Transient activation of potent progenitor cells is required for spinal cord regeneration

Chase A. Weinholtz[1-3], Lili Zhou[1-3], Vishnu Muraleedharan Saraswathy[1-3], Yuxiao Xu[1-3], Dana Klatt Shaw[1-3], Anthony R. McAdow[1-3], Dongkook Park[1], Jimann Shin[1], Lila Solnica-Krezel[1,2] Aaron N Johnson[1] and Mayssa H. Mokalled[1-4]

[1] Department of Developmental Biology, Washington University School of Medicine, Saint Louis, MO, USA

[2] Center of Regenerative Medicine, Washington University School of Medicine, Saint Louis, MO, USA

[3] Hope Center for Neurological Disorders, Washington University School of Medicine, Saint Louis, MO, USA

[4] Author for correspondence: mmokalled@wustl.edu

Keywords: spinal cord injury, regeneration, progenitor cell potency, restored quiescence, single-cell analysis, lineage tracing, dual-function transcription factors, *sox2*, *bach1*.

ABSTRACT

Adult zebrafish exhibit full recovery following spinal cord injury. Transient expansion of stem cell-like progenitors is thought to underlie their regenerative capacity. Yet, our understanding of the identities and contributions of the crucial stem cell populations that direct spontaneous neural repair remains limited. Moreover, while most neural regeneration research is centered on promoting proliferative repair, the regulatory mechanisms that reinstate quiescence post-repair are unknown. Here, we determined the molecular identities and cellular contributions of *sox2*$^{+}$ progenitors during spinal cord repair. Genetic lineage tracing shows zebrafish spinal progenitors, while quiescent in uninjured tissues, self-renew and differentiate into neurons and glia after injury. By single-cell sequencing, *sox2*$^{+}$ cells are heterogeneous and biased towards neuronal or glial fates in both homeostatic and regenerating tissues. By screening for transcription factors that are differentially expressed in acute versus chronic spinal cord injury, we find the Bach1 transcription factors control transient progenitor cell activation by acting as dual activators and repressors of *sox2* expression. This study elucidates the molecular diversity and contributions of *sox2* expressing cells during spinal cord repair and identifies a transcriptional regulatory switch by which progenitor cells expand after injury and restore quiescence after regeneration is completed.

INTRODUCTION

Spinal cord injury (SCI) triggers distinct injury-responsive genetic and cellular programs that underlie differential regenerative capacities across vertebrates (Silver, 2016; Edwards-Faret et al., 2017; Tazaki et al., 2017; Sofroniew, 2018; Maden and Varholick, 2020; Phipps et al., 2020; Walker and Echeverri, 2022; Zheng and Tuszynski, 2023; Tendolkar and Mokalled, 2025). In regenerative vertebrates including zebrafish, spontaneous spinal cord (SC) repair is driven by potent stem cell-like progenitors that surround the central canal and undergo extensive proliferation following injury (Johansson et al., 1999; Mothe and Tator, 2005; Meletis et al., 2008; Barnabe-Heider et al., 2010; Hui et al., 2010; Rodrigo Albors et al., 2023). In non-regenerative mammalian SCs, adult ependymal cells proliferate and differentiate into astrocytes and oligodendrocytes, but their ability to contribute to neuronal differentiation *in vivo* is limited (Martens et al., 2002; Horky et al., 2006; Meletis et al., 2008; Barnabe-Heider et al., 2010; Karimi-Abdolrezaee et al., 2012; McDonough and Martinez-Cerdeno, 2012; Sabelström et al., 2013; Su et al., 2014; Paniagua-Torija et al., 2018; Shah et al., 2018; Xue et al., 2022). Notably, mammalian ependymal cells can be reprogrammed to enhance their potency, suggesting their regenerative potential is transcriptionally dormant (Shihabuddin et al., 2000; Ohori et al., 2006; Lee et al., 2013; Xu et al., 2017; Tai et al., 2021). We propose that elucidating the transcriptional programs that underscore the potency of endogenous stem cells in regenerative animals will guide protocols to improve reprogramming strategies and SCI outcomes.

Injury-induced expression of *sex determining region Y box 2 (sox2)* is a defining feature of neural progenitors in zebrafish (Adolf et al., 2006; Grandel et al., 2006; Kroehne et al., 2011; Barbosa et al., 2015). Sox2 is a highly conserved transcription factor capable of inducing and maintaining pluripotency (Avilion et al., 2003; Graham et al., 2003; Takahashi and Yamanaka, 2006; Schaefer and Lengerke, 2020). In mouse SCI, exogenous *sox2* expression is sufficient to reprogram NG2 glia into neurons *in vivo* (Su et al., 2014). In zebrafish, *sox2*$^+$ progenitors elicit compartmentalized responses after SCI (Reimer et al., 2008; Klatt Shaw et al., 2021; Saraswathy et al., 2022; Zhou et al., 2023). For instance, progenitors within the progenitor motor neuron (pMN) domain express *olig2* and regenerate motor neurons (Reimer et al., 2008). A distinct neurogenic niche dorsal to the ependyma upregulates TGF-β signaling to control the rates of self-renewal and interneuron neurogenesis after injury (Saraswathy et al., 2022). On the other hand, ventral *sox2*$^+$ progenitors express *connective tissue growth factor a* (*ctgfa*), undergo epithelial-to-mesenchymal transition (EMT) and direct regenerative glial bridging after SCI (Mokalled et al., 2016; Klatt Shaw et al., 2021; Zhou et al., 2023). As these studies revealed niches of *sox2*$^+$

48 progenitors underlying the regeneration of distinct cell types following SCI, several fundamental
49 questions remain unresolved: How do SC progenitor niches balance lineage restriction and
50 potency; how many progenitor domains are needed to regenerate spinal cell types; and does the
51 adult zebrafish SC harbor multipotent progenitors? Comprehensive molecular and transcriptional
52 characterization is thus required to determine the potency and contributions of *sox2*$^{+}$ progenitors
53 during innate SC repair.

54 *sox2* expression and the proliferation of *sox2*$^{+}$ progenitors are transiently activated after SCI.
55 Zebrafish spinal progenitors upregulate *sox2* and proliferate as early as 1 day post-injury (dpi)
56 (Ogai et al., 2014; Hui et al., 2015; Gorsuch et al., 2017). For decades, their acute activation post-
57 injury and their neurogenic potential during sub-acute SCI have been the primary focus of pre-
58 clinical regeneration studies (Becker et al., 2004; Reimer et al., 2013; Barreiro-Iglesias et al.,
59 2015; Hui et al., 2015; Mokalled et al., 2016; Ribeiro et al., 2017; Cavone et al., 2021; Klatt Shaw
60 et al., 2021; Vandestadt et al., 2021; Saraswathy et al., 2022; Cigliola et al., 2023; Zhou et al.,
61 2023; de Sena-Tomas et al., 2024). However, successful SC repair also relies on restricting
62 progenitor cell proliferation and regaining homeostasis as regeneration concludes. Understanding
63 how to re-establish progenitor cell quiescence is also necessary for the development of
64 regenerative therapies. However, the regulatory mechanisms that suppress *sox2* expression and
65 reinstate progenitor cell quiescence as SC repair concludes remain undefined.

66 BTB and CNC homology 1 (Bach1) is a basic leucine zipper transcription factor that plays
67 important and diverse roles in stem cell pluripotency and differentiation (Itoh-Nakadai et al., 2014;
68 Sato et al., 2020; Suzuki et al., 2020). Bach1 is comprised of several functional domains including
69 an N-terminal BTB/POZ domain that facilitates protein-protein interactions, and a C-terminal basic
70 leucine zipper (bZIP) domain responsible for DNA binding. This structural organization allows
71 Bach1 to form various protein heterodimers that carry out context-dependent transcriptional
72 activation or repression functions (Turner and Tjian, 1989; Albagli et al., 1995; Oyake et al., 1996;
73 Katsani et al., 1999; Sun et al., 2002; Jia et al., 2022; He et al., 2023). Previous studies reported
74 *bach1* expression in ventral spinal progenitors and showed that *bach1a/b* paralogues are required
75 for functional recovery following SCI (Klatt Shaw et al., 2021). However, the timing, transcriptional
76 activity and mechanisms of Bach1 functions during SC repair are still unknown.

77 This study determines the molecular identities, stem cell properties and contributions of *sox2*$^{+}$
78 progenitors during SC repair. Using a new *sox2-CreER*^{T2} knock-in line for genetic lineage tracing,

we show that *sox2*$^{+}$ cells self-renew and give rise to neurons and glia after SCI. Single nuclear RNA sequencing (snRNA-seq) and *in vivo* validation show *sox2*$^{+}$ cells are heterogeneous and biased towards distinct neuronal or glial fates in both homeostatic and regenerating SC tissues. To identify transcriptional regulators that play repressive roles in chronic SCI, we screen for transcription factors that are differentially expressed in acute versus chronic SCI. Loss- and gain-of-function studies reveal Bach1a/b as early activators and late repressors of *sox2* expression and progenitor cell expansion. Using a series of luciferase, chromatin immunoprecipitation and genetic rescue assays, we establish Bach1 as a direct transcriptional regulator of *sox2* that de-represses or represses *sox2* expression depending on cofactor availability. These findings delineate the molecular diversity and contributions of *sox2* expressing progenitors during spontaneous SC repair and reveal a regulatory switch can promote or restrain progenitor cell activation at different stages of regeneration.

RESULTS

Sox2^{+} cells mount an acute proliferative response in sub-acute SCI

To examine the stem cell properties of Sox2^{+} cells during SC regeneration, wild-type zebrafish adults were subjected to complete SC transections followed by daily intraperitoneal EdU injections after injury. This EdU labeling regimen establishes a cumulative profile of injury-induced cell proliferation. SC tissues at 7, 14 and 28 dpi were compared to tissues collected from uninjured siblings after 6 days of daily EdU injection **(Fig. 1A)**. SC sections 750, 450 and 150 µm rostral and caudal to the lesion were analyzed. Compared to uninjured controls where 2-3% of SC nuclei were EdU^{+}, cell proliferation robustly increased after SCI and peaked at 7 dpi (67% EdU incorporation at the lesion) **(Fig. S1A, S1B)**. Mirroring global cell proliferation, the proportion of Sox2^{+} EdU^{+} cells increased from 1-3% in uninjured controls to 72% at 7 dpi, 86% at 14 dpi and began to decline at 28 dpi **(Fig. 1B, 1C)**. These findings suggested Sox2^{+} cells are relatively quiescent in homeostatic adult SCs, and mount an acute and transient proliferative response within 7-14 days of SCI.

Sox2^{+} cells self-renew and differentiate following SCI

To determine the potency and contribution of *sox2*$^{+}$ cells after SCI, we employed a genetic lineage tracing system that combines a new (*sox2-2a-CreER*^{T2}*)*^{stl374} knock-in allele with *ubi:loxP-H2b-mCherry-STOP-loxP-sfGFP*^{stl345}, referred to hereafter as *sox2-Tracer* fish (Dean et al., 2025). This

109 system enabled us to induce Tamoxifen- and Cre-dependent recombination to permanently label
110 *sox2*$^{+}$ cells and their progenies during regeneration. To test the specificity of GFP expression after
111 recombination, *sox2-Tracer* fish were subjected to SCI, treated with 5 mM Tamoxifen (TAM) for
112 15 hours at 3 dpi, and SC tissues were collected at 7 dpi. GFP^{+} cells were only detected in TAM-
113 treated Cre^{+} SCs compared to TAM-treated Cre^{-}, vehicle-treated Cre^{+} or vehicle-treated Cre^{-}
114 controls **(Fig. S1C, S1D)**. These findings confirmed our TAM treatment induces specific
115 recombination and GFP expression in *sox2*$^{+}$ cells.

116 To determine the potency of *sox2*$^{+}$ cells, we performed the same regimen of SCI and TAM
117 treatment at 3 dpi **(Fig. 1D)**. SC tissues were collected at 5, 7, 14 and 28 dpi and stained for *sox2-*
118 *Tracer* (GFP) and Sox2 **(Fig. 1E, S1E)**. By quantifying *sox2*-derived cells within Sox2^{+} cells (*sox2-*
119 *Tracer*$^{+}$ Sox2^{+} cells normalized to Sox2^{+} cells), the proportion of recombined cells within Sox2^{+}
120 cells significantly increased from 5 dpi to 14 dpi **(Fig. 1F).** These findings indicated recombined
121 *sox2*-derived cells undergo self-renewal to expand the pool of Sox2^{+} cells after SCI. On the other
122 hand, by quantifying Sox2+ cells within *sox2*-derived cells (*sox2-Tracer*$^{+}$Sox2^{+} cells normalized to
123 *sox2-Tracer*$^{+}$ cells), the proportion of recombined cells that maintained Sox2 expression showed
124 a significant decrease from 5 dpi to 14 and 28 dpi **(Fig. 1G).** These findings indicate a subset of
125 *sox2*-derived cells downregulate Sox2 expression over the regeneration timeline, ostensibly as
126 they differentiate into regenerating cell types. Together, our genetic lineage tracing showed *sox2*$^{+}$
127 cells undergo proliferative self-renewal in sub-acute SCI and downregulate *sox2* expression at
128 later stages of SC regeneration.

129 ***sox2* is a comprehensive marker of SC progenitors**

130 We performed snRNA-seq to determine the molecular signatures that underlie the potency and
131 transient activation of progenitor cells after SCI. Multiple genes including *sox2, gfap*, *nes* and
132 *foxj1a* have been previously used to label and study zebrafish SC progenitors (März et al., 2010;
133 Kroehne et al., 2011; Ogai et al., 2014; Ribeiro et al., 2017; Than-Trong et al., 2018; Than-Trong
134 et al., 2020). Postulating that *sox2* serves as one of the most comprehensive markers of SC
135 progenitors, we performed snRNA-seq using a *sox2-2a-sfGFP*^{stl84} knock-in fish that faithfully
136 recapitulate Sox2 protein expression in adult SC tissues **(Fig. 2A)** (Shin et al., 2014). For snRNA-
137 seq, *sox2-2a-sfGFP* fish were subjected to SCI and SC tissues spanning 3 mm around the lesion
138 site were collected at 7 and 14 dpi **(Fig. 2B)**. Comparable SC tissue sections were collected from
139 uninjured controls. For each time point, 2 replicate pools of nuclei were sequenced, and 50 fish
140 were pooled within each replicate. A total of 63,804 nuclei were sequenced **(Fig. 2B)**. After quality

control analysis **(Fig. S2A, S2B)**, unsupervised Seurat clustering yielded 33 clusters including neurons, glia, oligodendrocytes and OPCs **(Fig. 2C)**. Gene expression analysis of various progenitor markers (*sox2, gfap, foxj1a, cspg4, prom2* and *nes*) in this integrated dataset showed pronounced differences that reflect progenitor cell diversity in SC tissues **(Fig. 2D, S2C)**. For instance, *cspg4*, which is expressed in bipotent OPCs that give rise to neurons and oligodendrocytes during embryonic development, and *foxj1a*, which labels ependymal progenitors were expressed in distinct clusters by snRNA-seq **(Fig. S2C)** (Ribeiro et al., 2017; Bromley-Coolidge et al., 2024). Compared to and encompassing other progenitor markers, *sox2* was comprehensively expressed in *nes*$^+$ neural stem cells, *cspg4*$^+$ OPCs and *foxj1*$^+$ ependymal cells **(Fig. 2D, S2C)**. *sox2* was also strongly upregulated across clusters after injury **(Fig. 2E)**. Aiming to broadly capture and characterize SC progenitors, we subclustered *sox2*$^+$ cells for subsequent analysis.

sox2$^+$ cells are heterogeneous and lineage-biased before and after SCI

Unsupervised Seurat clustering of *sox2-2a-sfGFP* expressing cells yielded 19 clusters highlighting the transcriptional diversity of *sox2*$^+$ cells **(Fig. 2F, S2D-F)**. By assessing changes in cell proportions within each cluster between 0 and 14 dpi, only 3 clusters (8, 12 and 16) were classified to be injury-induced, suggesting they may have regenerative functions during SC repair **(Fig. 2F, 2G)**. Surprisingly, 16 out of 19 *sox2*$^+$ cell clusters were present in both injured and uninjured SC, suggesting *sox2*$^+$ cells are heterogeneous and lineage-biased in adult zebrafish regardless of injury status. To test the heterogeneity of *sox2*$^+$ cells in homeostatic adults, we performed unsupervised Seurat clustering for *sox2-2a-sfGFP* expressing cells in uninjured animals without integrating post-injury datasets. This analysis revealed 14 clusters comprised of diverse cell types, mirroring the cluster identities obtained by integrating uninjured and injured time points **(Fig. 2H)**. These naïve *sox2*$^+$ cells displayed differential gene expression congruent with lineage biases towards various glial and neuronal cell types **(Fig. 2I, S2G).** As the SC regeneration field has been focused on studying genes that are induced in different progenitor domains after SCI, the underlying assumption has been that niches of specialized progenitors emerge after SCI and that quiescent progenitors represent a more homogenous transcriptional profile prior to injury. Contradicting this model, our unbiased snRNA-seq indicates naïve *sox2*$^+$ cells are heterogenous and transcriptionally biased towards specific lineages in homeostatic adult zebrafish.

172 We performed Hybridization Chain Reaction (HCR) *in situ* hybridization to validate select
173 progenitor cell clusters identified *in silico*. By snRNA-seq, cluster 0 exhibited elevated expression
174 of cilia-related genes, suggesting it is likely to represent ependymal cells **(Fig. S2E)**. *wnt4a* was
175 identified as a unique marker of c0 and plays an important role in cellular differentiation in adult
176 SCs **(Fig. 3A)** (Gao et al., 2016; Rodrigo Albors et al., 2023). Indeed, HCR labeling showed
177 elevated *wnt4a* expression in the ventral and lateral ependymal cells that encircle the central
178 canal **(Fig. 3B)**. *wnt4a* transcripts co-localized *sox2-2a-sfGFP* and were upregulated proximal
179 and rostral to the lesion at 7 and 14 dpi **(Fig. 3C)**. Further validation showed that *pcdh18b*,
180 predicted to label ependymal cluster 11 **(Fig. S2E)**, was also expressed in a subset of central
181 canal lining cells at 7 dpi **(Fig. S3A)**. On the other hand, the transcription factor *lhx4*, responsible
182 for V2a interneuron differentiation in mice, was identified as a unique marker of neuron-biased
183 c16 **(Fig. 3D)** (Renaux et al., 2024). By HCR *in situ* hybridization, *lhx4* expression was detected
184 in domains consistent with neuronal progenitors and co-localized with both *sox2* and the post-
185 mitotic neuronal marker HuC/D **(Fig. 3E-G)**. Further validation mapped neurogenic clusters 4, 13,
186 15, 16 and 17 to distinct neurogenic domains **(Fig. S3B).** Together, these scRNA-seq and HCR
187 gene expression studies mapped ependymal and neurogenic progenitors within zebrafish SC.

188 **Sox2$^+$ progenitors give rise to neurons and glia during SC repair**

189 We used *sox2-CreER^{T2}*-based genetic lineage tracing to track the fates of *sox2$^+$* progenitors and
190 their progenies during SC regeneration **(Fig. 3H)**. GFP labeling was co-stained with neuronal
191 (HuC/D) and pan-glial (Gfap) markers **(Fig. 3I-L)**. The proportions of *Tracer$^+$* neurons (GFP$^+$
192 HuC/D$^+$) increased 3-fold between 7 and 14 dpi, with progenitor-derived neurons comprising 27-
193 42% of all neurons 750 µm from the lesion **(Fig. 3J)**. Consistent with ongoing neurogenesis of
194 *sox2*-derived cells, HuC/D expression within recombined *Tracer$^+$* cells increased 3-fold from 7 to
195 14 dpi **(Fig. 3J)**. Similarly, the proportions of *Tracer$^+$* glia (GFP$^+$ Gfap$^+$) increased 2.6-fold between
196 7 and 14 dpi **(Fig. 3L).** Progenitor-derived glia comprised 23-34% of all glia 750 µm from the
197 lesion and accounted for 26-38% of total traced cells at 14 dpi **(Fig. 3L)**. Together, our *in silico*
198 and *in vivo* studies underscore a stepwise process whereby lineage-biased quiescent progenitors
199 transiently upregulate *sox2* and self-renew between 0 and 14 dpi**,** begin to differentiate into
200 neurons and glia starting at 7 dpi, and regain quiescence at later stages of regeneration.

201 ***bach1a/b* are transcriptionally enriched in *sox2$^+$* cells after SCI**

202 While progenitor cell activation has been widely studied, we aimed to examine how progenitor
203 cells return to quiescence during later regeneration stages. Postulating *sox2* regulation is a major

204 tipping point between quiescence and activation **(Fig. 1)**, we set out to identify putative *sox2*
205 activators and repressors **(Fig. 4A)**. To this end, we performed pseudobulk analysis to identify
206 transcription factors that are enriched at 7 versus 42 dpi **(Fig. 4B)**. Pseudobulk analysis enhanced
207 our ability to capture transcription factors, which exert biologically significant effects but elicit low
208 gene expression levels in single-cell datasets (Pokhilko et al., 2021). This analysis identified 21
209 transcription factors enriched at 7 dpi, coinciding with peak *sox2* expression. Compared to 7 dpi,
210 16 transcription factors were identified at 42 dpi, when *sox2* expression returns to pre-injury levels
211 and regeneration concludes. Intriguingly, while activators account for 71% of sub-acutely
212 expressed transcription factors at 7 dpi (15 out of 21), they decrease to 56% of transcription
213 factors expressed at 42 dpi (6 out of 16) **(Fig. 4B).** Notably, we identified *bach1b* as the only dual-
214 function transcription factor that has canonical activator and repressor activity and whose
215 transcripts were detected at both 7 and 42 dpi **(Fig. 4C)**. This analysis identified activators and
216 repressors of gene transcription during sub-acute and chronic SCI, and suggested Bach1 may
217 play dual roles in the initiation and termination of SC regeneration.

218 By snRNA-seq, *bach1b* and its paralogue *bach1a* were broadly expressed in multiple *sox2*$^{+}$
219 subclusters **(Fig. S4A, S4B)**. Consistent with sustained *bach1b* expression at 7 and 42 dpi,
220 *bach1b* was upregulated at 7 and 14 dpi in pooled subclusters of *sox2*$^{+}$ cells **(Fig. 4D)**. On the
221 other hand, *bach1a* was only transiently upregulated in *sox2*$^{+}$ cells at 7 dpi **(Fig. 4D)**. *in situ*
222 hybridization confirmed *bach1a* and *bach1b* were markedly upregulated at 7 dpi compared to
223 uninjured controls **(Fig. 4E-H, S4C, S4D)**. *bach1a* maintained elevated expression between 7 and
224 21 dpi and was downregulated at 28 dpi **(Fig. 4E, 4F)**. Compared to uninjured controls, *bach1b*
225 showed a 5-fold increase at 7 and 14 dpi and a further increase to 9- and 8-fold at 21 and 28 dpi
226 **(Fig. 4G, 4H)**. Notably, *bach1a* and *bach1b* were strongly induced around the central canal, in a
227 pattern consistent with *sox2*$^{+}$ progenitors. HCR *in situ* hybridization for *bach1*a/b in *sox2-2A-*
228 *sfGFP* zebrafish confirmed *bach1a/b* are preferentially upregulated in *sox2*$^{+}$ cells, **(Fig. S4E,**
229 **S4F)**. These findings identified *bach1a/b as* putative regulators of progenitor cell dynamics after
230 SCI.

231 ***bach1a/b* are required for functional and anatomical regeneration after SCI**
232 To examine the role of *bach1a*/b during functional recovery, we performed SC transections on
233 *bach1a*^{stl666}*;bach1b*^{stl668} double mutant zebrafish, referred to hereafter as *bach1a/b* mutants or
234 *bach1a/b*$^{-/-}$ (Klatt Shaw and Mokalled, 2021). We then evaluated functional recovery at 14, 28 and
235 42 dpi **(Fig. 5A)(Mokalled et al., 2016)**. Siblings that are double heterozygous for *bach1a* and

236 *bach1b* (*bach1a/b*$^{+/-}$) were used as controls. Swim endurance was comparable across cohorts of
237 uninjured fish and declined in both groups at 14 dpi. *bach1a/b*$^{+/-}$ animals regained swim function
238 at 28 and 42 dpi. In contrast, *bach1a/b* mutants did not recover any swim function after 14 dpi
239 and showed significantly lower swim endurance compared to controls at 28 and 42 dpi **(Fig. 5B)**.
240 To assess anatomical regeneration, rostral neurons were anterogradely traced with biocytin and
241 visualized caudal to the lesion (Klatt Shaw et al., 2021). At 42 dpi, the proportion of caudally traced
242 axons proximal and distal to the lesion was significantly reduced in *bach1a/b*$^{-/-}$ compared to
243 *bach1a/b*$^{+/-}$ siblings **(Fig. 5C)**. By Gfap staining, glial bridging was deficient in *bach1a/b* mutants
244 relative to controls **(Fig. 5D)** (Klatt Shaw et al., 2021). These results confirmed *bach1a/b* are
245 required for functional and anatomical recovery after SCI.

246 ***bach1a/b* are required for activation and quiescence of Sox2^{+} progenitors**

247 To examine the role of *bach1a/b* in progenitor activation, we assayed the proliferation of Sox2^{+}
248 cells in *bach1a/b* mutants **(Fig. 5E)**. Sox2 and PCNA staining were performed at 7 and 42 dpi,
249 providing snapshots of actively dividing cells at the time of collection **(Fig. 5F)**. Sox2^{+} PCNA^{+} cells,
250 which accounted for 20% of Sox2^{+} cells in SC sections of *bach1a/b*$^{+/-}$ controls at 7 dpi, were 50%
251 depleted in *bach1a/b* mutants **(Fig. 5G, S5A, S5B)**. Consistent with decreased progenitor cell
252 proliferation in sub-acute SCI, daily EdU injections showed the cumulative proportions of Sox2^{+}
253 EdU^{+} cells decreased by 71% in *bach1a/b* mutants at 7 dpi **(Fig. S5C-G)**. Intriguingly, loss of
254 *bach1a/b* caused an inverted phenotype at 42 dpi, whereby the proportions of proliferating Sox2^{+}
255 PCNA^{+} cells increased 3.7-fold in *bach1a/b* mutants compared to controls **(Fig. 5F-G)**. One
256 possibility is that SC regeneration is slower or delayed in the absence of *bach1a/b*, accounting
257 for an extended timeline of Sox2^{+} cell proliferation in *bach1a/b* mutants. Alternatively, as Bach1
258 can act as a transcriptional repressor and is expressed in chronic SCI, it is plausible that Bach1
259 is required for both early induction and late repression of Sox2^{+} cell proliferation. To determine
260 whether *sox2* expression is transcriptionally impacted in *bach1a/b* mutants, we performed RT-
261 qPCR for *sox2* on *bach1a/b* mutants at 7 and 42 dpi **(Fig. S5H)**. Mirroring our quantifications of
262 Sox2^{+} cells **(Fig. 5E-G)**, loss of *bach1a/b* depleted *sox2* transcripts at 7 dpi, but significantly
263 enhanced *sox2* expression at 42 dpi **(Fig. S5H)**. These results indicated *bach1a/b* are required
264 to induce Sox2^{+} cells in sub-acute SCI, and to restore their homeostatic quiescence at later stages
265 of SC repair.

266 **Bach1a/b are direct transcriptional regulators of *sox2***

267 To examine whether Bach1 is a direct transcriptional regulator of *sox2*, we searched for the
268 consensus binding motif of Bach1 (TGACTCAGC) within a 5 kb region upstream of the *sox2* start
269 codon (Fornes et al., 2020). This analysis identified 8 putative binding regions with a tolerance of
270 1 nucleotide mismatch **(Fig. 6A)**. We used chromatin immunoprecipitation (ChIP) to assess
271 Bach1 binding at 7 and 42 dpi **(Fig. 6B)**. In the absence of a reliable antibody targeting zebrafish
272 Bach1a/b, we generated a new *hsp70:hBACH1-2A-EGFP*^{stl692} transgenic line expressing human
273 BACH1 under control of an inducible heat shock promoter, referred to hereafter as *hBACH1*^{OE}
274 fish **(Fig. S6A)**. Expression of hBACH1 and antibody specificity were confirmed by western blot
275 **(Fig. S6B)**. Noting that the *hBACH1*^{OE} line did not induce excessive levels of BACH1
276 overexpression, we proceeded with ChIP. *hBACH1*^{OE} fish were subjected to SCI and 3 daily heat
277 shocks preceding collection of SC tissues at either 7 or 42 dpi **(Fig. 6B)**. ChIP-qPCR revealed
278 robust enrichment of hBACH1 at 3 sites (Binding elements 1, 5 and 6) upstream of *sox2* at 7 dpi
279 **(Fig. 6C)**. Interestingly, 1 out the 3 sites (Binding element 5) also enriched for hBACH1 binding at
280 42 dpi **(Fig. 6C)**. These results showed Bach1 binds distinct regulatory elements upstream of
281 *sox2* at different stages of SC regeneration, and suggested Bach1 regulates *sox2* transcription in
282 sub-acute and chronic SCI.

283 **Bach1a/b exhibit dual activation and repression activity**

284 We performed a series of luciferase assays to examine whether Bach1 activates and/or represses
285 *sox2* expression after SCI. First, we characterized the inherent transcriptional activity of zebrafish
286 Bach1a and Bach1b using a synthetic enhancer that contains 6 tandem copies of the BACH1
287 consensus motif (TGACTCAGC) (Fornes et al., 2020). To this end, *6xBACH1* was cloned into the
288 *sv40* minimum promoter-driven luciferase construct and co-transfected into HEK-293T cells with
289 constructs encoding Bach1a or Bach1b. In this experiment, Bach1a and Bach1b repressed
290 luciferase activity in a dose dependent manner **(Fig. 6D)**. In a second set of luciferase assays, a
291 402 bp region encompassing binding elements 5 and 6 was cloned into the *sv40* minimum
292 promoter-driven luciferase construct, and co-transfected with Bach1a and Bach1b into HEK-293T
293 cells **(Fig. 6E)**. Recapitulating Bach1 repressive activity in the *6xBACH1* luciferase assay, either
294 Bach1a or Bach1b alone significantly repressed baseline luciferase activity **(Fig. 6E)**. As Bach1
295 is known to partner with small MAF proteins and numerous other cofactor complexes to regulate
296 gene expression (Oyake et al., 1996), we tested whether the transcriptional activity of Bach1 can
297 be modulated by co-expressing zebrafish Mafga, Mafgb or Maff **(Fig. 6E)**. In these assays, Mafga
298 and Mafgb were sufficient to overcome the repressive effects of Bach1a and Bach1b. On the other
299 hand, co-expression with Maff exacerbated Bach1-mediated repression. These findings

300 established context-dependent regulation of *sox2* expression by Bach1a/b and their Maf binding
301 partners.

302 **BACH1 overexpression alters the expansion and quiescence of *sox2*$^{+}$ cells after SCI**
303 Our loss-of-function and molecular studies support a model in which Bach1 performs context-
304 dependent regulatory roles after SCI, activating *sox2* expression in sub-acute SCI and repressing
305 *sox2* at late stages of regeneration. One prediction of this model is that Bach1 overexpression
306 can either enhance or repress progenitor cell activation at different SCI stages. To test this
307 prediction, we performed SCI and daily heat shocks on *hBACH1*^{OE} and wild-type siblings **(Fig.**
308 **6F)**. We then evaluated the profiles of Sox2^{+} cells at 7 and 42 dpi. Supporting a dual regulatory
309 role for Bach1, *hBACH1*^{OE} fish showed a greater proportion of Sox2^{+} cells at 7 dpi and fewer
310 Sox2^{+} cells at 42 dpi **(Fig. 6G)**. These findings showed hBACH1 overexpression has an opposite
311 phenotype to Bach1a/b loss-of-function, enhancing early activation and promoting late repression
312 of Sox2^{+} cells after SCI. However, like *bach1a/b* mutants, *hBACH1*^{OE} fish showed significantly
313 lower swim endurance compared to controls **(Fig. 6H)**, suggesting tight regulation of *sox2* is
314 critical for successful SC repair.

315 ***sox2* overexpression rescues SC regeneration in *bach1a/b* mutants**
316 Our genetic and molecular studies indicated Bach1 regulates *sox2* expression and that *sox2*
317 dysregulation in *bach1* mutants may underlie their regeneration defects. To test this hypothesis,
318 we examined whether deficient regeneration in *bach1a/b* mutants could be rescued by *sox2*
319 overexpression. To this end, we used an established *hsp70:sox2*^{x21} transgenic line, referred to
320 hereafter as *sox2*^{OE}, to induce *sox2* expression under control of a heat shock inducible promoter
321 (Millimaki et al., 2010) **(Fig. 6I, 6J)**. We crossed the *sox2*^{OE} transgene into *bach1a/b* mutants to
322 generate *sox2*^{OE};*bach1a/b*$^{-/-}$ fish. *bach1a/b*$^{-/-}$ and *sox2*^{OE};*bach1a/b*$^{-/-}$ fish were subjected to SC
323 transections and daily heat shocks to induce transgene expression. Compared to *bach1a/b*$^{-/-}$ fish
324 where functional recovery is impaired, swim endurance and axon regrowth were significantly
325 higher in *sox2*^{OE};*bach1a/b*$^{-/-}$ fish at 42 dpi **(Fig. 6K, S6C)**. Importantly, *sox2* overexpression did
326 not enhance functional recovery in wild-type fish **(Fig. S6D)**, indicating the phenotype observed
327 in *sox2*^{OE};*bach1a/b*$^{-/-}$ animals is a specific rescue of the *bach1a/b* mutant background. These
328 findings indicated that exogenous *sox2* expression partially rescues the regeneration defects in
329 *bach1a/b* mutants and support a model in which tightly controlled Bach1-mediated *sox2*
330 expression is required for successful SC repair.

DISCUSSION

Adult zebrafish offer a valuable platform to elucidate transcriptional mechanisms of innate SC repair in regenerative vertebrates. By generating a multidimensional map of SC progenitors in zebrafish, this study provides a much-needed resource to determine the molecular underpinnings of progenitor cell potency and to advance translational studies. Anatomically and molecularly, we found *sox2*$^{+}$ cells are compartmentalized into niches of ependymal and non-ependymal progenitors that exhibit specific lineage biases in homeostatic and lesioned SC tissues. Temporally, our results support a model in which a single transcription factor like Bach1 can either activate or repress *sox2* expression to fine tune progenitor cell expansion and quiescence during the time course of regeneration.

SC progenitors are thought to maintain elevated potency in adult zebrafish, but not in mammals (Horky et al., 2006; Meletis et al., 2008; Reimer et al., 2008; Barnabe-Heider et al., 2010; Karimi-Abdolrezaee et al., 2012; Su et al., 2014; Becker and Becker, 2015; Ribeiro et al., 2017; Paniagua-Torija et al., 2018; Shah et al., 2018; Klatt Shaw et al., 2021; Saraswathy et al., 2022; Xue et al., 2022; Zhou et al., 2023). Yet, our understanding of the identities, heterogeneity and potency of these cells remains limited. Consequently, spinal progenitors have been termed ependymal, ependymo-radial glial, glial or neural stem cells, and interchangeably labeled with markers such as *sox2, gfap*, *hey, nes,* and *foxj1a* (März et al., 2010; Kroehne et al., 2011; Ogai et al., 2014; Ribeiro et al., 2017; Than-Trong et al., 2018; Than-Trong et al., 2020). By genetic lineage tracing, we showed that *sox2*$^{+}$ cells give rise to most regenerating cell types after SCI. Our single-cell analysis and *in vivo* validations revealed ependymal and non-ependymal niches of progenitor cells in both homeostatic and lesioned SC tissues. Further studies will delineate the lineage relationships between different domains of *sox2* expressing cells. Consistent with prior knowledge, markers such as *foxj1a* and *nes* are expressed in restricted progenitor domains comprising ependymal and neural stem cells, respectively (Ribeiro et al., 2017; Xue et al., 2022). On the other hand, *gfap* is more broadly expressed in SC progenitors, but also in non-progenitor glia and other cell types (Lam et al., 2009; Goldshmit et al., 2012; Johnson et al., 2016). Here, we used *sox2* to comprehensively characterize SC progenitors, but it is important to note that *sox2* is also expressed in additional cell types after SCI. As there is no single definitive marker that is both inclusive and specific to all SC progenitors, we propose designating cell populations by the markers they express (e.g. *sox2*$^{+}$ cells), rather than relying on alternate nomenclature.

362 At the core of zebrafish regenerative capacity and this study is the ability of SC tissues to
363 curtail regenerative processes and restore quiescence as SC regeneration concludes. Restricting
364 the proliferation of progenitor cells is essential to reestablish homeostasis and prevent the
365 detrimental consequences associated with sustained cellular proliferation. Yet, most tissue
366 regeneration research has centered on promoting proliferative repair studies (Becker et al., 2004;
367 Reimer et al., 2013; Barreiro-Iglesias et al., 2015; Hui et al., 2015; Mokalled et al., 2016; Ribeiro
368 et al., 2017; Cavone et al., 2021; Klatt Shaw et al., 2021; Vandestadt et al., 2021; Saraswathy et
369 al., 2022; Cigliola et al., 2023; Zhou et al., 2023; de Sena-Tomas et al., 2024), while the regulatory
370 mechanisms that reinstate quiescence post-regeneration are largely unknown. Our findings
371 support a model in which specific Bach1-Maf interactions can either activate or repress *sox2*
372 expression and progenitor cell expansion, ostensibly in response to extracellular cues that relay
373 the tissue's regenerative state. Insights from salamander midbrain regeneration revealed a
374 negative feedback loop from regenerating dopaminergic neurons to neurogenic progenitors to re-
375 establish homeostasis following regenerative neurogenesis (Berg et al., 2011). In this model,
376 dopamine signaling from regenerated neurons or administration of a dopamine receptor agonist
377 are sufficient to limit progenitor cell proliferation (Berg et al., 2011). Intriguingly, in mammals,
378 application of either dopamine receptor agonist or antagonist have been shown to promote neural
379 stem cell proliferation, but downstream mechanisms remain to be elucidated (Kippin et al., 2005;
380 Yang et al., 2008). Further studies into the extrinsic signals that regulate the expression and
381 activity of diverse Bach1 and Maf transcription factors are needed to better understand
382 mechanisms of initiation and resolution of SC regeneration.

383 Dual-function transcription factors enable context-dependent transcriptional regulation during
384 tissue regeneration. Although Bach1 is best known for its canonical repressive functions, it is
385 increasingly recognized for maintaining stem-like states through transcriptional activation (Sun et
386 al., 2002; Fuse et al., 2015; Sato et al., 2020; Niu et al., 2021). Our findings suggest Bach1's
387 canonical repressive functions downregulate *sox2* at late stages of regeneration, but that cofactor
388 availability can override this repressive role at earlier time points. Unlike pioneer factors like Oct4
389 and Sox2, which simultaneously activate and repress different sets of genes to control stem cell
390 maintenance versus differentiation, our results suggest Bach1 can directly repress or derepress
391 the same target gene at different stages of regeneration (Zhang et al., 2019). We propose
392 differential binding to Maff versus Mafg as one mechanism that underlies Bach1 dual functions.
393 However, as an epigenetic regulator, Bach1 may also act by recruiting chromatin remodeling
394 complexes to regulate *sox2* expression (Guo et al., 2023). Further characterization of Bach1

395 binding partners as well as chromatin accessibility at different stages of regeneration are needed
396 to better understand the dynamics of gene regulation during SC regeneration.

ACKNOWLEDGMENTS

We thank V. Cavalli, H. Gabel, P. Williams, L. Sheets (Washington University) for discussion; S. Nandagopal, T. Tsai (Washington University), S. Megason (Harvard) for the HCR probe design script; B. Riley (Texas A&M) for zebrafish lines. We thank the Zebrafish Consortium Facility at Washington University for animal care, and the Genome Access Technology Center at Washington University for single-cell sequencing. This research was supported by grants from the NIH (R35 GM118179 to LSK, R01 AR070299 to ANJ, R01 NS113915 and R01 NS123708 to MHM).

COMPETING INTERESTS

The authors declare no competing interests.

METHODS

Zebrafish

Adult zebrafish of the Ekkwill, Tubingen and AB strains were maintained at the Washington University Zebrafish Consortium Facility. All animal experiments were performed in compliance with institutional animal protocols. Fish stocks were tracked and catalogued using VMS-Fish software (Muraleedharan Saraswathy and Mokalled, 2025). Adult male and female animals of ~2 cm in length (4-6 months of age) were used. Experimental fish and control siblings of similar size and equal sex distribution were used for all experiments. SC transection surgeries and regeneration analyses were performed in a blinded manner and 2 to 4 independent experiments were performed using different clutches of animals. *bach1a*^{stl664}, *bach1b*^{666} and *sox2-2A-sfGFP*^{stl84} were generated as previously described (Shin et al., 2014; Klatt Shaw and Mokalled, 2021). *Tg(hspl:sox2)*^{x21} fish were generously provided by Dr. Bruce B. Riley (Texas A&M)(Millimaki et al., 2010).

Generation of *hsp70:hBACH1-2A-EGFP* zebrafish

The following primers were used to amplify hBACH1 cDNA: ClaI forward primer 5'-ccATCGATatgtctctgagtgagaactcggtttt-3' and ClaI reverse primer 5'-ccatcgat atgactgataaatgtactactgatgagt-3'. The genomic fragments were cloned into PCR2.1-TOPO vector, then subcloned into ClaI digested PCS2-hsp70-2A-EGFP plasmid. Clones were co-injected into one-cell stage wild-type embryos with I-SceI. A minimum of 3 founders were isolated and propagated. The transgenic line was established and designated *Tg(hsp70:hBACH1-2A-EGFP)*^{stl692}. Animals were analyzed as hemizygotes.

Generation of *sox2-2a-CreER*^{T2} zebrafish

The DNA sequence for *CreER*^{T2} was obtained from pME minprom_creER^{T2} (PMID 36975217) and assembled with the previously established *sox2-2a* targeting vector (M2) (Shin et al., 2014), by Gibson assembly. The *sox2-2a-CreER*^{T2} targeting vector was linearized by NaeI. *sox2* TALEN RNAs were previously published (Shin et al., 2014) and synthesized with the SP6 mMessage mMachine Kit. Briefly, one-cell embryos were injected with *sox2* TALEN RNAs at 35 ng/µl each and linearized *sox2-2a-CreER*^{T2} targeting vector at 10 ng/µl in 1x injection buffer (0.1 M KCL, 0.003% Phenol Red). Injected embryos were raised to adulthood and *sox2-2a-CreER*^{T2} founders were identified by F1 progeny screen using PCR-based screen with *CreER*^{T2} specific primers (cre_geno_F1 TGCAGGGAGAGGAGTTTGTGTGC and cre_geno_R1 TGTGGGAGAGGATGAGGAGGAGC). The obtained F1 progeny were further analyzed and

428 confirmed that *2A-CreER*^{T2} is precisely inserted by sanger sequencing. The knock-in line was
429 established and designated *(sox2-2a-CreERT2)*^{stl374}*.* Animals were analyzed as hemizygotes.

430 **Spinal cord transection and treatments**

431 Zebrafish were anaesthetized in 0.2 g/l of MS-222 buffered to pH 7.0. Fine scissors were used to
432 make a small incision that transects the spinal cord 4 mm caudal to the brainstem region.
433 Complete transection was visually confirmed at the time of surgery. Injured animals were also
434 assessed at 2 dpi to confirm loss of swim capacity post-surgery.

435 For EdU administration, zebrafish were anaesthetized and subjected to intraperitoneal EdU
436 injections. 12.5 mM EdU (Sigma 900584) diluted in PBS was used. Daily injections were
437 performed and paraformaldehyde-fixed cryosections were generated.

438 **Histology**

439 Transverse cryosections of 16 µm thickness were generated from paraformaldehyde-fixed SC
440 tissues. Tissue sections were imaged using a Zeiss AxioVision compound microscope or a Zeiss
441 Axioscan.Z1 slide scanner for *in situ* hybridization, and a Zeiss LSM 800 confocal microscope or
442 a Zeiss Axioscan.Z1 slide scanner for immunofluorescence.

443 For EdU staining, sections were incubated with freshly prepared EdU staining solution
444 (100 mM Tris at pH 8.5, 1 mM CuSO4, 200 µM Fluorescent azide and 100 mM Ascorbic) for
445 30 min in the dark at room temperature.

446 For immunohistochemistry, tissue sections were hydrated in PBT then blocked with 5% goat
447 serum for 1 h at room temperature. For nuclear antigens, sections were treated with 0.2% TritonX-
448 100 for 5 min prior to the blocking step. Sections were incubated overnight at 4 °C with primary
449 antibodies diluted in blocking agent, then washed in PBT, and treated for 1 h in secondary
450 antibodies diluted in blocking agent at room temperature. Primary antibodies used in this study
451 were: rabbit anti-Sox2 (GeneTex, 124477, 1:250), mouse anti-GFAP (ZIRC, ZRF1; 1:500),
452 chicken anti-GFP (Aves Labs, GFP-1020; 1:1000) and rabbit anti-PCNA (GeneTex, GTX124496;
453 1:500). Secondary antibodies used in this study were: Alexa Fluor 488, Alexa Fluor 594, and
454 Alexa 647 goat anti-rabbit or anti-mouse antibodies (Jackson ImmunoResearch, 1:200, 103-545-
455 155, 111-605-003, 111-585-003, 115-605-003 and 115-585-003). Hoechst staining was
456 performed according to the manufacturer's recommendation (Thermo Fisher Scientific, H3570).

457 The HCR RNA *in situ* hybridization protocol was adapted from Molecular Instruments (Choi et
458 al., 2018; Schwarzkopf et al., 2021). Briefly, tissue sections were hydrated and pre-treated either
459 with 0.2% TritonX-100 for 5 min or boiled in Citrate Buffer (10 mM Citric Acid, 0.05% Tween-20,
460 pH 6.0) for 10 min. For blocking, sections were incubated in Hybridization buffer (Molecular
461 Instruments) for 1 h at 37 °C, then incubated in pre-warmed sets of DNA probes diluted to
462 0.0015 pmol/uL in Hybridization buffer for 48 h at 37 °C. Washes were then performed with Wash
463 Buffer (Molecular Instruments) at 37 °C followed by 5x SSCT (3 M NaCl, 0.3 M Sodium Citrate,
464 0.1% Tween-20, pH 7.0). For signal amplification, sections were incubated in Amplification buffer
465 (Molecular Instruments) for 1 h at room temperature. Prior to amplification, h1 and h2 hairpins
466 were snap-cooled and diluted 1:50 in Amplification buffer. Amplification proceeded overnight at
467 room temperature in the dark. Samples were then washed in 5x SSCT, 5x SSC and PBT before
468 proceeding to immunohistochemistry. HCR RNA probes used in this study (*bach1a* and *bach1b*)
469 were designed using a python script and ordered as 50 pmol opools from IDT **(Table S1)**.

470 RNAscope *in situ* hybridization for *bach1a* and *bach1b* was performed according to the
471 manufacturer' protocol (Advanced Cell Diagnostics, 322300).

472 **Histology quantification**

473 Counting of Edu$^+$, Sox2$^+$ or PCNA$^+$ cells were performed using a customized Fiji script (adapting
474 ITCN: Image based Tool for counting nuclei; https://imagej.nih.gov/ij/plugins/itcn.html). The Fiji
475 script incorporated user-defined inputs to define channels (including Hoechst) and to outline SC
476 perimeters. To quantify nuclei, the following parameters were set in ITCN counter: width, 15;
477 minimal distance, 7.5; threshold, 0.2. Once nuclei were identified, user-defined thresholds of
478 individual cell markers were used to mask the image and identify nuclei located inside the masked
479 regions. *xy* coordinates were extracted for each nucleus for cell counting. Raw counts and *xy*
480 coordinates from Fiji were processed using a customized R script. Two markers were considered
481 overlapping if they shared nuclei with same *xy* coordinates.

482 Colocalization analyses were performed using the JACoP plugin in Fiji (Bolte and
483 CordeliÈRes, 2006). The polygon selection tool was used to outline SC perimeters.

484 HCR *in situ* hybridization signals were quantified using a customized Fiji script (Saraswathy
485 et al., 2024). SC perimeters were defined and user-defined thresholds were set to create an

486 inverted mask. Background noise was either manually corrected or defined by thresholding a
487 dedicated, unstained background channel. Integrated density was quantified using the "Analyze
488 Particles" command. To quantify RNA signals inside *sox2*$^+$ cells, an inverted mask was created
489 after thresholding *sox2*$^+$ cells, then added to the final quantification mask. Fluorescence signals
490 inside this newly generated mask were considered *sox2*$^+$ cell specific.

491 For all statistical data, data points represent individual animals and N numbers are indicated.
492 Statistical tests are reported in figure legends for each panel. Multiple comparisons were
493 performed within groups and significant observations were reported.

494 **Isolation of spinal cord nuclei from adult zebrafish**

495 Nuclear isolation from adult SC tissues was performed as previously described (Matson et al.,
496 2018; Saraswathy et al., 2024). Briefly, 3 mm SC tissue sections flanking the lesion site were
497 collected from 50 adult zebrafish at 7 and 14 dpi. Corresponding 3 mm tissue segments were
498 collected from uninjured control siblings. For tissue lysis, the detergent mechanical lysis protocol
499 described by Matson et al. was performed. SC tissues were homogenized at low setting for 15 s.
500 Density gradient separation using sucrose solution was used to sediment the nuclei from the
501 supernatant. Final nuclear lysates were resuspended using 100 µl of resuspension solution (1x
502 PBS + 2% BSA + 0.2 U/µl RNase inhibitor). Hoechst staining was performed to assess the quality
503 of isolated nuclei based on their shape. Samples, in which more than 70% of nuclei were scored
504 as 'healthy', were submitted for single nuclear RNA-sequencing.

505 **Single nuclear RNA-sequencing**

506 For snRNA-seq, 30 µl resuspension of isolated nuclei at a concentration of ~1000 nuclei/µl was
507 submitted to the Genome Technology Access Center in the McDonnell Genome Institute of
508 Washington University. Two biological replicates were used for each time point (0, 7 and 14 dpi).
509 cDNA was prepared after GEM generation and barcoding, followed by the GEM-RT reaction and
510 bead cleanup steps. cDNA was amplified for 11−13 cycles then purified using SPRIselect beads.
511 Purified cDNA samples were then run on a Bioanalyzer to determine cDNA concentration. GEX
512 libraries were prepared as recommended by the 10x Genomics Chromium Single Cell 3' Reagent
513 Kits User Guide (v3.1 Chemistry Dual Index) with appropriate modifications to the PCR cycles
514 based on the calculated cDNA concentration. For sample preparation on the 10x Genomics
515 platform, the Chromium Next GEM Single Cell 3' Kit v3.1, 16 rxns (PN-1000268), Chromium Next
516 GEM Chip G Single Cell Kit, 48 rxns (PN-1000120), and Dual Index Kit TT Set A, 96 rxns (PN-

517 1000215) were used. The concentration of each library was accurately determined through qPCR
518 utilizing the KAPA library Quantification Kit according to the manufacturer's protocol (KAPA
519 Biosystems/Roche) to produce cluster counts appropriate for the Illumina NovaSeq6000
520 instrument. Normalized libraries were sequenced on a NovaSeq6000 S4 Flow Cell using the XP
521 workflow and a 50 × 10 × 16 × 150 sequencing recipe according to manufacturer protocol. A
522 median sequencing depth of 50,000 reads/cell was targeted for each Gene Expression Library.

523 **Aligning snRNA-seq reads**

524 After sequencing, the Illumina output was processed using the CellRanger (v6.0.0) recommended
525 pipeline to generate gene-barcode count matrices. A custom reference genome was made with
526 the "cellranger mkref" command, using the fasta file of zebrafish reference genome GRCz11
527 constructed from the Ensemble genome build (https://useast.ensembl.org/Danio_rerio/Info/Index)
528 and the sorted Gene Transfer Format file (v4.3.2) from the improved zebrafish transcriptome
529 annotation(Lawson et al., 2020). Base call files for each sample from Illumina were demultiplexed
530 into FASTQ reads. Then, the "cellranger count" pipeline was used to align sequencing reads in
531 FASTQ files to the custom reference genome. Both exon and intron sequences were included
532 during the alignment. The filtered gene-barcode count matrices generated by "cellranger count"
533 was used for downstream analysis.

534 **Quality control**

535 DecontX package (Yang et al., 2020) (celda v1.16.1) was used to remove droplets containing
536 aberrant counts of ambient mRNA. The function "decontx" was used on the raw "RNA" counts to
537 obtain "decontX" count data. Contamination score was assigned to each cell after running
538 "decontx". Cells with contamination scores greater than 0.75 were filtered out and only cells with
539 contamination scores less than 0.75 were included for downstream analysis. Furthermore,
540 decontX counts were used as default for pre-processing and normalization. After ambient RNA
541 removal, DoubletFinder (McGinnis et al., 2019) (v2.0.3) was used to identify doublets formed from
542 transcriptionally distinct cells. Optimal pK values were calculated from the outputs of function
543 "paramsweep_v3". Then, "doubletFinder_v3" function was used to predict doublet cells, where 50
544 principal components and pN value of 0.25 were given as input along with the previously
545 calculated pK values (0.26). All the cells predicted as doublets were removed.

546 **Integrated analysis of snRNA-seq dataset**

547 Datasets were integrated and analyzed using Seurat (v4.1.1) package with R (v4.2.1) (Team,
548 2018; Stuart et al., 2019). Each sample count matrix was filtered for genes that were expressed
549 in at least 3 cells and cells expressing at least 200 genes, followed by cell quality assessment
550 using commonly used QC matrices (Ilicic et al., 2016). Cells having a unique number of genes
551 between 200 and 4000 and a mitochondrial gene percentage < 5% were used for downstream
552 processing. Each dataset was independently normalized and scaled using the "SCTransform"
553 function, which is an improved method for normalization, that performs a variance-stabilizing
554 transformation using negative binomial regression (Hafemeister and Satija, 2019). Standard
555 integration workflow of Seurat was used to identify shared sources of variation across
556 experiments as well as mutual nearest neighbors(Butler et al., 2018; Haghverdi et al., 2018).
557 Integration features were selected based on the top 4000 highly variable features using
558 "SelectIntegrationFeatures" function (nfeatures = 4000), which was used as input for the
559 "anchor.features" argument of the "FindIntegrationAnchors" function. PCA analysis was
560 performed on the 4000 variable features and the top 50 principal components selected based on
561 the elbow plot heuristic, which measures the contribution of variation in each component. These
562 50 principal components were used in "FindNeighbors" and "FindClusters" functions to perform
563 graph-based clustering on a shared nearest neighbor graph(Levine et al., 2015; Xu and Su, 2015).
564 Louvain algorithm was used for modularity optimization in cell clustering using "FindClusters"
565 function. The resolution parameter (res = 0.4) that determines the granularity of clustering was
566 selected by visually inspecting clusters with resolutions ranging between 0.1 and 2.0 as well as
567 clustree graphs (Zappia and Oshlack, 2018). Uniform Manifold Approximation and Reduction
568 (UMAP) was used for non-linear dimensional reduction of the first 50 principal components and
569 visualize the data using "RunUMAP" function(Becht et al., 2018). Data was graphed using
570 different plot functions, such as "DimPlot", "VlnPlot", "FeaturePlot", "Dotplot" and "DoHeatmap",
571 to view the cell cluster identity and marker gene expression. Cell proportions were extracted using
572 the "table" and "prop.table" functions. Differential gene expression for individual clusters was
573 identified using Wilcoxon rank sum test in the "FindAllMarkers" function. Marker genes detected
574 in at least 25% of the clustered cells and with a logFC threshold of 0.25 were selected. Only
575 positive markers were reported.

576 **Subset of *sox2*$^{+}$ clusters**

577 *sox2*$^{+}$ cells identified from the complete dataset were subclustered using the "subset" function for
578 subcluster analysis. The *sox2*$^{+}$ subset was again normalized and scaled using the "SCTransform"
579 function with glmGamPoi method (Ahlmann-Eltze and Huber, 2021). 50 principal components

580 were used, and the resolution parameter was set to 0.7. Downstream analysis was done as
581 described above for the integrated analysis.

582 **Cluster identification using differentially expressed markers**

583 A “CNS markers” database of previously published markers of the different cell types that
584 comprise the vertebrate brain and/or SC tissues was compiled(Saraswathy et al., 2024). For each
585 cell cluster, every marker gene identified as a top differentially expressed (DE) marker of that
586 cluster was cross-referenced with our compiled VNM database using our “DE-Marker-Scoring”
587 algorithm (Saraswathy et al., 2024). For every matching marker gene, one point was given to the
588 respective cluster under the column name with matching cell identity. Iteration over every marker
589 gene was performed to generate a scoring matrix with varying points for each cluster against the
590 different cell identities compiled in the VNM database. The “phyper” function in R was then used
591 to calculate one-tailed binomial probabilities using hypergeometric distribution for the total score
592 obtained from each cluster against each cell identity in the database. −log10 of probability values
593 were obtained for plotting the heatmap. The resulting values were scaled from 0 to 100 and plotted
594 as a heatmap using GraphPad prism. Each cluster was given an identity based on the maximum
595 −log10 P score obtained in the heatmap. The top DE markers of clusters with ambiguous scores
596 were manually searched in the literature (Zhang and Cui, 2014) (https://brainrnaseq.org/) to
597 confirm cluster identity. “RenameIdents” function was used to assign identity to each cluster. To
598 confirm the assigned cluster identities, enrichment of classical markers of respective cell types
599 were tested using Dot plot.

600 **Selection of cluster markers**

601 To identify gene expression markers for each *sox2*$^{+}$ cluster, top DE markers were filtered for
602 specificity to such cluster. To this end, top DE markers were first filtered by their unique enrichment
603 in one *sox2*$^{+}$ cluster **(Table S2, S3)**. Enrichment and percent expression of these unique genes
604 were then confirmed using Dot plots. These top DE markers were ranked based on their specificity
605 and cluster-specific expression enrichment. Then, each gene marker was visualized in the whole
606 spinal cord dataset (including *sox2*$^{-}$ cells) using Dotplot and Featureplot to validate that they were
607 indeed a specific marker for a *sox2*$^{+}$ cluster. At least two markers were selected for downstream
608 HCR *in vivo* validation.

609 **Pseudobulk analysis**

610 To identify enriched transcription factors, snRNA Seq data was aggregated using the
611 “AggregateExpression” (Seurat v4.4) function for each replicate from uninjured, 7, 21 or 42 dpi
612 time points (Saraswathy et al., 2024). Transcription factors were then filtered and classified by
613 GO terms “TF.Activator” and “pos.regulation.tf” for transcriptional activators and “TF.Repressor”
614 and “neg.regulation.tf” for transcriptional repressors **(Table S4)**. Transcription factors that are
615 present in both classifications were categorized as “activator and repressor”. Differential
616 expression analysis, compared to uninjured controls, was performed for each time point with
617 DESeq2 using a p-value threshold of $p<0.05$ (Love et al., 2014). Only transcription factors with -
618 $1>Log2(FC)>1$ aggregate expression were considered down- or upregulated.

619 **Swim endurance assays**

620 Zebrafish were exercised in groups of 8-12 in a 5 l swim tunnel device (Loligo, SW100605L, 120
621 V/60 Hz). After 10 min of acclimation inside the enclosed tunnel, water current velocity was
622 increased every 2 min and fish swam against the current until they reached exhaustion.
623 Exhausted animals were removed from the chamber without disturbing the remaining fish. Swim
624 times at exhaustion were recorded for each fish. Results are expressed as mean ± SEM. One-
625 way ANOVA and multiple comparisons were performed using the Prism software to determine
626 statistical significance of swim times between groups.

627 **Axon tracing**

628 Anterograde axon tracing was performed on adult fish at 42 dpi. Fish were anaesthetized using
629 MS-222 and fine scissors were used to transect the cord 4 mm rostral to the lesion site. Biocytin-
630 soaked Gelfoam Gelatin Sponge was applied at the new injury site (Gelfoam, Pfizer, 09-0315-08;
631 biocytin, saturated solution, Sigma-Aldrich, B4261). Fish were euthanized 6 h post-treatment and
632 biocytin was detected histologically using Alexa Fluor 594-conjugated streptavidin (Molecular
633 Probes, S-11227). Biocytin-labeled axons were quantified using the ‘threshold’ and ‘particle
634 analysis’ tools in the Fiji software. Four sections per fish at 600 µm (proximal) and 1500 µm (distal)
635 caudal to the lesion core, and two sections 450 and 750 µm rostral to the lesion, were analyzed.
636 Axon growth was normalized to the efficiency of biocytin labeling rostral to the lesion for each fish.
637 Percentage axon growth was then normalized to the rostral level of the control group.

638 **Glial bridging**

639 GFAP immunohistochemistry was performed on serial transverse sections from adult fish at 14
640 dpi. The cross-sectional area of the glial bridge (at the lesion site) and the area of the intact SC

641 (750 µm rostral to the lesion) were measured using ImageJ software. Bridging was calculated as
642 a ratio of these measurements.

643 **Quantitative RT-PCR (RT-qPCR)**

644 RNA collections were performed using Trizol. 1 µg of RNA was converted into cDNA using the
645 Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, K1672) according to the
646 manufacturer's specifications. Quantitative PCR was performed using the Luna polymerase
647 master mix (NEB, M3003) using gene-specific primers and the Bio-Rad CFX Connect Real-Time
648 System **(Table S5)**. To determine primer efficiency, a standard curve was generated for each
649 primer set using cDNA pooled from 7 dpi SC tissues. Only primer sets with a calculated
650 amplification efficiency of 1.8-2.2 per cycle were used. Expression fold change for each gene was
651 calculated using the ΔCq method and dual normalized to *eef1a1l1* gene expression and to target
652 gene expression in control samples. For all RT-qPCR experiments, 2-3 biological replicates were
653 used per genotype and a minimum of 3 SC tissues were pooled within each replicate. Each
654 biological replicate was run in technical triplicate, and the average $\log_2$ fold change of these
655 technical replicates is shown.

656 **Chromatin immunoprecipitation (ChIP)**

657 *hsp70:hBACH1-2A-EGF*= zebrafish were subjected to SC transections and daily heat shocks for
658 1h at 33°C for 3 days prior to tissue collection. 4 mm SC tissues containing the lesion were
659 collected at 5 or 42 dpi. ChIP was performed using the SimpleChIP Plus Sonication Chip Kit (Cell
660 Signaling Technology #56383) according to the manufacturer's instructions. Briefly, input SC
661 tissues were cross-linked with 1% formaldehyde for 30 min, homogenized and sonicated.
662 Chromatin fragment sizes were confirmed by gel electrophoresis. Immunoprecipitation was
663 performed with either BACH1 polyclonal antibody (Proteintech #14018-1-AP) or rabbit IgG
664 negative control antibody (Cell Signaling Technology #2729). Samples were incubated with
665 primary antibodies overnight at 4°C, then with magnetic beads for 3h, washed and eluted.
666 Candidate BACH1 binding sites were assessed by RT-qPCR using primers targeting candidate
667 genomic regions containing the consensus BACH1 binding motif **(Table S5)**. Reported values
668 represent BACH1 immunoprecipitated enrichment in *Tg*$^{+}$ animals normalized to Rabbit IgG control
669 enrichment. A minimum of 12 SCs were pooled in each replicate.

670 **Luciferase Assay**

671 Human Embryonic Kidney 293 (HEK293) cells (atcc #CRL-1573) were cultured in DMEM/F-12
672 supplemented with 10% fetal bovine serum, 1x L-Glutamine and 500 U/mL of penicillin-
673 streptomycin. $15x10^3$ cells were seeded per well in a 96-well flat-bottom microplate (Corning
674 #07200566) and allowed to adhere for 36 h at 37ºC. Cells were transiently transfected with
675 different combinations pGL3-sv40-Firefly luciferase constructs encoding 6XBACH1 or the *sox2*
676 enhancer regions identified by ChIP. Cells were co-transfected with pcDNA3 constructs encoding
677 zebrafish Bach1a, Bach1b, Mafga, Mafgb or Maff, along with a Renilla luciferase control. Cells
678 were transfected using Lipofectamine-2000 following the manufacturer's protocol (Invitrogen
679 #15338100). Luciferase activity was evaluated using the Dual-Glo Luciferase Assay System
680 (Promega #E2920) 36 hours after transfection. Firefly and Renilla Luciferase activity were
681 measured using a Biotek Synergy HT plate reader. Reported values represent Firefly luciferase
682 activity normalized to Renilla activity and pcDNA3 vector controls.

REFERENCES

Adolf, B, Chapouton, P, Lam, CS, Topp, S, Tannhäuser, B, Strähle, U, Götz, M and Bally-Cuif, L (2006) 'Conserved and acquired features of adult neurogenesis in the zebrafish telencephalon', *Developmental biology* 295(1).

Ahlmann-Eltze, C. and Huber, W. (2021) 'glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data', *Bioinformatics* 36(24): 5701-5702.

Albagli, O., Dhordain, P., Deweindt, C., Lecocq, G. and Leprince, D. (1995) 'The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins', *Cell Growth & Differentiation: The Molecular Biology Journal of the American Association for Cancer Research* 6(9): 1193-1198.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003) 'Multipotent cell lineages in early mouse development depend on SOX2 function', *Genes Dev* 17(1): 126-40.

Barbosa, JS, Sanchez-Gonzalez, R, Di Giaimo, R, Baumgart, EV, Theis, FJ, Götz, M and Ninkovic, J (2015) 'Neurodevelopment. Live imaging of adult neural stem cell behavior in the intact and injured zebrafish brain', *Science (New York, N.Y.)* 348(6236).

Barnabe-Heider, F., Goritz, C., Sabelstrom, H., Takebayashi, H., Pfrieger, F. W., Meletis, K. and Frisen, J. (2010) 'Origin of new glial cells in intact and injured adult spinal cord', *Cell Stem Cell* 7(4): 470-82.

Barreiro-Iglesias, A., Mysiak, K. S., Scott, A. L., Reimer, M. M., Yang, Y., Becker, C. G. and Becker, T. (2015) 'Serotonin Promotes Development and Regeneration of Spinal Motor Neurons in Zebrafish', *Cell Rep* 13(5): 924-32.

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., Ginhoux, F. and Newell, E. W. (2018) 'Dimensionality reduction for visualizing single-cell data using UMAP', *Nat Biotechnol*.

Becker, C. G. and Becker, T. (2015) 'Neuronal regeneration from ependymo-radial glial cells: cook, little pot, cook!', *Developmental cell* 32(4): 516-27.

Becker, C. G., Lieberoth, B. C., Morellini, F., Feldner, J., Becker, T. and Schachner, M. (2004) 'L1.1 is involved in spinal cord regeneration in adult zebrafish', *J Neurosci* 24(36): 7837-42.

Berg, D. A., Kirkham, M., Wang, H., Frisen, J. and Simon, A. (2011) 'Dopamine controls neurogenesis in the adult salamander midbrain in homeostasis and during regeneration of dopamine neurons', *Cell Stem Cell* 8(4): 426-33.

Bolte, S. and CordeliÈRes, F. P. (2006) 'A guided tour into subcellular colocalization analysis in light microscopy', *Journal of Microscopy* 224(3): 213-232.

Bromley-Coolidge, S., Iruegas, D. and Appel, B. (2024) 'Cspg4 sculpts oligodendrocyte precursor cell morphology', *bioRxiv*.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018) 'Integrating single-cell transcriptomic data across different conditions, technologies, and species', *Nat Biotechnol* 36(5): 411-420.

Cavone, L., McCann, T., Drake, L. K., Aguzzi, E. A., Oprisoreanu, A. M., Pedersen, E., Sandi, S., Selvarajah, J., Tsarouchas, T. M., Wehner, D. et al. (2021) 'A unique macrophage subpopulation signals directly to progenitor cells to promote regenerative neurogenesis in the zebrafish spinal cord', *Developmental cell* 56(11): 1617-1630 e6.

Choi, Harry M. T., Schwarzkopf, Maayan, Fornace, Mark E., Acharya, Aneesh, Artavanis, Georgios, Stegmaier, Johannes, Cunha, Alexandre and Pierce, Niles A. (2018) 'Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust', *Development* 145(12).

Cigliola, V., Shoffner, A., Lee, N., Ou, J., Gonzalez, T. J., Hoque, J., Becker, C. J., Han, Y., Shen, G., Faw, T. D. et al. (2023) 'Spinal cord repair is modulated by the neurogenic factor Hb-egf under direction of a regeneration-associated enhancer', *Nat Commun* 14(1): 4857.

de Sena-Tomas, C., Rebola Lameira, L., Rebocho da Costa, M., Naique Taborda, P., Laborde, A., Orger, M., de Oliveira, S. and Saude, L. (2024) 'Neutrophil immune profile guides spinal cord regeneration in zebrafish', *Brain Behav Immun* 120: 514-531.

Dean, H. C., Saraswathy, V. M., Saini, A., Ou, T., McAdow, J., Tendolkar, A., Mokalled, M. H. and Johnson, A. N. (2025) 'Hb-EGF directs systemic muscle repair', *bioRxiv*.

Edwards-Faret, G., Munoz, R., Mendez-Olivos, E. E., Lee-Liu, D., Tapia, V. S. and Larrain, J. (2017) 'Spinal cord regeneration in Xenopus laevis', *Nat Protoc* 12(2): 372-389.

Fornes, O., Castro-Mondragon, J. A., Khan, A., van der Lee, R., Zhang, X., Richmond, P. A., Modi, B. P., Correard, S., Gheorghe, M., Baranasic, D. et al. (2020) 'JASPAR 2020: update of the open-access database of transcription factor binding profiles', *Nucleic Acids Res* 48(D1): D87-D92.

Fuse, Y., Nakajima, H., Nakajima-Takagi, Y., Nakajima, O. and Kobayashi, M. (2015) 'Heme-mediated inhibition of Bach1 regulates the liver specificity and transience of the Nrf2-dependent induction of zebrafish heme oxygenase 1', *Genes Cells* 20(7): 590-600.

Gao, Yuhua, Bai, Chunyu, Zheng, Dong, Li, Changli, Zhang, Wenxiu, Li, Mei, Guan, Weijun and Ma, Yuehui (2016) 'Combination of melatonin and Wnt-4 promotes neural cell differentiation in bovine amniotic epithelial cells and recovery from spinal cord injury', *Journal of Pineal Research* 60(3): 303-312.

751 Goldshmit, Yona, Sztal, Tamar E., Jusuf, Patricia R., Hall, Thomas E., Nguyen-Chi, Mai and
752 Currie, Peter D. (2012) 'Fgf-Dependent Glial Cell Bridges Facilitate Spinal Cord Regeneration in
753 Zebrafish', *The Journal of Neuroscience* 32(22): 7477.

754 Gorsuch, R. A., Lahne, M., Yarka, C. E., Petravick, M. E., Li, J. and Hyde, D. R. (2017) 'Sox2
755 regulates Muller glia reprogramming and proliferation in the regenerating zebrafish retina via
756 Lin28 and Ascl1a', *Exp Eye Res* 161: 174-192.

757 Graham, V., Khudyakov, J., Ellis, P. and Pevny, L. (2003) 'SOX2 functions to maintain neural
758 progenitor identity', *Neuron* 39(5): 749-65.

759 Grandel, H, Kaslin, J, Ganz, J, Wenzel, I and Brand, M (2006) 'Neural stem cells and
760 neurogenesis in the adult zebrafish brain: origin, proliferation dynamics, migration and cell fate',
761 *Developmental biology* 295(1).

762 Guo, X., Jiang, C., Zhang, Y., Chen, Z., Hao, D. and Zhang, H. (2023) 'Spermatogonial stem-
763 cell-derived neural-like cell transplantation enhances the functional recovery of a rat spinal cord
764 injury model: characterization of evoked potentials', *Front Neurosci* 17: 1289581.

765 Hafemeister, C. and Satija, R. (2019) 'Normalization and variance stabilization of single-cell
766 RNA-seq data using regularized negative binomial regression', *Genome Biol* 20(1): 296.

767 Haghverdi, L., Lun, A. T. L., Morgan, M. D. and Marioni, J. C. (2018) 'Batch effects in single-cell
768 RNA-sequencing data are corrected by matching mutual nearest neighbors', *Nat Biotechnol*
769 36(5): 421-427.

770 He, Y., Guo, J., Yu, Y., Jin, J., Jiang, Q., Li, Q., Ma, S., Pan, Q., Lin, J., Jiang, N. et al. (2023)
771 'BACH1 regulates the differentiation of vascular smooth muscle cells from human embryonic
772 stem cells via CARM1-mediated methylation of H3R17', *Cell Rep* 42(12): 113468.

773 Horky, L. L., Galimi, F., Gage, F. H. and Horner, P. J. (2006) 'Fate of endogenous
774 stem/progenitor cells following spinal cord injury', *J Comp Neurol* 498(4): 525-38.

775 Hui, S. P., Dutta, A. and Ghosh, S. (2010) 'Cellular response after crush injury in adult zebrafish
776 spinal cord', *Dev Dyn* 239(11): 2962-79.

777 Hui, S. P., Nag, T. C. and Ghosh, S. (2015) 'Characterization of Proliferating Neural Progenitors
778 after Spinal Cord Injury in Adult Zebrafish', *PloS one* 10(12): e0143595.

779 Ilicic, T., Kim, J. K., Kolodziejczyk, A. A., Bagger, F. O., McCarthy, D. J., Marioni, J. C. and
780 Teichmann, S. A. (2016) 'Classification of low quality cells from single-cell RNA-seq data',
781 *Genome Biol* 17: 29.

782 Itoh-Nakadai, A., Hikota, R., Muto, A., Kometani, K., Watanabe-Matsui, M., Sato, Y., Kobayashi,
783 M., Nakamura, A., Miura, Y., Yano, Y. et al. (2014) 'The transcription repressors Bach2 and

Bach1 promote B cell development by repressing the myeloid program', *Nat Immunol* 15(12): 1171-80.

Jia, Mengping, Li, Qinhan, Guo, Jieyu, Shi, Weihao, Zhu, Lei, Huang, Yijun, Li, Yongbo, Wang, Li, Ma, Siyu, Zhuang, Tao et al. (2022) 'Deletion of BACH1 Attenuates Atherosclerosis by Reducing Endothelial Inflammation', *Circulation Research* 130(7): 1038-1055.

Johansson, C. B., Momma, S., Clarke, D. L., Risling, M., Lendahl, U. and Frisén, J. (1999) 'Identification of a neural stem cell in the adult mammalian central nervous system', *Cell* 96(1): 25-34.

Johnson, K., Barragan, J., Bashiruddin, S., Smith, C. J., Tyrrell, C., Parsons, M. J., Doris, R., Kucenas, S., Downes, G. B., Velez, C. M. et al. (2016) 'Gfap-positive radial glial cells are an essential progenitor population for later-born neurons and glia in the zebrafish spinal cord', *Glia* 64(7): 1170-89.

Karimi-Abdolrezaee, S., Schut, D., Wang, J. and Fehlings, M. G. (2012) 'Chondroitinase and growth factors enhance activation and oligodendrocyte differentiation of endogenous neural precursor cells after spinal cord injury', *PLoS One* 7(5): e37589.

Katsani, K. R., Hajibagheri, M. A. and Verrijzer, C. P. (1999) 'Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology', *The EMBO journal* 18(3): 698-708.

Kippin, T. E., Kapur, S. and van der Kooy, D. (2005) 'Dopamine specifically inhibits forebrain neural stem cell proliferation, suggesting a novel effect of antipsychotic drugs', *J Neurosci* 25(24): 5815-23.

Klatt Shaw, D. and Mokalled, M. H. (2021) 'Efficient CRISPR/Cas9 mutagenesis for neurobehavioral screening in adult zebrafish', *G3 (Bethesda)* 11(8).

Klatt Shaw, D., Saraswathy, V. M., Zhou, L., McAdow, A. R., Burris, B., Butka, E., Morris, S. A., Dietmann, S. and Mokalled, M. H. (2021) 'Localized EMT reprograms glial progenitors to promote spinal cord repair', *Dev Cell* 56(5): 613-626 e7.

Kroehne, V, Freudenreich, D, Hans, S, Kaslin, J and Brand, M (2011) 'Regeneration of the adult zebrafish brain from neurogenic radial glia-type progenitors', *Development (Cambridge, England)* 138(22).

Lam, C. S., Marz, M. and Strahle, U. (2009) 'gfap and nestin reporter lines reveal characteristics of neural progenitors in the adult zebrafish brain', *Developmental dynamics : an official publication of the American Association of Anatomists* 238(2): 475-86.

Lawson, N. D., Li, R., Shin, M., Grosse, A., Yukselen, O., Stone, O. A., Kucukural, A. and Zhu, L. (2020) 'An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes', *Elife* 9.

Lee, Hyun Joon, Wu, Junfang, Chung, Jumi and Wrathall, Jean R. (2013) 'SOX2 expression is upregulated in adult spinal cord after contusion injury in both oligodendrocyte lineage and ependymal cells', *Journal of Neuroscience Research* 91(2): 196-210.

Levine, J. H., Simonds, E. F., Bendall, S. C., Davis, K. L., Amir el, A. D., Tadmor, M. D., Litvin, O., Fienberg, H. G., Jager, A., Zunder, E. R. et al. (2015) 'Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis', *Cell* 162(1): 184-97.

Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol* 15(12): 550.

Maden, M. and Varholick, J. A. (2020) 'Model systems for regeneration: the spiny mouse, Acomys cahirinus', *Development* 147(4).

Martens, David J., Seaberg, Raewyn M. and Van Der Kooy, Derek (2002) 'In vivo infusions of exogenous growth factors into the fourth ventricle of the adult mouse brain increase the proliferation of neural progenitors around the fourth ventricle and the central canal of the spinal cord', *European Journal of Neuroscience* 16(6): 1045-1057.

März, M, Chapouton, P, Diotel, N, Vaillant, C, Hesl, B, Takamiya, M, Lam, CS, Kah, O, Bally-Cuif, L and Strähle, U (2010) 'Heterogeneity in progenitor cell subtypes in the ventricular zone of the zebrafish adult telencephalon', *Glia* 58(7).

Matson, K. J. E., Sathyamurthy, A., Johnson, K. R., Kelly, M. C., Kelley, M. W. and Levine, A. J. (2018) 'Isolation of Adult Spinal Cord Nuclei for Massively Parallel Single-nucleus RNA Sequencing', *Journal of visualized experiments : JoVE*(140).

McDonough, A. and Martinez-Cerdeno, V. (2012) 'Endogenous proliferation after spinal cord injury in animal models', *Stem Cells Int* 2012: 387513.

McGinnis, C. S., Murrow, L. M. and Gartner, Z. J. (2019) 'DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors', *Cell Syst* 8(4): 329-337 e4.

Meletis, K., Barnabe-Heider, F., Carlen, M., Evergren, E., Tomilin, N., Shupliakov, O. and Frisen, J. (2008) 'Spinal cord injury reveals multilineage differentiation of ependymal cells', *PLoS Biol* 6(7): e182.

Millimaki, B. B., Sweet, E. M. and Riley, B. B. (2010) 'Sox2 is required for maintenance and regeneration, but not initial development, of hair cells in the zebrafish inner ear', *Dev Biol* 338(2): 262-9.

Mokalled, M. H., Patra, C., Dickson, A. L., Endo, T., Stainier, D. Y. and Poss, K. D. (2016) 'Injury-induced ctgfa directs glial bridging and spinal cord regeneration in zebrafish', *Science* 354(6312): 630-634.

Mothe, A. J. and Tator, C. H. (2005) 'Proliferation, migration, and differentiation of endogenous ependymal region stem/progenitor cells following minimal spinal cord injury in the adult rat', *Neuroscience* 131(1): 177-187.

Muraleedharan Saraswathy, Vishnu and Mokalled, Mayssa H. (2025) 'VMS-Fish: A Comprehensive Application for Zebrafish Husbandry Management', *Zebrafish* 22(2): 59-63.

Niu, Cong, Wang, Siqing, Guo, Jieyu, Wei, Xiangxiang, Jia, Mengping, Chen, Zhaoxiong, Gong, Wenxuan, Qin, Yue, Wang, Xinhong, Zhi, Xiuling et al. (2021) 'BACH1 recruits NANOG and histone H3 lysine 4 methyltransferase MLL/SET1 complexes to regulate enhancer–promoter activity and maintains pluripotency', *Nucleic Acids Research* 49(4): 1972-1986.

Ogai, K., Nakatani, K., Hisano, S., Sugitani, K., Koriyama, Y. and Kato, S. (2014) 'Function of Sox2 in ependymal cells of lesioned spinal cords in adult zebrafish', *Neurosci Res* 88: 84-7.

Ohori, Y., Yamamoto, S., Nagao, M., Sugimori, M., Yamamoto, N., Nakamura, K. and Nakafuku, M. (2006) 'Growth factor treatment and genetic manipulation stimulate neurogenesis and oligodendrogenesis by endogenous neural progenitors in the injured adult spinal cord', *J Neurosci* 26(46): 11948-60.

Oyake, T., Itoh, K., Motohashi, H., Hayashi, N., Hoshino, H., Nishizawa, M., Yamamoto, M. and Igarashi, K. (1996) 'Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site', *Molecular and Cellular Biology* 16(11): 6083-6095.

Paniagua-Torija, Beatriz, Norenberg, Michael, Arevalo-Martin, Angel, Carballosa-Gautam, Melissa M., Campos-Martin, Yolanda, Molina-Holgado, Eduardo and Garcia-Ovejero, Daniel (2018) 'Cells in the adult human spinal cord ependymal region do not proliferate after injury', *The Journal of Pathology* 246(4): 415-421.

Phipps, L. S., Marshall, L., Dorey, K. and Amaya, E. (2020) 'Model systems for regeneration: Xenopus', *Development* 147(6).

Pokhilko, A., Handel, A. E., Curion, F., Volpato, V., Whiteley, E. S., Bostrand, S., Newey, S. E., Akerman, C. J., Webber, C., Clark, M. B. et al. (2021) 'Targeted single-cell RNA sequencing of transcription factors enhances the identification of cell types and trajectories', *Genome Res* 31(6): 1069-1081.

Reimer, M. M., Norris, A., Ohnmacht, J., Patani, R., Zhong, Z., Dias, T. B., Kuscha, V., Scott, A. L., Chen, Y. C., Rozov, S. et al. (2013) 'Dopamine from the brain promotes spinal motor neuron generation during development and adult regeneration', *Developmental cell* 25(5): 478-91.

Reimer, M. M., Sorensen, I., Kuscha, V., Frank, R. E., Liu, C., Becker, C. G. and Becker, T. (2008) 'Motor neuron regeneration in adult zebrafish', *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28(34): 8510-6.

Renaux, E., Baudouin, C., Marchese, D., Clovis, Y., Lee, S. K., Gofflot, F., Rezsohazy, R. and Clotman, F. (2024) 'Lhx4 surpasses its paralog Lhx3 in promoting the differentiation of spinal V2a interneurons', *Cell Mol Life Sci* 81(1): 286.

Ribeiro, A., Monteiro, J. F., Certal, A. C., Cristovao, A. M. and Saude, L. (2017) 'Foxj1a is expressed in ependymal precursors, controls central canal position and is activated in new ependymal cells during regeneration in zebrafish', *Open Biol* 7(11).

Rodrigo Albors, A., Singer, G. A., Llorens-Bobadilla, E., Frisen, J., May, A. P., Ponting, C. P. and Storey, K. G. (2023) 'An ependymal cell census identifies heterogeneous and ongoing cell maturation in the adult mouse spinal cord that changes dynamically on injury', *Developmental cell* 58(3): 239-255 e10.

Sabelström, Hanna, Stenudd, Moa, Réu, Pedro, Dias, David O., Elfineh, Marta, Zdunek, Sofia, Damberg, Peter, Göritz, Christian and Frisén, Jonas (2013) 'Resident Neural Stem Cells Restrict Tissue Damage and Neuronal Loss After Spinal Cord Injury in Mice', *Science (New York, N.Y.)* 342(6158): 637-640.

Saraswathy, V. M., Zhou, L., McAdow, A. R., Burris, B., Dogra, D., Reischauer, S. and Mokalled, M. H. (2022) 'Myostatin is a negative regulator of adult neurogenesis after spinal cord injury in zebrafish', *Cell Rep* 41(8): 111705.

Saraswathy, V. M., Zhou, L. and Mokalled, M. H. (2024) 'Single-cell analysis of innate spinal cord regeneration identifies intersecting modes of neuronal repair', *Nat Commun* 15(1): 6808.

Sato, Masaki, Matsumoto, Mitsuyo, Saiki, Yuriko, Alam, Mahabub, Nishizawa, Hironari, Rokugo, Masahiro, Brydun, Andrey, Yamada, Shinji, Kaneko, Mika K., Funayama, Ryo et al. (2020) 'BACH1 Promotes Pancreatic Cancer Metastasis by Repressing Epithelial Genes and Enhancing Epithelial-Mesenchymal Transition', *Cancer Research* 80(6): 1279-1292.

Schaefer, T. and Lengerke, C. (2020) 'SOX2 protein biochemistry in stemness, reprogramming, and cancer: the PI3K/AKT/SOX2 axis and beyond', *Oncogene* 39(2): 278-292.

Schwarzkopf, M., Liu, M. C., Schulte, S. J., Ives, R., Husain, N., Choi, H. M. T. and Pierce, N. A. (2021) 'Hybridization chain reaction enables a unified approach to multiplexed, quantitative, high-resolution immunohistochemistry and in situ hybridization', *Development* 148(22).

Shah, Prajay T., Stratton, Jo A., Stykel, Morgan Gail, Abbasi, Sepideh, Sharma, Sandeep, Mayr, Kyle A., Koblinger, Kathrin, Whelan, Patrick J. and Biernaskie, Jeff (2018) 'Single-Cell Transcriptomics and Fate Mapping of Ependymal Cells Reveals an Absence of Neural Stem Cell Function', *Cell* 173(4): 1045-1057.e9.

Shihabuddin, Lamya S., Horner, Philip J., Ray, Jasodhara and Gage, Fred H. (2000) 'Adult Spinal Cord Stem Cells Generate Neurons after Transplantation in the Adult Dentate Gyrus', *The Journal of Neuroscience* 20(23): 8727-8735.

923 Shin, J., Chen, J. and Solnica-Krezel, L. (2014) 'Efficient homologous recombination-mediated
924 genome engineering in zebrafish using TALE nucleases', *Development* 141(19): 3807-18.

925 Silver, Jerry (2016) 'The glial scar is more than just astrocytes', *Experimental Neurology* 286:
926 147-149.

927 Sofroniew, Michael V. (2018) 'Dissecting spinal cord regeneration', *Nature* 557(7705): 343-350.

928 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, Hao, Y.,
929 Stoeckius, M., Smibert, P. and Satija, R. (2019) 'Comprehensive Integration of Single-Cell Data',
930 *Cell* 177(7): 1888-1902 e21.

931 Su, Z., Niu, W., Liu, M. L., Zou, Y. and Zhang, C. L. (2014) 'In vivo conversion of astrocytes to
932 neurons in the injured adult spinal cord', *Nat Commun* 5: 3338.

933 Sun, J., Hoshino, H., Takaku, K., Nakajima, O., Muto, A., Suzuki, H., Tashiro, S., Takahashi, S.,
934 Shibahara, S., Alam, J. et al. (2002) 'Hemoprotein Bach1 regulates enhancer availability of
935 heme oxygenase-1 gene', *The EMBO journal* 21(19): 5216-24.

936 Suzuki, Katsushi, Matsumoto, Mitsuyo, Katoh, Yasutake, Liu, Liang, Ochiai, Kyoko, Aizawa,
937 Yuta, Nagatomi, Ryoichi, Okuno, Hiroshi, Itoi, Eiji and Igarashi, Kazuhiko (2020) 'Bach1
938 promotes muscle regeneration through repressing Smad-mediated inhibition of myoblast
939 differentiation', *PLoS One* 15(8): e0236781.

940 Tai, W., Wu, W., Wang, L. L., Ni, H., Chen, C., Yang, J., Zang, T., Zou, Y., Xu, X. M. and Zhang,
941 C. L. (2021) 'In vivo reprogramming of NG2 glia enables adult neurogenesis and functional
942 recovery following spinal cord injury', *Cell Stem Cell* 28(5): 923-937 e4.

943 Takahashi, Kazutoshi and Yamanaka, Shinya (2006) 'Induction of pluripotent stem cells from
944 mouse embryonic and adult fibroblast cultures by defined factors', *Cell* 126(4): 663-676.

945 Tazaki, A., Tanaka, E. M. and Fei, J. F. (2017) 'Salamander spinal cord regeneration: The
946 ultimate positive control in vertebrate spinal cord regeneration', *Dev Biol* 432(1): 63-71.

947 Team, R Core (2018) 'R: A language and environment for statistical computing'.

948 Tendolkar, A. and Mokalled, M. H. (2025) 'Mechanisms underpinning spontaneous spinal cord
949 regeneration', *Development* 152(20).

950 Than-Trong, E, Kiani, B, Dray, N, Ortica, S, Simons, B, Rulands, S, Alunni, A and Bally-Cuif, L
951 (2020) 'Lineage hierarchies and stochasticity ensure the long-term maintenance of adult neural
952 stem cells', *Science advances* 6(18).

Than-Trong, E, Ortica-Gatti, S, Mella, S, Nepal, C, Alunni, A and Bally-Cuif, L (2018) 'Neural stem cell quiescence and stemness are molecularly distinct outputs of the Notch3 signalling cascade in the vertebrate adult brain', *Development (Cambridge, England)* 145(10).

Turner, R. and Tjian, R. (1989) 'Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers', *Science (New York, N.Y.)* 243(4899): 1689-1694.

Vandestadt, C., Vanwalleghem, G. C., Khabooshan, M. A., Douek, A. M., Castillo, H. A., Li, M., Schulze, K., Don, E., Stamatis, S. A., Ratnadiwakara, M. et al. (2021) 'RNA-induced inflammation and migration of precursor neurons initiates neuronal circuit regeneration in zebrafish', *Developmental cell* 56(16): 2364-2380 e8.

Walker, S. E. and Echeverri, K. (2022) 'Spinal cord regeneration - the origins of progenitor cells for functional rebuilding', *Curr Opin Genet Dev* 75: 101917.

Xu, Bai, Zhao, Yannan, Xiao, Zhifeng, Wang, Bin, Liang, Hui, Li, Xing, Fang, Yongxiang, Han, Sufang, Li, Xiaoran, Fan, Caixia et al. (2017) 'A Dual Functional Scaffold Tethered with EGFR Antibody Promotes Neural Stem Cell Retention and Neuronal Differentiation for Spinal Cord Injury Repair', *Advanced Healthcare Materials* 6(9).

Xu, C. and Su, Z. (2015) 'Identification of cell types from single-cell transcriptomes using a novel clustering method', *Bioinformatics* 31(12): 1974-80.

Xue, Xiaoyu, Shu, Muya, Xiao, Zhifeng, Zhao, Yannan, Li, Xing, Zhang, Haipeng, Fan, Yongheng, Wu, Xianming, Chen, Bing, Xu, Bai et al. (2022) 'Lineage tracing reveals the origin of Nestin-positive cells are heterogeneous and rarely from ependymal cells after spinal cord injury', *Science China Life Sciences* 65(4): 757-769.

Yang, P., Arnold, S. A., Habas, A., Hetman, M. and Hagg, T. (2008) 'Ciliary neurotrophic factor mediates dopamine D2 receptor-induced CNS neurogenesis in adult mice', *J Neurosci* 28(9): 2231-41.

Yang, S., Corbett, S. E., Koga, Y., Wang, Z., Johnson, W. E., Yajima, M. and Campbell, J. D. (2020) 'Decontamination of ambient RNA in single-cell RNA-seq with DecontX', *Genome Biol* 21(1): 57.

Zappia, L. and Oshlack, A. (2018) 'Clustering trees: a visualization for evaluating clusterings at multiple resolutions', *Gigascience* 7(7).

Zhang, S., Bell, E., Zhi, H., Brown, S., Imran, S. A. M., Azuara, V. and Cui, W. (2019) 'OCT4 and PAX6 determine the dual function of SOX2 in human ESCs as a key pluripotent or neural factor', *Stem Cell Res Ther* 10(1): 122.

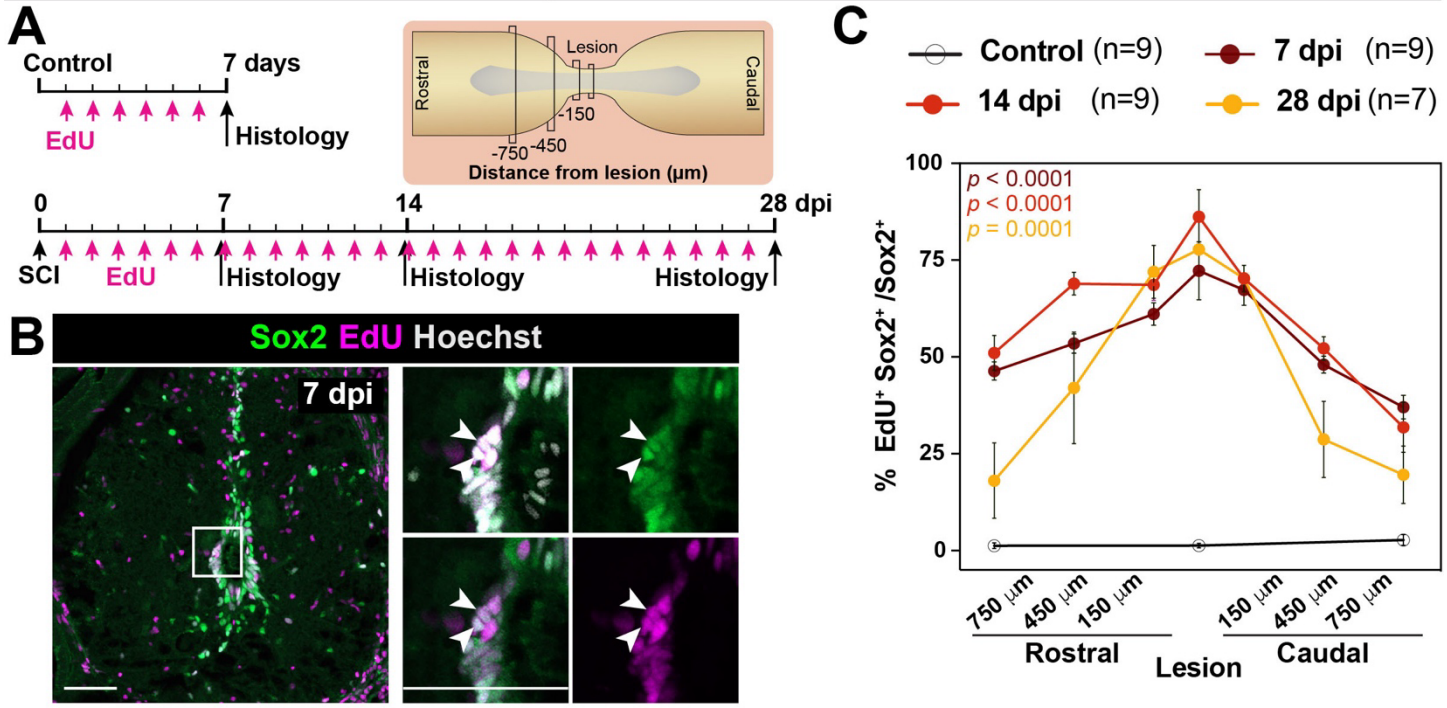
Zhang, S. and Cui, W. (2014) 'Sox2, a key factor in the regulation of pluripotency and neural differentiation', *World J Stem Cells* 6(3): 305-11.

988 Zheng, B. and Tuszynski, M. H. (2023) 'Regulation of axonal regeneration after mammalian
989 spinal cord injury', *Nat Rev Mol Cell Biol* 24(6): 396-413.

990 Zhou, L., McAdow, A. R., Yamada, H., Burris, B., Klatt Shaw, D., Oonk, K., Poss, K. D. and
991 Mokalled, M. H. (2023) 'Progenitor-derived glia are required for spinal cord regeneration in
992 zebrafish', *Development* 150(10).

Figure 1

Transient proliferation of Sox2⁺ cells after SCI



Self-renewal of *sox2*⁺ cells after SCI

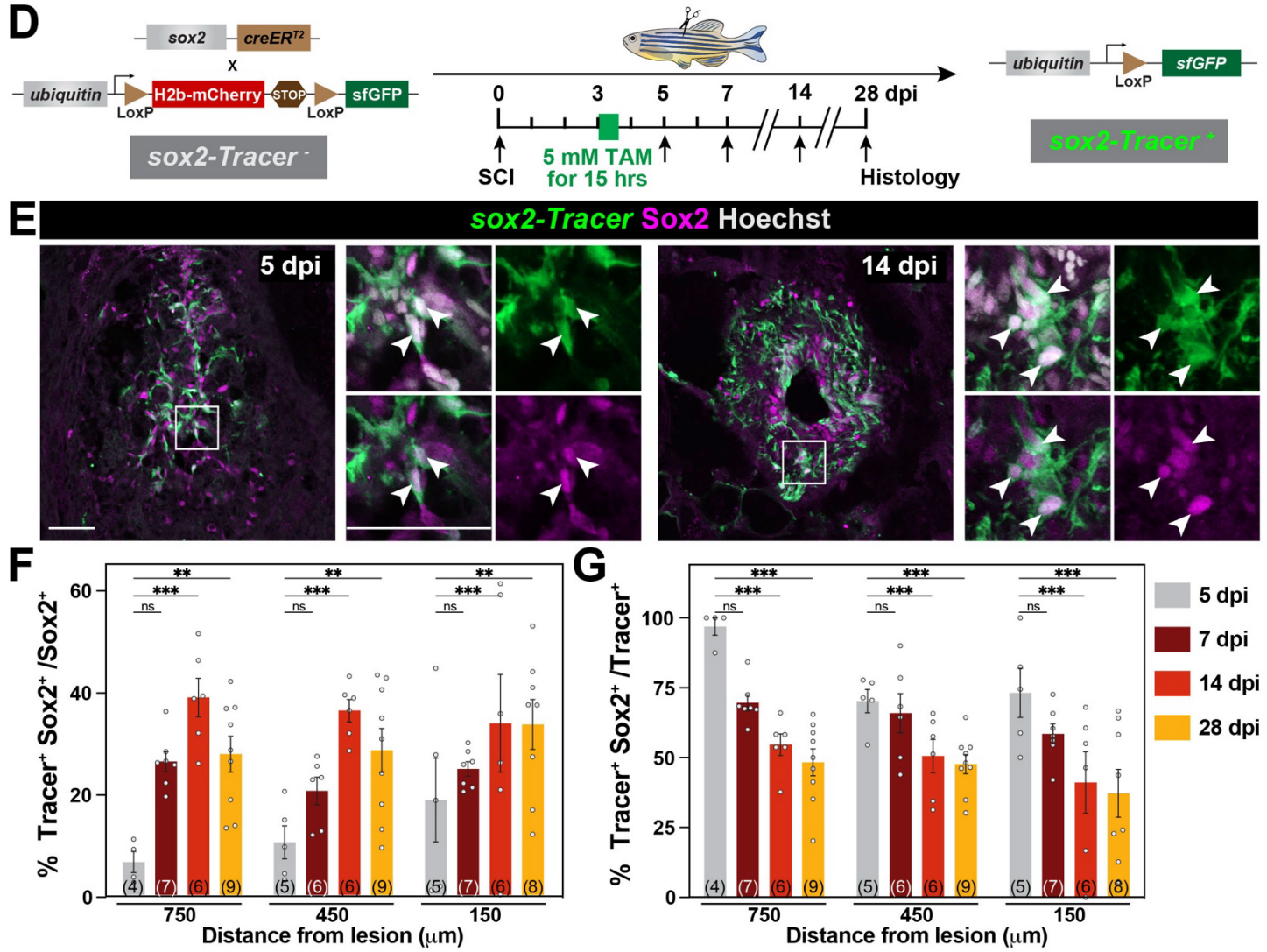


Figure 1. Sox2$^+$ cells undergo transient self-renewal after SCI. (A) Experimental timeline to evaluate the proliferation of Sox2$^+$ cells after SCI. Uninjured SC tissues were collected 6 days after daily EdU injections. Injured tissues were collected at 7, 14 and 28 dpi for histological examination. **(B)** Immunostaining for Sox2 (green), EdU (magenta) and Hoechst (grey) at 7 dpi. SC cross sections 750 µm from the lesion are shown. High-magnification insets show select Sox2$^+$ EdU$^+$ cells in triple, double or single channel views. Arrowheads indicate Sox2$^+$ EdU$^+$ cells. **(C)** Quantification of the numbers of Sox2$^+$ EdU$^+$ cells at 0, 7, 14 and 28 dpi. Percent Sox2$^+$ EdU$^+$ cells were normalized to Sox2$^+$ cells for each section. Cross sections 150, 450 and 750 µm rostral and caudal to the lesion were analyzed. **(D)** Experimental timeline for genetic lineage tracing of Sox2$^+$ cells after SCI. *sox2-Tracer* animals were treated with 5 mM Tamoxifen (TAM) for 15 hrs at 3 dpi. SC tissue sections were analyzed at 5, 7, 14 and 28 dpi. **(E)** Immunostaining for *sox2-Tracer*$^+$ (GFP, green), Sox2 (magenta) and Hoechst (grey). SC sections 450 µm from the lesion are shown at 5 and 14 dpi. Arrowheads indicate *sox2*-Tracer$^+$ Sox2$^+$ cells. **(F,G)** Quantification of *sox2-Tracer*$^+$ Sox2$^+$ cells at 5, 7, 14 and 28 dpi. For each section, the numbers of *sox2-Tracer*$^+$ Sox2$^+$ cells were normalized to either Sox2$^+$ cells (F), or to *sox2-Tracer*$^+$ cells (G). Quantifications were performed 150, 450 and 750 µm from the lesion. Data points indicate individual animals and sample sizes are indicated in parentheses. Two-way ANOVA was performed in C. Two-way ANOVA with Holm-Šidák's multiple comparisons were performed in J and L. Error bars represent SEM. ns, $p \geq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$. Scale bars: 50 µm, inset 10 µm.

Figure 2

snRNA-seq of *sox2-2a-sfGFP* zebrafish

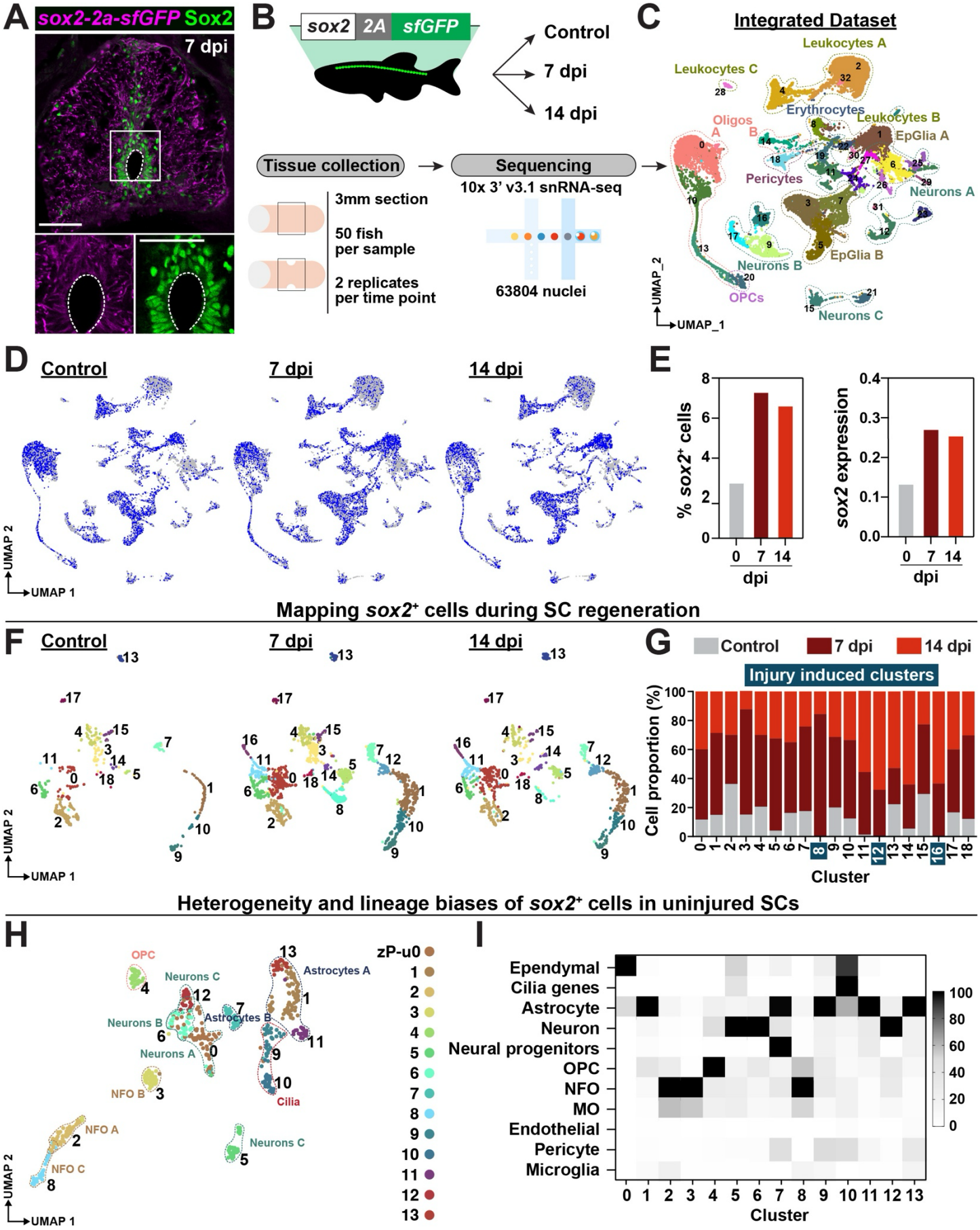
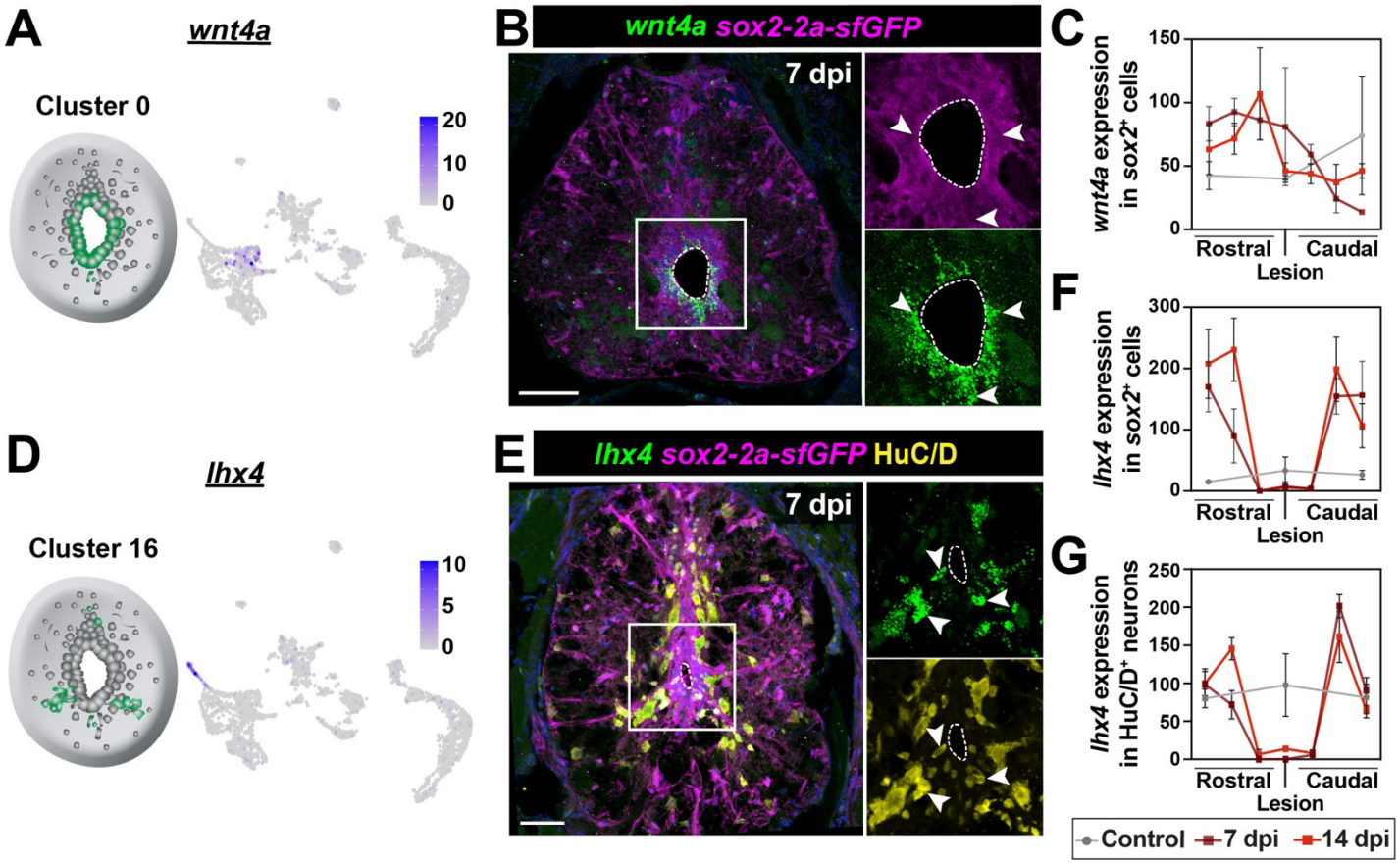


Figure 2. *sox2*$^+$ cells are heterogeneous and lineage-biased in homeostatic and lesioned SCs. **(A)** Sox2 staining in *sox2-2a-sfGFP* zebrafish at 7 dpi. SC cross sections 450 μm from the lesion are shown and dotted lines delineate central canal edges. **(B)** Experimental pipeline to generate a single-cell atlas of zebrafish SC progenitors. SC tissue collection, nuclear isolation and single nuclear RNA-seq (snRNA-seq) were performed on *sox2-2a-sfGFP* knock-in fish at 0, 7 and 14 dpi. **(C)** Merged UMAP representation of the integrated dataset. Two biological replicates, 3 time points and 63,804 nuclei were clustered into major spinal cell populations. **(D,E)** Feature plots of *sox2*$^+$ cells during SC regeneration. For each time point, cell proportions were normalized to the total number of cells at that time point. **(F)** Split UMAP representation shows 19 subclusters of *sox2*$^+$ cells at 0.6 resolution parameter. Two biological replicates and 3,620 *sox2*$^+$ cells were analyzed. **(G)** Distribution of *sox2*$^+$ cells per cluster during SC regeneration. For each time point and cluster, the numbers of *sox2*$^+$ cells were normalized to the total number of cells within that cluster. **(H)** UMAP representation of *sox2*$^+$ cells in uninjured SC tissues. Cluster identities and numbers were matched with their corresponding clusters from the integrated dataset in F. **(I)** Lineage biases of *sox2*$^+$ cells in uninjured SC tissues. The top DE markers for each cluster were cross-referenced with our assembled database of vertebrate nervous system markers (VNM marker database). Heatmap depicting -$\log_{10}$ p-value for the marker scoring output from the “DEMarkerScoring” algorithm is shown. One-tailed hypergeometric probability test was used to calculate the p-value. -$\log_{10}$ p-value for each cluster was scaled between 0-100. Cluster identity was assigned based on the highest -$\log_{10}$ p-value obtained for each cluster. Scale bars: 50 μm.

Figure 3

Mapping ependymal- and neuron-biased *sox2*⁺ cells *in vivo*



sox2⁺ cells give rise to neurons and glia after SCI

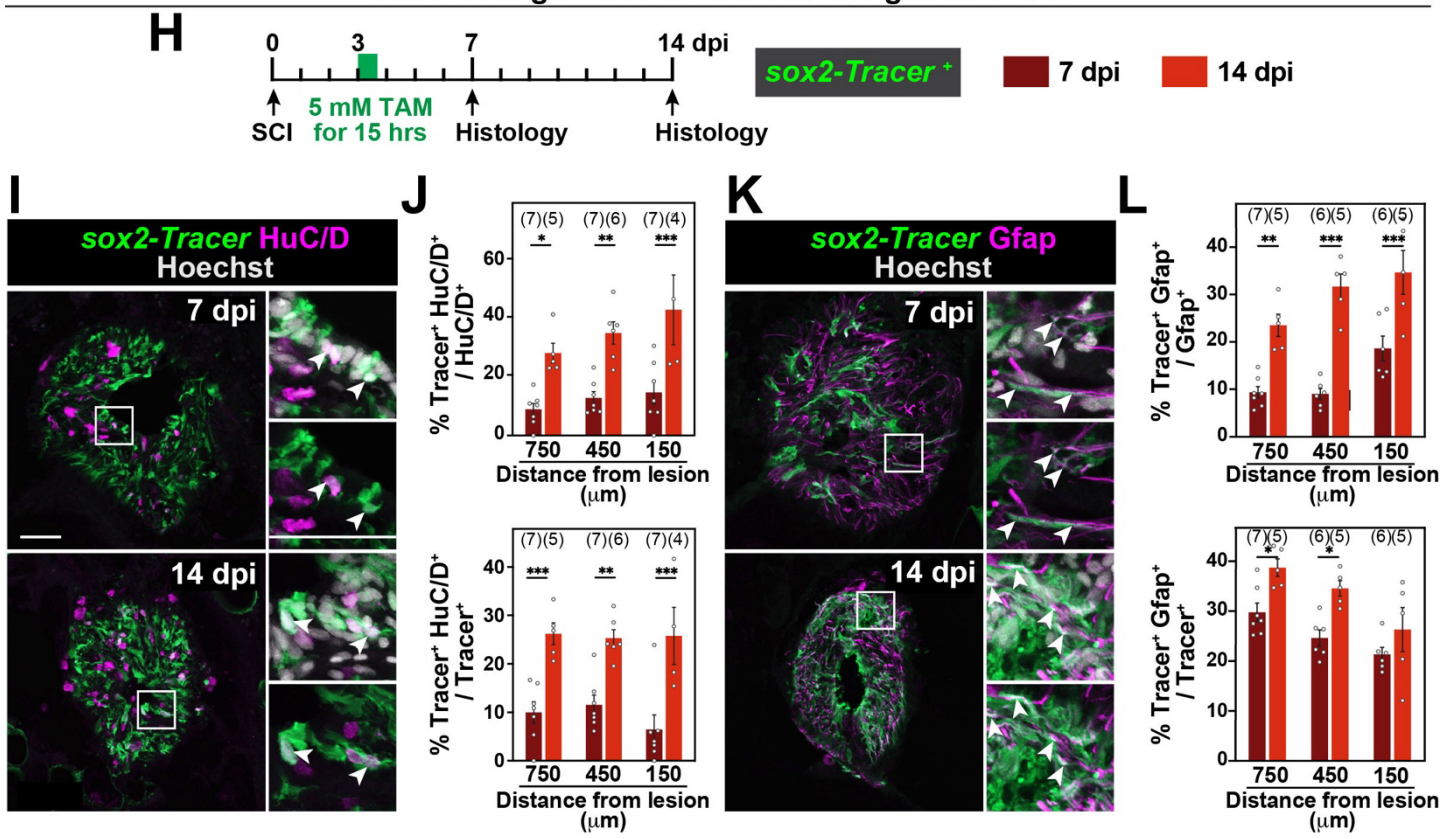
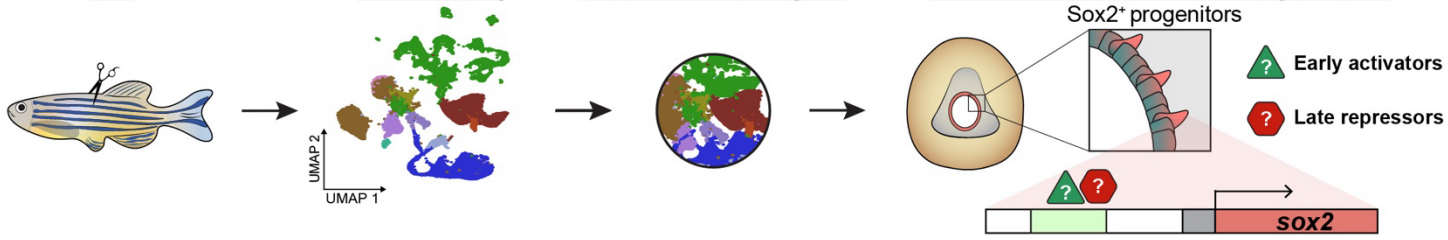
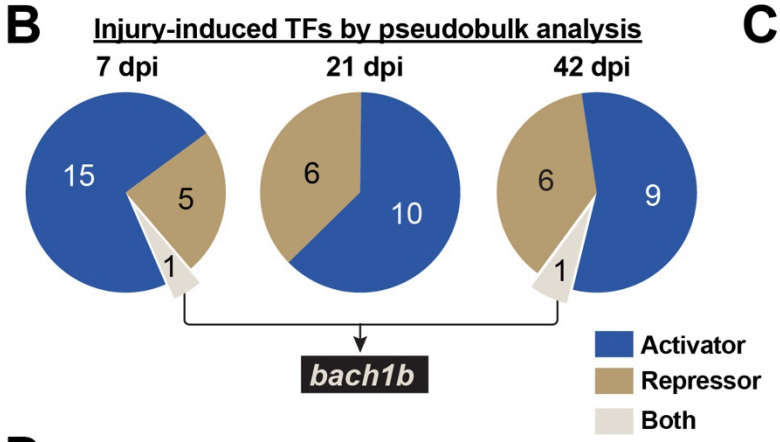


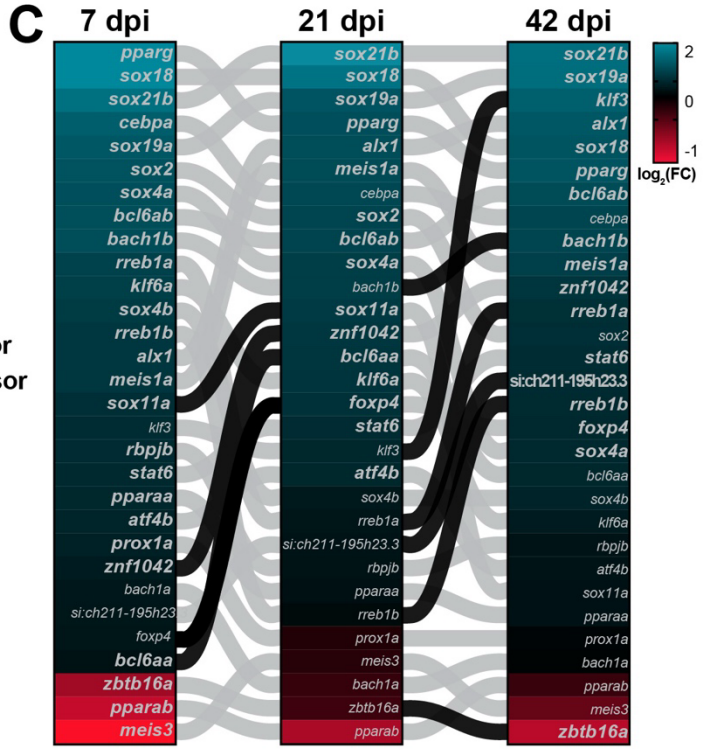
Figure 3. Compartmentalized niches of *sox2*$^+$ cells give rise to neurons and glia after SCI. **(A-C)** *in vivo* validation of ependymal-biased *sox2*$^+$ cell cluster 0. Feature plot shows *wnt4a* expression in c0. HCR *in situ* hybridization shows *wnt4a* transcripts (green) at 7 dpi. Arrowheads point to *wnt4a*$^+$*sox2-2a-sfGFP*$^+$ cells. Colocalization of *wnt4a* with *sox2-2a-sfGFP* was quantified at 0, 7 and 14 dpi. **(D-G)** *in vivo* validation of neuron-biased *sox2*$^+$ cell cluster 16. Feature plot shows *lhx4* expression in c16. HCR *in situ* hybridization shows *lhx4* transcripts (green) at 7 dpi. Arrowheads point to *lhx4*$^+$*sox2-2a-sfGFP*$^+$ cells. Colocalization of *wnt4a* with *sox2-2a-sfGFP* (F) and with HuC/D (G) were quantified at 0, 7 and 14 dpi. **(H)** Experimental timeline for genetic lineage tracing of Sox2$^+$ cells after SCI. *sox2-Tracer* animals were treated with 5 mM Tamoxifen (TAM) for 15 hrs at 3 dpi. SC tissue sections were analyzed at 7 and 14 dpi. **(I)** Immunostaining for *sox2*-Tracer$^+$ (GFP, green), HuC/D (magenta) and Hoechst (grey). SC sections from TAM-treated *sox2*-Tracer (Cre$^+$) animals are shown. Insets show high magnification *sox2*-Tracer$^+$ cells marked with white arrowheads. **(J)** Quantification of *sox2*-Tracer$^+$ and HuC/D$^+$ colocalization at 7 and 14 dpi. The numbers of *sox2*-Tracer$^+$ and Gfap$^+$ cells were quantified. The numbers of s*ox2*-Tracer$^+$ HuC/D$^+$ cells were normalized to the number of HuC/D$^+$ cells (upper panel), and to the total number of *sox2*-Tracer$^+$ cells (lower panel). **(K)** Immunostaining for *sox2*-Tracer$^+$ (GFP, green), Gfap (magenta) and Hoechst (grey). SC sections from TAM-treated *sox2*-Tracer (Cre$^+$) animals are shown. Insets show high magnification *sox2*-Tracer$^+$ cells with white arrowheads indicating *sox2*-Tracer$^+$ Gfap$^+$ staining. **(L)** Quantification of *sox2*-Tracer$^+$ and Gfap$^+$ colocalization at 7 and 14 dpi. *sox2*-Tracer$^+$ and Gfap$^+$ fluorescence area was quantified. s*ox2*-Tracer$^+$ HuC/D$^+$ expression was normalized to Gfap expression (upper panel), and to *sox2*-Tracer expression (lower panel). Data points indicate individual animals and sample sizes are indicated in parentheses. Two-way ANOVA with Holm-Šidák's multiple comparisons were performed in J and L. Error bars represent SEM. ns, $p \geq 0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$. Scale bars: 50 µm, inset 10 µm.

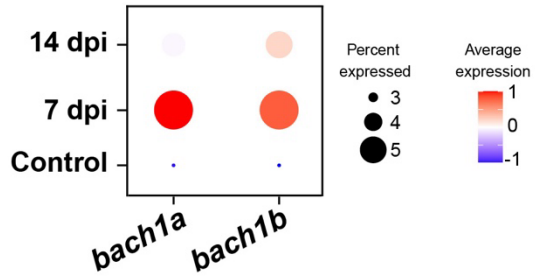
Figure 4

Screening for transcriptional activators and repressors in early vs late SC repair

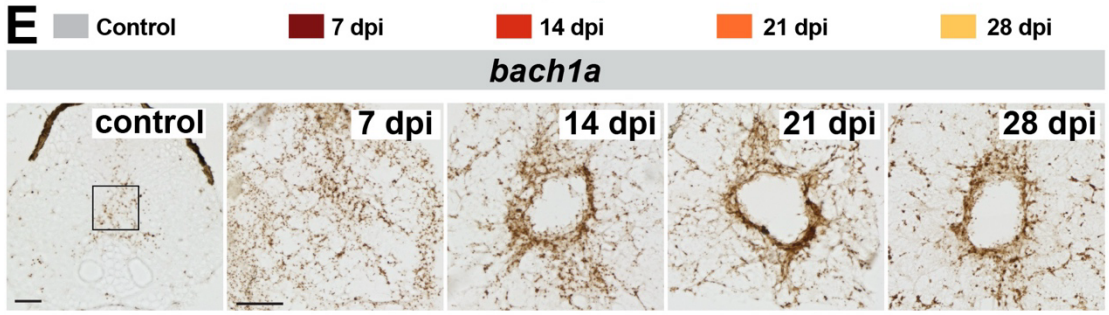


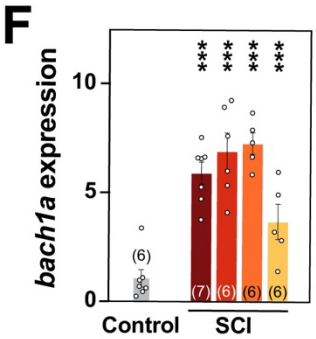


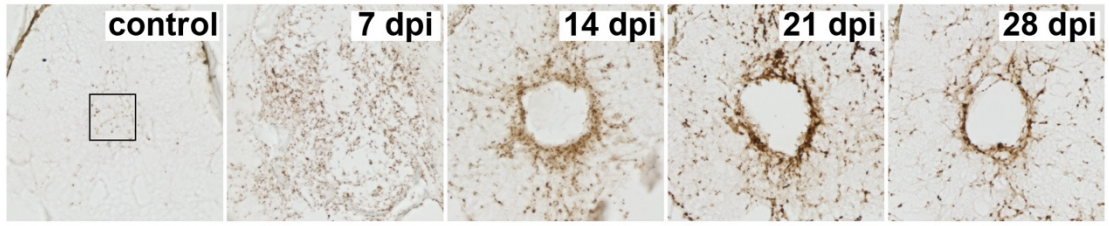




bach1a/b are upregulated in sub-acute and chronic SCI







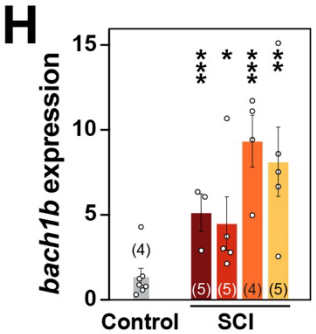
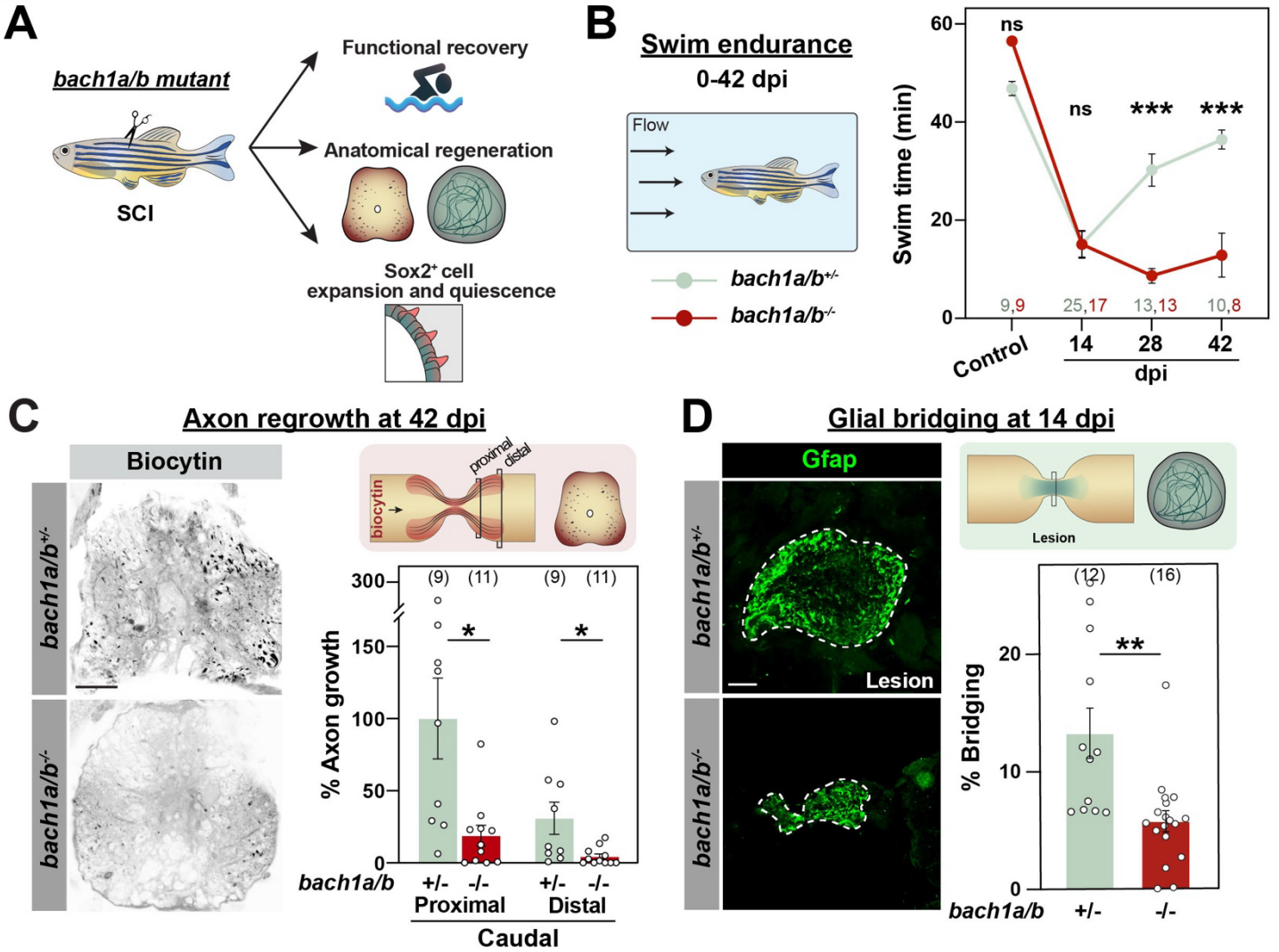


Figure 4. Screening for activators and repressors during SC regeneration. (A) Experimental pipeline to identify transcriptional activators and repressors in sub-acute and chronic SCI. Pseudobulk analysis was performed to overcome the limitation of detecting transcription factors by snRNA-seq. **(B)** Proportions of activators, repressors or dual-function transcription factors identified from pseudobulk analysis at 7, 21 and 42 dpi compared to uninjured controls. **(C)** Heatmap depicting the expression of differentially expressed transcription factors at 7, 21 and 42 dpi. Transcription factors that are differentially expressed compared to uninjured at any given time point are bolded. Dark connector lines indicate genes that gained significance between time points. **(D)** Dotplots represent *bach1a* and *bach1b* expression in subclustered *sox2*$^+$ cells at 0, 7 and 14 dpi. **(E)** RNAscope *in situ* hybridization for *bach1a* during SC regeneration. Cross sections from wild type zebrafish at 0, 7, 14, 21 and 28 dpi are shown. **(F)** Quantification of *bach1a* expression during SC regeneration. *bach1a*$^+$ area was quantified and normalized to uninjured controls. Sections 450 µm from the lesion were quantified. **(G)** RNAscope *in situ* hybridization for *bach1b* during SC regeneration. Cross sections from wild type zebrafish at 0, 7, 14, 21 and 28 dpi are shown. **(H)** Quantification of *bach1b* expression during SC regeneration. *bach1b*$^+$ area was quantified and normalized to uninjured controls. Sections 450 µm from the lesion were quantified. Data points indicate individual animals and sample sizes are indicated in parentheses. Two-way ANOVA with Holm-Šidák's multiple comparisons were performed in F and H. Error bars represent SEM. *$p<0.05$; **$p<0.01$; ***$p<0.001$. Scale bars: 50 µm, inset 10 µm.

Figure 5

SC regeneration in *bach1a/b* mutants



sox2+ cell expansion and quiescence in *bach1a/b* mutants

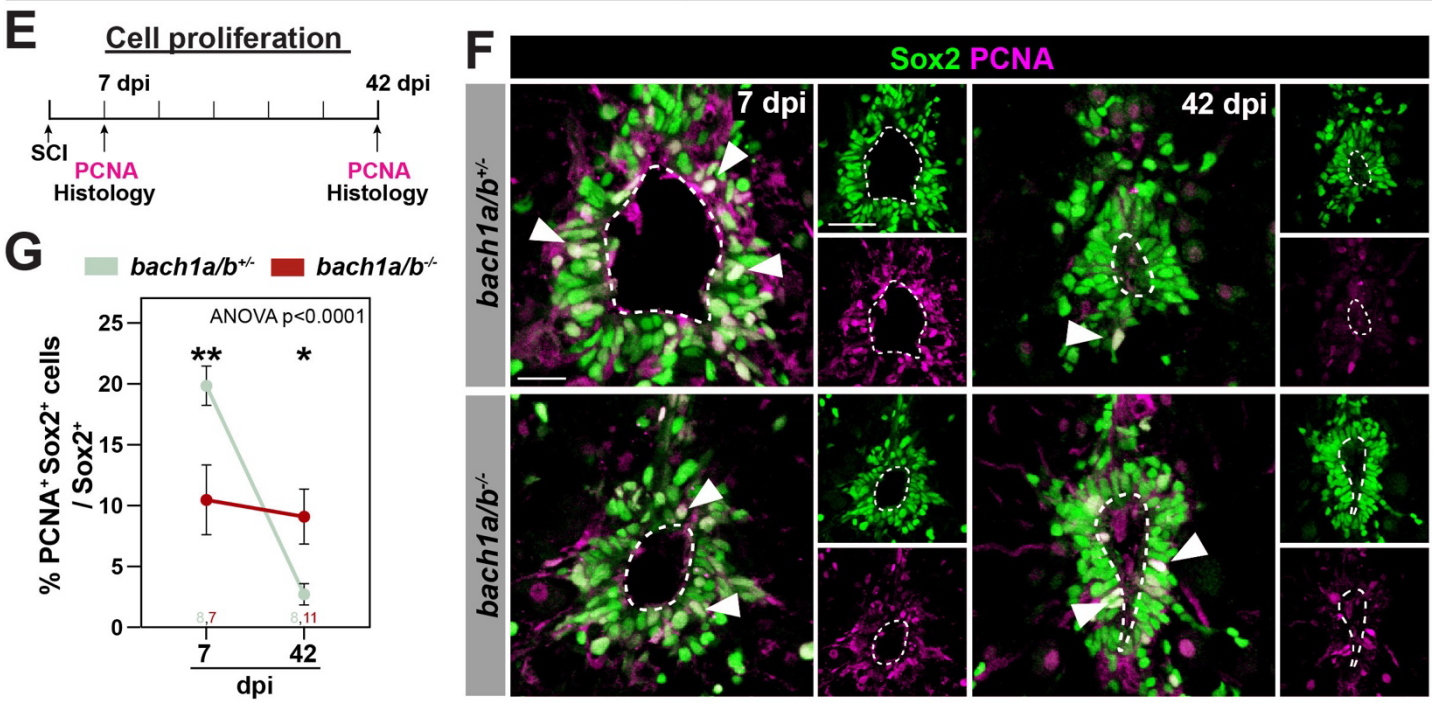


Figure 5. Loss of *bach1* impairs the expansion and quiescence of *sox2*$^{+}$ cells after SCI. (A) *bach1a/b* double mutant fish were subjected to SCI. Standard regeneration assays and the profiles of *sox2*$^{+}$ cells were assessed at 7 and 42 dpi. **(B)** Swim endurance for *bach1a/b*$^{-/-}$ fish and *bach1a/b*$^{+/-}$ control siblings at 14, 28 and 42 dpi. This behavioral test assessed the capacity of *bach1a/b*$^{-/-}$ animals to swim in an enclosed swim tunnel under increasing water current velocities. The time at exhaustion for each group is shown. **(C)** Anterograde axon tracing in *bach1a/b*$^{-/-}$ and *bach1a/b*$^{+/-}$ controls at 42 dpi. Biocytin axon tracer was applied rostrally and analyzed at 600 μm (proximal) and 1500 μm (distal) caudal to the lesion. Representative images show the extent of Biocytin labeling 600 μm (proximal) caudal to the lesion. Axon growth was first normalized to the extent of Biocytin rostral to the lesion, and then to labeling in *bach1a/b*$^{+/-}$ controls. **(D)** Glial bridging in *bach1a/b*$^{-/-}$ fish and *bach1a/b*$^{+/-}$ control siblings at 14 dpi. Representative immunohistochemistry shows the Gfap^{+} bridge at the lesion site. Percent bridging represents the cross-sectional area of the glial bridge at the lesion site relative to the intact SC 750 μm rostral to the lesion. **(E)** Experimental timeline to assess Sox2^{+} cell proliferation at 7 and 42 dpi. **(F)** PCNA (magenta) and Sox2 (green) staining in *bach1a/b*$^{-/-}$ and *bach1a/b*$^{+/-}$ controls. Representative micrographs 450 μm from the lesion are shown. Arrowheads indicate Sox2^{+} PCNA^{+} cells. **(G)** Quantification of Sox2^{+} cell proliferation at 7 and 42 dpi. For each section, the numbers of PCNA^{+} Sox2^{+} cells were quantified and normalized to Sox2^{+} cells. Quantifications were performed 450 μm from the lesion. For all quantifications, data points indicate individual animals and sample sizes are indicated in parentheses. Two-way ANOVA with Holm-Šidák's multiple comparisons were performed in B, C and G. Unpaired t test was performed in D. Error bars represent SEM. ns, $p \geq 0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$. Scale bars: 50 μm.

Figure 6

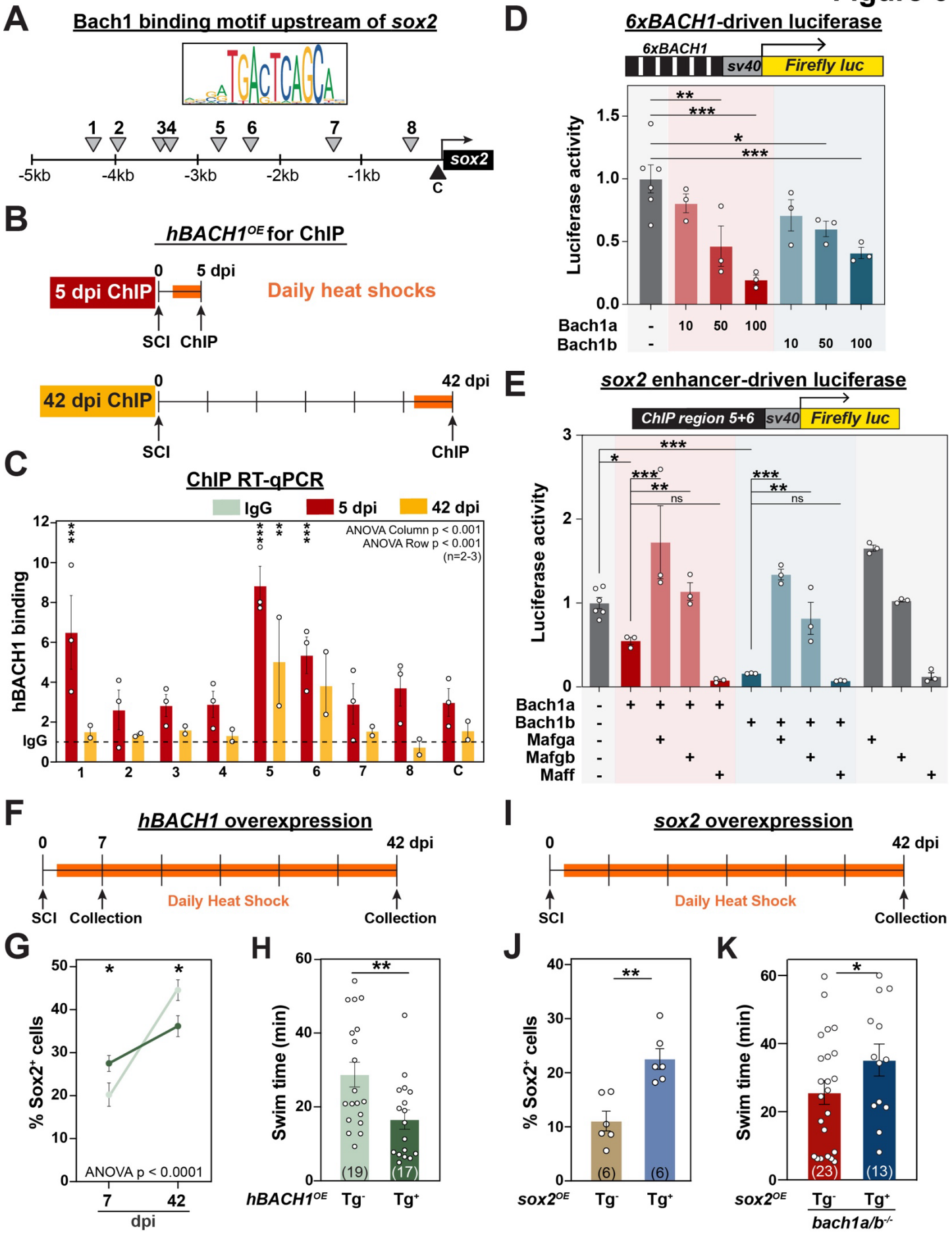

A
Bach1 binding motif upstream of sox2
1 2 34 5 6 7 8
-5kb -4kb -3kb -2kb -1kb
sox2
C
B
hBACH1OE for ChIP
5 dpi ChIP
0 5 dpi
Daily heat shocks
SCI ChIP
42 dpi ChIP
0 42 dpi
SCI ChIP
C
ChIP RT-qPCR
IgG 5 dpi 42 dpi
ANOVA Column p < 0.001
ANOVA Row p < 0.001
(n=2-3)
hBACH1 binding
IgG
1 2 3 4 5 6 7 8 C
D
6xBACH1-driven luciferase
6xBACH1 sv40 Firefly luc
Luciferase activity
Bach1a - 10 50 100
Bach1b - 10 50 100
E
sox2 enhancer-driven luciferase
ChIP region 5+6 sv40 Firefly luc
Luciferase activity
ns
Bach1a
Bach1b
Mafga
Mafgb
Maff
F
hBACH1 overexpression
0 7 42 dpi
SCI Collection Collection
Daily Heat Shock
I
sox2 overexpression
0 42 dpi
SCI Collection
Daily Heat Shock
G
% Sox2+ cells
ANOVA p < 0.0001
7 42
dpi
H
Swim time (min)
(19) (17)
hBACH1OE Tg- Tg+
J
% Sox2+ cells
(6) (6)
sox2OE Tg- Tg+
K
Swim time (min)
(23) (13)
sox2OE Tg- Tg+
bach1a/b-/-

Figure 6. Bach1 is a dual-regulator of *sox2* expression in sub-acute and chronic SCI. (A) Schematic representation of the genomic sequence upstream of *sox2*. Grey arrowheads indicate putative binding sites for hBACH1 +/-1 mismatch. The control region marked with black arrowhead was used to detect background binding by ChIP. **(B)** Experimental timeline for ChIP depicting 3 daily heat shocks and SC tissue collection at either 5 or 42 dpi. **(C)** qRT-PCR analysis of hBACH1 enrichment at genomic regions corresponding to those marked in A at 5 and 42 dpi. Y axis represents hBach1 binding. Numbers on the X axis represent putative Bach1 binding sites from 6A, C represents a control genomic region that does not contain Bach1 binding sites. hBach1 binding was calculated using the ΔCq method and normalized to IgG immunoprecipitation (dotted line). **(D)** Luciferase assays using *6xBACH1*-driven Firelfy luciferase. Quantification of *6xBACH1*-Firefly luciferase after co-expression of Bach1a or Bach1b. Y axis represents that activity of Firefly luciferase normalized to control Renilla luciferase and to control wells transfected with empty pcDNA3 vector. **(E)** Luciferase assays using *sox2* enhancer-driven Firelfy luciferase. A 402 bp genomic region containing Bach1 binding regions 5 and 6 upstream of *sox2* was used. Bach1a, Bach1b with or without Maf transcription factors were co-expressed. Firefly luciferase activity was normalized to Renilla luciferase and to control wells transfected with empty pcDNA3 vector. **(F-H)** Experimental timeline for *hBACH1* overexpression. $Sox2^{+}$ cells were quantified 450 μm from the lesion at 7 and 42 dpi (G). For each section, the numbers of $Sox2^{+}$ cells were normalized to the total number of nuclei. Swim endurance for $hBACH1^{OE}$ fish and control siblings was performed at 42 dpi (H). **(I-K)** Experimental timeline for *sox2* overexpression in *bach1a/b* mutants. $Sox2^{+}$ cells were quantified 450 μm from the lesion (J). Swim endurance for $sox2^{OE}$;$bach1a/b^{-/-}$ fish and $bach1a/b^{-/-}$ controls was performed at 42 dpi (K). Data points indicate individual animals or separate biological replicates, and sample sizes are indicated in parentheses. Two-way ANOVA with Holm-Šidák's multiple comparisons were performed in C and G. Ordinary one-way ANOVA with Holm-Šidák's multiple comparisons was performed in D and E. Unpaired t tests were performed in J and H. Error bars represent SEM. ns, $p \geq 0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$.