Multiple cis-regulatory elements collaborate to control *mdka* expression in zebrafish adult telencephalic neural stem cells during constitutive and regenerative neurogenesis

Jincan Chen[1], Agnes Hendriks[1], Masanari Takamiya[1], Tanja Beil[1], Nicolas Diotel[2] and Sepand Rastegar[1]*

1 Institute of Biological and Chemical Systems-Biological Information Processing (IBCS-BIP), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
2 Université de La Réunion, INSERM, UMR 1188, Diabète athérothrombose Thérapies Réunion Océan Indien (DéTROI), Saint-Pierre, La Réunion, France

*Correspondence to:
S. Rastegar, Institute of Biological and Chemical Systems-Biological Information Processing (IBCS-BIP), Karlsruhe Institute of Technology, Karlsruhe 76344, Germany. Tel: +49 721 608 22507. E-mail: sepand.rastegar@kit.edu.

Running title: Transcriptional regulation of *mdka* in the nervous system

Keywords: Neurogenesis; Radial glial cells; Regeneration; Zebrafish; *mdka*; Cis-regulatory elements; Gene transcription regulation

Abbreviations: CRE: cis-regulatory element; dpf: days post-fertilization; dpl: days post-lesion; EGFP: enhanced green fluorescent protein; FISH: fluorescent *in situ* hybridization; hpf: hours post-fertilization; IHC: immunohistochemistry; kb: kilo bases; PCNA: proliferating cell nuclear antigen; RGCs: radial glial cells; mRNA: messenger RNA; *mdka*: *midkine a*; NSCs: neural stem cells; S100β: S100 calcium-binding protein β; *Tg*: transgenic line; TSS: transcription start site; VZ: ventricular zone;

Abstract

Zebrafish is a powerful animal model for studying nervous system regeneration due to its remarkable regenerative abilities and the availability of diverse molecular tools. After telencephalic brain injury, neural stem cells (NSCs) in the ventricular zone (VZ) become activated, proliferate, and generate new neurons essential for brain repair. However, the molecular mechanisms regulating these processes remain unclear. Here, we investigate the transcriptional regulation of *midkine-a* (*mdka*), a heparin-binding growth factor gene encoding the secreted protein Midkine-a (Mdka), which is upregulated after injury in radial glial cells (RGCs), the bona fide NSCs of the adult zebrafish telencephalon. Using genome-wide bioinformatic analysis, we identified six putative cis-regulatory elements (CREs) associated with *mdka*. Transgenic assays showed that these CREs coordinate *mdka* expression during both development and regeneration. In the zebrafish embryo, CRE2, CRE3, CRE4, and CRE6 are required for EGFP expression in the nervous system, with CRE3 showing the strongest activity. In the adult telencephalon, CRE2, CRE4, and CRE6 are active in NSCs, with CRE2 best mimicking *mdka* expression at the ventricular zone. Importantly, individual CREs could not fully reproduce endogenous *mdka* expression, especially under regenerative conditions. In contrast, a combined CRE2/3/4/6 construct recapitulated *mdka* expression in both the embryo and adult telencephalon. These results suggest that *mdka* expression is controlled by a modular and cooperative cis-regulatory architecture that enables precise gene regulation during development, telencephalon homeostasis, and regeneration.

Introduction

Understanding the molecular mechanisms underlying tissue and organ regeneration is critical for addressing the growing challenges posed by neurodegenerative diseases, such as dementia, Parkinson's, and Alzheimer's. These diseases place a significant medical and societal burden on an aging population, underscoring the urgent need for effective therapeutic strategies to address these conditions. Zebrafish has emerged as a key vertebrate model for studying regeneration due to its exceptional ability to regenerate tissues throughout both embryonic and adult stages (Chen et al. 2024; Marques, Lupi, and Mercader 2019; Becker and Becker 2022; Gemberling et al. 2013; Kizil et al. 2012; Alunni and Bally-Cuif 2016). Unlike most mammals, zebrafish can regenerate organs like the heart, fin, retina, spinal cord, and specific brain regions without scarring and obvious disabilities, which makes them invaluable for studying regenerative processes (Marques, Lupi, and Mercader 2019). With over 70% genetic similarity to humans (Howe et al. 2013) and a vast number of mutants (Driever et al. 1996), transgenic lines (Ellingsen et al. 2005; Choe et al. 2021; Parinov et al. 2004), and genomic resources such as the DANIO-CODE Data Co-Ordination Center (DCC) (https://danio-code.zfin.org/), zebrafish provide a unique platform to study gene function, regulation, and cellular dynamics during embryonic development, adulthood under homeostatic and regenerative conditions (Hörtenhuber et al. 2020; Yang et al. 2020; Baranasic et al. 2022). In zebrafish, the adult telencephalon regenerates by activating dormant RGCs in the ventricular zone (VZ), which produce new neurons to replace those lost from injury (Diotel et al. 2020; Ghaddar et al. 2021). Recent transcriptomic studies have identified several genes significantly upregulated following injury (Rodriguez Viales et al. 2015; Demirci et al. 2020; Gourain et al. 2021; Lange et al. 2020), one of the most notable being *midkine-a* (*mdka*), a heparin-binding growth factor with multiple roles in development, repair and diseases (Lübke et al. 2022; Winkler and Yao 2014; Aynacioglu, Bilir, and Tuna 2019). Interestingly, *mdka* expression increases in quiescent RGCs after telencephalic injury, and its expression pattern is consistent under both physiological and regenerative conditions, resembling the behaviour of *id1* (*inhibitor of differentiation 1*) (Lübke et al. 2022), a transcriptional regulator downstream of the BMP signalling pathway (Zhang et al. 2021; Zhang et al. 2020). Both *mdka* and *id1* are predominantly expressed in quiescent type 1 cells (RGCs) and are absent in type 2 cells

corresponding to actively dividing (Lübke et al. 2022). Both of type 1 and type 2 cells express glial markers such as S100β and GFAP, while type 2 cells in addition express the proliferation marker PCNA (proliferating cell nuclear antigen). Lineage tracing experiments and neuroanatomical analysis have demonstrated that type 2 cells differentiate into type 3 neuroblasts, which migrate and mature into specific neuronal types (Adolf et al. 2006; März et al. 2011).

Beyond the brain, *mdka* has been implicated in the regeneration of other tissues in zebrafish and other vertebrates, including the heart, fin, and retina, suggesting that it may play a common role in the regenerative process (Grivas, Gonzalez-Rajal, and de la Pompa 2021; Gramage et al. 2015; Ang et al. 2020; Nagashima et al. 2020; Kikuchi-Horie et al. 2004; Lien et al. 2006; Thompson et al. 2020; Weinberger et al. 2024). The zebrafish *midkine* family includes two genes: *mdka* and its paralog *midkine b* (*mdkb*), the result of a gene duplication event (Winkler et al. 2003). However, only *mdka* is significantly upregulated after injury, highlighting its specific role in regenerative responses (Lübke et al. 2022). This study investigates the transcriptional regulation of *mdka* in the zebrafish telencephalon, focusing on its regulation in NSCs during neurogenesis and regeneration. We aim to uncover the molecular mechanisms behind its transcription by identifying and characterizing cis-regulatory elements (CREs) that control *mdka* expression in both homeostatic and regenerative contexts.

Our findings, using genome-wide bioinformatic approaches and validation by transgenesis, reveal that multiple CRMs work together to regulate *mdka* expression. In embryos, CRE2, CRE3, CRE4, and CRE6 are critical for nervous system expression, while in the adult telencephalon, CRE2, CRE4, and CRE6 are essential for expression in NSCs. Notably, none of the CREs alone were sufficient to fully recapitulate *mdka* expression or its strength in the embryonic nervous system or adult telencephalon. These findings highlight the complexity of *mdka* regulation, where individual CRMs contribute differently to expression patterns and intensity depending on molecular and environmental signals. Consistently, the EGFP reporter line containing CRE2, CRE3, CRE4, and CRE6 best mimicked endogenous *mdka* expression in both zebrafish embryos and the adult telencephalon under homeostatic and regenerative conditions, emphasizing the role of multiple molecular cues in *mdka* regulation.

Results

Identification and characterization of *mdka* CREs in the zebrafish embryo

To investigate the transcriptional regulation of *mdka*, we conducted a systematic search for CREs controlling *mdka* expression in zebrafish embryos at 24 hours post-fertilization (hpf) and in the ventricular zone of the adult telencephalon under both physiological and injury-induced conditions. Using the DANIO-CODE repository (Gonzalez, Baranasic, and Müller 2023; Baranasic et al. 2022), we identified six potential CREs located upstream, downstream, or in intergenic regions of the *mdka* coding sequence (**Figure 1**). These regions were selected based on chromatin accessibility, assessed by ATAC-seq (Assay for Transposase-Accessible Chromatin with sequencing), across zebrafish developmental stages from dome to long-pec (**Figure 1**). Additionally, we analysed sequence conservation with the carp genome (Chen et al. 2019) (data not shown). To determine the functional properties of these CREs, we classified their regulatory roles as promoters or enhancers and categorized their activation states (repressed, quiescent, or activated) and strength (weak or strong) based on epigenetic features, including DNA methylation and histone modifications.

To validate their regulatory activity, each CRE was cloned upstream of a *gata2* minimal promoter (Meng et al. 1997) driving EGFP expression (Navratilova et al. 2009; Zhang et al. 2020) (**Figure 2A**). These constructs were introduced into zebrafish germline cells to generate stable transgenic lines, *Tg(CREx:gata2aPR:EGFP)*, where CREx represents one of the identified CREs (**Figure 2A**). Five of the six CREs (CRE2-6) drove expression in 24 hpf zebrafish embryos, displaying partially overlapping patterns with endogenous *mdka* expression (**Figure 2B-G; Figure S1A, B**). Among them, *Tg(CRE3:gata2aPR:EGFP)* most closely recapitulated *mdka* expression (**Figure 2C, C'** and **Table 1**). In this transgenic line, EGFP expression was detected along the spinal cord, tail/fin, and brain (**Figure 2C, C'; Table 1; Figure S1A, B**). However, none of the individual CREs fully replicated the endogenous *mdka* expression pattern or intensity, indicating that multiple regulatory elements cooperate to establish proper *mdka* expression during embryogenesis (**Table 1**). Furthermore, each CRE was tested in combination with three different promoters: the *ctgf* promoter (Rastegar et al. 2008), the zebrafish synthetic promoter (zsp) (Begeman et al. 2022), and the *mdka* intrinsic

promoter (mdkaPR, this study). The expression of EGFP in these transient transgenics closely resembled that observed using the *gata2a* promoter (**Figure S2**), confirming that EGFP expression is attributed to the activity of the CREs rather than the specific promoter used.

Characterization of CREs driving *mdka* expression in adult NSCs

In the adult zebrafish telencephalon, EGFP expression was detected in four out of six transgenic lines generated (**Figure 3B, B'; C, C'; D, D'; F, F'**). No expression was observed for CRE1 and CRE5 (**Figure 3A, A'; E, E'**). CRE3 drove strong ectopic EGFP expression in the dorsal pallium (Dp) and lateral pallium (Dl) (**Figure 3C**), while no expression was seen in the VZ (**Figure 3C, C'**). In contrast, CRE2, CRE4, and CRE6 directed EGFP expression in the VZ of the telencephalon, where RGCs reside (**Figure 3B, B'; D, D'; F, F'**, and **Table 2**), reflecting the endogenous *mdka* expression pattern. In *Tg(CRE2:gata2aPR:EGFP)*, EGFP-positive cells were present along the VZ (**Figure 3B, B'**) and co-expressed with *mdka* mRNA (**Figure 3G-G''**). This transgenic line exhibited robust EGFP expression, primarily in quiescent S100β+/PCNA− type 1 RGCs, closely mirroring endogenous *mdka* expression (**Figure 3G-G''**, **H; Figure S1F**). These findings suggest that CRE2 is important for *mdka* regulation in the NSCs of adult telencephalon. In contrast, *Tg(CRE4:gata2aPR:EGFP)* (**Figure 3I-I''**) and *Tg(CRE6:gata2aPR:EGFP)* (**Figure 3K-K''**) displayed only partial overlap with endogenous *mdka* expression in the pallial VZ. Co-expression of *mdka* mRNA and EGFP was detected in the subpallial VZ, predominantly in type 1 RGCs (**Figure 3I–I'', K–K''**, white arrowheads; **Figure 3J, L**). Furthermore, CRE6 drove strong ectopic EGFP expression in neuron-like cells adjacent to RGCs in the pallial VZ (red arrowheads in **Figure 3F', K-K''**), a region where *mdka* is not normally expressed, indicating a lack of specificity. Despite capturing certain features of endogenous *mdka* expression, none of the individual CREs fully recapitulated its complete spatial distribution or expression strength in the adult telencephalon. A similar discrepancy was also observed during embryonic stages at 24 hpf (**Figure 2)**. While *Tg(CRE2:gata2aPR:EGFP)*, *Tg(CRE4:gata2aPR:EGFP)*, and *Tg(CRE6:gata2aPR:EGFP)* reflect aspects of *mdka* expression, their activity remains incomplete in both location and intensity. This suggests that single CREs are insufficient to fully regulate *mdka* in the adult

telencephalon. Interestingly, CREs that were strongly active in embryos (**Figure 2C**, **C'**) also exhibited activity in the adult telencephalon (**Figure 3C, C'**, yellow arrowheads). However, their expression patterns differed from that of the endogenous *mdka* mRNA, indicating that distinct regulatory elements likely control *mdka* expression at different developmental stages. These findings support the idea that *mdka* is regulated by stage-specific mechanisms involving combinatorial CRE activity.

Identification of injury-responsive CREs regulating *mdka* expression in the adult telencephalon

Since *mdka* expression is significantly upregulated during telencephalon regeneration following brain injury, particularly in type 1 RGCs (Lübke et al. 2022), we subjected all six *mdka* CRE transgenic lines to telencephalon injury (**Figure 4A–F, A'–F'**). Stab injury was induced in the right hemisphere using a 30G syringe needle inserted through the skull (März et al. 2011; Schmidt et al. 2014), while the uninjured left hemisphere served as an internal control (**Figure S1E**). Following injury, EGFP reporter expression was absent in the VZ of the Dm region of the adult telencephalon in transgenic lines CRE1, CRE3, and CRE5, in both injured and uninjured hemispheres (**Figure 4A, A', C, C', F,** and **F'**), consistent with the lack of expression observed under homeostatic conditions (**Figure 3A, A', C, C', E,** and **E'**). In contrast, transgenic lines carrying CRE2 (**Figure 4B, B', G–G", H, I**), CRE4 (**Figure 4D, D', J–J", K,** and **L**), and CRE6 (**Figure 4F, F', M–M", N,** and **O**) exhibited EGFP expression in the VZ, with a trend toward increased expression in the injured hemisphere. However, quantification of type 1 and type 2 RGCs, along with the number of EGFP-positive RGCs in injured versus uninjured hemispheres, did not confirm this trend statistically, as adjust *p*-values did not reach significance (**Figure 4H, I, K, L, N,** and **O**). Among CRE2, CRE4, and CRE6, only the reporter expression driven by CRE2 closely resembled the pattern of *mdka* mRNA expression and showed a more consistent trend toward upregulation after telencephalic stab injury (**Figure 4H, I; Figure S1G**) (Lübke et al. 2022). This suggests that none of the individual CREs possess a strong injury-responsive regulatory element capable of upregulating *mdka* expression during regeneration. Like the regulation under physiological conditions, it is possible that multiple CREs need to act in combination to respond to regenerative cues.

Cooperative CRE activity recapitulates endogenous *mdka* expression in embryos, adult RGCs, and in response to injury

The individual analysis of stable transgenic lines indicates that none of the six CREs alone fully restores the intensity and expression pattern of *mdka* genes throughout embryonic development and adulthood, under homeostatic or regenerative conditions. To explore potential cooperative interactions among these CREs, we generated a transgenic reporter construct (**Figure 5A**). This construct integrated CRE3, which exhibited strong regulatory activity in embryo, and CRE2, which proved most effective in adult telencephalon. Additionally, we included CRE4 and CRE6, both showing minor activity in embryos and the adult brain. This combined construct drove EGFP expression patterns closely mimicking endogenous *mdka* expression in embryos (**Figure 5B-B''**). In the adult telencephalon (**Figure 5C, D-D''**), a combination of CRE2, 3, 4, and 6 (referred to as CRE2346) mediated specific EGFP expression in the VZ (white arrowheads, **Figure 5D**), resembling the morphology of RGCs and co-expressing the RGC marker S100β. EGFP predominantly marked S100β+/PCNA− type 1 cells (83.76 ± 1.71%, **Figure 5E-E''', F-F''', G** and **H**), with a smaller fraction of EGFP+/S100β+ cells representing PCNA+ type 2 cells (16.24 ± 1.71%), indicating proliferative activity. Following stab injury, EGFP expression and intensity significantly increased in the right telencephalic hemisphere (**Figure 6G,** 27.20 ± 2.15 EGFP positive cells) compared to the control non-injured left hemisphere (20.99 ± 2.25, $p < 0.05$, **Figure 6A, B-B'', C-C'', F, G**). However, EGFP expression remained predominantly observed in quiescent cells post-injury (**Figure 6H**), the proportion of EGFP+/S100β+/PCNA− type 1 and EGFP+/S100β+/PCNA+ type 2 cells remained unchanged in the injured right hemisphere relative to the uninjured left side, consistent with endogenous *mdka* behaviour (Lübke et al. 2022).

Our findings demonstrate that the *mdka* CRE2346 construct induces EGFP expression specifically in RGCs, closely resembling the endogenous *mdka* expression pattern and responsive to injury. This highlights the necessity of multiple regulatory elements for precise spatial and temporal control of *mdka* in both physiological and regenerative contexts. Interestingly, all four CREs contain multiple transcription factor binding sites, including robust AP-1 binding sites within their evolutionarily conserved regions (**Figure S3**). AP-1 is a well-known regulator of gene expression during

regeneration across various tissues and species, including the zebrafish heart, fin, and tail (Beisaw et al. 2020; Nakamura et al. 2020; Wang et al. 2020; Tamaki et al. 2023; Weinberger et al. 2024; Dong et al. 2024), further supporting their potential role in the regenerative response to injury.

Discussion

Our study reveals that *mdka* expression in the zebrafish central nervous system is regulated by a complex and highly context-dependent network of cis-regulatory elements (CREs). These elements act in a modular, combinatorial, and sometimes redundant fashion, reflecting a multifaceted regulatory architecture that governs *mdka* transcription across different developmental stages and physiological states. Consistent with previous reports, *mdka* displays dynamic, tissue-specific expression patterns. For example, in the adult zebrafish telencephalon, *mdka* is expressed in RGCs under homeostatic conditions (Lübke 2022), while in the retina, it is absent in resting stem cells but becomes induced in proliferating Müller glia following injury (Gramage et al. 2015; Gramage, Li, and Hitchcock 2014). Interestingly, in the injured telencephalon, *mdka* expression remains restricted to quiescent RGCs, underscoring its tight, tissue-specific transcriptional control (Lübke 2022).

Modular and context-specific regulation of *mdka*

Our transgenic analyses demonstrate that *mdka* expression is orchestrated by multiple CREs, each contributing distinct spatiotemporal aspects to its regulatory program. Some elements, such as CRE5 and CRE3, are active primarily during embryogenesis. For example, CRE3 drives strong expression in the embryonic brain and spinal cord but fails to recapitulate specific adult brain patterns, instead causing ectopic and nonspecific activation. In contrast, elements like CRE2 are active mainly in the adult telencephalon, particularly within ventricular zone RGCs. This modularity is not limited to the brain. In the heart, previous studies showed that a transgene containing *mdka* intronic sequences (overlapping our CRE3, **Table S3**) was active in the developing epicardium but inactive after adult injury (Weinberger et al. 2024). Meanwhile, a downstream element (mdka_e2, overlapping CRE6, **Table S3**) was inactive during development but became induced in a subset of adult epicardial cells following injury (Weinberger et al. 2024). These examples reinforce the idea that distinct CREs mediate *mdka* expression in different developmental or physiological contexts.

Combinatorial input cooperative and redundant CRE activity

While individual CREs contribute to specific features of *mdka* expression, none alone were sufficient to fully recapitulate endogenous expression, particularly in response to injury. For instance, CRE2, CRE4, and CRE6 each drove weak expression in adult telencephalic RGCs, but none showed injury responsiveness on their own. However, when combined in a multi-CREs construct, these same elements collectively drove expression that closely mimicked endogenous *mdka*, including robust injury-induced activation. This supports a model in which full *mdka* expression depends on the cooperative action of multiple CREs. This combinatorial regulation may reflect the integration of distinct signaling inputs across separate regulatory elements. One possibility is that different signaling pathways act through specific CREs, with their additive or synergistic activation required to achieve the correct spatiotemporal expression profile. Alternatively, or in addition a single signaling pathway may target several CREs in parallel, with expression only reaching functional levels when all responsive elements are engaged. The presence of conserved AP-1 binding motifs within CRE2, CRE3, CRE4, and CRE6 (**Figure S3A-D**), coupled with the established role of the AP-1 complex in regeneration (Tamaki et al. 2023; Thompson et al. 2020; Beisaw et al. 2020), supports both scenarios. Importantly, our observations also reveal redundancy among CREs. Although each element has specific functions, some show overlapping activity. For instance, despite weak individual responses to injury, CRE2, CRE4, and CRE6 all target the same ventricular zone region. Furthermore, CRISPR/Cas9-mediated deletion of a regulatory element in the regenerating fin did not abolish *mdka* expression, and regeneration proceeded without detectable patterning defects (Thompson et al. 2020). This suggests that redundant or compensatory enhancer activity may preserve gene expression when individual elements are disrupted. Such redundancy adds robustness to the regulatory system, ensuring consistent *mdka* activation even in fluctuating or damaged environments. Together, our results highlight a regulatory logic in which enhancer modularity, cooperation, and redundancy work in concert to control *mdka* expression with high precision and resilience.

Broader implications for gene regulation and regenerative biology

The modular and distributed regulatory architecture of *mdka* is consistent with classical models of enhancer function. The *even-skipped (eve)* gene in *Drosophila*, for

example, is expressed in a seven-stripe pattern during embryogenesis, with each stripe controlled by a separate enhancer integrating local signals (Goto, Macdonald, and Maniatis 1989; Berrocal et al. 2020). Similarly, in mammals, the neural crest gene *sox10* is regulated by multiple enhancers with overlapping and redundant functions, enabling robust expression across tissues and stages (Werner et al. 2007; Betancur, Bronner-Fraser, and Sauka-Spengler 2010). These cases, like *mdka*, illustrate how enhancer modularity enables nuanced, context-dependent expression and buffering against perturbation. Our *mdka* reporter lines provide a useful system for dissecting transcriptional programs that govern adult neurogenesis and injury-induced gene activation. These tools could be employed to identify upstream regulators of regenerative gene expression and screen for factors that modulate regenerative capacity *in vivo*.

Regulatory complexity vs. simplicity: A comparative perspective

Intriguingly, the complexity of *mdka* regulation contrasts with the simpler enhancer architecture observed for other genes expressed in the same cellular context. For example, *id1*, which is co-expressed with *mdka* in quiescent RGCs under both physiological and regenerative conditions, appears to be controlled by a single enhancer responsive primarily to BMP signaling (Zhang et al. 2021; Zhang et al. 2020). Despite their overlapping expression domains, *id1* and *mdka* rely on very different regulatory strategies. This contrast raises important questions: Why do some genes require complex enhancer networks while others are regulated through simpler mechanisms? Comparative studies between genes like *mdka* and *id1* could illuminate how enhancer complexity evolves and functions in the context of regeneration and plasticity.

Conclusion

In summary, *mdka* expression in zebrafish is regulated by a network of cis-regulatory elements that act in a modular, cooperative, and partially redundant manner. This intricate architecture enables precise gene regulation during both developmental and regenerative neurogenesis. The inability of individual CREs to fully recapitulate endogenous patterns highlights the importance of enhancer synergy and redundancy. Our findings shed light on the sophisticated regulatory logic underlying regeneration-associated gene expression and provide a foundation for future efforts to manipulate these pathways for therapeutic purposes.

Materials and methods

Zebrafish husbandry

All zebrafish lines were maintained at the European Zebrafish Resource Center (EZRC). Experiments were performed on 6- to 12-month-old wild-type fish or transgenic reporter lines (**Table S1**). Animal husbandry was carried out in accordance with permit guidelines under §11 TSchG (Regierungspräsidium Karlsruhe, Aktenzeichen 35–9185.64/BH KIT). Fish were kept according to the European Society for Fish Models in Biology and Medicine (EuFishBioMed) guidelines at the Karlsruhe Institute of Technology (KIT) (Aleström et al. 2020), and animal experimentation adhered to German animal protection standards, approved by the Government of Baden-Württemberg (Regierungspräsidium Karlsruhe, Reference Numbers: 35-9185.81/G-214/21 and 35-9185.81/G-215/21). Efforts were made to minimize animal distress and reduce their number.

Plasmids construction

Plasmids for enhancer/promoter reporter assays were generated using Gateway cloning and TOPO/TA cloning strategies. The primers used for PCR amplification of putative CREs, and promoter are listed in **Table S2,** while the coordinates and sizes of the CREs and the promoter are provided in **Table S3**. Most plasmids (**Table S4**) were constructed using commonly available vectors, with detailed protocols available on the Chien Lab website: http://tol2kit.genetics.utah.edu/index.php/Protocols. The pTol2:mdkaCRE2346:gata2aPR:egfp:pA plasmid was created by combining four *mdka* CREs (CRE2, CRE3, CRE4, and CRE6) into the pT2KHGPzGATA2C1 destination vector and was commercially synthesized by GenScript Biotech Corporation.

Microinjection to generate transient and stable transgenesis

One-cell stage embryos were injected with 1 µl of 20 ng/µl Tol2 transposase mRNA, 1 µl of 10 % phenol red, 50 ng/µl plasmid DNA, and nuclease-free water (final volume 10 µl) directly into the yolk through the chorion. After injection, embryos were incubated at 28 °C in E3 medium with methylene blue. At 24 hpf, embryos were sorted based on EGFP fluorescence. Positive F0 founders were identified by EGFP

expression, raised to sexual maturity, and outcrossed with wildtype zebrafish. Stable EGFP expression in ~20 % of F1 progeny confirmed successful stable transgenesis.

Fluorescent *in situ* hybridization (FISH) of whole-mount embryos

Embryos were collected at the desired developmental stage, and chorions were removed manually or with pronase (1:100 w/v, 28.5 °C). The *mdka* probe (Molecular Instruments Inc.) was used, following the HCR™ RNA-FISH protocol with minor modifications, available on the Molecular Instruments website.

Stab injury of adult telencephalon and preparation of sections

The stab injury experiment and telencephalon sections were performed following the protocol described by Schmidt et al. (2014).

FISH on adult telencephalon sections and embryos

For experiments using probes from Molecular Instruments Inc., the protocol "HCR™ RNA-FISH protocol for fresh frozen or fixed frozen tissue sections was followed, with minor modifications. Specifically, for 24 hpf embryos, proteinase K treatment was conducted at a concentration of 10 µg/ml (diluted from a 1000 × stock to 1 × working concentration) for 10 min to improve probe penetration. The *mdka* probe used for fluorescent *in situ* hybridization was purchased from the Molecular Instruments Inc. with a dilution of 1:500.

Immunohistochemistry (IHC) on adult telencephalon sections

Vibratome sections were blocked with 1 ml blocking buffer for 1 hour at RT. After removing the buffer, sections were incubated overnight at 4 °C with primary antibodies in 300 µl blocking buffer. The next day, sections were washed 3 × 10 min with PTW (1 × PBS containing 0.1% Tween-20), then incubated with secondary antibodies (diluted in PTW) for 2 hours at RT in the dark. DAPI staining (1:2000 in PTW) was performed for 20 min if needed, followed by 3 × 5 min washes in PTW. Sections were mounted and imaged. All antibodies used are listed in **Table S5**.

Mounting for imaging

For embryos, dechorionated embryos were embedded in a 1.6 % low melting-point agarose (Sigma) and 0.016 % Tricaine (Sigma) mixture in E3 medium, oriented within 1 min using a needle or fine hair, left at room temperature for 3 min, cooled at 4 °C for 3 min, supplemented with water drops to prevent evaporation, and stored at 4 °C until imaging. While for adult telencephalon sections, fluorescently labelled brain sections were mounted on glass slides with 0.17 mm coverslips using Aqua Polymount (PolySciences Inc., USA) and stored at 4 °C in the dark before and after imaging (Schmidt et al. 2014).

Confocal imaging

Fluorescently labelled embryos and telencephalon sections were imaged using Leica TCS SP5 DM5000 and SP8/DLS confocal fluorescence microscope (Leica Microsystems, Wetzlar, Germany) equipped with 405 nm, 488 nm, 561 nm and 633 nm lasers. Excitation (Ex) wavelengths and emission (Em) ranges were as follows: Ex 488 nm for AlexaFluor 488 (EGFP, Em 492–550 nm), Ex 561 nm for AlexaFluor 561 (PCNA, Em 565–605 nm), Ex 633 nm for AlexaFluor 633 (S100β, Em 650–740 nm), Ex 633 nm for AlexaFluor 633-labeled *mdka* RNA probe (EM 655–724 nm), and Ex 405 nm for DAPI (Em 409–480 nm). Imaging was performed with sequential scanning at 16-bit color depth with a 1.0 AU pinhole aperture. Images were acquired at a resolution of 1024 × 1024-pixel resolution, with a pixel size of approximately 1.55 µm × 1.55 µm (field of view: ~1.55 mm × 1.55 mm), using a 63 ×/1.20 NA Water Immersion objective (Leica Microsystems) at a scanning frequency of 400 Hz. Laser intensity, PMT gain, detector settings, AOTF (Acousto-Optic Tuneable Filters) attenuation and other imaging settings were kept constant across experimental groups to ensure comparability. Imaging was performed using a wide-field compound fluorescence microscope (e.g., the images in Figure S2) or a confocal microscope (for all other images), with specific imaging details provided per case.

Image processing and arrangement

Images were processed with consistent exposure times and gain settings across all series. Z-stack maximum projections were generated and cropping or rotation was performed in ImageJ/Fiji. For wide-field images in **Figure S2**, deconvolution using the Iterative Deconvolve 3D and Diffraction PSF 3D plugins were applied to remove out-of-focus blur and enhance clarity. All figures, including flowcharts, were created and arranged using Microsoft PowerPoint (version 16.78) or Adobe Illustrator 2022 (version 26.4.1).

Statistical analysis

Type 1 and type 2 RGCs expressing EGFP were quantified in 1 µm steps across 50 µm-thick z-stacks for *Tg(CRE2:gata2aPR:EGFP), Tg(CRE4:gata2aPR:EGFP)*, and *Tg(CRE6:gata2aPR:EGFP),* while for *Tg(CRE2346:gata2aPR:EGFP)*, only single slices were analysed. Imaging was performed using a 63 ×/1.20 NA Water Immersion objective on a TCS-SP5 microscope (Leica Microsystems). Manual counting of positive and negative cells of PCNA, *mdka* EGFP, and S100β was performed in Fiji ImageJ. Three sections per brain from at least three individuals were analysed (*n* provided in figure legends). All statistical analyses were performed and visualized using GraphPad Prism (v10.2.0). Normality of the data was assessed using the Shapiro-Wilk test. For comparisons between two groups, an unpaired t-test (Welch's t-test) was used if both groups followed a normal distribution, while the Mann-Whitney U test was applied if normality assumptions were violated. For multiple-group comparisons, a one-way ANOVA was conducted when normality assumptions were met. If the normality assumption was not satisfied, a Kruskal-Wallis test was performed instead. To account for multiple pairwise comparisons, *p*-values were adjusted using the Tukey correction. Statistical significance was categorized as APA style: ns ($p > 0.05$, not significant), * ($p < 0.05$, significant), ** ($p < 0.01$, significant), and *** ($p < 0.001$, significant). Data are presented as mean ± S.E.M.

Acknowledgements

We would like to sincerely thank Anne Schröck for her invaluable technical support throughout this project. Special thanks to Luisa Lübke, Gaoqun Zhang and Barbara Schornick for their contributions during the early phase of this study, and to Fushun Chen for his initial bioinformatics support. Research in Sepand Rastegar's laboratory is supported by the BioInterfaces in Technology and Medicine and Natural, Artificial, and Cognitive Information Processing (NACIP) programmes of the Helmholtz Association, as well as project grants from the German Research Foundation (DFG GRK2039, STR 439/17-1, RA 3469/5-1) and the Marie Sklodowska-Curie Initial Training Network (ITN) ZenCode and Danio-Recode. Jincan Chen is supported by a grant from the China Scholarship Council (CSC), grant number: 202106310017. We would also like to thank the KIT Publication Fund of the Karlsruhe Institute of Technology for its support. Funding for open access was made possible and organised by Projekt DEAL.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

SR designed, supervised the experiments, analysed the data, and co-wrote the manuscript with JC and ND. JC, AH, and TB conducted experiments and analysed data. ND and MT contributed to data analysis. JC and MT analysed and quantified the data. All authors have reviewed and approved the final version of the manuscript for publication.

Data availability statement

The data that supports the findings of this study are available in the main text and the supporting information of this article.

References

Adolf, B., P. Chapouton, C. S. Lam, S. Topp, B. Tannhäuser, U. Strähle, M. Götz, and L. Bally-Cuif. 2006. 'Conserved and acquired features of adult neurogenesis in the zebrafish telencephalon', *Developmental Biology*, 295: 278-93.

Aleström, P., L. D'Angelo, P. J. Midtlyng, D. F. Schorderet, S. Schulte-Merker, F. Sohm, and S. Warner. 2020. 'Zebrafish: Housing and husbandry recommendations', *Laboratory Animals*, 54: 213-24.

Alunni, A., and L. Bally-Cuif. 2016. 'A comparative view of regenerative neurogenesis in vertebrates', *Development*, 143: 741-53.

Ang, N. B., A. Saera-Vila, C. Walsh, P. F. Hitchcock, A. Kahana, R. Thummel, and M. Nagashima. 2020. 'Midkine-a functions as a universal regulator of proliferation during epimorphic regeneration in adult zebrafish', *Plos One*, 15: e0232308.

Aynacioglu, A. S., A. Bilir, and M. Y. Tuna. 2019. 'Involvement of midkine in autoimmune and autoinflammatory diseases', *Modern Rheumatology*, 29: 567-71.

Baranasic, D., M. Hörtenhuber, P. J. Balwierz, T. Zehnder, A. K. Mukarram, C. Nepal, C. Várnai, Y. Hadzhiev, A. Jimenez-Gonzalez, N. Li, J. Wragg, F. M. D'Orazio, D. Relic, M. Pachkov, N. Díaz, B. Hernández-Rodríguez, Z. L. Chen, M. Stoiber, M. Dong, I. Stevens, S. E. Ross, A. Eagle, R. Martin, O. Obasaju, S. Rastegar, A. C. McGarvey, W. Kopp, E. Chambers, D. Wang, H. R. Kim, R. D. Acemel, S. Naranjo, M. Lapinski, V. Chong, S. Mathavan, B. Peers, T. Sauka-Spengler, M. Vingron, P. Carninci, U. Ohler, S. A. Lacadie, S. M. Burgess, C. Winata, F. van Eeden, J. M. Vaquerizas, J. L. Gómez-Skarmeta, D. Onichtchouk, B. Brown, O. Bogdanovic, E. van Nimwegen, M. Westerfield, F. C. Wardle, C. O. Daub, B. Lenhard, and F. Müller. 2022. 'Multiomic atlas with functional stratification and developmental dynamics of zebrafish regulatory elements', *Nature Genetics*, 54: 1037-50.

Becker, T., and C. G. Becker. 2022. 'Regenerative neurogenesis: the integration of developmental physiological and immune signals', *Development*, 149.

Begeman, I. J., B. Emery, A. Kurth, and J. S. Kang. 2022. 'Regeneration and developmental enhancers are differentially compatible with minimal promoters', *Developmental Biology*, 492: 47-58.

Beisaw, A., C. Kuenne, S. Guenther, J. Dallmann, C. C. Wu, M. Bentsen, M. Looso, and D. Y. R. Stainier. 2020. 'AP-1 contributes to chromatin accessibility to promote sarcomere disassembly and cardiomyocyte protrusion during zebrafish heart regeneration', *Circulation Research*, 126: 1760-78.

Berrocal, Augusto, Nicholas C. Lammers, Hernan G. Garcia, and Michael B. Eisen. 2020. 'Kinetic sculpting of the seven stripes of the Drosophila even-skipped gene', *Elife*, 9: e61635.

Betancur, P., M. Bronner-Fraser, and T. Sauka-Spengler. 2010. 'Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest', *Proc Natl Acad Sci U S A*, 107: 3570-5.

Chen, J. C., H. Sanchez-Iranzo, N. Diotel, and S. Rastegar. 2024. 'Comparative insight into the regenerative mechanisms of the adult brain in zebrafish and mouse: highlighting the importance of the immune system and inflammation in successful regeneration', *Febs Journal*, 291: 4193-205.

Chen, Z. L., Y. Omori, S. Koren, T. Shirokiya, T. Kuroda, A. Miyamoto, H. Wada, A. Fujiyama, A. Toyoda, S. Y. Zhang, T. G. Wolfsberg, K. Kawakami, A. M. Phillippy, J. C. Mullikin, S. M. Burgess, and NISC Comparative Sequencing

Prog. 2019. 'De novo assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication', *Science Advances*, 5.

Choe, C. P., S. Y. Choi, Y. Kee, M. J. Kim, S. H. Kim, Y. Lee, H. C. Park, and H. Ro. 2021. 'Transgenic fluorescent zebrafish lines that have revolutionized biomedical research', *Lab Anim Res*, 37: 26.

Demirci, Y., G. Cucun, Y. K. Poyraz, S. Mohammed, G. Heger, I. Papatheodorou, and G. Ozhan. 2020. 'Comparative Transcriptome Analysis of the Regenerating Zebrafish Telencephalon Unravels a Resource With Key Pathways During Two Early Stages and Activation of Wnt/beta-Catenin Signaling at the Early Wound Healing Stage', *Frontiers in Cell and Developmental Biology*, 8: 584604.

Diotel, N., L. Lübke, U. Strähle, and S. Rastegar. 2020. 'Common and distinct features of adult neurogenesis and regeneration in the telencephalon of zebrafish and mammals', *Frontiers in Neuroscience*, 14.

Dong, Yanhan, Yuchen Yang, Haofei Wang, Dong Feng, Elizabeth Nist, Nicholas Yapundich, Brian Spurlock, Madison Craft, Li Qian, and Jiandong Liu. 2024. 'Single-cell chromatin profiling reveals genetic programs activating proregenerative states in nonmyocyte cells', *Science Advances*, 10: eadk4694.

Driever, W., L. SolnicaKrezel, A. F. Schier, S. C. F. Neuhauss, J. Malicki, D. L. Stemple, D. Y. R. Stainier, F. Zwartkruis, S. Abdelilah, Z. Rangini, J. Belak, and C. Boggs. 1996. 'A genetic screen for mutations affecting embryogenesis in zebrafish', *Development*, 123: 37-46.

Ellingsen, S., M. A. Laplante, M. König, H. Kikuta, T. Furmanek, E. A. Hoivik, and T. S. Becker. 2005. 'Large-scale enhancer detection in the zebrafish genome', *Development*, 132: 3799-811.

Gemberling, M., T. J. Bailey, D. R. Hyde, and K. D. Poss. 2013. 'The zebrafish as a model for complex tissue regeneration', *Trends in Genetics*, 29: 611-20.

Ghaddar, B., L. Lübke, D. Couret, S. Rastegar, and N. Diotel. 2021. 'Cellular mechanisms participating in brain repair of adult zebrafish and mammals after injury', *Cells*, 10.

Gonzalez, A. J., D. Baranasic, and F. Müller. 2023. 'Zebrafish regulatory genomic resources for disease modelling and regeneration', *Disease Models & Mechanisms*, 16.

Goto, T., P. Macdonald, and T. Maniatis. 1989. 'Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues', *Cell*, 57: 413-22.

Gourain, V., O. Armant, L. Lubke, N. Diotel, S. Rastegar, and U. Strahle. 2021. 'Multi-Dimensional Transcriptome Analysis Reveals Modulation of Cholesterol Metabolism as Highly Integrated Response to Brain Injury', *Front Neurosci*, 15: 671249.

Gramage, E., T. D'Cruz, S. Taylor, R. Thummel, and P. F. Hitchcock. 2015. 'Midkine-a protein localization in the developing and adult retina of the zebrafish and its function during photoreceptor regeneration', *Plos One*, 10: e0121789.

Gramage, E., J. Li, and P. Hitchcock. 2014. 'The expression and function of midkine in the vertebrate retina', *British Journal of Pharmacology*, 171: 913-23.

Grivas, D., A. Gonzalez-Rajal, and J. L. de la Pompa. 2021. 'Midkine-a regulates the formation of a fibrotic scar during zebrafish heart regeneration', *Frontiers in Cell and Developmental Biology*, 9: 669439.

Hörtenhuber, M., A. K. Mukarram, M. H. Stoiber, J. B. Brown, and C. O. Daub. 2020. '*-DCC: A platform to collect, annotate, and explore a large variety of sequencing experiments', *Gigascience*, 9.

Howe, K., M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J. E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J. C. Barrett, R. Koch, G. J. Rauch, S. White, W. Chow, B. Kilian, L. T. Quintais, J. A. Guerra-Assunçao, Y. Zhou, Y. Gu, J. Yen, J. H. Vogel, T. Eyre, S. Redmond, R. Banerjee, J. X. Chi, B. Y. Fu, E. Langley, S. F. Maguire, G. K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Eliott, G. Threadgold, G. Harden, D. Ware, B. Mortimer, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C. M. Dooley, Z. Ersan-Ürün, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberländer, S. Rudolph-Geiger, M. Teucke, K. Osoegawa, B. L. Zhu, A. Rapp, S. Widaa, C. Langford, F. T. Yang, N. P. Carter, J. Harrow, Z. M. Ning, J. Herrero, S. M. J. Searle, A. Enright, R. Geisler, R. H. A. Plasterk, C. Lee, M. Westerfield, P. J. de Jong, L. I. Zon, J. H. Postlethwait, C. Nüsslein-Volhard, T. J. P. Hubbard, H. Roest Crollius, J. Rogers, and D. L. Stemple. 2013. 'The zebrafish reference genome sequence and its relationship to the human genome', *Nature*, 496: 498-503.

Kikuchi-Horie, K., E. Kawakami, M. Kamata, M. Wada, J. G. Hu, H. Nakagawa, K. Ohara, K. Watabe, and K. Oyanagi. 2004. 'Distinctive expression of midkine in the repair period of rat brain during neurogenesis: immunohistochemical and immunoelectron microscopic observations', *J Neurosci Res*, 75: 678-87.

Kizil, C., J. Kaslin, V. Kroehne, and M. Brand. 2012. 'Adult neurogenesis and brain regeneration in zebrafish', *Developmental Neurobiology*, 72: 429-61.

Lange, C., F. Rost, A. Machate, S. Reinhardt, M. Lesche, A. Weber, V. Kuscha, A. Dahl, S. Rulands, and M. Brand. 2020. 'Single cell sequencing of radial glia progeny reveals the diversity of newborn neurons in the adult zebrafish brain', *Development*, 147.

Lien, C. L., M. Schebesta, S. Makino, G. J. Weber, and M. T. Keating. 2006. 'Gene expression analysis of zebrafish heart regeneration', *Plos Biology*, 4: e260.

Lübke, L., G. Q. Zhang, U. Straehle, and S. Rastegar. 2022. '*mdka* expression is associated with quiescent neural stem cells during constitutive and reactive neurogenesis in the adult zebrafish telencephalon', *Brain Sciences*, 12.

Marques, I. J., E. Lupi, and N. Mercader. 2019. 'Model systems for regeneration: zebrafish', *Development*, 146.

März, M., R. Schmidt, S. Rastegar, and U. Strähle. 2011. 'Regenerative response following stab injury in the adult zebrafish telencephalon', *Developmental Dynamics*, 240: 2221-31.

Meng, A. M., H. Tang, B. A. Ong, M. J. Farrell, and S. Lin. 1997. 'Promoter analysis in living zebrafish embryos identifies a *cis*-acting motif required for neuronal expression of GATA-2', *Proceedings of the National Academy of Sciences of the United States of America*, 94: 6267-72.

Nagashima, M., T. S. D'Cruz, A. E. Danku, D. Hesse, C. Sifuentes, P. A. Raymond, and P. F. Hitchcock. 2020. 'Midkine-a is required for cell cycle progression of Müller glia during neuronal regeneration in the vertebrate retina', *Journal of Neuroscience*, 40: 1232-47.

Nakamura, M., H. Yoshida, E. Takahashi, M. Wlizla, K. Takebayashi-Suzuki, ME. Horb, and A. Suzuki. 2020. 'The AP-1 transcription factor JunB functions in Xenopus tail regeneration by positively regulating cell proliferation', 522: 990-95.

Navratilova, P., D. Fredman, T. A. Hawkins, K. Turner, B. Lenhard, and T. S. Becker. 2009. 'Systematic human/zebrafish comparative identification of *cis*-regulatory activity around vertebrate developmental transcription factor genes', *Developmental Biology*, 327: 526-40.

Parinov, S., I. Kondrichin, V. Korzh, and A. Emelyanov. 2004. 'Tol2 transposon-mediated enhancer trap to identify developmentally regulated zebrafish genes in vivo', *Dev Dyn*, 231: 449-59.

Rastegar, S., I. Hess, T. Dickmeis, J. C. Nicod, R. Ertzer, Y. Hadzhiev, W. G. Thies, G. Scherer, and U. Strähle. 2008. 'The words of the regulatory code are arranged in a variable manner in highly conserved enhancers', *Developmental Biology*, 318: 366-77.

Rodriguez Viales, R., N. Diotel, M. Ferg, O. Armant, J. Eich, A. Alunni, M. März, L. Bally-Cuif, S. Rastegar, and U. Strahle. 2015. 'The helix-loop-helix protein Id1 controls stem cell proliferation during regenerative neurogenesis in the adult zebrafish telencephalon', *Stem Cells*, 33: 892-903.

Schmidt, R., T. Beil, U. Strähle, and S. Rastegar. 2014. 'Stab wound injury of the zebrafish adult telencephalon: A method to investigate vertebrate brain neurogenesis and regeneration', *Jove-Journal of Visualized Experiments*.

Tamaki, T., T. Yoshida, E. Shibata, H. Nishihara, H. Ochi, and A. Kawakami. 2023. 'Splashed E-box and AP-1 motifs cooperatively drive regeneration response and shape regeneration abilities', *Biol Open*, 12.

Thompson, J. D., J. H. Ou, N. Lee, K. Shin, V. Cigliola, L. Y. Song, G. E. Crawford, J. S. Kang, and K. D. Poss. 2020. 'Identification and requirements of enhancers that direct gene expression during zebrafish fin regeneration', *Development*, 147.

Wang, W., CK. Hu, A. Zeng, D. Alegre, D. Hu, K. Gotting, A. Ortega Granillo, Y. Wang, S. Robb, R. Schnittker, S. Zhang, D. Alegre, H. Li, E. Ross, N. Zhang, A. Brunet, and A. Sánchez Alvarado. 2020. 'Changes in regeneration-responsive enhancers shape regenerative capacities in vertebrates', *Science*, 369: eaaz3090.

Weinberger, M., F. C. Simoes, T. Gungoosingh, T. Sauka-Spengler, and P. R. Riley. 2024. 'Distinct epicardial gene regulatory programs drive development and regeneration of the zebrafish heart', *Developmental Cell*, 59: 351-67.

Werner, T., A. Hammer, M. Wahlbuhl, M. R. Bösl, and M. Wegner. 2007. 'Multiple conserved regulatory elements with overlapping functions determine Sox10 expression in mouse embryogenesis', *Nucleic Acids Research*, 35: 6526-38.

Winkler, C., M. Schäfer, J. Duschl, M. Schartl, and J. N. Volff. 2003. 'Functional divergence of two zebrafish midkine growth factors following fish-specific gene duplication', *Genome Research*, 13: 1067-81.

Winkler, C., and S. Yao. 2014. 'The midkine family of growth factors: diverse roles in nervous system formation and maintenance', *Br J Pharmacol*, 171: 905-12.

Yang, H., Y. Luan, T. Liu, H. J. Lee, L. Fang, Y. Wang, X. Wang, B. Zhang, Q. Jin, K. C. Ang, X. Xing, J. Wang, J. Xu, F. Song, I. Sriranga, C. Khunsriraksakul, T. Salameh, D. Li, M. N. K. Choudhary, J. Topczewski, K. Wang, G. S. Gerhard, R. C. Hardison, T. Wang, K. C. Cheng, and F. Yue. 2020. 'A map of cis-regulatory elements and 3D genome structures in zebrafish', *Nature*, 588: 337-43.

Zhang, G. Q., M. Ferg, L. Lübke, M. Takamiya, T. Beil, V. Gourain, N. Diotel, U. Strähle, and S. Rastegar. 2020. 'Bone morphogenetic protein signaling regulates Id1-mediated neural stem cell quiescence in the adult zebrafish brain via a phylogenetically conserved enhancer module', *Stem Cells*, 38: 875-89.

Zhang, G. Q., L. Lübke, F. S. Chen, T. Beil, M. Takamiya, N. Diotel, U. Strähle, and S. Rastegar. 2021. 'Neuron-radial glial cell communication via BMP/Id1 signaling is key to long-term maintenance of the regenerative capacity of the adult zebrafish telencephalon', *Cells*, 10.

Table 1. Summary of EGFP expression patterns driven by individual CREs in 24 hpf embryos.

	Endogenous	CRE1	CRE2	CRE3	CRE4	CRE5	CRE6
Brain	++	-	+	+	+	+	+
Retina	++	-	+	++	++	+	++
Spinal cord	++	-	+	++	+	+	+
Notochord	+	-	+	-	-	-	-
Fin fold/Tail	++	-	-	++	+	-	+
Somite/Muscle	++	-	+	-	++	-	++

Note: -, no detectable expression; +, weak expression; ++, strong expression.

Table 2. Summary of EGFP expression patterns driven by individual CREs in the adult telencephalon.

	Endogenous	CRE1	CRE2	CRE3	CRE4	CRE5	CRE6
Dm	++	-	++	-	+	-	+
Dl	++	-	+	+	+	-	+
Dp	+	-	-	++	-	-	-
Vd	+	-	+	-	-	-	-
Vv	+	-	+	-	++	-	-

Note: -, no detectable expression; +, weak expression; ++, strong expression. Abbreviations: Dm, medial pallium; Dp, posterior pallium; Dl, lateral pallium; VZ, ventricular zone; Vd, dorsal nucleus of ventral telencephalon; Vv, ventral nucleus of ventral telencephalon.

Table 3. Summary of EGFP expression patterns observed in distinct *mdka* CRE transgenic lines.

CRE(s)	Activity summary
CRE2	Adult regulatory sequence; active in physiological conditions; closest to *mdka* expression
CRE3	Mostly active in the embryo; overlaps with *mdka* expression; ectopic expression in the adult telencephalon in Dp and Dl
CRE4	Adult regulatory sequence; active in NSCs; no significant response to induced injury
CRE6	Adult regulatory sequence; active in NSCs with ectopic activity in neurons; no significant response to induced injury
CRE2346	Fully recapitulates *mdka* expression; EGFP reporter overlaps with *mdka* mRNA in time and space; injury inducible

Note: Dp, posterior pallium; Dl, lateral pallium. NSCs, neural stem cells.

Figures legends

Figure 1. Identification of potential *mdka* regulatory sequences based on embryonic developmental ATAC-seq data. Snapshot of UCSC genome browser tracks displaying ATAC-seq data from different embryonic developmental stages (Dome, 75%-epiboly, 5-9 somites, Prime-5 and Long-pec) over a representative genomic region of 45 kilobase pairs (chr7: 38,875,000-38,920,000). Color-coded horizontal bars indicate the genomic position of putative promoters (upstream/downstream from transcription start site [TSS], including 1_TssA and 2_TssA2 [blue]), cis-regulatory elements (CREs, such as 5_EnhA1 [red] and 7_EnhWk1 [orange]), and other elements (e.g., 10_Quies [grey]). These annotations are based on PADRES (Predicted ATAC-seq-supported developmental regulatory elements) and reflect ATAC-seq signal intensity, with distal ATAC-seq peaks indicated by green profiles The grey box highlights six putative *mdka* CREs (CRE1 to CRE6) identified in this study, while the magenta box marks the putative *mdka* promoter region. The relative positions of all CREs are shown in reference to the *mdka* gene, with its TSS indicated.

Figure 2. Stable EGFP transgenic line generated using potential *mdka* CREs. A, Schematic of the plasmid pTol2:CREx:gata2aPR:egfp:pA, where CREx represents one of the six putative CREs (CRE1 to CRE6). **B-G**, Stable EGFP reporter expression driven by *mdka* CRE1 (**B**), CRE2 (**C**), CRE3 (**D**), CRE4 (**E**), CRE5 (**F**), CRE6 (**G**) in embryo at 24 hours post-fertilization (hpf). **B'-G'**, Merged images showing *mdka* mRNA (magenta) and EGFP (green) expression in the corresponding CRE transgenic lines. *Tg(CRE3:gata2aPR:EGFP)* embryo exhibited strong co-localization of *mdka* mRNA and EGFP signals in the spinal cord (yellow arrowheads, **D**, **D'**) and fin fold region (white arrowheads, **D**, **D'**). *Tg(CRE4:gata2aPR:EGFP)* and *Tg(CRE6:gata2aPR:EGFP)* embryos showed strong activity in the retina (green arrowheads) and brain regions (red arrowheads, **E, E', G, G'**), respectively. *Tg(CRE1:gata2aPR:EGFP)* showed no detectable EGFP expression (*B, B'*). Scale bar: 200 μm.

Figure 3. Analysis of EGFP expression in the adult telencephalon of *mdka* transgenic lines during constitutive neurogenesis. A-F', EGFP reporter expression in the adult telencephalon of different *mdka* CRE transgenic lines: CRE1 (**A, A'**), CRE2 (**B, B'**), CRE3 (**C, C'**), CRE4 (**D, D'**), CRE5 (**E, E'**), and CRE6 (**F, F'**). Yellow arrowheads indicate EGFP expression in the dorsal pallium (Dp) and lateral pallium (Dl) (**C**), while red arrowheads indicate ectopic expression in the medial pallium (Dm) near but not in the ventricular zone (VZ) (**F, F'**). Rectangles in **A-F** indicate magnified regions shown in **A'-F'**, respectively. Dashed lines outline the telencephalon boundaries (**A**, **E**). **G-G'', I-I'', K-K''**, Magnified transverse views of the Dm VZ showing *mdka* mRNA and EGFP expression in the adult telencephalon of transgenic lines *Tg(CRE2:gata2aPR:EGFP)*(**G-G''**), *Tg(CRE4:gata2aPR:EGFP)*(**I-I''**), and *Tg(CRE6:gata2aPR:EGFP)* (**K-K''**). White arrowheads indicate co-localized regions of *mdka* mRNA and EGFP signals, confirming CRE activity. **H, J, L** Quantification of S100β+/PCNA− type 1 and S100β+/PCNA+ type 2 radial glial cells (RGCs) in CRE2 (**H**), CRE4 (**J**), and CRE6 (**L**) transgenic lines. Each circle represents data from an individual brain; floating bars represent the mean ± S.E.M. Statistical significance was determined using unpaired t-tests. Asterisks denote p-values: ** $p < 0.01$; *** $p < 0.001$; ns, not significant. Scale bar: 200 µm (A-F), 25 µm (**A'-F'**, G-G'', **I-I''**, and **K-K''**). Sample sizes: $n = 3$ brains (**A, A'**), $n = 4$ (**B, B', G-G''**), $n = 6$ (**C, C'**), $n = 3$ (**D, D'**, **I-I''**), $n = 4$ (**E, E'**), $n = 3$ (**F, F', K-K''**), $n = 6$ (**H**), $n = 3$ (**J**), $n = 3$ (**L**).

Figure 4. EGFP expression in the adult telencephalon of *mdka* transgenic lines during regenerative neurogenesis. A-F', The *mdka* CREs drive EGFP expression in response to stab injury at 5 days post-lesion (dpl) in the adult telencephalon. Stable EGFP reporter expression is shown for CRE1 (**A, A'**), CRE2 (**B, B'**), CRE3 (**C, C'**), CRE4 (**D, D'**), CRE5 (**E, E'**), and CRE6 (**F, F'**). Rectangles in **A-F** indicate magnified regions in **A'-F'**, respectively. Dashed lines mark the telencephalon boundaries (**A, E**). White asterisks indicate injury sites. **G-G'', J-J'', M-M''**, Magnified transverse sections of the injured adult telencephalon (5 dpl) showing co-localization of *mdka* mRNA and EGFP in transgenic lines *Tg(CRE2:gata2aPR:EGFP)* (**G-G''**), *Tg(CRE4:gata2aPR:EGFP)* (**J-J''**), and *Tg(CRE6:gata2aPR:EGFP)* (**M-M''**), validating CRE activity. White arrowheads indicate regions of overlapping *mdka* mRNA and EGFP signals. White asterisks mark injury sites. **H, K, N,** Quantification of EGFP-positive cells in the uninjured (control) and injured (lesion) hemispheres in

CRE2 (**H**), CRE4 (**K**), and CRE6 (**N**) transgenic lines. **I, L, O,** Relative proportions of type 1 (S100β$^+$/PCNA$^-$) and type 2 (S100β$^+$/PCNA$^+$) radial glial cells (RGCs) in CRE2 (**I**), CRE4 (**L**), and CRE6 (**O**) transgenics, in both control and lesion hemispheres. Each circle represents an individual brain; floating bars show the mean ± S.E.M. Statistical analysis was performed using unpaired t-tests for two-group comparisons (**H, K, N**) and one-way ANOVA for multiple-group comparisons (**I, L, O**). Asterisks denote p-values: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, not significant. Scale bars: 200 μm (**A–F**); 25 μm (**A'–F', G–G'', J–J'', M–M''**).Sample sizes: $n = 3$ brains (**A, A'**), $n = 4$ (**B, B', G–G''**), $n = 5$ (**C, C'**), $n = 3$ (**D, D', J–J''**), $n = 4$ (**E, E'**), $n = 3$ (**F, F', M–M''**), $n = 3$ (**H, I, K, L, N, O**).

Figure 5. EGFP expression pattern driven by CRE2346 in 24 hpf embryo and the adult telencephalon. A, Schematic of the pTol2:mdkaCRE2346: gata2aPR:egfp:pA construct in which a combination of CRE2, CRE3, CRE4, and CRE6 drives EGFP expression under the control of the gata2a promoter region (gata2aPR). **B–B'',** Colocalization of EGFP (**B**) and *mdka* mRNA (**B'**) in 24 hpf embryos; merged in (**B''**). **C,** Merged image showing EGFP and *mdka* mRNA expression in a transverse section of the adult telencephalon. **D-D'',** Magnified views of the section shown in (**C**): EGFP (**D**), *mdka* mRNA (**D'**), and the merged image (**D''**). **E-E''', F-F'''**, Transverse telencephalon sections showing EGFP expression (**E'**, magnified in **F'**), PCNA (**E''**, magnified in **F''**), **S100β (E''',** magnified in **F''').** merged in (**E,** magnified in **F). G,** Relative proportions of type 1 (S100β$^+$/PCNA$^-$) and type 2 (S100β$^+$/PCNA$^+$) radial glial cells (RGCs) in the uninjured *CRE2346* transgenic brain. **H,** Quantification of EGFP-positive cells in the left and right hemispheres of the uninjured *CRE2346* transgenic brain. Each data point represents an individual brain (plotted as circles), and floating bars indicate the mean ± S.E.M. Statistical significance was assessed using an unpaired t-test for two-group comparisons. Asterisks denote p-values: ** $p < 0.001$; ns, not significant. Scale bars: 200 μm (**B–B'', C, E–E'''**); 25 μm (**D–D'', F–F'''**). Sample sizes: $n = 3$ brains (**C, D–D'', G, H**); $n = 5$ brains (**E–E''', F–F'''**).

Figure 6. EGFP expression driven by CRE2346 after brain injury. A, B-B'', EGFP expression driven by a combination of *mdka* CREs (CRE2, CRE3, CRE4, and CRE6) in the adult telencephalon. Magnified views are shown in **B** (merged), **B'** (EGFP), and **B''** (*mdka* mRNA). White asterisks indicate injury sites. **C-C'**, Magnified views of the

upregulated regions indicated in **A**. **D-D’’’, E-E’’’**, Transverse sections of the injured telencephalon at 5 days post-lesion (dpl) showing EGFP (**D’**, magnified in **E’**), PCNA (**D’’**, magnified in **E’’**; proliferation marker), and S100β (**D’’’**, magnified in **E’’’**; glial marker). Merged images are shown in **D** and **E**, respectively. White asterisks indicate injury sites. **F,** Quantification shows significantly more EGFP-positive cells in the injured hemisphere (27.20 ± 2.15 cells) compared to the uninjured hemisphere (20.99 ± 2.25 cells; ***p < 0.001***). **G,** The EGFP intensity ratio between hemispheres (left/right in uninjured brains and injured/uninjured in injured brains) is significantly higher in injured brains (1.41 ± 0.03) than in uninjured brains (1.17 ± 0.04; **p < 0.01**). **H**, Relative proportions of type 1 ($S100\beta^{+}/PCNA^{-}$) and type 2 ($S100\beta^{+}/PCNA^{+}$) radial glial cells (RGCs) in the uninjured versus injured hemispheres. Each data point represents an individual brain (plotted as circles), and floating bars indicate the mean ± S.E.M. Statistical significance was assessed using an unpaired t-test for two-group comparisons (**F, G**) and one-way ANOVA for multiple-group comparisons (**H**). Asterisks denote p-values (**F, G**) and adjusted p-values (**H**): *p < 0.05; **p < 0.01; ***p < 0.001*. Scale bars: 200 μm (**A, D–D’’’**); 25 μm (**B–B’’, C–C’, E–E’’’**). Sample sizes: *n* = 4 brains (**A, B–B’’, C–C’**); *n* = 6 brains (**D–D’’’, E–E’’’, F, H**); *n* = 3 brains (**G**).

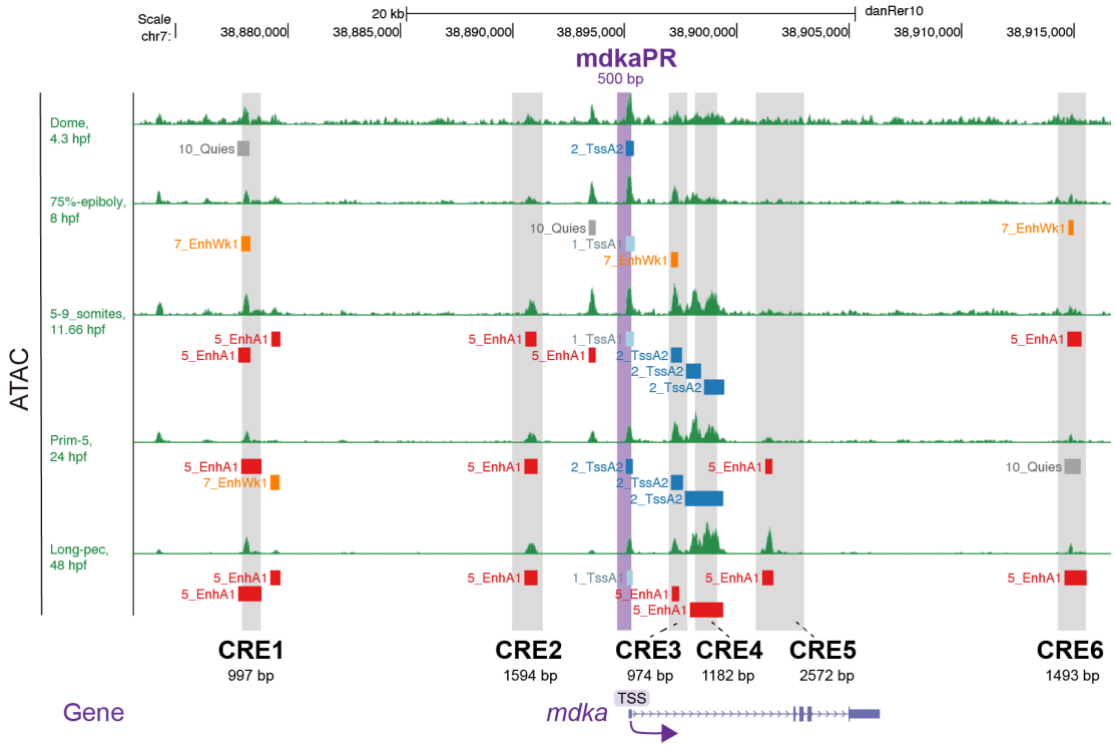

Figure 1. Identification of potential *mdka* regulatory sequences based on embryonic developmental ATAC-seq data.

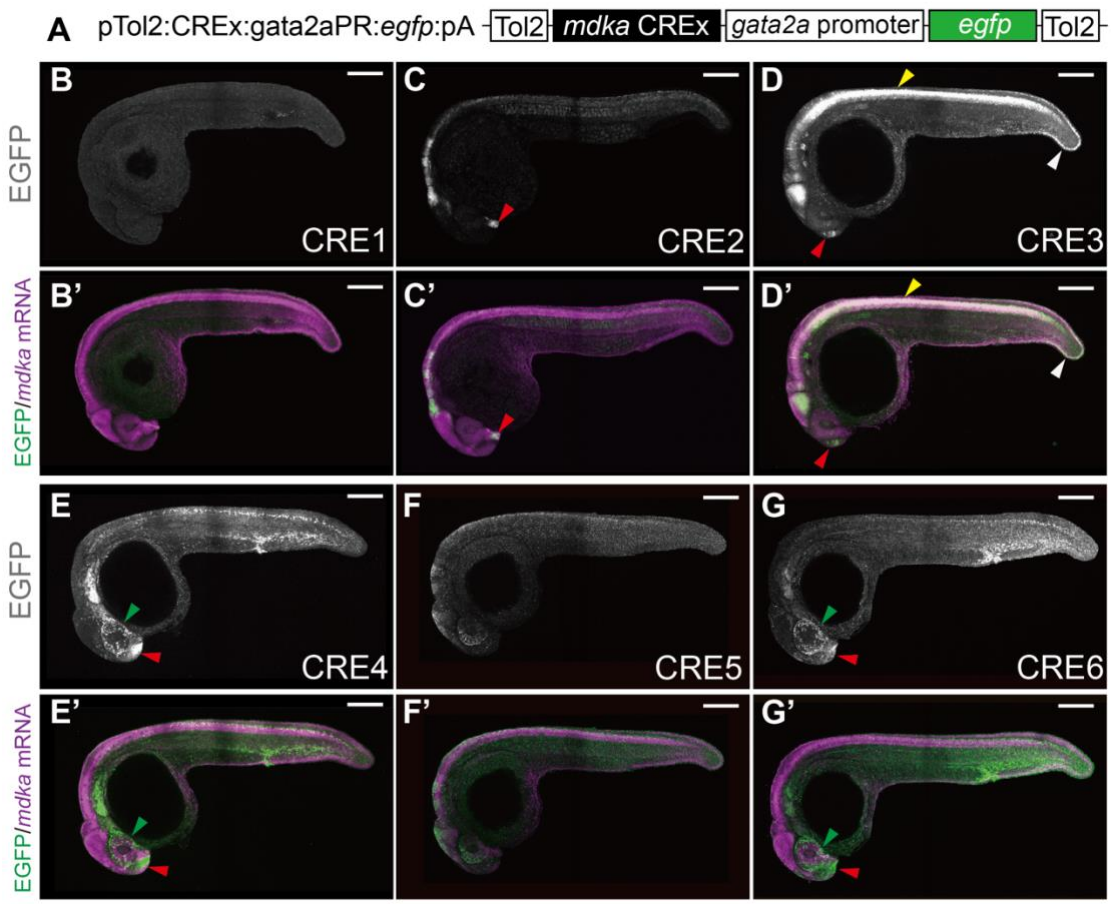

Figure 2. Stable EGFP transgenic line generated using potential *mdka* CREs.

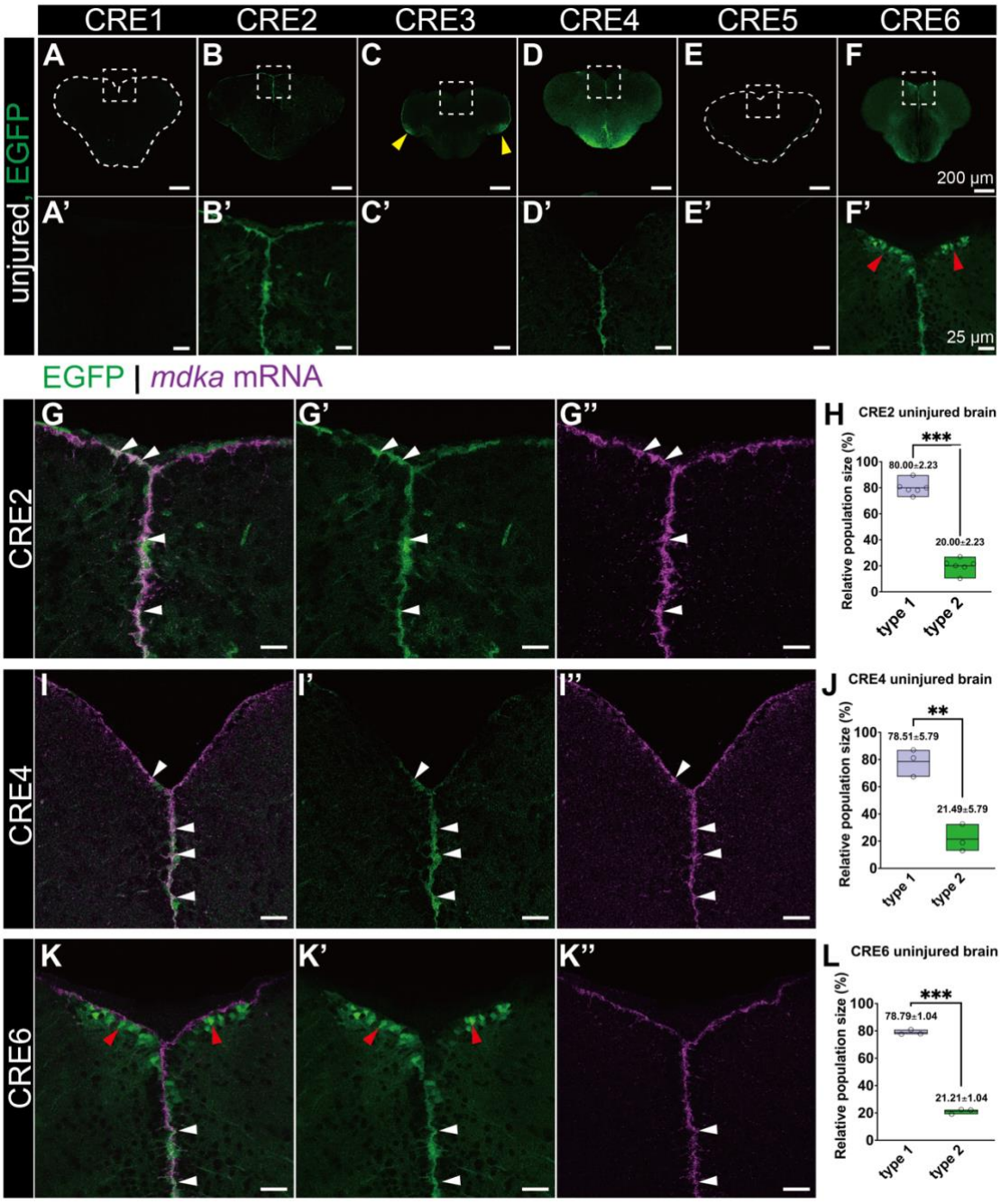

Figure 3. Analysis of EGFP expression in the adult telencephalon of *mdka* transgenic lines during constitutive neurogenesis.

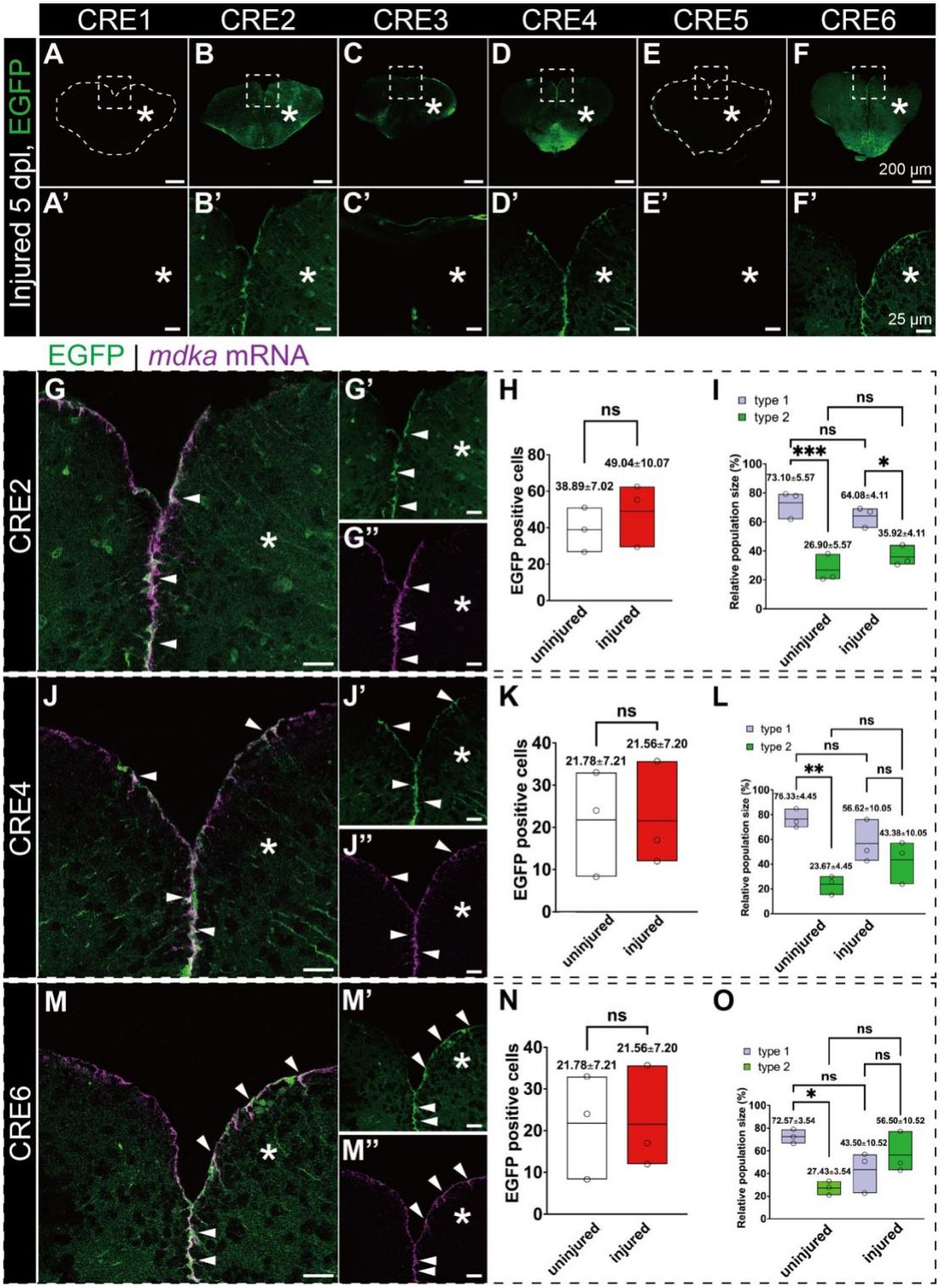

Figure 4. EGFP expression in the adult telencephalon of *mdka* transgenic lines during regenerative neurogenesis.

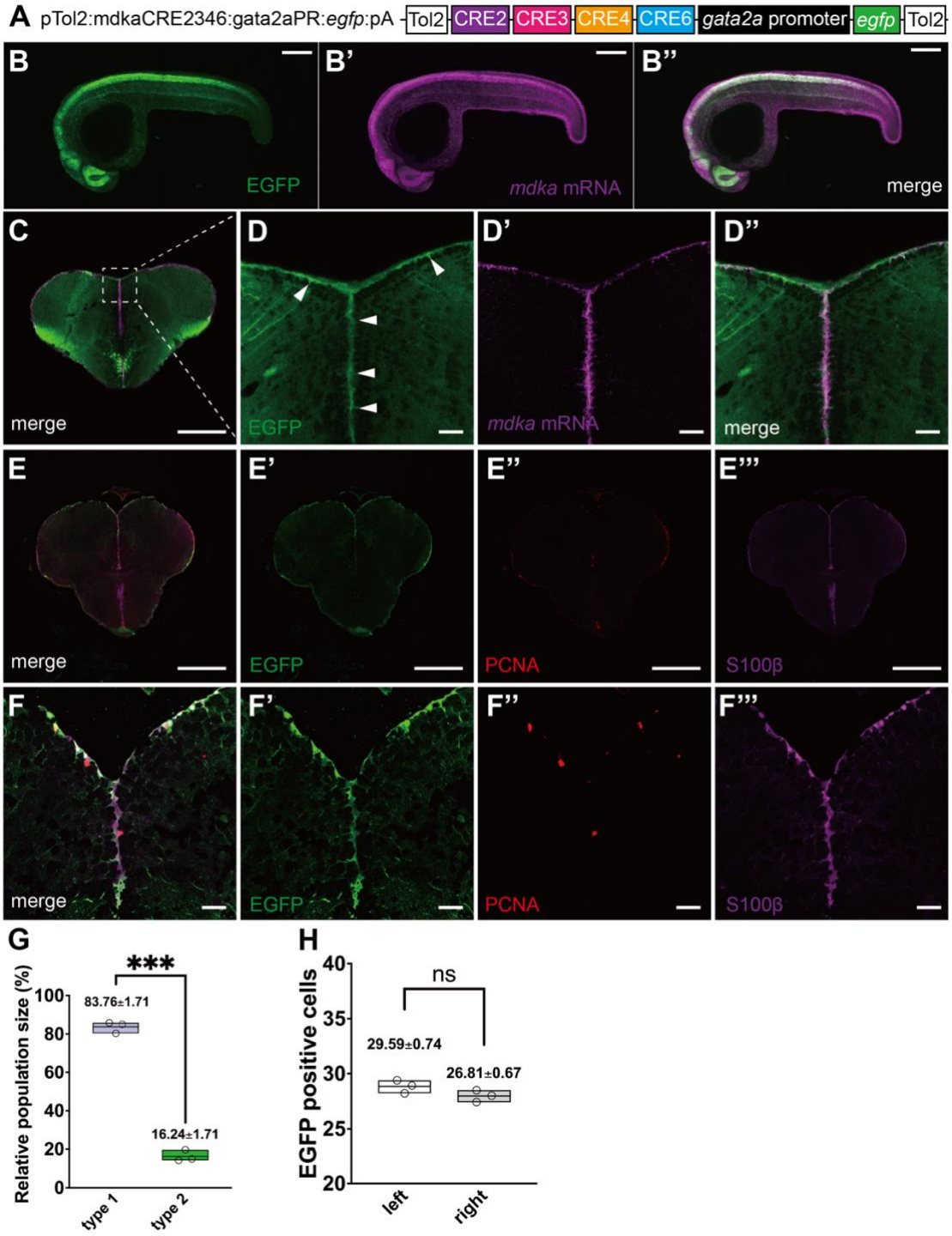

Figure 5. EGFP expression pattern driven by CRE2346 in 24 hpf embryo and the adult telencephalon.

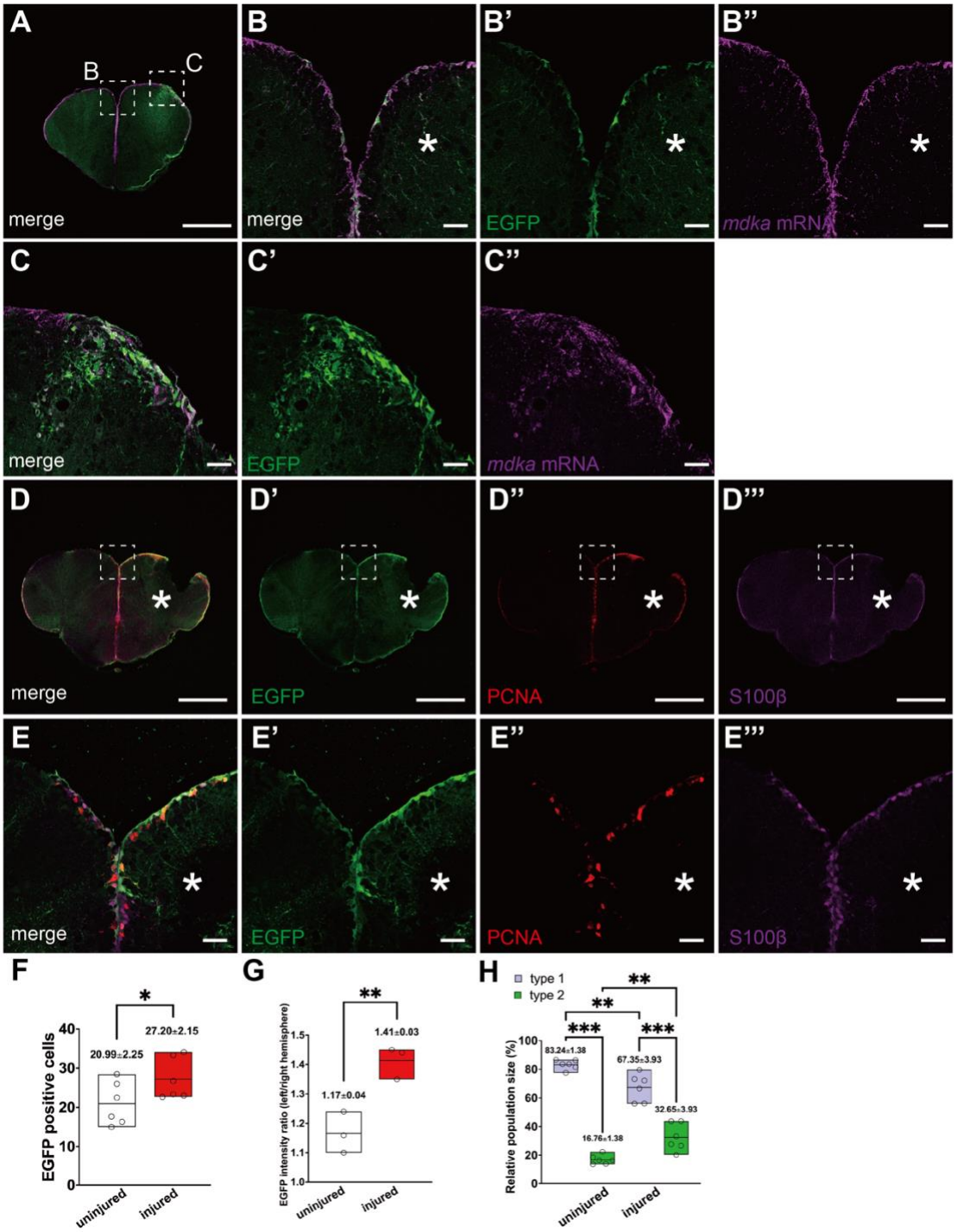

Figure 6. EGFP expression driven by CRE2346 after brain injury.