respect to ATG. All morpholinos were conjugated to fluorescein at the 3' end to allow for electroporation and visualization. Morpholinos were reconstituted to 1 mM in 2x PBS and diluted to a working concentration of 500 µM in 1x PBS.

For *cirbp*, all four limbs of juvenile axolotls (6.0 – 8.0 cm. in length) were amputated proximally and electroporated with either PBS, standard control morpholino, or the *cirbp*-targeting morpholino. For *kazald1*, all four limbs of juvenile axolotls (5.0 – 8.0 cm. in length) were amputated proximally and electroporated with either *kazald1* targeting morpholinos or inverted control morpholinos at 10 days post-amputation (early to medium bud stage blastema). Each blastema was injected with ~0.5 - 1.0 µL of morpholino (500 µM), and all four limbs were used. For confirmation of Kazd1 and Cirbp protein knockdown, juvenile axolotls (5.0 – 6.5 cm. in length) were amputated proximally and electroporated at the early to medium bud stage (11 days post-electroporation). The vectors used for morpholino knockdown confirmation experiments (i.e. pCAG-tdTomato, pCAG-*kazald1*-myc, and pCAG-*cirbp*-myc) were diluted to a final concentration of 200 ng/µL prior to injection, and the morpholinos used were diluted to a final concentration of 500 µM. Our mock solution consisted of injection solution without vectors or morpholinos (i.e. 1x PBS with Fast Green dye to aid visualization). Following injection, electroporation was performed while animals were immersed in 1x PBS using a NepaGene Super Electroporator NEPA21 Type II electroporator. Our poring pulse consisted of 3 pulses at 150 Volts with a pulse length of 5 milliseconds, a pulse interval of 10 milliseconds, a decay rate of 0 %, and a positive (+) polarity. Our transfer pulse consisted of 5 pulses at 50 Volts with a pulse length of 50 milliseconds, a pulse interval of 950 milliseconds, a decay rate of 0 %, and a positive (+) polarity. The distance between electrodes for all electroporations was 3 millimeters.

Western blotting

Protein samples were run on a 4 – 12% Bis-Tris gel and transferred to PVDF. Membranes were blocked in 5% milk (dissolved in 0.1% Tween 20 in Tris-buffered saline) and incubated with primary antibodies overnight at 4 °C. Following overnight incubation, secondary antibodies conjugated to horseradish peroxidase were added, and reactions were visualized using the ECL method (PerkinElmer: NEL103001EA). The primary antibodies used in this study were mouse anti-c-myc (1:500 dilution; Santa Cruz Biotechnology: sc-40) and rabbit anti-dsred (1:1000; Clontech: 632496). The secondary antibodies used were goat anti-rabbit HRP (1:5000; Jackson ImmunoResearch: 111-035-003) and goat anti-mouse HRP (1:5000; Bio-Rad: 1721019).

Cell culture

Human embryonic kidney cells (HEK 293T) were used for experiments and maintained at 37°C in Dulbecco's Modified Eagle Medium (Life Technologies :11965-092) supplemented with 10% Fetal Bovine Serum. Cells were seeded approximately 24 hours prior to transfection. Immediately before transfection, cells were washed with PBS and placed into serum-free DMEM. Cells were then transfected with 1 µg of either pCAG-tdTomato, pCAG-*kazald1*-myc, or pCAG-*cirbp*-myc using polyethylenimine (PEI). Protein for western blotting was harvested at 24 hours post-transfection as described above.

Cloning and vector construction

The *kazald1* ORF was amplified from cDNA using primers 5'-CGAACAGAATTCCGCGCAATGGCTGTCACTGTGAGCTGCC -3' and 5' TGCCAAGCGGCCGCTTACAGATCTTCTTCAGAAATAAGTTTTTGTTCATCTCTATCCAGAGTTCCTTGCTC 3' (underscored letters bind *kazald1* ORF) and cloned into pCAG with EcoRI and NotI to create a version of pCAG-*kazald1*-myc that is targeted by *kazald1*-MO1. The *kazald1* ORF was amplified using primers 5' -

CGAACAGAATTCTGGTCACCTAAACTTGTGGCACTGTCCCAGAATCCTTTGCTCTACACGCGCAATGGC TGTCACTGTGAGCTGCC -3' and 5' TGCCAAGCGGCCGCTTACAGATCTTCTTCAGAAATAAGTTTTTGTTCATCTCTATCCAGAGTTCCTTGCT C 3' (underscored letters bind *kazald1* ORF) and cloned into pCAG with EcoRI and NotI to create a version of pCAG-*kazald1*-myc that is targeted by *kazald1*-MO2. The *cirbp* ORF was amplified from cDNA using primers 5'-CGAACAGAATTCAGTGCTCCTGATCCGAGAGAACTAAGTCATGTCTTCGTCAGATGAAGG -3' and 5' TGCCAAGCGGCCGCTTACAGATCTTCTTCAGAAATAAGTTTTTGTTCGTTGTCATAATAGGATCCTC -3' (underscored letters bind *cirbp* ORF) and cloned into pCAG with EcoRI and NotI to create pCAG-*cirbp*-myc.

TUNEL assay

Tissue was harvested and fixed in 4% paraformaldehyde (PFA) in PBS for 1 hr, washed in PBS, and taken through a sucrose gradient to 30% sucrose in PBS. Specimens were embedded in OCT and frozen over dry ice/ethanol bath. Sections were cut at 16 μm on the cryostat, collected on Superfrost Plus slides (Fisher), and stored at -80ºC. Slides were later equilibrated to room temperature, rehydrated in PBS, and fixed in 4% PFA for 20 min, followed by 3 x 5 min washes in PBS. They were then incubated in cold permeabilization solution (PBS with 0.1% Triton X-100, 0.1% sodium citrate) at 4ºC for two min followed by immersion in 200 mL 0.1M sodium citrate pH 6.0 in the microwave for 1 min. Slides were cooled by adding 80 mL deionized water and then rinsed in PBS. They were blocked for 30 min at room temperature with 0.1M Tris-HCl (pH 7.5) and 3% BSA and 20% FBS. They were rinsed twice in PBS and then exposed to 100 μL of the TUNEL reaction enzyme mix (1:10) at 37oC for 1 hr. Slides were washed 3 x 5 min in PBS and mounted with Prolong Gold containing DAPI. Tissue sections were imaged at 4x on a Nikon Eclipse Ni microscope using NIS-Elements software. They were blindly scored by hand using ImageJ.

Alcian blue/alizarin red staining

Staining was performed as in (Whited et al., 2013).

CRISPR design, ribonucleoprotein complex preparation, and delivery

More details on the type II CRISPR-Cas9 system can be found in (Cong et al., 2013; Doudna and Charpentier, 2014). We designed our CRISPR RNA (crRNA) constructs by using the sgRNAcas9 software package to identify crRNAs, along with potential off-target sites, against the *kazald1* open reading frame based on our assembled transcriptome ((Xie et al., 2014)). The sgRNAcas9 software was run in "single" mode using default parameters. We identified two crRNAs that target the region spanning base pairs 73 to 95 of the *kazald1* open reading frame (region includes the PAM sequence). The sequence targeted by the first *kazald1* crRNA (crRNA-1) including the PAM region (underlined) is: 5' CCCCGGCGTTTGCCCTTCAAGGG 3'. The sequence targeted by crRNA-1 is on the sense strand of the *kazald1* open reading frame. The sequence targeted by the second *kazald1* crRNA (crRNA-2) including the PAM region (underlined) is: 5' CCCTTGAAGGGCAAACGCCGGGG 3'. The sequence targeted by crRNA-2 is on the anti-sense strand of the *kazald1* open reading frame.

Recombinant Cas9 protein was purchased from PNA Bio Inc. (Catalog number: CP02) and reconstituted at a concentration of 5 μg/μL in UltraPure™ DNase/RNase-free distilled water (ThermoFisher Scientific; Catalog number: 10977023). The recombinant Cas9 protein had a 6 His tag and nuclear localization signal (NLS) from Simian virus 40 (SV40). Alt-R™ CRISPR Trans-activating RNA (tracrRNA), *kazald1* crRNA-1, and *kazald1* crRNA-2 were purchased from Integrated DNA Technologies and reconstituted at 100 μM in UltraPure™ DNase/RNase-free distilled water.

For injection of the CRISPR Cas9 ribonucleoprotein complex into embryos, tracrRNA and crRNAs were mixed at a 1:1 ratio and heated to 95 °C for 5 minutes. Following the 95 °C incubation, the tracrRNA:crRNA solution was kept on the thermocycler block and slowly cooled to room temperature. The two crRNAs used in this study were prepared separately, and control reactions consisted of tracrRNA mixed with an equal volume of UltraPure™ DNase/RNase-free distilled water. Once cooled, the tracrRNA:crRNA solution was mixed with recombinant Cas9 protein and incubated at 37 °C for 10 minutes and then cooled on ice to generate a working solution of the CRISPR Cas9 ribonucleoprotein complex. The working solution injected into embryos consisted of Cas9 at a final concentration of 500 ng/μL, tracrRNA at a concentration of 17.6 μM, and crRNA at a concentration of 17.6 μM. Wild type one-cell stage axolotl embryos (~90 % of injected embryos) and two-cell stage embryos

(~10% of injected embryos) were collected and processed for injection as described previously (Khattak et al., 2009). Prior to injection, embryos across multiple wild-type batches were mixed together and randomly assigned to groups in order to minimize potential batch effects. Embryos at the one-cell stage were injected once with approximately 50 nanoLiters of working ribonucleoprotein complex solution or control solution (ribonucleoprotein complex without crRNA), while two-cell stage embryos received two 50 nanoLiter injections (one for each cell). Non-injected embryos were set aside as controls for batch viability. Injections were performed with a glass capillary needle.

More information on the electroporation of CRISPR-Cas9 RNPs into tissues and cell in culture can be found in (Chen et al., 2016; Kalebic et al., 2016; Liang et al., 2015; Shinmyo et al., 2016; Xu et al., 2016). For electroporation of the CRISPR Cas9 ribonucleoprotein complex into intact limbs, the ribonucleoprotein complex was assembled as described above. Following assembly, the ribonucleoprotein complex was mixed with the pCAG-EGFP plasmid, Fast Green dye (for visualization), and Dulbecco's Phosphate Buffered Saline solution (Life Technologies™; Catalog number: 14200-075). The working solution injected into intact axolotl forelimbs consisted of recombinant Cas9 at a final concentration of 1.67 µg/µL, tracrRNA at a concentration of 20 µM, crRNA at a concentration of 20 µM, pCAG-EGFP at a concentration of 209 ng/µL, and 1X Dulbecco's Phosphate Buffered Saline. The right forelimbs of juvenile axolotls (~ 6 – 7 cm. in length) were electroporated with 3 – 4 microliters of working ribonucleoprotein complex, and the left forelimbs of these axolotl were injected with 3 – 4 microliters of control solution (ribonucleoprotein complex without crRNA). Following injection, electroporation was performed while animals were immersed in 1x PBS using a NepaGene Super Electroporator NEPA21 Type II electroporator. Our poring pulse consisted of 3 pulses at 150 Volts with a pulse length of 5 milliseconds, a pulse interval of 10 milliseconds, a decay rate of 0 %, and a positive (+) polarity. Our transfer pulse consisted of 5 pulses at 50 Volts with a pulse length of 50 milliseconds, a pulse interval of 950 milliseconds, a decay rate of 0 %, and a positive (+) polarity. The distance between electrodes for all electroporations was 3 millimeters.

For studies that have pioneered the use of CRISPR-Cas9 and other genome editing technologies in the axolotl, please refer to (Fei et al., 2014; Flowers and Crews, 2015; Flowers et al., 2014; Kuo, 2015).

Genomic DNA extraction and T7 Endonuclease I digestion

We took tail clips from axolotl hatchlings (~ 3.5 weeks old) that were injected with the CRISRP Cas9 ribonucleoprotein complex or control hatchlings and extracted genomic DNA using the Qiagen DNeasy® Blood & Tissue Kit (Qiagen; Catalog number: 69506). We amputated intact limbs that were electroporated with the CRISPR Cas9 ribonucleoprotein complex at 25 days post-electroporation and isolated tissue that was EGFP positive using a Leica M165 FC microscope to visualize fluorescence. We then extracted genomic DNA from the EGFP positive tissue using the Qiagen DNeasy® Blood & Tissue Kit.

We amplified genomic DNA samples using One *Taq*® DNA Polymerase (NEB; Catalog number: M0480L). PCR conditions were as follows: 94 °C for 3 minutes, then 35 cycles of 94 °C for 30 seconds, 57.2 °C for 30 seconds, and 68 °C for 15 seconds; the 35 cycles were followed by a 68 °C incubation for 5 minutes. The following primers were used to produce an amplicon of approximately 250 base pairs surrounding the region targeted by the *kazald1* crRNAs:

Primer 1: 5' CAGATCCTTGCCACTCTCCTGC 3'
Primer 2: 5' GCCATAGACTCGGTCAGACCTTGG 3'

PCR products were purified using Agencourt AMPure XP Beads (Beckman Coulter; Catalog number: A63881).

Following purification, 400 nanograms of amplified DNA were mixed with NEB Buffer 2 and re-annealed as follows: DNA was denatured at 95 °C for 5 minutes, cooled to 85 °C at -2 °C/second, and then cooled to 25 °C at -0.1 °C/second. Next, 20 units of T7 Endonuclease I (NEB; Catalog number: M0302L) was added to the reaction, and samples were incubated at 37 °C for 15 minutes. The digestion reaction was halted through the addition of EDTA to a final concentration of 17 mM and run on a 4% agarose gel. The expected size of the cleavage products produced from edited samples were approximately 200 base pairs and 50 base pairs for each crRNA. To calculate editing efficiency, ImageJ was used to quantify the band intensities of the wild-type product, cleavage products produced by T7 digestion, and background levels. Bands corresponding to apparent insertions or large deletions were also quantified. Following subtraction of background intensity from all measured band intensities for each lane, fractional cleavage was calculated as the sum of the cleaved product intensities divided by the sum of the intensities for all DNA bands for a given sample. Editing efficiency (Indel %) was calculated as:

$$\text{Indel } (\%) = 100\ x\ (1 - (1 - fraction\ cleaved)^{\frac{1}{2}})$$

Primers for *in situ* probes and RT-PCR

Genes and their associate primers are: *spatc1* (5' GCGTAAGATGGCCAGCGAAG 3', *5'* CAGAGGTCTTTTAATGATTTGTGATCTGG 3'), *tpm1* (5' GACCATTGACGATCTGGAAGATGAGC 3', 5' CCGACACAAAGCAAGAGGAATTGAG 3'), *ttn* (5' CAAGTGGATCAAATGGACTAAACTATC 3', 5' GCTTCCATGAGGGTAGTTGGAAAC 3'), *klhl41* (*5'* CATAAAGTGTGACTGCTGCTTCCAG 3', 5' CCTTTCATTGGGTCCTTTAATTACAACT 3'), *acta1* (*5'* CCAGAGCGCAAGTACTCTGTCT 3', *5'* CGGGTAGCATATTTAAGGTTTAT 3'), *col5a1* (*5'* CACTACCTCCCTCTACTGCCAC 3', *5'* CAGAGTGTGAGTGGAATGCTGC 3'), *ctsk* (*5'* GTGCAGAACCGACCCGATG 3', *5'* CAGCTGGACTCGGAGTGATGC 3'), *gspb1* (5' CCGCTTTGCTTCATCAACATATTGGGAG 3', 5' GGCAGGGCTGCTACTACAGCTAG 3'), *muc1* (*5'* GGGGGTTGAAGGTGTTGCTGATATTTC 3', 5' CTCATGATGCAACTGCAGAGCTTTCC 3'), *ef1a* (5' CGGGCACAGGGATTTCATC 3', 5' TGCCGGCTTCAAACTCTCC 3'), *cirbp* (5' CGACCCAGATGTCAAGATCCTCTTTAC 3', 5' CACCAGTCAAGCAAGTCACAAGCAAG 3'), *ptma* (5' GCTCTTGCGACTCGTCTTAGGCTTTG 3', 5' CAGGATGATCTTCAGACTCTTTCTCGGC 3'), *kazald1* (5' GAAAATGGATAAGGTGGTGGGGAGGG 3', 5' CTCGTGACATCCTGAGCCTGGAAG 3'), *sfrs1* (5' GCTATCCAGAGTAATGCAATTTTGTAGAGCC 3', 5' GACCATGACTAAAACAGTTTAACGGCACC 3'), *fus* (5' CCGATCTACCCAGCAAGCTGCC 3', 5' GTTGGCCATAACCACTCTGCTGCGC 3'), *hnrnpa1* (5' GGGCCCTGCAGAAGGTACGTTTATG 3', 5' GGGTTCTGGTTCAGTGCTACACAAGG 3'), *tecta* (5' GGCATAAGATGAAGAGGACTGATTGACAAC 3', 5' GAATTGGCAGTGCTGTTCAGTGCATATC 3'), *ef1a* (alternative, (Shaikh et al., 2011)) (5' AACATCGTGGTCATCGGCCAT 3', 5' GGAGGTGCCAGTGATCATGTT 3', *krt17* (5' GCCCAGTAACACCCTTGTCGGAG 3', 5' GGCGGTCGTTCAGGGTTTGCATG 3'), *shox2* (5' GCCATTGGCATTCACTCTCCAGGC 3', 5' CCCCATGACATAGAGGTGAAGGTCC 3'), and *cd38* (5' CAACAGGTAGATAGAGTATTTAACATTGTAGGC 3', 5' GGTGTGGTACATGTGATGCTGAATGG 3').

Supplemental References

Chen, S., Lee, B., Lee, A.Y., Modzelewski, A.J., and He, L. (2016). Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. J Biol Chem.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A*., et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Fei, J.F., Schuez, M., Tazaki, A., Taniguchi, Y., Roensch, K., and Tanaka, E.M. (2014). CRISPR-mediated genomic deletion of Sox2 in the axolotl shows a requirement in spinal cord neural stem cell amplification during tail regeneration. Stem Cell Reports *3*, 444-459.

Flowers, G.P., and Crews, C.M. (2015). Generating and identifying axolotls with targeted mutations using Cas9 RNA-guided nuclease. Methods Mol Biol *1290*, 279-295.

Flowers, G.P., Timberlake, A.T., McLean, K.C., Monaghan, J.R., and Crews, C.M. (2014). Highly efficient targeted mutagenesis in axolotl using Cas9 RNA-guided nuclease. Development *141*, 2165-2171.

Kalebic, N., Taverna, E., Tavano, S., Wong, F.K., Suchold, D., Winkler, S., Huttner, W.B., and Sarov, M. (2016). CRISPR/Cas9-induced disruption of gene expression in mouse embryonic brain and single neural stem cells in vivo. EMBO Rep *17*, 338-348.

Khattak, S., Richter, T., and Tanaka, E.M. (2009). Generation of transgenic axolotls (Ambystoma mexicanum). Cold Spring Harb Protoc *2009*, pdb prot5264.

Kuo, T.-H., Kowalko, J. E., DiTommaso, T., Nyambi, M., Montoro, D. T., Essner, J. J., and Whited, J. L. (2015). TALEN-mediated gene editing of the thrombospondin-1 locus in axolotl. Regeneration *2*, 37-43.
Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035-1043.
Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics *12*, 323.
Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S*., et al.* (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol *208*, 44-53.
Miller, J.R., Koren, S., and Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. Genomics *95*, 315-327.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Shaikh, N., Gates, P.B., and Brockes, J.P. (2011). The Meis homeoprotein regulates the axolotl Prod 1 promoter during limb regeneration. Gene *484*, 69-74.
Shinmyo, Y., Tanaka, S., Tsunoda, S., Hosomichi, K., Tajima, A., and Kawasaki, H. (2016). CRISPR/Cas9-mediated gene knockout in the mouse brain using in utero electroporation. Sci Rep *6*, 20611.
Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc Natl Acad Sci U S A *100*, 9440-9445.
Turro, E., Astle, W.J., and Tavare, S. (2014). Flexible analysis of RNA-seq data using mixed effects models. Bioinformatics *30*, 180-188.
Turro, E., Su, S.Y., Goncalves, A., Coin, L.J., Richardson, S., and Lewin, A. (2011). Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. Genome biology *12*, R13.
Whited, J.L., Tsai, S.L., Beier, K.T., White, J.N., Piekarski, N., Hanken, J., Cepko, C.L., and Tabin, C.J. (2013). Pseudotyped retroviruses for infecting axolotl in vivo and in vitro. Development *140*, 1137-1146.
Xie, S., Shen, B., Zhang, C., Huang, X., and Zhang, Y. (2014). sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PLoS One *9*, e100448.
Xu, L., Park, K.H., Zhao, L., Xu, J., El Refaey, M., Gao, Y., Zhu, H., Ma, J., and Han, R. (2016). CRISPR-mediated Genome Editing Restores Dystrophin Expression and Function in mdx Mice. Mol Ther *24*, 564-569.
Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol *11*, R14.