SoxB1-Mediated Chromatin Remodeling Promotes Sensory Neuron Differentiation in Planarians

Mallory L. Cathell[1], Mohamad A. Auwal[1], Sarai Alvarez Zepeda[1], Kelly G. Ross[1], and Ricardo M. Zayas[1]*

[1]Department of Biology, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182-4614, USA

***Correspondence:**
Dr. Ricardo M. Zayas
E-mail: rzayas@sdsu.edu

Keywords: ATAC-seq, epigenetics, *mecom*, *castor,* stem cells

Abstract

Understanding how adult stem cells generate neurons is critical for advancing regenerative medicine. However, few in vivo models enable studying how stem cell fates are specified as neurons in an adult body. The planarian *Schmidtea mediterranea* provides a powerful system for investigating these mechanisms, owing to its abundant adult pluripotent stem cells, termed neoblasts, and its capacity to regenerate a molecularly complex nervous system. The SoxB1 family of transcription factors is broadly implicated in ectodermal lineage commitment. In planarians, the SoxB1 homolog *soxB1-2* has been shown to promote neural and epidermal differentiation. However, the mechanisms by which *soxB1-2* influences chromatin dynamics and transcriptional programs during adult neurogenesis remain unknown. To address this, we performed ATAC-seq and RNA-seq on neural-rich head tissues to assess how *soxB1-2* RNAi knockdown alters chromatin accessibility and gene expression. Disrupting *soxB1-2* resulted in reduced chromatin accessibility and transcriptional downregulation at neural and epidermal loci, consistent with a pioneer-like role in chromatin priming. We identified 31 candidate downstream targets with concordant accessibility and expression changes, including the transcription factors *castor* and *mecom*, which regulate mechanosensory and ion transport genes. Head tissue sampling enabled the detection of *soxB1-2*-responsive genes within rare neural subtypes that were missed in our previous whole-body RNA-seq experiments. These findings offer mechanistic insight into adult ectodermal lineage specification and establish a framework for understanding chromatin-mediated neurogenesis in regenerative systems.

Introduction

The nervous system is necessary for most of the physiological and behavioral activities of metazoans. However, many adult organisms lack the ability to regenerate lost neurons *de novo*. Understanding the molecular programming of stem cells during adult neurogenesis may inform strategies to induce neural fates and restore damaged or lost neurons. One highly conserved class of transcription factors that has roles in neural programming during development includes the SoxB1 group proteins. Diversification of the SoxB family into SoxB1 and SoxB2 subfamilies occurred before diversification of vertebrates, and the ancestral bilaterian had both a SoxB1 and SoxB2 gene (Zhong et al., 2011). Thus, members of the SoxB1 family are conserved across bilaterians, including SoxB1 family members Sox1, 2, and 3 in mammals, of which Sox2 is critical for maintaining stem cell potency and directing the specification of multiple neural progenitor lineages (Zhang and Cui, 2014). Beyond development, SoxB1 genes are expressed in adult tissues and contribute to regenerative processes across diverse organisms (Kamachi and Kondoh, 2013; Reiprich and Wegner, 2015; She and Yang, 2015).

SoxB1 homologs regulate neurogenesis across various neural lineages by modulating transcription factor networks that drive terminal fate decisions (Miyagi et al., 2009). They are thought to function as pioneer transcription factors, capable of binding to closed chromatin and altering accessibility, thus facilitating the binding of downstream factors and initiating broad regulatory cascades (Stevanovic et al., 2021). Despite their well-characterized roles in development, the genomic activity and function of SoxB1 genes influencing cell fate specification and differentiation during adult neural regeneration remain poorly understood.

Integrating epigenomic and transcriptomic approaches offers a path to dissect how SoxB1 genes orchestrate adult stem cell-driven neurogenesis. This approach enables identification of

SoxB1 regions targeted for chromatin alteration and links the changes in chromatin dynamics to gene regulation, offering mechanistic insight into SoxB1-driven cell fate specification. One such strategy involves comparing chromatin accessibility and gene expression profiles in systems where SoxB1 activity is perturbed to controls. Characterization of these target loci can illuminate how chromatin-level regulation underlies SoxB1-mediated neurogenesis.

The planarian *Schmidtea mediterranea* presents an ideal model for investigating neural regeneration. *S. mediterranea* harbors a large population of pluripotent stem cells, enabling complete tissue regeneration, including a molecularly complex nervous system (Cebrià, 2007; Rink, 2013, 2018; Ross et al., 2017). Previous work has shown that inhibition of the planarian SoxB1 gene, *soxB1-2*, results in severe phenotypes, including seizure-like behaviors and loss of mechanosensation (Ross et al., 2018). Co-expression analyses, conducted via in situ hybridization with cell-type-specific markers, along with pseudotime analysis of single-cell RNA sequencing data, suggested that *soxB1-2*$^{+}$ stem cells constitute ectodermal progenitors destined for epidermal or neuronal fates (Ross et al., 2018).

Functional screens from our lab have identified *soxB1-2* downstream targets, including the BRN3/POU4 homolog *pou4-2*, which plays a role in specifying a subset of mechanosensory neurons (McCubbin et al., 2025). Knockdown of *soxB1-2* leads to depletion of the *pou4-2*$^{+}$ mechanosensory population, suggesting a positive regulatory relationship (McCubbin et al., 2025). Additionally, RNAi of *pou4-2* results in loss of the sensory neuron terminal gene *pkd1L-2*, a phenotype also observed following *soxB1-2* disruption (McCubbin et al., 2025; Ross et al., 2018). However, the mechanisms by which *soxB1-2* regulates sensory neural fate trajectories remain unclear. One hypothesis is that *soxB1-2* acts as a pioneer transcription factor in planarians, binding to silent chromatin and modifying the chromatin landscape at target loci to

permit binding of additional transcription factors that direct the differentiation of terminal sensory and epidermal lineages.

Despite the expanding use of transcriptomic tools in planarians, genome-scale chromatin accessibility datasets remain scarce. Applying ATAC-seq in this system not only enables mechanistic investigation of *soxB1-2* function but also broadens the epigenomic resources available for this emerging regenerative model. By generating high-quality chromatin accessibility profiles from control and *soxB1-2(RNAi)* animals, our study establishes a foundational dataset for investigating adult neural fate specification and offers a framework for integrating chromatin and transcriptional regulation in planarian stem cell biology. To maximize accessibility, all datasets will be publicly available and formatted for easy browsing by the planarian research community, creating a resource to support hypothesis generation in regeneration and cell fate regulation studies.

In this study, we tested the hypothesis that *soxB1-2* primes chromatin for the acquisition of sensory and epidermal progenitor fates through the integration of epigenetic and transcriptomic approaches focused on isolated planarian head tissues to enrich for neuronal cell populations. We assessed changes in chromatin accessibility (via ATAC-seq) and gene expression (via RNA-seq) following *soxB1-2* RNAi knockdown to identify direct and functional targets. Our findings revealed that loss of *soxB1-2* leads to reduced chromatin accessibility at genomic regions associated with sensory neuronal and epidermal genes, some of which encode transcription factors. Among these, we identified *mecom* and *castor* transcription factor genes, which we found are involved in the regulation of mechanosensory signaling genes. Together, these results provide mechanistic evidence that *soxB1-2* functions as a pioneer factor in priming chromatin for the emergence of sensory neuronal and epidermal progenitor fates.

Results

Neural genes exhibit reduced chromatin accessibility following *soxB1-2* RNAi

To investigate how *soxB1-2* activity influences chromatin accessibility, we performed ATAC-seq (Buenrostro et al., 2015) on head tissues (amputated in the "neck" region anterior to the pharynx and posterior to the auricles; see Supplemental Figure 1) from *soxB1-2(RNAi*) and *gfp*(*RNAi*) (control) animals to enrich for neural cell populations. The optimal RNAi feeding regimen and timepoint for obtaining *soxB1-2* knockdown samples were determined based on prior studies conducted in our laboratory (Ross et al., 2018). Sequencing data were aligned to the *S. mediterranea* genome assembly (S3h1) (Ivankovic et al., 2024), and peaks were mapped to the nearest annotated gene model (see Methods). Peak calling identified 27,560 reproducible accessible chromatin regions in control animals and 26,679 regions in *soxB1-2*(RNAi) animals with 66% and 64% overlap between biological replicates, respectively ($p \leq 0.01$, minimum peak length of 150 bp; see fragment size distributions from each ATAC-seq replicate in Supplemental Figure 2A). Replicate accessibility profiles were highly correlated (Pearson $r = 0.96$ for both conditions; Supplemental Figure 2B-C and Supplemental Table 1), confirming strong reproducibility across samples.

Subsequent differential accessibility analysis using csaw (see Methods) identified 264 significantly differentially accessible (DA) peaks. The majority of these (257 peaks) were present in controls but significantly reduced in *soxB1-2(RNAi)* animals (down). In comparison, only seven peaks were significantly increased in *soxB1-2(RNAi)* animals relative to controls (up) (Supplemental Table 1). This overall reduction in accessible peaks in the *soxB1-2(RNAi)* animals was also observed as a cumulative decrease in signal across all gene bodies and promoter regions

in the *soxB1-2* RNAi groups (Figure 1A-D). Thirteen genes associated with transcription factor binding were identified within DA chromatin regions, such as *castor* (Figure 1D-E), which has been linked to neurodevelopmental processes in both invertebrate and vertebrate models (Isshiki et al., 2001). Most of the peaks that were shared between the *gfp* and *soxB1-2* RNAi groups were located in distal intergenic regions (54.7% and 57.54%, respectively; Supplemental Figure 3). Promoter regions accounted for 22% of peaks in controls and 21.25% in *soxB1-2* RNAi samples, while internal intronic regions contained 17.53% and 16.23%, respectively. Only a small fraction mapped to internal exonic regions, 3% in controls and 2.56% in *soxB1-2* RNAi. In contrast, DA peaks showed a markedly different distribution: only 25.52% were in distal intergenic regions, while 36.39% mapped to internal intronic regions and 9.96% to internal exonic regions (Supplemental Figure 3).

Distinguishing direct chromatin-opening activity of SoxB1-2 from indirect effects mediated through downstream regulators is an important consideration. To evaluate this, we performed motif enrichment analysis following Neiro et al. (2022) using differentially accessible regions between control and *soxB1-2(RNAi)* animals. Sox-family binding motifs, including SoxB1-like sites, were significantly enriched in regions that remain accessible in control animals relative to *soxB1-2(RNAi)* (Supplemental Figure 4). This supports a model in which SoxB1-2 contributes to establishing or maintaining chromatin accessibility at these loci, consistent with pioneer-like activity as described for other SoxB1 factors (Neiro et al., 2022).

To validate ATAC-seq results and determine whether genes that lose chromatin accessibility are downstream targets of *soxB1-2*, we selected two genes, *dd_26266* and *castor,* for i*n situ* hybridization following *soxB1-2* RNAi. These genes are expressed in the same neural or ciliated neural clusters as *soxB1-2* (Fincher et al., 2018). *In situ* analysis revealed reduced

expression of *dd_26266* and *castor* in the head margin (Figure 1E), a region enriched in sensory neurons (Okamoto et al., 2005; Reho et al., 2022; Wang et al., 2016). Additionally, *castor* exhibited further expression loss in the dorsal epidermis and dorsal ciliated stripe (an organ associated with water flow sensing in planarians) (Farnesi and Tei, 1980; Hyman, 1951; Ross et al., 2018) (Figure 1E). The loss of chromatin accessibility and gene expression in sensory neuron-rich regions supports a key role of *soxB1-2* in regulating downstream genes with potential roles in neuronal specification or terminal differentiation, including other transcription factors.

soxB1-2 RNAi results in loss of sensory and neural gene expression

To examine the relationship between gene regulation and chromatin accessibility in the *soxB1-2* RNAi amputation scheme, we repeated *soxB1-2* RNAi experiments and performed RNA-seq on RNA extracted from animal heads (Supplemental Figure 1). RNA-seq analysis identified 533 downregulated genes meeting our significance cut-offs (fold change ≥ 1.4, *p-adj* ≤ 0.1) (Figure 2A, Supplemental Table 2). Among these, we observed a significant reduction in genes previously shown to be downstream targets of *soxB1-2* that are involved in neural and mechanosensory function. These included *pou4-2, pkd1L-1, pkd1L-2, pkd1L-3, pkd2L-1, pkd2L-2, pkd2-1,* and *pkd2-4* (Figure 2A) (McCubbin et al., 2025; Ross et al., 2024).

To validate the differential expression results, we isolated all genes that were significantly differentially expressed between *gfp* controls and *soxB1-2* RNAi-treated animals and generated a heatmap (Supplemental Figure 5). This visualization confirmed consistent expression patterns across replicates, with upregulated genes in the control group showing corresponding downregulation in the RNAi group, and vice versa, indicating robust sample

clustering for this gene set. To further investigate these expression changes and validate patterns observed in our dataset, we selected *no mechanoreceptor protein channel c* (*nompc*)*, dd_17839, formimidoyltransferase cyclodeaminase* (*ftcd*)*, ms310138,* and *receptor-type tyrosine-protein phosphatase S* (*ptprs*) as representative genes downregulated in the *soxB1-2*(*RNAi*) group. These genes were co-expressed with *soxB1-2* specifically within neural and ciliated neural populations in published single-cell RNA-seq datasets (Wurtzel et al., 2015) (Supplemental Figure 6). To assess whether they were potential downstream targets of *soxB1-2*, we performed *in situ* hybridization on *gfp* and *soxB1-2* RNAi animals. Consistent with the RNA-seq results, these genes exhibited reduced expression following *soxB1-2* RNAi, particularly in sensory neuron-enriched regions concentrated in the head rim and dorsal ciliated stripe (Figure 2B).

The observed transcriptional changes following *soxB1-2* RNAi prompted us to examine how these patterns compared to those in previous RNA-seq datasets. Given that our prior studies were conducted on RNA extracted from whole worms (Ross et al., 2018), we sought to determine whether broader tissue sampling may have masked gene expression changes in rare neural populations in the whole-worm datasets. Comparing this head fragment dataset to previous whole-worm *soxB1-2* RNAi RNA-seq results identified 134 shared downregulated genes (Supplemental Table 2). Gene Ontology (GO) enrichment analysis indicated that shared genes are primarily associated with 'ion homeostasis', 'transmembrane transport regulation', and 'cell communication', with key processes including 'negative regulation of ion and cation transport', 'sodium ion transmembrane transport', and 'response to metal ions' (Supplemental Figure 7A). Additionally, the presence of electrical coupling-related terms suggests a role in direct cell-to-cell communication, particularly in excitable tissues such as neurons and muscle.

In contrast, 396 genes were uniquely identified in the RNA-seq dataset from head tissue compared to our previously published day-14 whole worm dataset (Supplemental Table 2). GO analysis of these genes revealed enrichment in 'ciliary and flagellar dependent cell motility', 'movement', 'axoneme assembly', and 'microtubule bundle formation' (Supplemental Figure 7B). These features are fundamental not only for locomotion and fluid movement in the ventral epidermis but also for sensory-neural functions given the presence of cilia in sensory neurons (Azimzadeh and Basquin, 2016; Rompolas et al., 2013).

To examine whether loss of *soxB1-2* causes neoblast progeny to adopt alternative differentiation fates, we also performed GO enrichment analysis on genes upregulated in *soxB1-2(RNAi)* animals. These genes were enriched for metabolic processes, xenobiotic detoxification, potassium ion homeostasis, and endocytic pathways. Rather than indicating a shift toward alternative lineages, including non-ectodermal fates, these signatures are consistent with stress-responsive and homeostatic programs activated following loss of *soxB1*-2 (Supplemental Figure 7C). We did not detect enrichment strongly associated with alternative cell fates. These findings support the role of *soxB1-2* in sensory neural gene expression and indicate that its loss leads to widespread transcriptional changes, particularly in mechanosensory genes. RNA-seq analysis of head tissue fragments revealed additional downregulated genes that were previously undetected in whole-worm RNA-seq, many of which are likely involved in ciliated sensory neuron function, suggesting a more specialized role for *soxB1-2* in these populations. Therefore, we further investigated the genome-wide overlap between the RNA-seq and ATAC-seq datasets.

Integrating ATAC-seq and RNA-seq identifies candidate downstream targets of *soxB1-2* in neural and ciliated neural cells

To uncover potential downstream targets of *soxB1-2*, we integrated our head region ATAC-seq and RNA-seq datasets. This combined approach allowed us to identify genes that not only exhibit changes in chromatin accessibility but also corresponding changes in transcriptional output. While we found consistent signals across both assays in some genes, others were detected only by one method, underscoring the complementary nature of chromatin and transcriptomic profiling.

While differences between ATAC-seq and RNA-seq datasets are expected due to the distinct regulatory layers they capture, their integration revealed a set of 31 genes exhibiting both decreased chromatin accessibility and transcriptional downregulation (Figure 3A), highlighting them as strong candidates for direct downstream targets of *soxB1-2*. We assessed enrichment of the 31 candidate genes in neural and ciliated neural cell clusters and examined their co-expression with *soxB1-2* using single-cell RNA-seq datasets. Single-cell databases indicated most of the 31 genes were either enriched in the same neural and ciliated neural clusters where *soxB1-2* is also enriched or co-expressed with *soxB1-2* within these populations (Fincher et al., 2018; Wurtzel et al., 2015). Exceptions included *dd_380,* enriched in epidermal clusters, and *dd_8471* and *dd_8567*, which lacked specific cluster enrichment and exhibited broad co-expression across all cell types (Supplemental Table 3).

Among the genes identified, we were particularly interested in transcription factors, as these are likely to act as intermediate regulators within the broader *soxB1-2* gene regulatory network. By integrating ATAC-seq and RNA-seq, we could prioritize transcription factors that were either losing accessibility and downregulated at the chromatin and transcriptional levels or uniquely detected through chromatin accessibility changes, suggesting potential regulatory

priming. This allowed us to expand the architecture of the *soxB1-2* pathway by highlighting candidate nodes acting downstream of *soxB1-2*.

Notably, transcription factors, including *castor*, were exclusively detected through ATAC-seq assays (Figure 1D); they were not identified as differentially expressed in RNA-seq following *soxB1-2* RNAi treatment (Table 1), potentially due to low transcript abundance or subtle expression changes. However, some transcription factors, such as *mecom* and *foxj1-5,* exhibited significant decreases in both chromatin accessibility and expression in *soxB1-2*(*RNAi*) animals (Figure 3B). The *mecom* gene is a neural transcription factor that is expressed in the same neural and ciliated neural clusters as *soxB1-2* (Fincher et al., 2018; Wurtzel et al., 2015). ATAC-seq and RNA-seq peaks at the *mecom* locus in both *soxB1-2*(*RNAi*) and control animals are shown in Figure 3B, demonstrating a significant reduction in chromatin accessibility following RNAi of *soxB1-2*.

GO analysis revealed distinct enrichment patterns between the ATAC-seq and the RNA-seq datasets (Figure 3C-D). ATAC-seq enrichment was associated with direct neural functions, with 'forebrain development' as the top term, alongside pathways such as 'potassium ion transport', which is critical for neural processes like action potential generation and neurotransmitter release (Meir et al., 1999; Wang et al., 2023). Additionally, ATAC-seq revealed enrichment for 'synaptic development and assembly' (Figure 3C). In contrast, RNA-seq data GO enrichment predominantly involved cilium-related processes, including movement, assembly, and development (Figure 3D), consistent with our previous findings on gene expression associated with ciliated neurons and the epidermis (Ross et al., 2018). The enrichment of cilium-related terms aligns with the observed loss of chromatin accessibility and gene downregulation following *soxB1-2* RNAi in genes that have sensory or mechanosensory-related functions.

Additionally, this gene enrichment was reflected in spatial expression data, as in situ hybridization of these potential *soxB1-2* target genes showed a decrease in expression in body regions strongly associated with ciliated sensory neurons (Figures 1E, 2B, 4, and Supplemental Table 3).

In addition to neural gene regulation, we observed overlapping chromatin accessibility loss and downregulation of multiple genes enriched in epidermal progenitor and late epidermal lineages, as well as genes shared by both neural and epidermal populations. These include candidates potentially involved in epidermal fate determination, such as *dd_12765*, *foxj1-5, dd_13134*, and *dd_380* (Supplemental Table 3) (Fincher et al., 2018; Pascual-Carreras et al., 2021; Vij et al., 2012). The presence of both neural and epidermal signatures suggests that *soxB1-2* plays a broader role in ectodermal lineage identity. However, because this study primarily focuses on its function in neuronal lineages, our downstream analyses were concentrated on genes enriched in neural populations.

To validate that genes showing overlap in both datasets are regulated by *soxB1-2*, we cloned 14 of the 31 candidate targets and assessed their expression using in situ hybridization in both control and *soxB1-2* RNAi animals. All tested genes were expressed in the head margin and other regions enriched for sensory neurons, and all showed reduced or absent expression following *soxB1-2* knockdown (Figure 4, Supplemental Figure 8). These results support the model that *soxB1-2* activity regulates these genes and provide strong functional validation of our integrated RNA-seq and ATAC-seq approach for identifying *soxB1-2* targets in neural and ciliated neural populations. Furthermore, the validated genes include several with known or predicted roles in neural development or function (Figure 3C). For example, *kcnka*, a potassium channel homolog, was previously shown to be downregulated following *coe* RNAi, a neural

transcription factor involved in nervous system development in *S. mediterranea* (Cowles et al., 2014).

Distinct and overlapping roles of *mecom* and *castor* in mechanosensory signaling, ion transport, and ciliary regulation

Based on the significant loss of chromatin accessibility at the *mecom* and *castor* loci and their downregulation following *soxB1-2* RNAi, we investigated these genes as candidate downstream transcription factor targets of *soxB1-2* activity. While only *mecom* was significantly reduced in the *soxB1-2(RNAi)* RNA-seq dataset, we observed a reduction in expression for both genes in sensory regions by in situ hybridization (Figures 1E and 4). Combined with their known roles in neural and sensory development in other model systems, these observations prompted us to define the regulatory targets of *mecom* and *castor* to better understand their contribution to the *soxB1-2*-dependent transcriptional network. To determine the optimal RNAi timepoints for transcriptomic analysis, we performed a time-course experiment across ten RNAi feedings and monitored *mecom* and *castor* expression levels. *mecom* transcript depletion was consistently observed by RNAi day 24 (Supplemental Figure 9A-B). *castor* transcript loss occurred by day 37 (Supplemental Figure 10A-B). RNA-seq analysis revealed that knockdown of *mecom* and *castor* resulted in the significant downregulation of 25 and 391 genes, respectively (Supplemental Tables 4 and 5). The relatively small number of *mecom*-dependent genes detected is consistent with its low baseline expression in a small number of cells (Figure 4). Among the downregulated genes, *nompc* was the only mechanosensory receptor reduced in both *mecom* and *castor* RNAi samples (Figure 5 and 6; Supplemental Figures 9 and 10). *pkd2-1*, another known *soxB1-2* target involved in mechanosensation, was downregulated exclusively in the *mecom* dataset (Figure 5A

and Supplemental Figure 9C). In contrast, *castor* knockdown led to downregulation of additional mechanosensory and ion channel-related genes, including *transient receptor potential ankyrin 1* (*trpa1*) and *pkd* family members *pkd1L-1, pkd1L-3, and pkd2L-1* (Figure 6A and Supplemental Figure 10C), which were all previously implicated in *soxB1-2*-dependent mechanosensory regulation (Ross et al., 2024). In addition to mechanosensory and ion channel-related genes, *castor* RNAi resulted in the downregulation of numerous genes associated with ciliary structure and function, including *dystonin* (*dst*) and multiple dynein component homologs. These findings suggest that while *mecom* and *castor* regulate distinct gene sets, both contribute to a mechanosensory and ion channel gene expression network downstream of *soxB1-2*.

GO enrichment analysis of downregulated genes in *mecom* RNAi revealed significant enrichment for 'ion transporter activity,' consistent with the functional roles of several affected genes. Some of these genes encode TRP channels, which detect mechanical stimuli and mediate ion flux across membranes, including *pkd2-1* and *nompc*. Additionally, genes involved in ion channel regulation, such as *cav3*, which encodes a homolog to human *caveolin 3*, were also downregulated (Supplemental Table 4), establishing a potential role for *mecom* in planarians in mechanosensory signaling and ion transport regulation. In contrast, although multiple mechanosensory and ion channel-related genes were downregulated in the *castor* RNA-seq dataset following RNAi, enriched GO terms were primarily associated with ciliary functions, including 'cilium movement', 'microtubule bundle formation', 'axoneme assembly', and 'microtubule-based movement' (Figure 6B). Given the role of cilia in planarian sensory neurons, this enrichment is consistent with *castor's* potential role in sensory regulation. GO terms related to the detection of mechanical stimuli were also enriched, supporting the observed downregulation of mechanosensory genes, such as *nompc* and members of the PKD gene family,

in planarians. Additionally, GO terms enriched among castor-regulated genes overlapped with those identified in the *soxB1-2* RNAi dataset (Figures 3D and 6B). Together, these findings suggest that *castor* plays a key role in regulating genes involved in mechanosensation, ion transport, and ciliary function, positioning it as a critical component of the *soxB1-2* downstream regulatory network in planarian mechanosensory biology.

Discussion

Evidence for a transcriptional activator role of *soxB1-2* in chromatin remodeling during neural and epidermal specification

Planarians provide a powerful *in vivo* model for investigating adult neurogenesis and lineage specification. Leveraging this system, we have begun to dissect the role of *soxB1-2* in directing adult stem cell fate toward ectodermal lineages, with a particular focus on neural differentiation. Although prior studies established that *soxB1-2* promotes ectodermal identity in planarians (Ross et al., 2018), the molecular mechanisms underlying this process had not been explored. In both vertebrate and invertebrate systems, SoxB1 family members are known to play critical roles in maintaining pluripotency, sustaining progenitor identity, and promoting ectodermal development (Buescher M., 2002; Kurtova et al., 2024; Meulemans and Bronner-Fraser, 2007). However, except for mammalian SOX2 homologs, few studies have addressed the mechanisms by which SoxB1 factors regulate these processes (Iwafuchi-Doi and Zaret, 2014; Soufi et al., 2015). Here, we show that loss of *soxB1-2* leads to a marked reduction in chromatin accessibility at numerous neuronal and epidermal loci, some of which encode transcription factors. These findings suggest that *soxB1-2* may function as a transcriptional activator, initiating chromatin remodeling to prime progenitor cells for neural and epidermal differentiation.

Although our findings indicate that *soxB1-2* modulates chromatin accessibility and regulates downstream transcription factors including *mecom*, *foxj1-5*, and *castor*, we have not yet provided direct biochemical evidence of SoxB1-2 binding to closed chromatin. However, motif enrichment analysis revealed that regions losing accessibility upon *soxB1-2* knockdown are significantly enriched for Sox-family motifs, particularly those resembling the SoxB1 subclass, consistent with a model in which SoxB1-2 directly establishes or maintains chromatin accessibility at its target loci. To definitively test this pioneer-like mechanism, future work will require generation of a SoxB1-2–specific antibody to enable chromatin profiling approaches such as CUT&RUN or ChIP–seq. Integrating these assays with ATAC–seq will allow us to distinguish direct versus indirect SoxB1-2 regulatory events and determine whether SoxB1-2 engages inaccessible chromatin *in vivo*.

Mechanistic insights into *soxB1-2* function via ATAC-seq and RNA-seq integration post-RNAi

ATAC-seq has enhanced our understanding of the chromatin landscape in planarian neoblasts and throughout regeneration (Ivankovic et al., 2024; Pascual-Carreras et al., 2023; Poulet et al., 2023). When combined with RNAi-mediated perturbations, this approach has enabled detailed investigation of the epigenomic effects resulting from disruption of key regulators, including the stem cell marker *smedwi-2*, the chromatin remodeler BPTF (a core subunit of the Nucleosome Remodeling Factor, NuRF), and BAF complex components such as *brg1* and *smarcc2* (Li et al., 2021; Verma et al., 2025; Wiggans et al., 2023).

More broadly, the integration of ATAC-seq with RNA-seq has become a widely adopted strategy for linking chromatin accessibility to gene expression, with applications spanning neural

development, cancer biology, and chromatin remodeling dynamics (Bai et al., 2024; Xu et al., 2023). Building on this, recent studies have begun applying this integrative strategy in planarians following RNAi (Bai et al., 2024). In our work, combining ATAC-seq and RNA-seq enabled a mechanistic analysis of *soxB1-2*–driven lineage specification and the regulation of its direct downstream targets. Although both assays were performed on head fragments, the number of biological replicates differed (ATAC-seq: N = 2 per condition; RNA-seq: N = 4 to 6 per condition), and we interpret cross-assay comparisons in light of these differences in statistical power. Nevertheless, the concordant loss of accessibility and gene expression at key neural regulators across assays supports a robust requirement for *soxB1-2* in neural lineage specification. We identified 31 genes that exhibited both differential chromatin accessibility and altered transcript levels following *soxB1-2* knockdown. These genes were either enriched in or co-expressed with *soxB1-2* across multiple cell types, including ciliated neural cells, general neural populations, early epidermal progenitors, and neoblasts (Supplemental Figure 6 and Supplemental Table 3). Among these, ciliated neural cells showed the highest levels of enrichment and co-expression (Supplemental Table 3). These patterns support a model in which *soxB1-2* activity regulates ectodermal progenitor states through coordinated control of chromatin accessibility and gene expression (Figure 7).

Furthermore, in situ hybridization of select candidate targets (Figure 4, Supplemental Figure 8) confirmed reduced or ablated expression of the putative target genes following *soxB1-2* RNAi, supporting their status as downstream, transcriptionally regulated genes. Interestingly, changes in chromatin accessibility following *soxB1-2* RNAi did not always correspond to changes in gene expression; some loci showed decreased accessibility without changes in transcript levels, while others exhibited altered expression despite stable accessibility.

Independent validation of several of these loci (Figure 1E and Figure 2B) supports a model in which *soxB1-2* initiates chromatin remodeling that may precede or occur independently of transcriptional activation. This model aligns with findings from *Xenopus* spinal cord regeneration, where chromatin changes have been shown to prime cells for future gene expression (Kakebeen et al., 2020). These findings highlight the power of combining ATAC-seq and RNA-seq to reveal regulatory mechanisms that underlie lineage specification and transcriptional control during planarian tissue homeostasis or regeneration.

Enhanced resolution of neural regulation via targeted analysis of head tissue

Similar to previous studies (Nakazawa et al., 2003; Roberts-Galbraith et al., 2016; Wang et al., 2016), using head tissue enriched for neural and ciliated neural cell populations improved our ability to detect expression changes in neurons. This spatially targeted sampling revealed gene expression changes in sensory neurons undetectable in previous whole-worm RNA-seq datasets (Ross et al., 2018), particularly within rare or spatially restricted cell types. For example, *mecom* exhibited significantly reduced chromatin accessibility and transcript levels following *soxB1-2* RNAi in the head tissue dataset. These changes were not observed in the whole-body RNA-seq dataset. In both our whole-body and head tissue RNA-seq datasets and head tissue ATAC-seq datasets, the samples represent heterogeneous cell populations, not isolated neural or *soxB1-2*$^{+}$ cells. This heterogeneity inherently introduces noise and likely obscures additional SoxB1-2 targets or changes in rare cell populations, even in our neural-enriched head tissue. Notably, we did not detect changes in chromatin accessibility or transcript levels in sensory populations marked by *pou4-2*, despite research in our lab showing strong downregulation following *soxB1-2* RNAi (McCubbin et al., 2025). This may reflect indirect regulation or a dilution effect due to the

limited expression of *pou4-2* and the reduced sensitivity of bulk assays to capture changes in rare cell types. Future studies employing single-cell RNA-seq or FACS-based enrichment of *soxB1-2*$^{+}$ populations, pending the development of a validated antibody, will be critical for resolving these questions.

***mecom* and *castor* are downstream of *soxB1-2* and are putative mechanosensory signaling, synapse, and ion transport regulators**

Chromatin accessibility at regulatory regions near *mecom* and *castor* was markedly reduced following *soxB1-2* RNAi. Consistent with this, in situ hybridization showed a loss of *mecom* expression in cells and a substantial reduction in *castor* expression. These findings indicate that *mecom* and *castor* act downstream of *soxB1-2* and suggest that the corresponding neural cell populations are very likely not produced in the absence of *soxB1-2*, although further functional or cell-type-specific analyses will be needed to formally confirm their absence.

As cilia mediate both motility and mechanical sensing in planarians, these findings suggest that *mecom* and *castor* regulate mechanosensory populations downstream of *soxB1-2*, which is consistent with known roles of these genes in other organisms. For example, Mecom promotes supporting cell proliferation and sensory hair cell differentiation in the mammalian cochlea, and Castor specifies early-born central complex neurons in Drosophila involved in sensory integration (Chen et al., 2022; Dillon and Doe, 2024; Plath and Barron, 2015). In the context of sensory neural development, *mecom* overexpression has been shown to promote supporting cell proliferation and hair cell differentiation by modulating Tgf-β, Notch, and Wnt signaling pathways (Chen et al., 2022). The identification of *soxB1-2* downstream targets exhibiting both chromatin accessibility loss and transcriptional downregulation highlights its

regulatory role in neural and ciliated neural populations, further supporting the involvement of genes like *mecom* in these developmental processes. Additional high-resolution analyses will be required to delineate the precise cellular relationships among these factors.

Although transcript knockdown was confirmed by in situ hybridization prior to RNA-seq (Supplemental Figures 9B and 10B), RNAi targeting *mecom* did not result in significant downregulation of *mecom* transcript levels in the RNA-seq data. This may be due to its low baseline expression or its localized enrichment at the head tip in both control and experimental samples. Because *mecom* is an uncharacterized gene in planarians, we performed RNA-seq on whole worms, which diluted the signal when sequencing whole worms. Future studies should employ spatially enriched sampling, such as head dissections, to improve detection sensitivity. Alternatively, the lack of statistical significance may reflect the use of stringent analytical thresholds. Similarly, *castor* did not meet RNA-seq significance cutoffs following *soxB1-2* RNAi, despite clear changes in chromatin accessibility. Such discrepancies between ATAC-seq and RNA-seq are not uncommon and may reflect regulatory complexity—chromatin accessibility may change prior to detectable transcript-level effects, or *castor* mRNA may be unusually stable and persist despite transcriptional repression. Despite these caveats, the data highlight *mecom* and *castor* as likely components of a conserved regulatory program linking *soxB1-2* to sensory specialization (summarized in Figure 7). Future work will be critical for understanding the specific functions and cell-type contexts of *mecom* and *castor* in this regulatory axis.

Conclusion

Our study demonstrates that *soxB1-2* plays a critical role in lineage specification in *S. mediterranea* by modulating chromatin accessibility and gene expression. Integration of ATAC-seq and RNA-seq revealed that *soxB1-2* knockdown reduces chromatin accessibility at neural and epidermal loci, supporting the idea that *soxB1-2* functions as a pioneer-like transcription factor that primes chromatin for differentiation. We identified *mecom* and *castor* as downstream targets implicated in mechanosensory signaling, ion transport, and synapse formation, thereby expanding the regulatory network governed by *soxB1-2*. In addition to its mechanistic insights, this work establishes one of the few chromatin accessibility datasets in planarians, offering a foundational resource for the regeneration community and a platform for future hypothesis generation. Beyond planarian biology, these findings contribute to a broader understanding of adult neurogenesis and chromatin remodeling in regeneration-competent systems. The regulatory principles uncovered here provide a conceptual framework that could inform stem cell reprogramming strategies in less regenerative organisms, including humans, where restoring neurogenic capacity remains a critical therapeutic challenge.

Methods

Animal husbandry

Asexual *S. mediterranea* (clonal strain CIW4) planarians were kept in 9-cup Ziploc reusable containers in 1x Montjuïc salts (1.6 mM NaCl, 1.0 mM $CaCl_2$, 1.0 mM $MgSO_4$, 0.1 mM $MgCl_2$, 0.1 mM KCl, 1.2 mM $NaHCO_3$) dissolved in ultrapure filtered water at 20°C. Animals were fed beef calf liver weekly, as previously described (Merryman et al., 2018). Animals were starved for 1 week prior to all experiments.

RNA interference

Double-stranded RNA (dsRNA) was expressed in HT115 *E. coli* containing cDNA inserts in either pJC.53.2 (Collins et al., 2010) or pPR-T4P (Liu et al., 2013) as described previously (Ross et al., 2018). Worms measuring 3-4 mm (for in situ hybridization) or 5-7 mm (for ATAC-seq and RNA-seq) were fed 0.5 µg dsRNA per worm, mixed 1:3 with a paste composed of 200 µL ultrapure water, 500 µL pureed calf liver, and 40 µL food dye. Animals were cleaned with 1X Montjuïc salt exchanges after each feeding and again the following day. Animals were fed dsRNA twice per week, for a total of five to ten feedings depending on the target gene. *soxB1-2* RNAi animals received five feedings, *mecom* RNAi animals received 6 feedings, and *castor* RNAi animals received 10 feedings. *Green fluorescent protein* (*gfp*) dsRNA was used as the negative control. For *soxB1-2* RNAi experiments, animals were amputated anterior to the pharynx three days after the fifth feeding for RNA-seq and ATAC-seq or fixed for in situ hybridization seven days after the fifth feeding. For *mecom* and *castor* RNAi experiments, whole worms (3-4 mm) were fixed for in situ hybridization or processed for RNA-seq seven days after the sixth feeding or tenth feeding, respectively.

ATAC-seq assay and analysis

Two independent replicates of both *soxB1-2* RNAi and *gfp* control groups (30 heads per replicate) were placed directly into ATAC-seq lysis buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl, 0.1% Digitonin, 0.1% Tween-20, 0.1% NP-40, in water) and pipetted up and down until homogeneous. The slurry was run through a 50-micron filter, and the ATAC-seq assay was performed as previously described (Corces et al., 2017) with the Illumina tagment DNA enzyme and buffer kit (cat# 20034197). Samples were purified using Zymo DNA Clean and

Concentrator-5 Kit (cat# D4014) and stored at -20°C until sequencing. Samples were quality-checked on an Agilent 2100 Bioanalyzer and sequenced to a read depth of 38 million paired-end reads on the Illumina NovaSeq 6000 at the University of California, San Diego IGM Genomics Center, Inc. (San Diego, CA). The ATAC-seq sequencing data have been deposited in NCBI under BioProject accession PRJNA1292954.

ATAC-seq libraries were assessed for quality using FastQC (v0.11.9) (Andrews, 2010), and fragment size distributions were evaluated with ATACseqQC (v1.26.0) (Ou et al., 2018). Adapter trimming and merging of forward and reverse reads were performed using NGmerge (v0.3) (Gaspar, 2018). Trimmed reads were aligned to both the S3h1 and S2F2 *Schmidtea mediterranea* genome assemblies using Bowtie2 (v2.4.4) (Langdon, 2015). Samtools (v1.2.0) software was used to remove mitochondrial and improperly paired reads and to generate BAM files (Li et al., 2009). To control for differences in sequencing depth across ATAC-seq libraries, reads were down-sampled prior to peak calling and footprinting to equalize the number of properly paired, uniquely mapped reads between conditions. Subsampling was performed using samtools (v1.17) with the -s flag, which performs Bernoulli sampling on alignments. For each replicate, BAM files were name-sorted and subsampled using a unique random seed and fraction to achieve matched read counts across samples. Picard (v2.26.4) was used to remove PCR duplicates. Deduplicated BAM files were used for peak calling with Genrich (BioGrids) using parameters -j -v -p 0.01 -a 150. Peaks were called for individual and pooled replicates. Genrich-generated BedGraph files were converted to BroadPeak format and processed as previously described (Reske et al., 2020). Consensus peaks were mapped to gene annotation models for both genomes using BEDTools (v2.31.1) (Quinlan and Hall, 2010). RPKM-normalized coverage tracks (bigWig) and profile plots were generated using deepTools (v3.5.1) (Ramirez et al., 2014).

Differential accessibility analysis between *gfp* controls and *soxB1-2* RNAi samples was performed using csaw (v1.36.0) (Lun and Smyth, 2016), as described previously (Reske et al., 2020). Gene Ontology enrichment was performed using enrichGO from clusterProfiler (v4.12.6) (Yu et al., 2012). The resulting datasets were compared to the human proteome using BLASTX (cut-off e-value < $1e^{-3}$). Human UniProt IDs (UniProt, 2015) were used as input for annotation and overrepresentation analysis. Statistical analyses were carried out in R (v4.4.0) (Team, 2024), and figures were generated using ggplot2 (Wickham, 2009).

Motif enrichment analysis

Differentially accessible (DA) regions were first identified independently using csaw and edgeR (Bioconductor v3.19), based on a unified peak set generated across all ATAC-seq samples. Peaks with FDR < 0.05 were considered significantly differentially accessible. Peaks with positive $\log_2$FC values were classified as more accessible in *soxB1-2*(*RNAi)*, while peaks with negative $\log_2$FC values were classified as more accessible in controls. Significant DA peaks were separated by direction and resized to a fixed width of 350 bp centered on the peak midpoint using GenomicRanges to standardize regions for footprinting. Footprint scores were computed across the unified peak set for each biological replicate using TOBIAS FootprintScores (Bentsen et al., 2020). Condition-level footprint signal tracks were generated by averaging replicate footprint bigWig files using bigwigCompare (deepTools 3.5.6; --operation mean) to produce a single representative footprint track for control and *soxB1-2* RNAi samples.

Differential TF binding analysis was performed using TOBIAS BINDetect (Bentsen et al., 2020), supplying the averaged footprint tracks, the unified resized peak set, and the S2F2

Schmidtea mediterranea genome. A curated set of transcription factor motifs derived from JASPAR (Castro-Mondragon et al., 2022) was used as input motifs. BINDetect was run with default settings to identify motif sites and infer differential motif occupancy between conditions (Bentsen et al., 2020; Quinlan and Hall, 2010; Ramirez et al., 2014)**.**

Gene identification and cloning

The *soxB1-2* cDNA used in all experiments was previously cloned into pJC.53.2 (Collins et al., 2010; Ross et al., 2018) in our lab. All other gene sequences were identified through PlanMine (Rozanski et al., 2019) using the dd_Smed_v6 transcriptome and cloned from *S. mediterranea* cDNA into pPR-T4P (Liu et al., 2013) using ligation-independent cloning and gene-specific primers as described in Adler and Alvarado (2018) (all primer sequences and gene identifiers listed in Supplemental Table 6). In the text, dd_Smed_v6 transcripts are referred to as *dd_transcript no.* for brevity.

Wholemount in situ hybridization

Antisense probes were generated from DNA templates amplified from pJC.53.2 (Collins et al., 2010; Ross et al., 2018) or pPR-T4P (Liu et al., 2013), incorporating digoxigenin-labeled UTPs, as previously described (Pearson B. J., 2009). Samples were incubated in 5% N-acetyl cysteine (NAC) in 1× PBS for five minutes to euthanize the animals, then fixed in 4% formaldehyde in 1× PBS with 0.1% Triton X-100 for 25 minutes. Animals were processed for in situ hybridization as described previously (King R. S., 2013), except for *mecom*, a low-abundance transcript, for which hybridization and heated washes were performed at 58°C.

RNA-seq assay and analysis

For *soxB1-2* RNAi experiments, head fragments from four independent *gfp* control and six independent *soxB1-2* RNAi replicates were homogenized in TRIzol (Invitrogen), and RNA was extracted and purified according to the manufacturer's protocol. For *mecom* and *castor* RNAi experiments, the same procedure was applied to whole worms, using five independent *gfp* controls and five *mecom* or castor RNAi replicates. RNA was treated with the TURBO DNA-free Kit and column-purified using the Zymoclean Miniprep Kit (cat. no. 11-330). RNA libraries were sequenced on an Illumina NovaSeq 6000 platform to a depth of 24 million reads per sample (12 million paired-end reads) at MedGenome, Inc. (Foster City, CA). Reads were pseudo-aligned to the S2F2 transcriptome using kallisto (Bray N., 2016), and differential gene expression analysis was performed using Bioconductor (Huber W., 2015) and DESeq2 (Love M., 2014) with an FDR cut-off value ≤ 0.05 and fold-change threshold of 1.4. The RNA sequencing data have been deposited in NCBI under BioProject accession numbers PRJNA1292856 and PRJNA1292967. Gene Ontology enrichment was conducted using enrichGO from clusterProfiler (v4.12.6) (Yu et al., 2012). Gene sets were compared to the human proteome using BLASTX (e-value cut-off $< 1e^{-3}$). Resulting UniProt human gene IDs were used for annotation and GO overrepresentation analysis. Statistical analyses were carried out in R (v4.4.0) (Team, 2024), and figures were generated using ggplot2 (Wickham, 2009).

Acknowledgements

This work is based in part on the doctoral dissertation of Mallory Cathell submitted to the University of California, San Diego, and San Diego State University in partial fulfillment of the requirements for the Doctor of Philosophy degree in Cell and Molecular Biology (Cathell, 2025).

This study was supported by the Rees-Stealy Research Foundation and California Institute for Regenerative Medicine Grant EDUC4-12813 predoctoral and postdoctoral fellowships to M.C. and M.A.A., respectively, and NIH R01GM135657 to RMZ. We thank Arianna Farrar, Carina Real, and Jennifer Severance for their help with RNAi experiments, and Elizabeth Duncan and Prince Verma for helpful discussions and advice regarding ATAC-seq analysis.

Tables

Table 1. Transcription factors with differential ATAC-seq accessibility but not differentially expressed in RNA-seq following *soxB1-2* RNAi.

Gene	Transcription factor domain classification
Pknox2 (dd_Smed_v6_8606)	Homeobox
Msh1 (dd_Smed_v6_18505)	Homeobox
Meox1 (dd_Smed_v6_12317)	Homeobox
Phtf2 (dd_Smed_v6_8465)	Homeobox
Castor (dd_Smed_v6_6778)	Zinc finger
Tcf-3 (dd_Smed_v6_8152)	T-cell specific & High mobility group
dd_Smed_v6_4717	Ets-1
Tcf-2 (dd_Smed_v6_9140)	T-cell specific & High mobility group
Smad9 (dd_Smed_v6_6050)	SMAD
Islet1 (dd_Smed_v6_8820)	LIM & Homeobox

Figure Legends

Figure 1. *soxB1-2* RNAi reduces chromatin accessibility in *S. mediterranea*. (A) Histogram of $\log_2$ fold changes ($\log_2$FC) in differentially accessible ATAC-seq peaks between *gfp* and *soxB1-2* RNAi groups, with 264 peaks exhibiting decreased accessibility (blue) and seven peaks showing increased accessibility (orange) (FDR < 0.05). Sample size: N = 2 for both control and *soxB1-2* RNAi groups. (B) Profile plot of global ATAC-seq signal across gene bodies in *gfp* control (blue) and *soxB1-2* RNAi (orange) groups. TSS, transcription start site; TES, transcription end site. (C) Profile plot of global ATAC-seq signal across promoter regions in *gfp* control (blue) and *soxB1-2* RNAi (orange) groups. TSS, transcription start site. (D) Genome tracks showing reduced chromatin accessibility at the *mecom* and dd_*26266* locus. (E) In situ hybridization of genes enriched in neural and ciliated neural cell populations that display reduced chromatin accessibility post-*soxB1-2* RNAi. Blue arrows indicate regions of in situ signal loss following *soxB1-2* RNAi, with specific anatomical areas marked using red arrows. Sample size: For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of each treatment group.

Figure 2. Genes identified by ATAC-seq are regulated by *soxB1-2*. (A) Volcano plot of differentially expressed genes in *gfp* control vs. *soxB1-2* RNAi groups, with downregulated (blue) and upregulated (red) genes (FC ≥ 1.4, p < 0.1). Sample size: N = 4 and N = 6 for control and *soxB1-2* RNAi groups, respectively. (B) In situ hybridization of putative sensory neural genes identified by RNA-seq following *soxB1-2* RNAi. Blue arrows indicate regions of in situ signal loss following *soxB1-2* RNAi. Red arrows mark anatomical regions with undetectable

signal; red asterisk denotes loss of expression in the tip of the head. Sample size: For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of each treatment group.

Figure 3. ATAC-seq and RNA-seq reveal coordinated loss of chromatin accessibility and downregulation of neural genes. (A) Heatmap of differentially expressed genes identified through RNA-seq (FC ≥ 1.4, $p < 0.1$) in *gfp* control vs. *soxB1-2* RNAi groups. Z-scores indicate upregulation (2), no change (0), and downregulation (-2). Labeled genes are putative neural genes identified from both ATAC-seq and RNA-seq experiments. (B) Genome tracks showing reduced chromatin accessibility and gene downregulation following *soxB1-2* RNAi at putative neural gene loci (*foxn(p)*, *foxj1-5*, and *stumble*) in *S. mediterranea*. (C) Gene Ontology (GO) term enrichment analysis of genes associated with reduced chromatin accessibility following *soxB1-2* knockdown. (D) Gene Ontology (GO) term enrichment analysis of downregulated genes in RNA-seq experiments following *soxB1-2* knockdown.

Figure 4. In situ hybridization validation of neural genes identified in ATAC-seq and RNA-seq. (A) In situ hybridization of neural genes expressed in the head rim, dorsal ciliated stripe, and periphery in *gfp* control (top) and *soxB1-2* RNAi (bottom) animals. Sample size: For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of each treatment group.

Figure 5. *mecom* RNAi disrupts mechanosensory and ion channel-like gene expression. (A) Volcano plot of differentially expressed genes in *gfp* control vs. *mecom* RNAi groups, with downregulated (blue) and upregulated (red) genes (FC ≥ 1.4, $p < 0.1$). Sample size: N = 5 for control and *mecom* RNAi groups. (B) Gene Ontology (GO) term enrichment analysis of downregulated genes following *mecom* knockdown.

Figure 6. *castor* RNAi disrupts genes involved in synapse and ciliary function. (A) Volcano plot of differentially expressed genes in *gfp* control vs. *castor* RNAi groups, with downregulated (blue) and upregulated (red) genes (FC ≥ 1.4, $p < 0.1$). Sample size: N = 5 for control and *castor* RNAi groups. (B) Gene Ontology (GO) term enrichment analysis of downregulated genes following *castor* knockdown.

Figure 7. Hypothetical model of SoxB1-2-regulated neuronal cell fates uncovered by integrating epigenomic and transcriptomic approaches. Previous work showed that *soxB1-2* is first expressed in a *piwi-1*$^{+}$ progenitor population that acquires epidermal or neuronal cell fates that terminally differentiate into ciliated sensory neurons or epidermis (Ross et al., 2018). Subsequent work has focused on understanding the mechanisms downstream of SoxB1-2 involved in specifying sensory neuron subclasses. (McCubbin et al., 2025) found that the Pou4 homolog, *pou4-2,* regulates the differentiation of ciliated sensory neurons in the ciliated stripes. In this study, we show that *mecom* and *castor* also function in the *soxB1-2* regulatory network by regulating neuronal gene expression. Created in https://BioRender.com.

Supplemental Figure Legends

Supplemental Figure 1. Schematic of the assay used to assess chromatin accessibility and transcriptional changes in *gfp* control and *soxB1-2* RNAi conditions.

Supplemental Figure 2. Fragment size distribution and Pearson correlation analysis of ATAC-seq replicates. (A) Fragment size distributions from each ATAC-seq replicate. Forward and reverse reads are shown in red and blue, respectively. (B) Biological replicates within control groups are highly correlated (r = 0.96; n = 2). (C) Biological replicates within *soxB1-2* RNAi groups are highly correlated (r = 0.96; n = 2).

Supplemental Figure 3. ATAC-seq analysis reveals differences in chromatin accessibility following *soxB1-2* knockdown. Distribution of ATAC-seq peaks across annotated genomic regions of the *S. mediterranea* genome, comparing all *gfp* control and *soxB1-2* RNAi replicates. "DA peaks" indicate regions with detectable accessibility in *gfp* controls that are absent following *soxB1-2* knockdown.

Supplemental Figure 4. Motif enrichment analysis of differential ATAC-seq footprints using BINDetect. (A) Raw score distributions for transcription factor binding site footprints in control and *soxB1-2* RNAi samples, showing the density of unscaled footprint scores across conditions. (B) Normalized score distributions after within-condition scaling, demonstrating comparable score distributions suitable for differential binding analysis. (C) Volcano plot of differential transcription factor binding generated by BINDetect. Positive differential binding scores indicate motifs with higher footprint occupancy in *soxB1-2* RNAi samples, whereas

negative scores indicate motifs more enriched in control samples. Notably, Sox family motifs (SoxB1, SoxB2, SoxC, SoxF) and several T-box factors (e.g., Tbx2/3, Brachyury) show strong differential binding.

Supplemental Figure 5. Heatmap of isolated differentially expressed genes (FC ≥ 1.4, p < 0.05) in *gfp* control vs. *soxB1-2* RNAi groups. Z-scores indicate upregulation (2), no change (0), and downregulation (-2). Labeled genes represent putative sensory neural genes selected for validation via in situ hybridization.

Supplemental Figure 6. Co-expression of *soxB1-2* and select genes identified by RNA-seq. Single-cell expression profiles of cells co-expressing *soxB1-2* and representative RNA-seq–identified genes that are significantly downregulated following *soxB1-2* RNAi. Cell type abbreviations: sNB, sigma neoblast; zNB, zeta neoblast; gNB, gamma neoblast; EEP, early epidermal progenitor; LEP, late epidermal progenitor; E1, epidermis 1; E2, epidermis 2; Gut, intestine; PN, protonephridia; PP, parapharyngeal; Mu, muscle; N, neural; N-C, neural-ciliated.

Supplemental Figure 7. Gene Ontology (GO) term enrichment of genes differentially expressed following soxB1-2 RNAi. (A) GO terms associated with 134 genes shared between RNA-seq datasets from whole-worm and head-only samples. (B) GO term enrichment of 148 downregulated genes uniquely identified in whole-worm RNA-seq datasets (Ross et al., 2018). (C) GO term enrichment of 416 upregulated genes in RNA-seq datasets of head-only samples.

Supplemental Figure 8. Expression patterns of genes with reduced chromatin accessibility or transcription following *soxB1-2* RNAi. Decreased or absent expression of genes identified through ATAC-seq or RNA-seq as downregulated following *soxB1-2* RNAi. *dd_11972* shows reduced expression in the head tip and across the epidermis, while *dd_7770*, *dd_45275*, *dnai1*, and s*lc25a-21* exhibit complete loss of expression. *pkd1L-2* serves as a positive control for confirming the knockdown of *soxB1-2* downstream targets via in situ hybridization. Sample size: For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of each treatment group.

Supplemental Figure 9. Experimental design for *mecom* RNA-seq and heatmap of candidate targets meeting differential expression cut-offs. (A) Schematic of the assay used to assess differential gene expression in *gfp* control and *mecom* RNAi conditions. (B) Validation of *mecom* transcript loss at the time point used for RNA-seq sample collection. For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of the treatment group. (C) Heatmap of differentially expressed genes (FC ≥ 1.4, $p < 0.1$) in *gfp* control vs. *mecom* RNAi groups. Z-scores indicate upregulation (2), no change (0), and downregulation (-2). Labeled genes are putative mechanosensory genes identified in RNA-seq experiments performed on *mecom* RNAi samples.

Supplemental Figure 10. Experimental design for *castor* RNA-seq and heatmap of candidate targets meeting differential expression cut-offs. (A) Schematic of the assay used to

assess differential gene expression in *gfp* control and *castor* RNAi conditions. (B) Validation of *castor* transcript loss at the time point used for RNA-seq sample collection. For controls, N ≥ 3 worms were analyzed, and all samples showed consistent expression patterns. For RNAi-treated animals, the number of biological replicates (N) is indicated at the top left of the treatment group. (C) Heatmap of differentially expressed genes (FC ≥ 1.4, $p < 0.1$) in *gfp* control vs. *castor* RNAi groups. Z-scores indicate upregulation (2), no change (0), and downregulation (-2). Labeled genes are putative mechanosensory genes identified in RNA-seq experiments performed on *castor* RNAi samples.

Supplemental Table Files

Supplemental Table 1. List of ATAC-seq peaks detected in control and RNAi groups, including motif enrichment analysis results.

Supplemental Table 2. Differentially expressed genes in soxB1-2 RNAi versus control groups and comparison of results between head-only and whole-body RNA-seq datasets.

Supplemental Table 3. Genes overlapping between ATAC-seq and RNA-seq datasets and associated single-cell cluster enrichment analysis.

Supplemental Table 4. DESeq2 differential expression analysis results for *Smed-mecom* RNAi.

Supplemental Table 5. DESeq2 differential expression analysis results for *Smed-castor* RNAi.

References

Adler, C.E., Alvarado, A.S., 2018. Systemic RNA Interference in Planarians by Feeding of dsRNA Containing Bacteria. Methods Mol Biol 1774, 445–454.

Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: [https://www.bioinformatics.babraham.ac.uk/projects/fastqc/], 0.11.9 ed.

Azimzadeh, J., Basquin, C., 2016. Basal bodies across eukaryotes series: basal bodies in the freshwater planarian Schmidtea mediterranea. Cilia 5, 15.

Bai, Y., Deng, X., Chen, D., Han, S., Lin, Z., Li, Z., Tong, W., Li, J., Wang, T., Liu, X., Liu, Z., Cui, Z., Zhang, Y., 2024. Integrative analysis based on ATAC-seq and RNA-seq reveals a novel oncogene PRPF3 in hepatocellular carcinoma. Clin Epigenetics 16, 154.

Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., Kim, J., Looso, M., 2020. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun 11, 4267.

BioGrids, Genrich: Detecting sites of genomic enrichment.

GitHub repository. Available at: https://github.com/jsh58/Genrich.

Bray N., P.H., Melsted P., Patcher L., 2016. Near-optimal Probabilistic RNA-Seq Quantification. Nat. Biotechnol. 34, 525–527.

Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol 109, 21 29 21–21 29 29.

Buescher M., H.F.S., Chia W., 2002. Formation of neuroblasts in the embryonic central nervous system of Drosophila melanogaster is controlled by SoxNeuro. Development (Cambridge, England) 129, 4193–4203.

Castro-Mondragon, J.A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., Fornes, O., Leung, Tiffany Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., Vandepoele, K., Wasserman, W.W., Parcy, F., Mathelier, A., 2022. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Research 50, Pages D165.

Cathell, M., 2025. SoxB1-2 Regulates Sensory Neuronal Differentiation Through Chromatin Remodeling in the Regenerative Flatworm, Schmidtea mediterranea. San Diego State University, United States -- California, p. 140.

Cebrià, F., 2007. Regenerating the central nervous system: how easy for planarians! Dev Genes Evol 217, 733–748.

Chen, P., Zuo, N., Wu, C., Ma, J., Li, Y., Gu, J., Li, W., Liu, S., 2022. MECOM promotes supporting cell proliferation and differentiation in cochlea. J Otol 17, 59–66.

Collins, J.J., 3rd, Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V., Newmark, P.A., 2010. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. PLoS Biol 8, e1000509.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., Kathiria, A., Cho, S.W., Mumbach, M.R., Carter, A.C., Kasowski, M., Orloff, L.A., Risca, V.I., Kundaje, A., Khavari, P.A., Montine, T.J., Greenleaf, W.J., Chang, H.Y., 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14, 959–962.

Cowles, M.W., Omuro, K.C., Stanley, B.N., Quintanilla, C.G., Zayas, R.M., 2014. COE loss-of-function analysis reveals a genetic program underlying maintenance and regeneration of the nervous system in planarians. PLoS genetics 10, e1004746.

Dillon, N.R., Doe, C.Q., 2024. Castor is a temporal transcription factor that specifies early born central complex neuron identity. Development (Cambridge, England) 151.

Farnesi, R.M., Tei, S., 1980. Dugesia lugubris s.l. auricles: research into the ultrastructure and on the functional efficiency. Riv Biol 73, 65–77.

Fincher, C.T., Wurtzel, O., de Hoog, T., Kravarik, K.M., Reddien, P.W., 2018. Cell type transcriptome atlas for the planarian Schmidtea mediterranea. Science 360.

Gaspar, J.M., 2018. NGmerge: merging paired-end reads via novel empirically-derived models of sequencing errors. BMC Bioinformatics 19, 536.

Huber W., C.V.J., Gentleman R., Anders S., Carlson M., Carvalho B. S., et al., 2015. Orchestrating High-Throughput Genomic Analysis with Bioconductor. Nat. Methods 12, 115–121.

Hyman, L.H., 1951. The Invertebrates: Platyhelminthes and Rhynchocoela. The Acoelomate Bilateria. McGraw Hill.

Isshiki, T., Pearson, B., Holbrook, S., Doe, C.Q., 2001. Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. Cell 106, 511–521.

Ivankovic, M., Brand, J.N., Pandolfini, L., Brown, T., Pippel, M., Rozanski, A., Schubert, T., Grohme, M.A., Winkler, S., Robledillo, L., Zhang, M., Codino, A., Gustincich, S., Vila-Farre, M., Zhang, S., Papantonis, A., Marques, A., Rink, J.C., 2024. A comparative analysis of planarian genomes reveals regulatory conservation in the face of rapid structural divergence. Nat Commun 15, 8215.

Iwafuchi-Doi, M., Zaret, K.S., 2014. Pioneer transcription factors in cell reprogramming. Genes Dev 28, 2679–2692.

Kakebeen, A.D., Chitsazan, A.D., Williams, M.C., Saunders, L.M., Wills, A.E., 2020. Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. Elife 9.

Kamachi, Y., Kondoh, H., 2013. Sox proteins: regulators of cell fate specification and differentiation. Development (Cambridge, England) 140, 4129–4144.

King R. S., N.P.A., 2013. In Situ hybridization Protocol for Enhanced Detection of Gene Expression in the Planarian Schmidtea Mediterranea. BMC Dev. Biol. 13.

Kurtova, A.I., Finoshin, A.D., Aparina, M.S., Gazizova, G.R., Kozlova, O.S., Voronova, S.N., Shagimardanova, E.I., Ivashkin, E.G., Voronezhskaya, E.E., 2024. Expanded expression of pro-neurogenic factor SoxB1 during larval development of gastropod Lymnaea stagnalis suggests preadaptation to prolonged neurogenesis in Mollusca. Front Neurosci 18, 1346610.

Langdon, W.B., 2015. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. BioData Min 8, 1.

Li, D., Taylor, D.H., van Wolfswinkel, J.C., 2021. PIWI-mediated control of tissue-specific transposons is essential for somatic cell differentiation. Cell Rep 37, 109776.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Liu, S.Y., Selck, C., Friedrich, B., Lutz, R., Vila-Farre, M., Dahl, A., Brandl, H., Lakshmanaperumal, N., Henry, I., Rink, J.C., 2013. Reactivating head regrowth in a regeneration-deficient planarian species. Nature 500, 81–84.

Love M., H.W., Anders S., 2014. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. Genome Biol 15.

Lun, A.T., Smyth, G.K., 2016. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res 44, e45.

McCubbin, R.A., Auwal, M.A., Wang, S., Alvarez Zepeda, S., Sasik, R., Zeller, R.W., Ross, K.G., Zayas, R.M., 2025. Smed-pou4-2 regulates mechanosensory neuron regeneration and function in planarians. Elife 14.

Meir, A., Ginsburg, S., Butkevich, A., Kachalsky, S.G., Kaiserman, I., Ahdut, R., Demirgoren, S., Rahamimoff, R., 1999. Ion channels in presynaptic nerve terminals and control of transmitter release. Physiol Rev 79, 1019–1088.

Merryman, M.S., Alvarado, A.S., Jenkin, J.C., 2018. Culturing Planarians in the Laboratory. Methods Mol Biol 1774, 241–258.

Meulemans, D., Bronner-Fraser, M., 2007. The amphioxus SoxB family: implications for the evolution of vertebrate placodes. International journal of biological sciences 3, 356–364.

Miyagi, S., Kato, H., Okuda, A., 2009. Role of SoxB1 transcription factors in development. Cellular and molecular life sciences : CMLS 66, 3675–3684.

Nakazawa, M., Cebrià, F., Mineta, K., Ikeo, K., Agata, K., Gojobori, T., 2003. Search for the evolutionary origin of a brain: planarian brain characterized by microarray. Mol Biol Evol 20, 784–791.

Neiro, J., Sridhar, D., Dattani, A., Aboobaker, A., 2022. Identification of putative enhancer-like elements predicts regulatory networks active in planarian adult stem cells. Elife 11.

Okamoto, K., Takeuchi, K., Agata, K., 2005. Neural projections in planarian brain revealed by fluorescent dye tracing. Zoolog Sci 22, 535–546.

Ou, J., Liu, H., Yu, J., Kelliher, M.A., Castilla, L.H., Lawson, N.D., Zhu, L.J., 2018. ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. BMC Genomics 19, 169.

Pascual-Carreras, E., Herrera-Ubeda, C., Rossello, M., Coronel-Cordoba, P., Garcia-Fernandez, J., Salo, E., Adell, T., 2021. Analysis of Fox genes in Schmidtea mediterranea reveals new families and a conserved role of Smed-foxO in controlling cell death. Sci Rep 11, 2947.

Pascual-Carreras, E., Marín-Barba, M., Castillo-Lara, S., Coronel-Córdoba, P., Magri, M.S., Wheeler, G.N., Gómez-Skarmeta, J.L., Abril, J.F., Saló, E., Adell, T., 2023. Wnt/β-catenin signalling is required for pole-specific chromatin remodeling during planarian regeneration. Nature Communications 14, 298.

Pearson B. J., E.G.T., Gurley K. A., Rink J. C., Miller D. E., Sánchez Alvarado A., 2009. Formaldehyde-based Whole-Mount In Situ Hybridization Method for Planarians. Dev. Dyn. 238, 433–450.

Plath, J.A., Barron, A.B., 2015. Current progress in understanding the functions of the insect central complex. Current Opinion in Insect Science 12, 11–18.

Poulet, A., Kratkiewicz, A.J., Li, D., van Wolfswinkel, J.C., 2023. Chromatin analysis of adult pluripotent stem cells reveals a unique stemness maintenance strategy. Sci Adv 9, eadh4887.

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., Manke, T., 2014. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42, W187–191.

Reho, G., Lelievre, V., Cadiou, H., 2022. Planarian nociception: Lessons from a scrunching flatworm. Front Mol Neurosci 15, 935918.

Reiprich, S., Wegner, M., 2015. From CNS stem cells to neurons and glia: Sox for everyone. Cell Tissue Res 359, 111–124.

Reske, J.J., Wilson, M.R., Chandler, R.L., 2020. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. Epigenetics Chromatin 13, 22.

Rink, J.C., 2013. Stem cell systems and regeneration in planaria. Dev Genes Evol 223, 67–84.

Rink, J.C., 2018. Stem Cells, Patterning and Regeneration in Planarians: Self-Organization at the Organismal Scale., in: Rink, J.C. (Ed.), Planarian Regeneration. Methods in Molecular Biology. Humana Press, New York, NY.

Roberts-Galbraith, R.H., Brubacher, J.L., Newmark, P.A., 2016. A functional genomics screen in planarians reveals regulators of whole-brain regeneration. Elife 5.

Rompolas, P., Azimzadeh, J., Marshall, W.F., King, S.M., 2013. Chapter Twelve - Analysis of Ciliary Assembly and Function in Planaria, in: Marshall, W.F. (Ed.), Methods in Enzymology. Academic Press, pp. 245–264.

Ross, K.G., Alvarez Zepeda, S., Auwal, M.A., Garces, A.K., Roman, S., Zayas, R.M., 2024. The Role of Polycystic Kidney Disease-Like Homologs in Planarian Nervous System Regeneration and Function. Integr Org Biol 6, obae035.

Ross, K.G., Currie, K.W., Pearson, B.J., Zayas, R.M., 2017. Nervous system development and regeneration in freshwater planarians. Wiley Interdiscip Rev Dev Biol 6.

Ross, K.G., Molinaro, A.M., Romero, C., Dockter, B., Cable, K.L., Gonzalez, K., Zhang, S., Collins, E.S., Pearson, B.J., Zayas, R.M., 2018. SoxB1 Activity Regulates Sensory Neuron Regeneration, Maintenance, and Function in Planarians. Dev Cell 47, 331–347 e335.

Rozanski, A., Moon, H., Brandl, H., Martin-Duran, J.M., Grohme, M.A., Huttner, K., Bartscherer, K., Henry, I., Rink, J.C., 2019. PlanMine 3.0-improvements to a mineable resource of flatworm biology and biodiversity. Nucleic Acids Res 47, D812–D820.

She, Z.Y., Yang, W.X., 2015. SOX family transcription factors involved in diverse cellular events during development. Eur J Cell Biol 94, 547–563.

Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., Zaret, K.S., 2015. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell 161, 555–568.

Stevanovic, M., Drakulic, D., Lazic, A., Ninkovic, D.S., Schwirtlich, M., Mojsin, M., 2021. SOX Transcription Factors as Important Regulators of Neuronal and Glial Differentiation During Nervous System Development and Adult Neurogenesis. Front Mol Neurosci 14, 654031.

Team, R.C., 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

UniProt, C., 2015. UniProt: a hub for protein information. Nucleic Acids Res 43, D204–212.

Verma, P., Allen, J.M., Sánchez Alvarado, A., Duncan, E.M., 2025. Chromatin remodeling protein BPTF mediates chromatin accessibility at gene promoters in planarian stem cells. BMC Genomics 26, 232.

Vij, S., Rink, J.C., Ho, H.K., Babu, D., Eitel, M., Narasimhan, V., Tiku, V., Westbrook, J., Schierwater, B., Roy, S., 2012. Evolutionarily ancient association of the FoxJ1 transcription factor with the motile ciliogenic program. PLoS genetics 8, e1003019.

Wang, I.E., Lapan, S.W., Scimone, M.L., Clandinin, T.R., Reddien, P.W., 2016. Hedgehog signaling regulates gene expression in planarian glia. Elife 5.

Wang, Z.W., Trussell, L.O., Vedantham, K., 2023. Regulation of Neurotransmitter Release by K(+) Channels. Adv Neurobiol 33, 305–331.

Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.

Wiggans, M., Zhu, S.J., Molinaro, A.M., Pearson, B.J., 2023. The BAF chromatin remodeling complex licenses planarian stem cells access to ectodermal and mesodermal cell fates. BMC Biology 21, 227.

Wurtzel, O., Cote, L.E., Poirier, A., Satija, R., Regev, A., Reddien, P.W., 2015. A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians. Dev Cell 35, 632–645.

Xu, F., Cong, P., Lu, Z., Shi, L., Xiong, L., Zhao, G., 2023. Integration of ATAC-Seq and RNA-Seq identifies key genes and pathways involved in the neuroprotection of S-adenosylmethionine against perioperative neurocognitive disorder. Comput Struct Biotechnol J 21, 1942–1954.

Yu, G., Wang, L.G., Han, Y., He, Q.Y., 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287.

Zhang, S., Cui, W., 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells 6, 305–311.

Zhong, L., Wang, D., Gan, X., Yang, T., He, S., 2011. Parallel expansions of Sox transcription factor group B predating the diversifications of the arthropods and jawed vertebrates. PloS one 6, e16570.

FIGURE 1

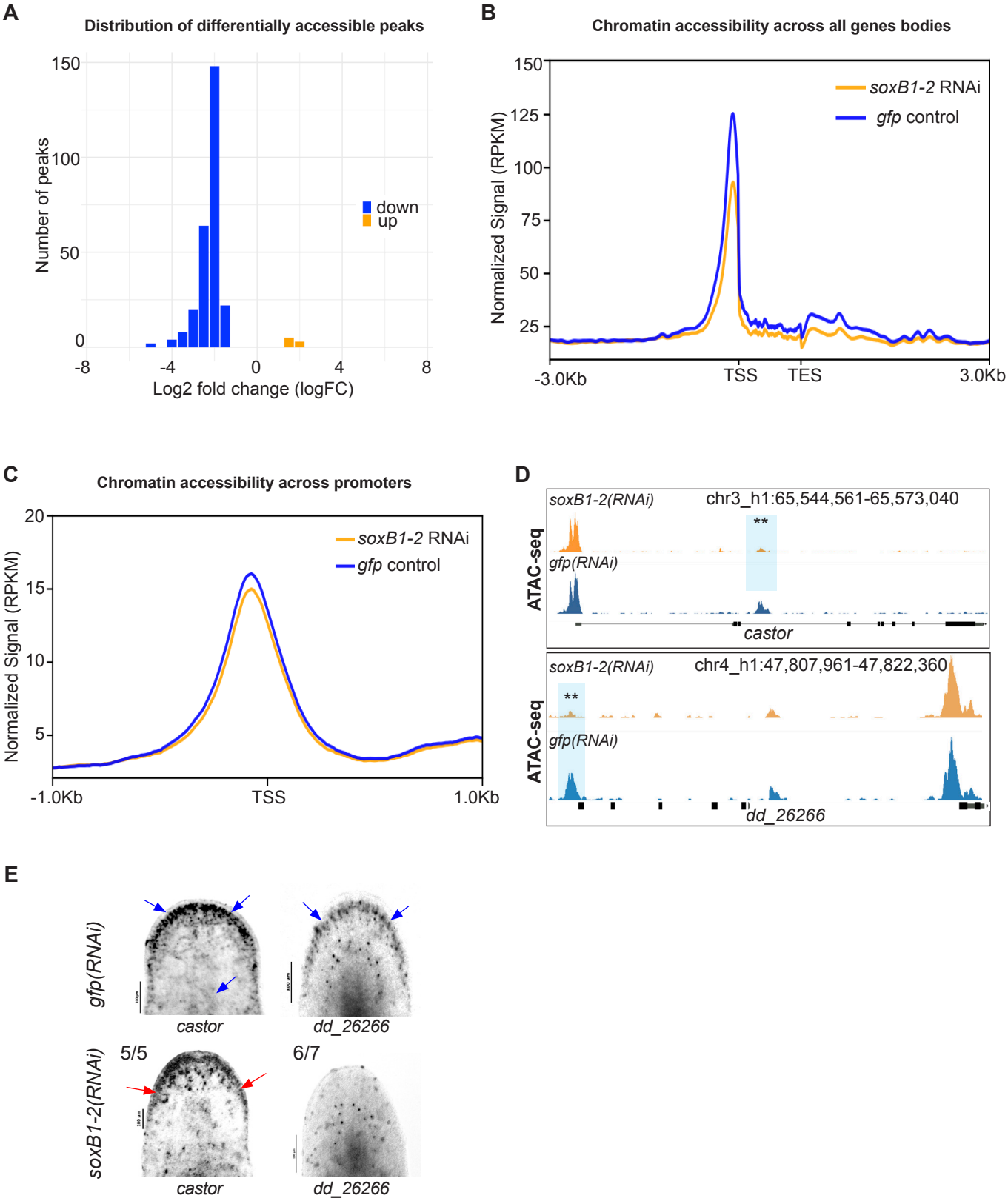

A
Distribution of differentially accessible peaks
Number of peaks
150
100
50
0
-8
-4
0
4
8
Log2 fold change (logFC)
down
up
B
Chromatin accessibility across all genes bodies
Normalized Signal (RPKM)
150
125
100
75
50
25
-3.0Kb
TSS
TES
3.0Kb
soxB1-2 RNAi
gfp control
C
Chromatin accessibility across promoters
Normalized Signal (RPKM)
20
15
10
5
-1.0Kb
TSS
1.0Kb
soxB1-2 RNAi
gfp control
D
soxB1-2(RNAi)
chr3_h1:65,544,561-65,573,040
**
gfp(RNAi)
ATAC-seq
castor
soxB1-2(RNAi)
chr4_h1:47,807,961-47,822,360
**
gfp(RNAi)
ATAC-seq
dd_26266
E
gfp(RNAi)
castor
dd_26266
5/5
6/7
soxB1-2(RNAi)
castor
dd_26266

FIGURE 2

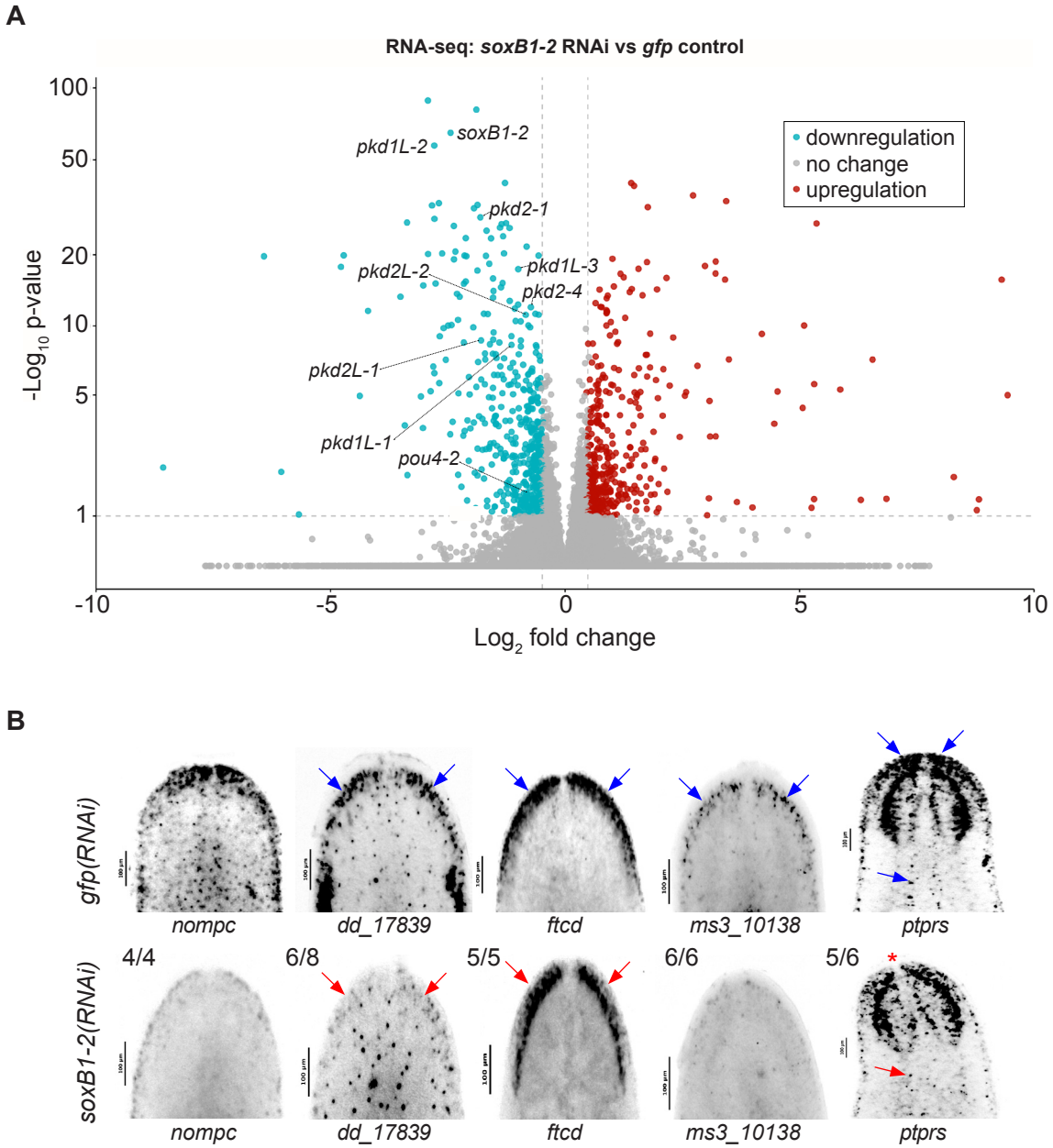

A
RNA-seq: soxB1-2 RNAi vs gfp control
downregulation
no change
upregulation
soxB1-2
pkd1L-2
pkd2-1
pkd1L-3
pkd2-4
pkd2L-2
pkd2L-1
pkd1L-1
pou4-2
-Log10 p-value
Log2 fold change
B
gfp(RNAi)
soxB1-2(RNAi)
nompc
dd_17839
ftcd
ms3_10138
ptprs
4/4
6/8
5/5
6/6
5/6

FIGURE 3

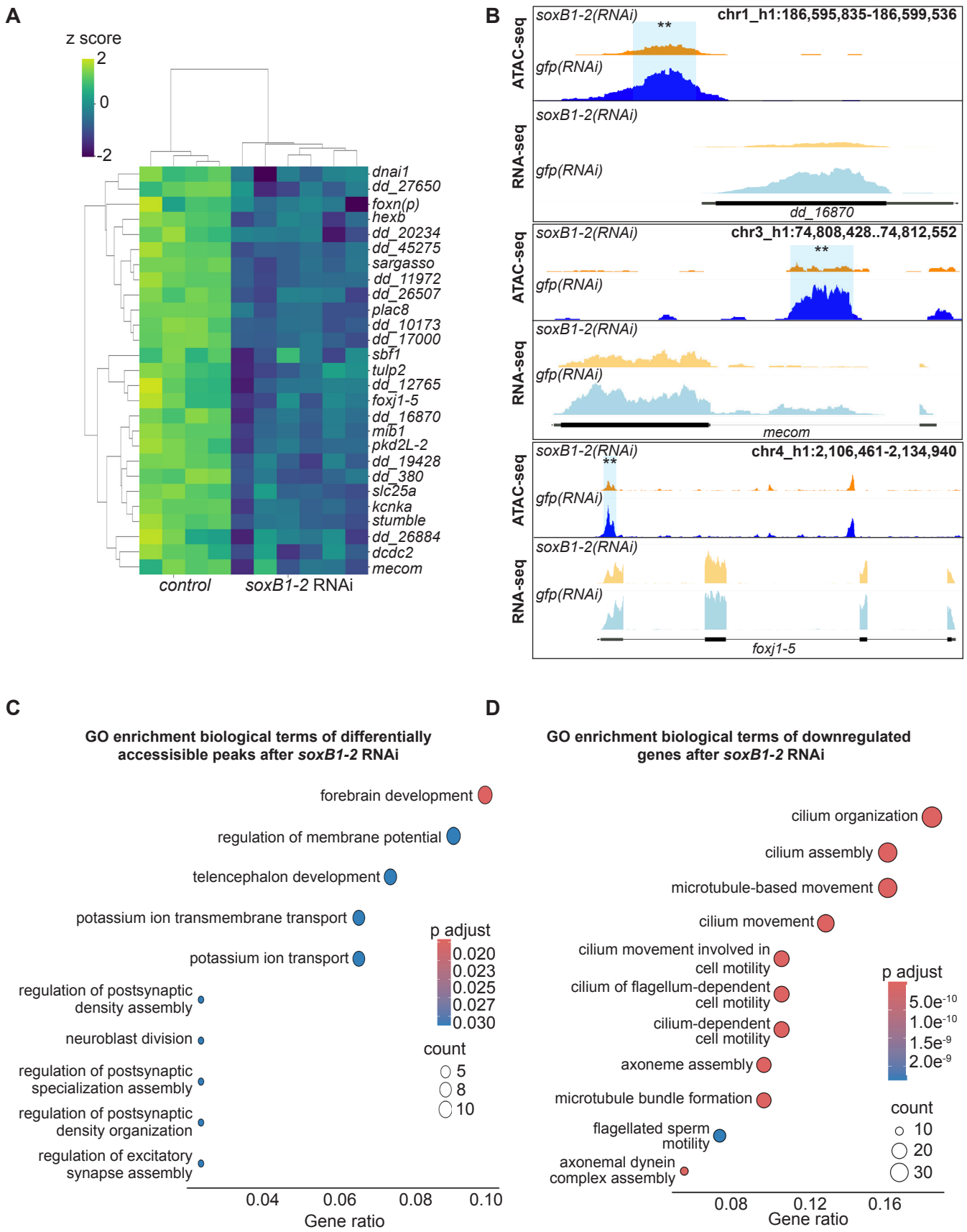

A
z score
2
0
-2
dnai1
dd_27650
foxn(p)
hexb
dd_20234
dd_45275
sargasso
dd_11972
dd_26507
plac8
dd_10173
dd_17000
sbf1
tulp2
dd_12765
foxj1-5
dd_16870
mib1
pkd2L-2
dd_19428
dd_380
slc25a
kcnka
stumble
dd_26884
dcdc2
mecom
control
soxB1-2 RNAi
B
ATAC-seq
RNA-seq
soxB1-2(RNAi)
gfp(RNAi)
chr1_h1:186,595,835-186,599,536
dd_16870
chr3_h1:74,808,428..74,812,552
mecom
chr4_h1:2,106,461-2,134,940
foxj1-5
C
GO enrichment biological terms of differentially accessible peaks after soxB1-2 RNAi
forebrain development
regulation of membrane potential
telencephalon development
potassium ion transmembrane transport
potassium ion transport
regulation of postsynaptic density assembly
neuroblast division
regulation of postsynaptic specialization assembly
regulation of postsynaptic density organization
regulation of excitatory synapse assembly
p adjust
0.020
0.023
0.025
0.027
0.030
count
5
8
10
0.04
0.06
0.08
0.10
Gene ratio
D
GO enrichment biological terms of downregulated genes after soxB1-2 RNAi
cilium organization
cilium assembly
microtubule-based movement
cilium movement
cilium movement involved in cell motility
cilium of flagellum-dependent cell motility
cilium-dependent cell motility
axoneme assembly
microtubule bundle formation
flagellated sperm motility
axonemal dynein complex assembly
p adjust
5.0e-10
1.0e-10
1.5e-9
2.0e-9
count
10
20
30
0.08
0.12
0.16
Gene ratio

FIGURE 4

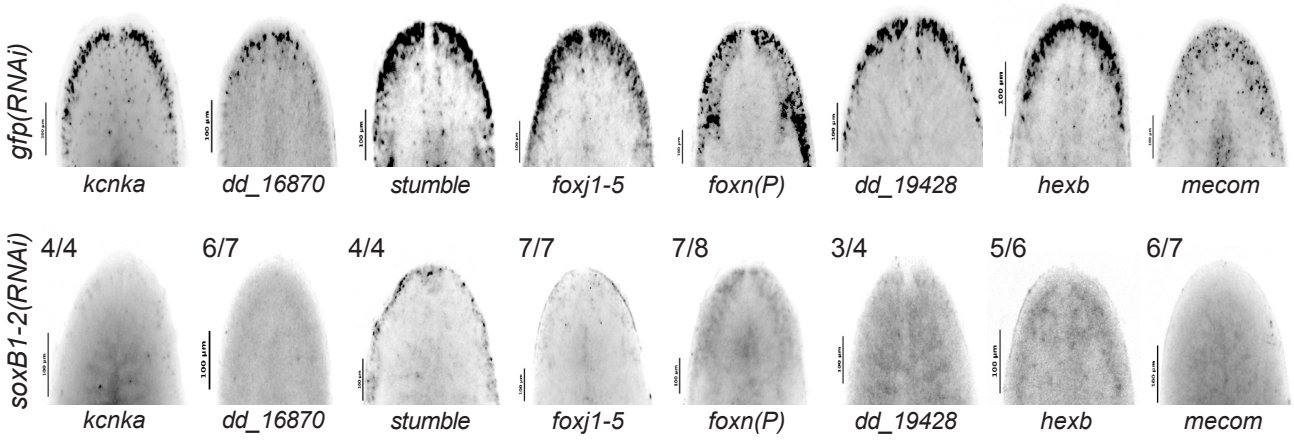
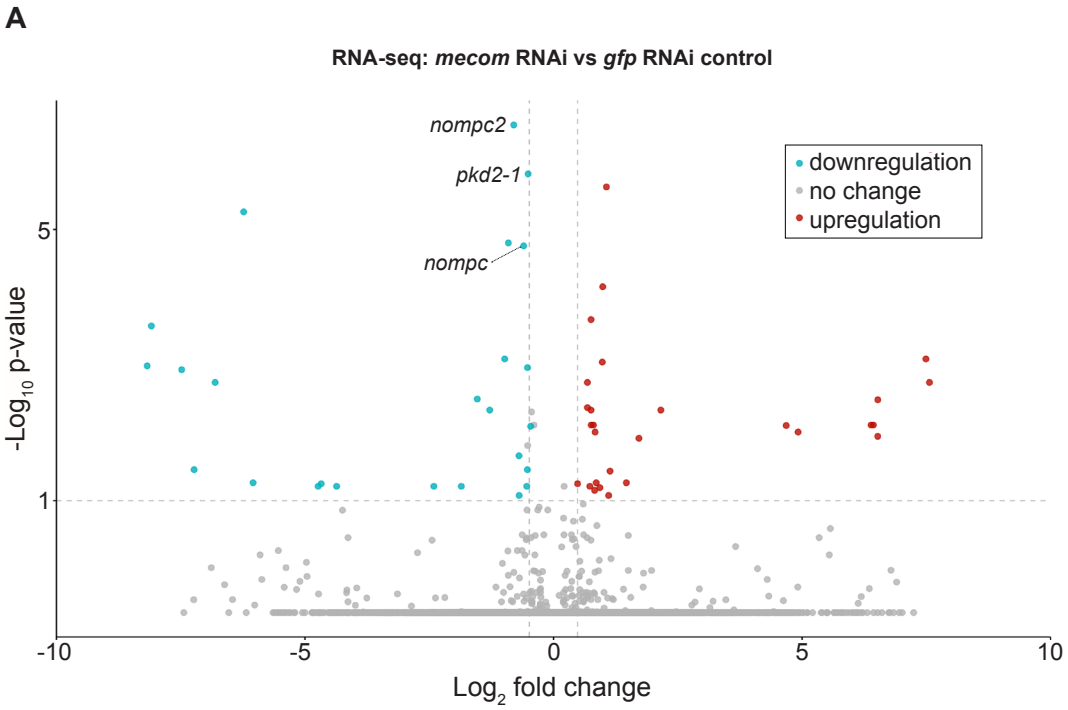

gfp(RNAi)
kcnka
dd_16870
stumble
foxj1-5
foxn(P)
dd_19428
hexb
mecom
soxB1-2(RNAi)
4/4
6/7
4/4
7/7
7/8
3/4
5/6
6/7
kcnka
dd_16870
stumble
foxj1-5
foxn(P)
dd_19428
hexb
mecom

FIGURE 5


A
RNA-seq: mecom RNAi vs gfp RNAi control
nompc2
pkd2-1
nompc
downregulation
no change
upregulation
-Log10 p-value
5
1
-10
-5
0
5
10
Log2 fold change

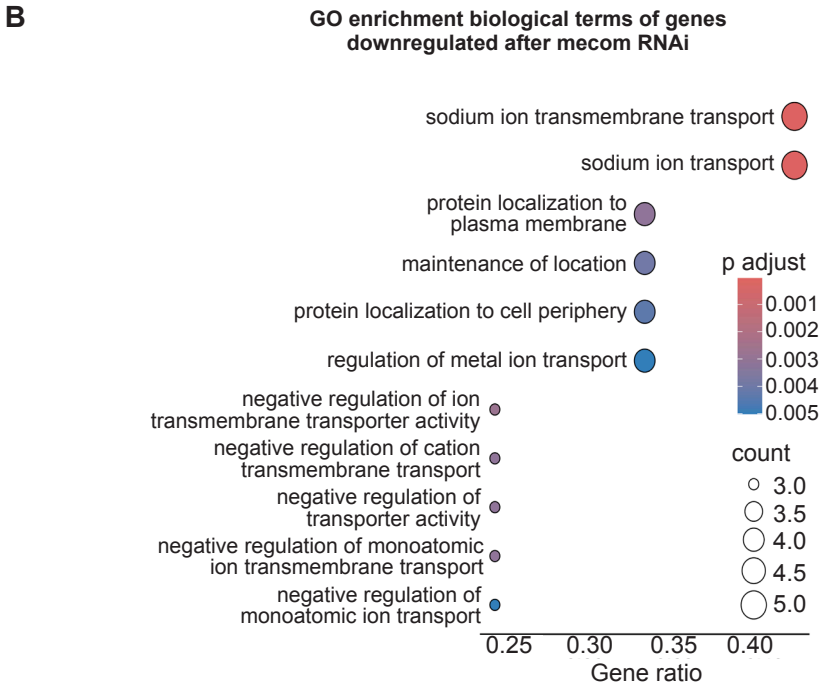
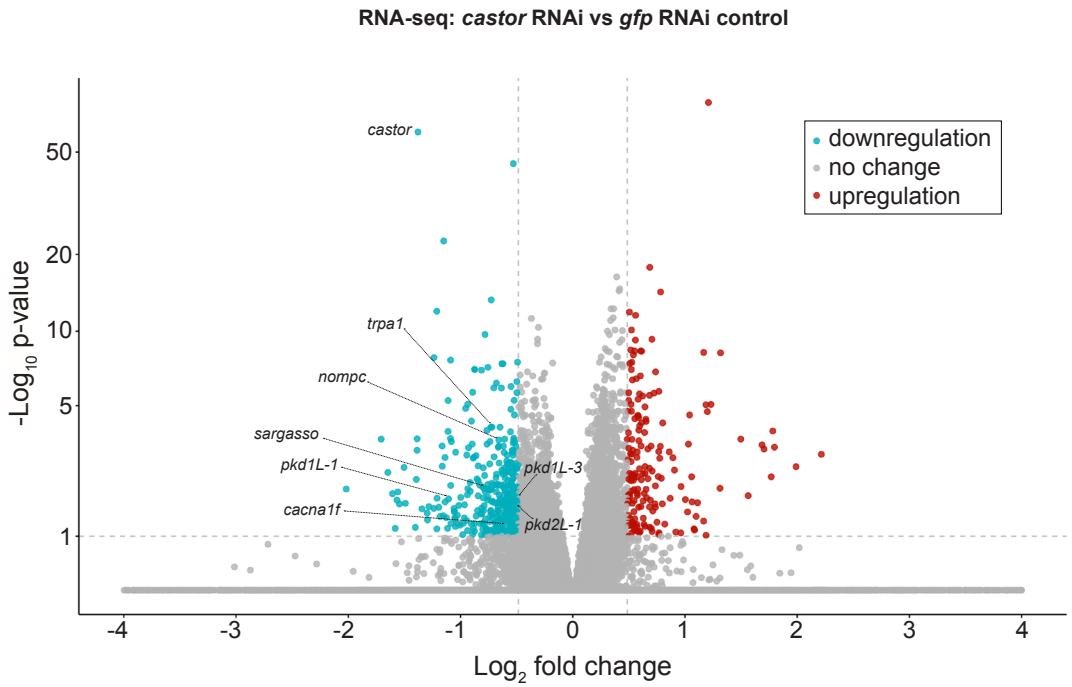

B
GO enrichment biological terms of genes downregulated after mecom RNAi
sodium ion transmembrane transport
sodium ion transport
protein localization to plasma membrane
maintenance of location
protein localization to cell periphery
regulation of metal ion transport
negative regulation of ion transmembrane transporter activity
negative regulation of cation transmembrane transport
negative regulation of transporter activity
negative regulation of monoatomic ion transmembrane transport
negative regulation of monoatomic ion transport
p adjust
0.001
0.002
0.003
0.004
0.005
count
3.0
3.5
4.0
4.5
5.0
0.25
0.30
0.35
0.40
Gene ratio

FIGURE 6

A


RNA-seq: *castor* RNAi vs *gfp* RNAi control
downregulation
no change
upregulation
castor
trpa1
nompc
sargasso
pkd1L-1
cacna1f
pkd1L-3
pkd2L-1
-Log10 p-value
50
20
10
5
1
-4
-3
-2
-1
0
1
2
3
4
Log2 fold change

B

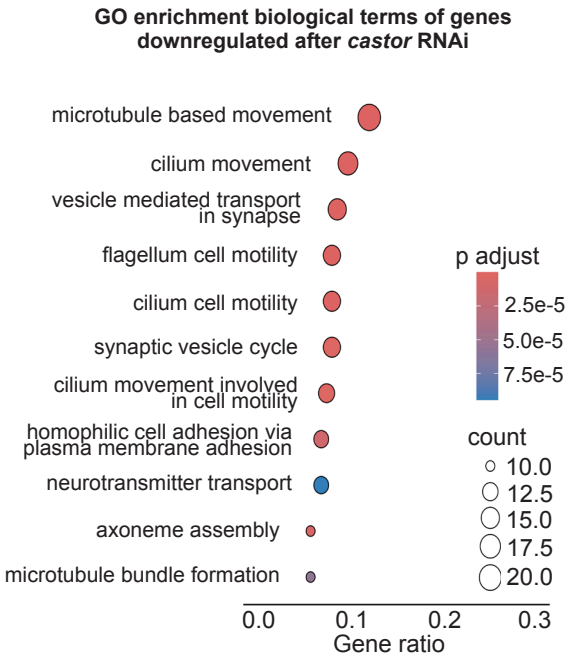
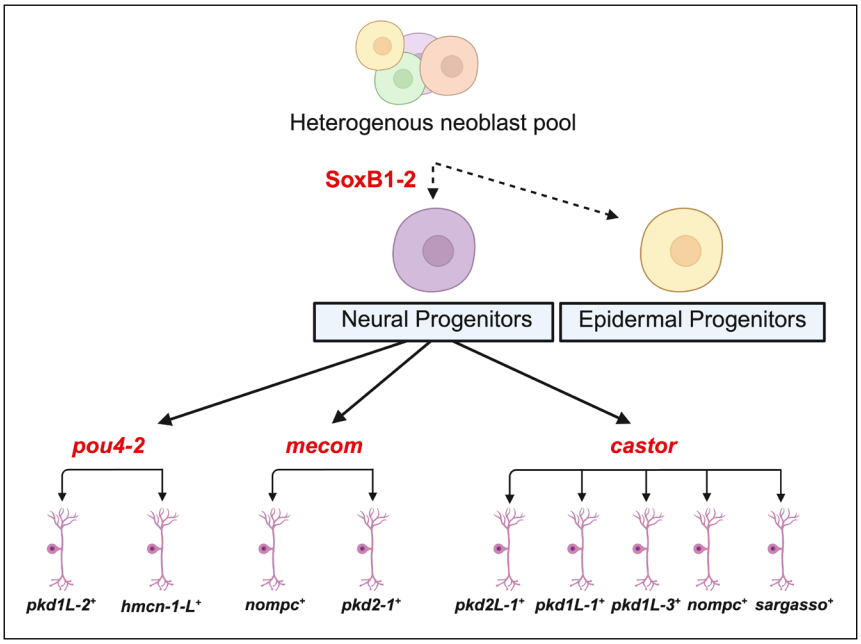
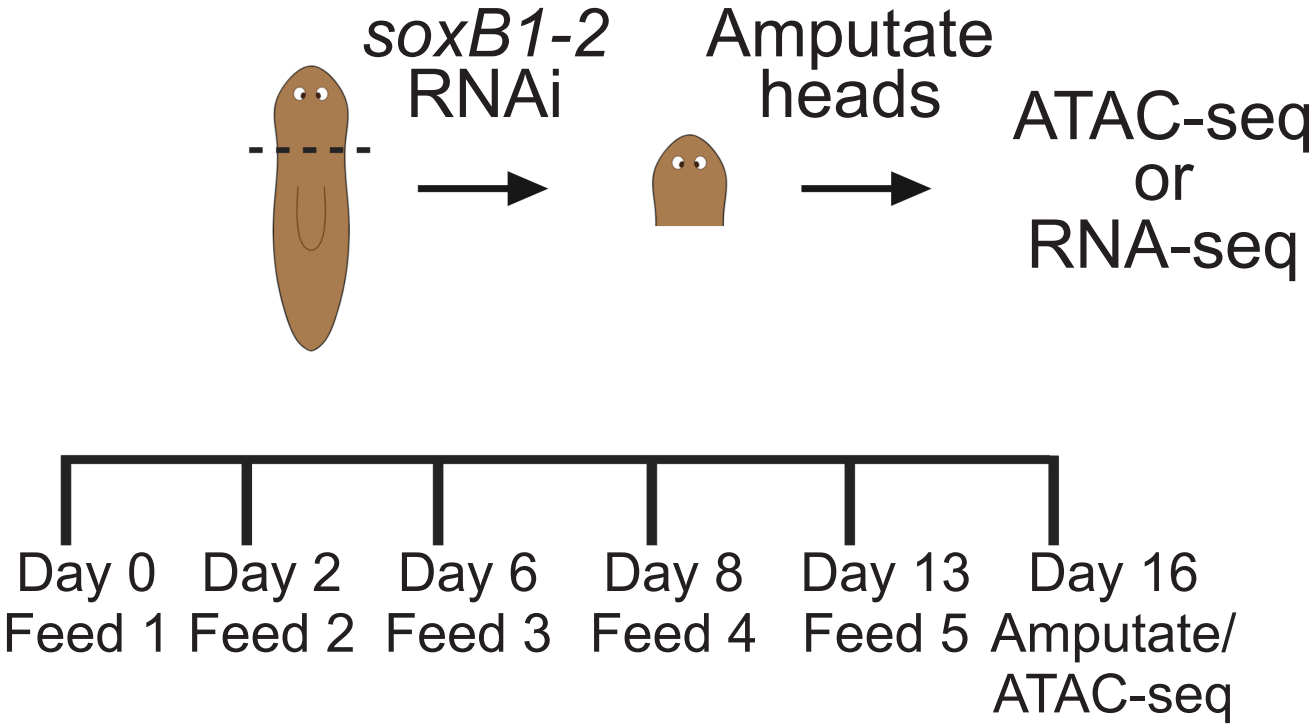

GO enrichment biological terms of genes downregulated after *castor* RNAi
microtubule based movement
cilium movement
vesicle mediated transport in synapse
flagellum cell motility
cilium cell motility
synaptic vesicle cycle
cilium movement involved in cell motility
homophilic cell adhesion via plasma membrane adhesion
neurotransmitter transport
axoneme assembly
microtubule bundle formation
p adjust
2.5e-5
5.0e-5
7.5e-5
count
10.0
12.5
15.0
17.5
20.0
0.0
0.1
0.2
0.3
Gene ratio

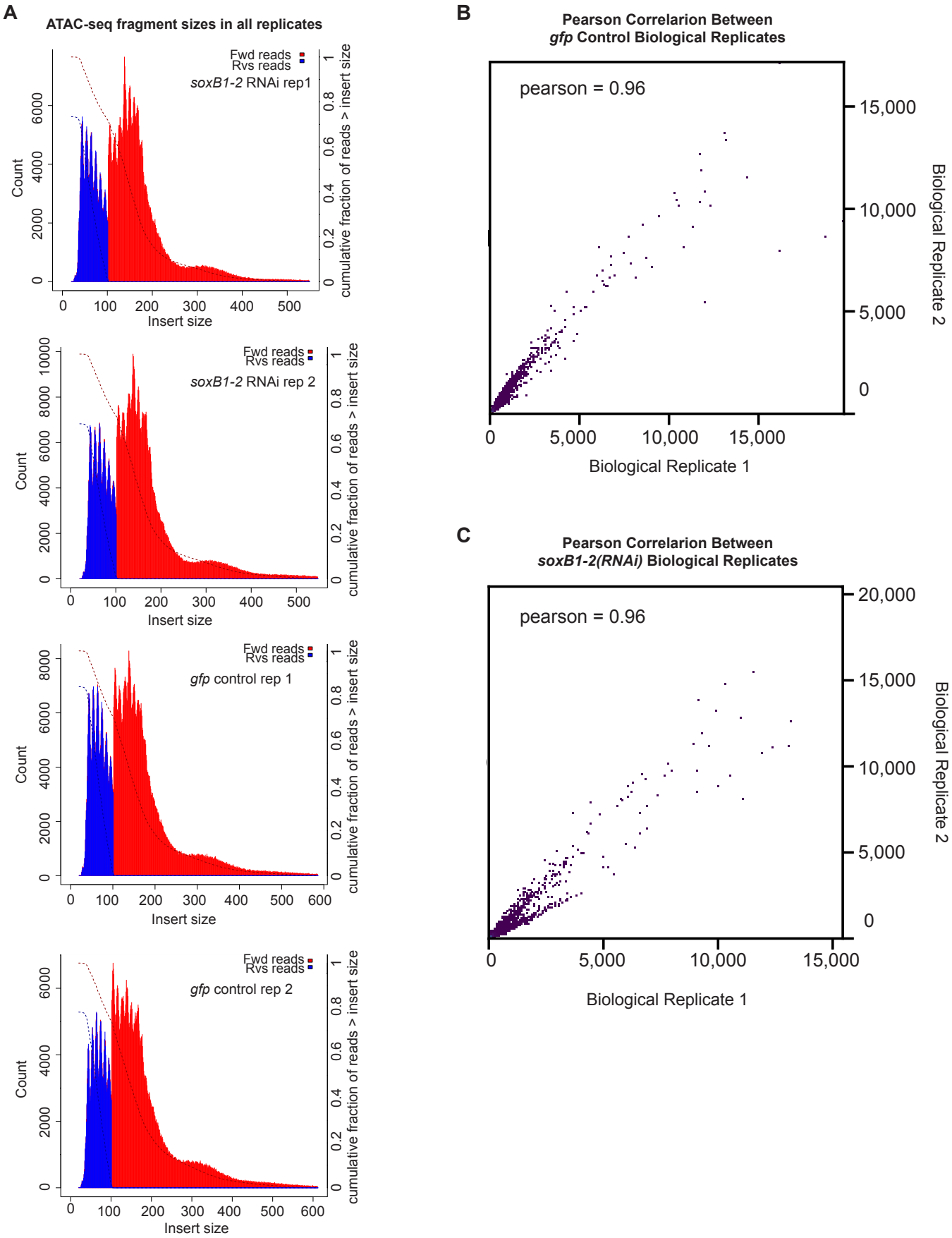
FIGURE 7


Heterogenous neoblast pool
SoxB1-2
Neural Progenitors
Epidermal Progenitors
pou4-2
mecom
castor
pkd1L-2+
hmcn-1-L+
nompc+
pkd2-1+
pkd2L-1+
pkd1L-1+
pkd1L-3+
nompc+
sargasso+

SUPPLEMENTAL FIGURE 1

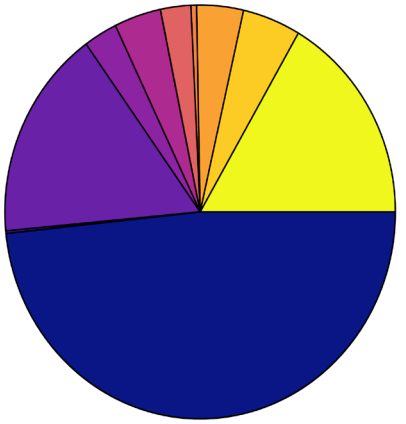


SUPPLEMENTAL FIGURE 2



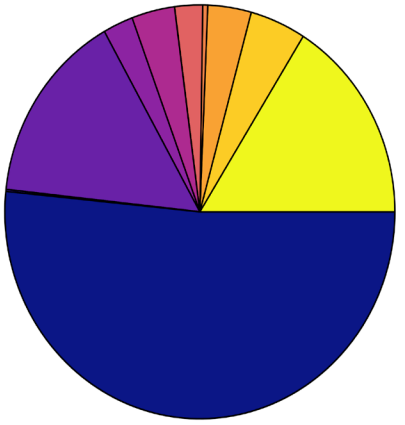
SUPPLEMENTAL FIGURE 3

Proportion of ATAC-seq peaks across the genome



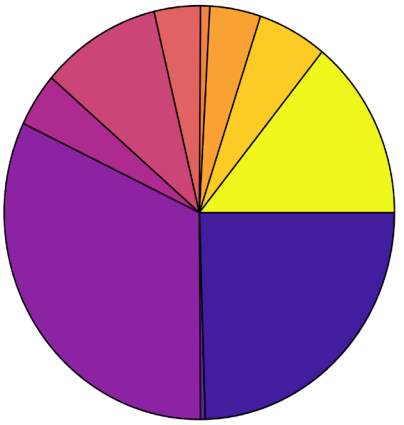
gfp control peaks

- Promoter (<=1kb) (14.66%)
- Promoter (<=1-2kb) (4.11%)
- Promoter (<=2-3kb) (3.23%)
- 5' UTR (0.47%)
- 3' UTR (2.51%)
- 1st Exon (0.01%)
- Other Exon (2.99%)
- 1st Intron (2.39%)
- Other Intron (15.14%)
- Downstream (<=300) (0.25%)
- Distal Iintergenic (54.7%)



soxB1-2 RNAi peaks

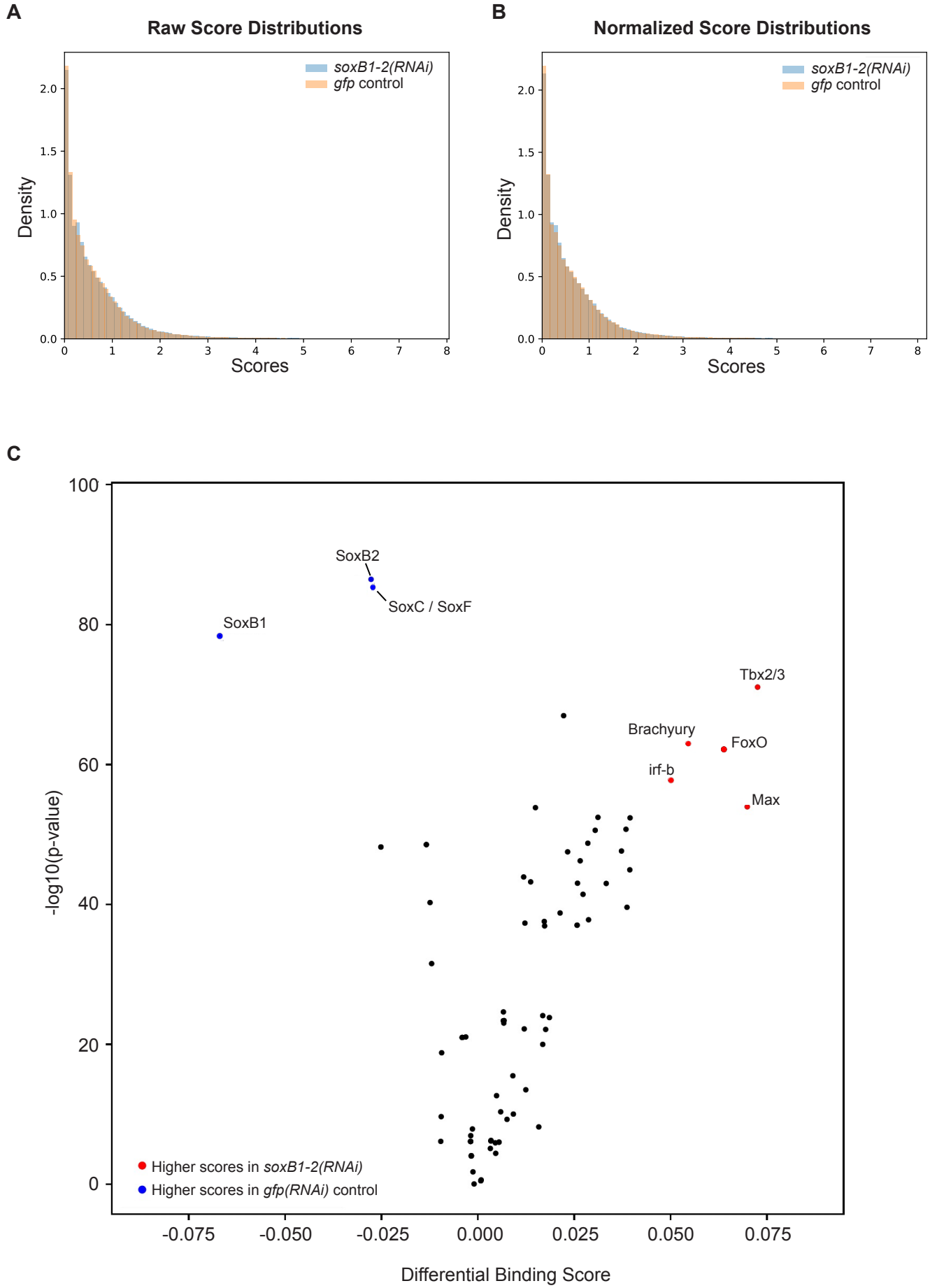
- Promoter (<=1kb) (14.15%)
- Promoter (<=1-2kb) (3.94%)
- Promoter (<=2-3kb) (3.16%)
- 5' UTR (0.27%)
- 3' UTR (1.96%)
- 1st Exon (0.02%)
- Other Exon (2.54%)
- 1st Intron (2.23%)
- Other Intron (14%)
- Downstream (<=300) (0.18%)
- Distal Iintergenic (57.54%)



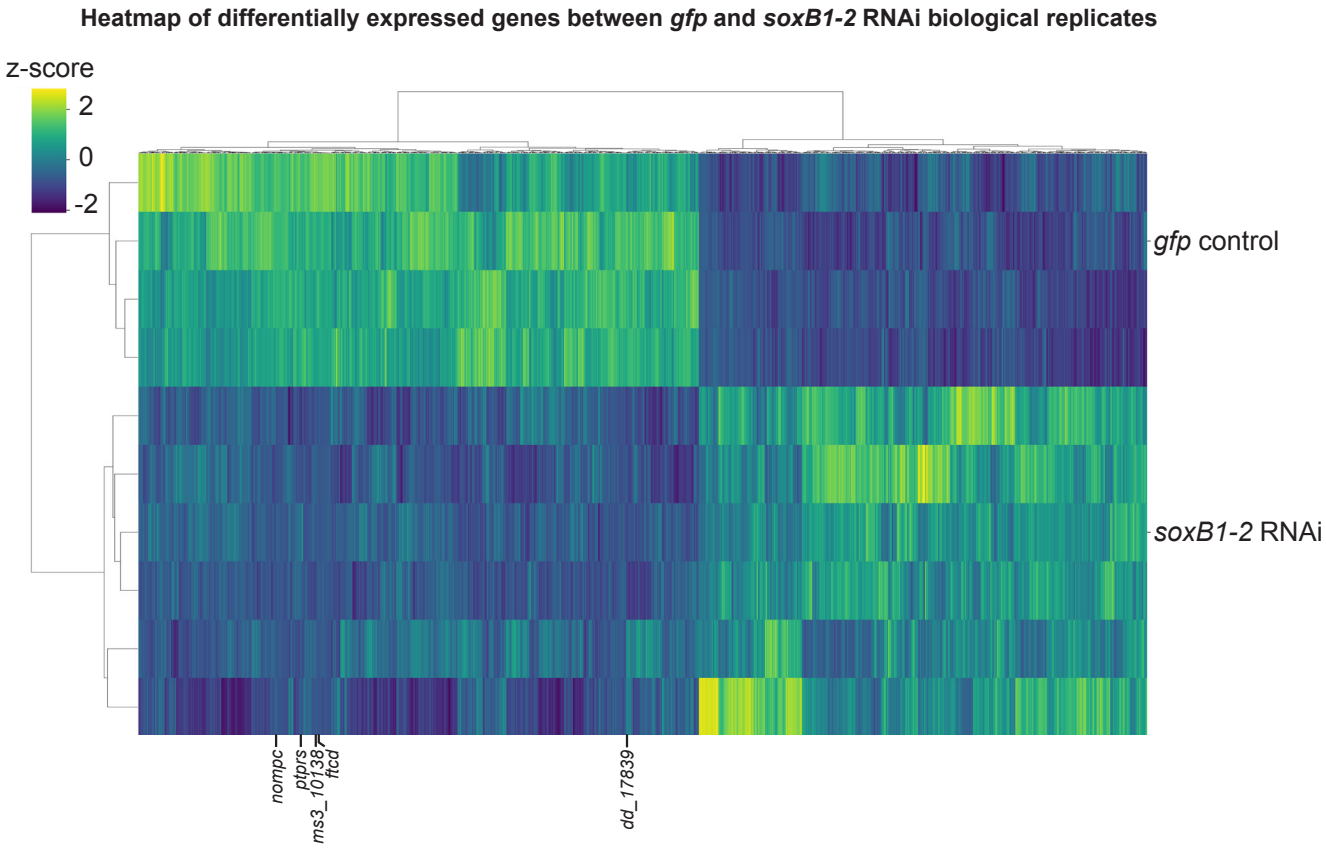
Differentially accessible peaks

- Promoter (<=1kb) (14.18%)
- Promoter (<=1-2kb) 5.75%)
- Promoter (<=2-3kb) (4.21%)
- 5' UTR (0.77%)
- 3' UTR (3.83%)
- Other Exon (9.96%)
- 1st Intron (4.21%)
- Other Intron (32.18%)
- Downstream (<=300) (0.38%)
- Distal Iintergenic (24.52%)

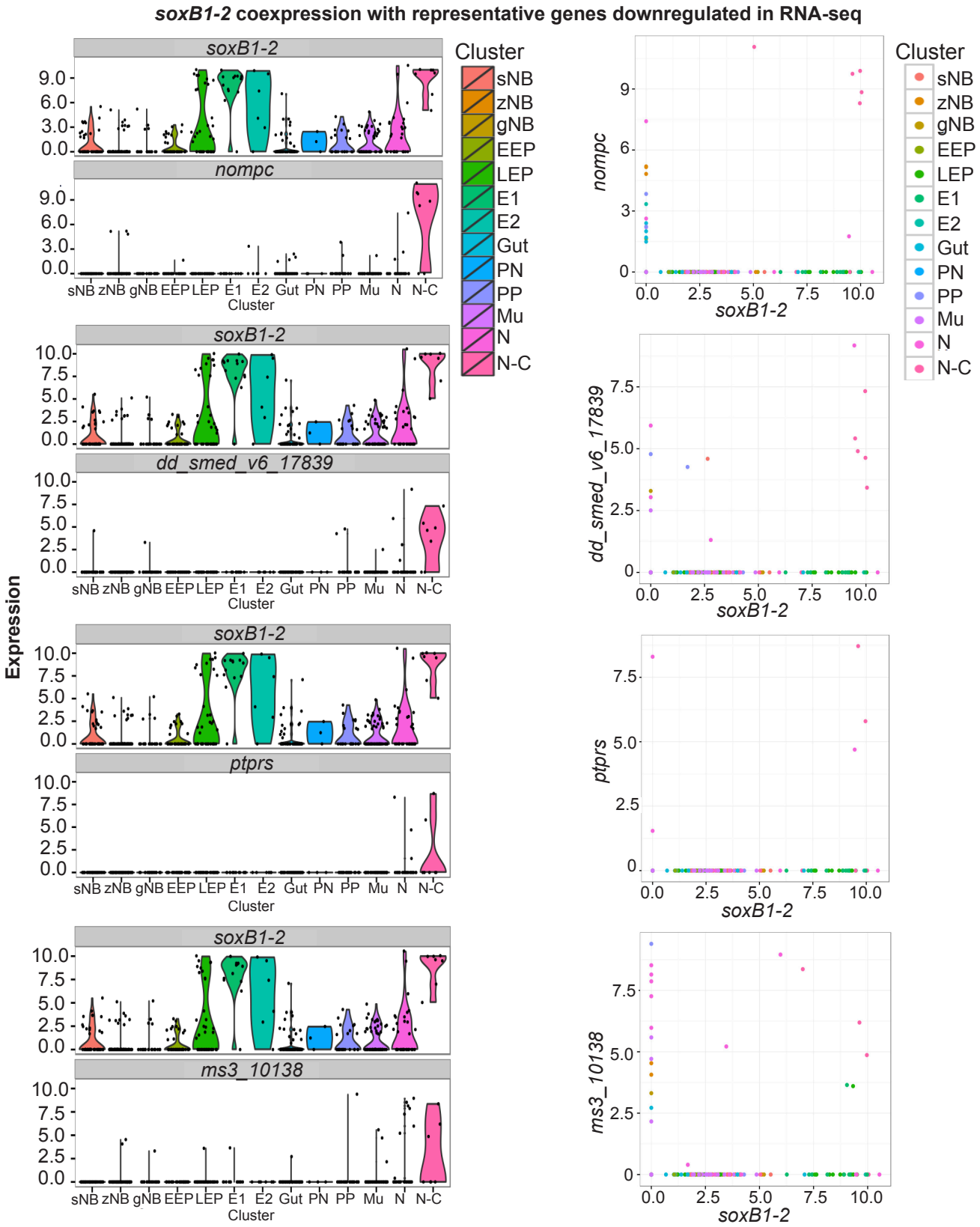
SUPPLEMENTAL FIGURE 4



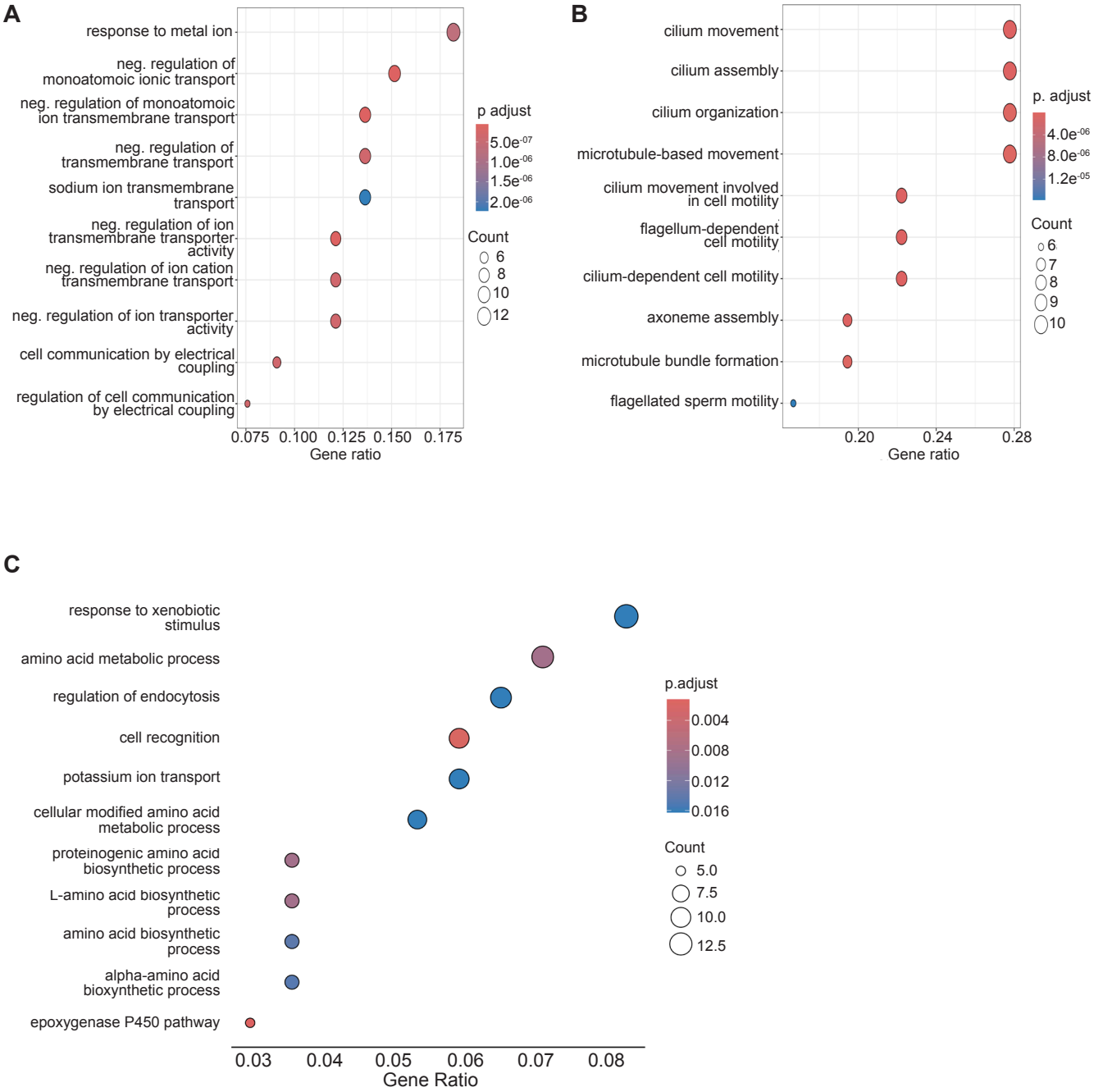
SUPPLEMENTAL FIGURE 5



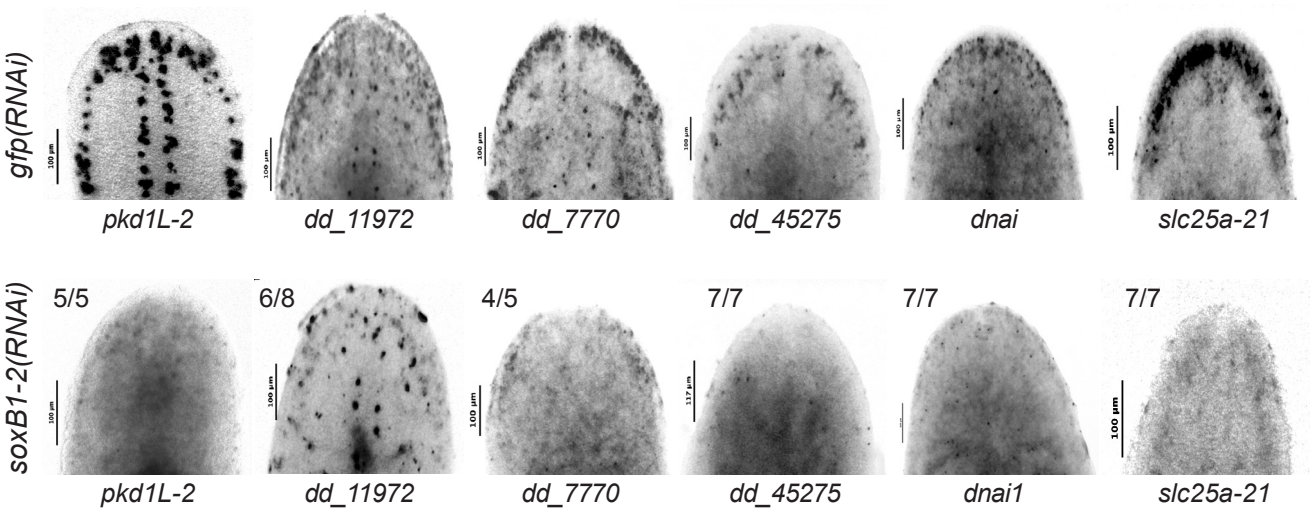
SUPPLEMENTAL FIGURE 6



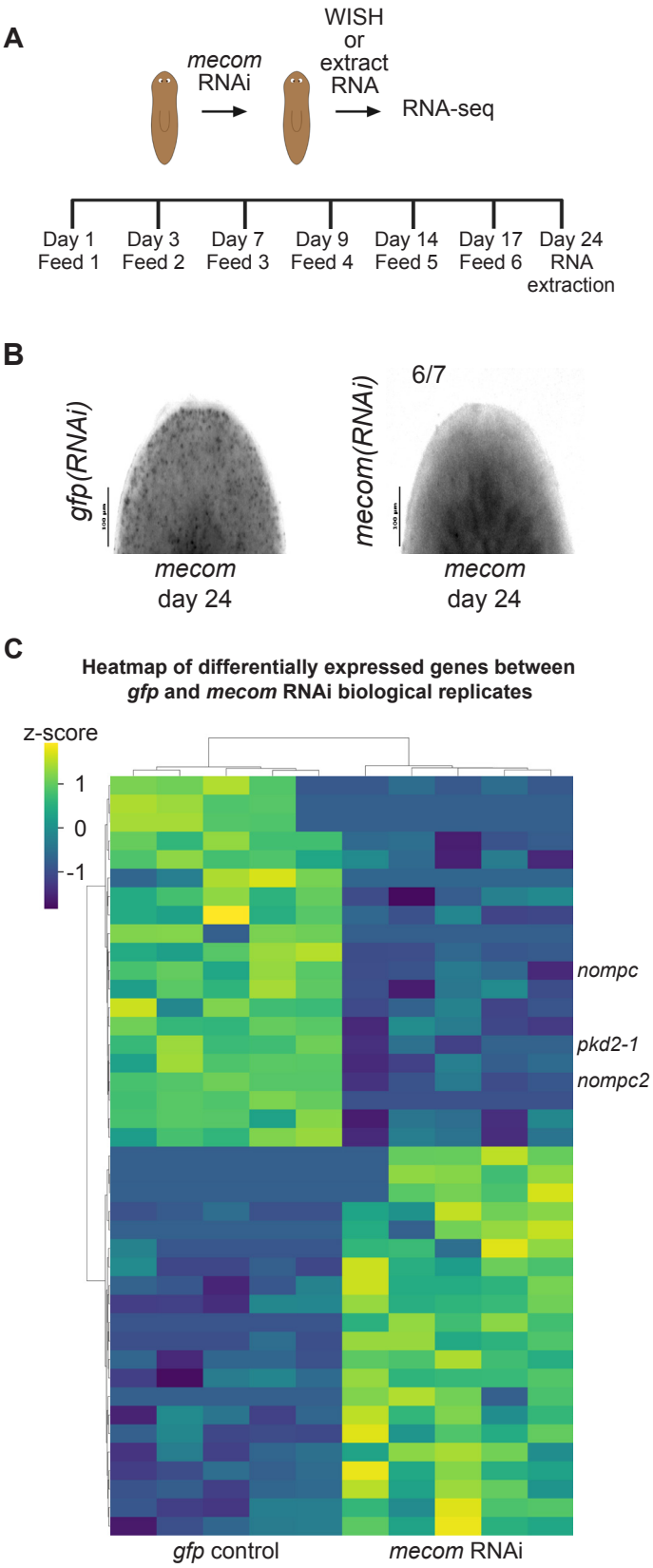
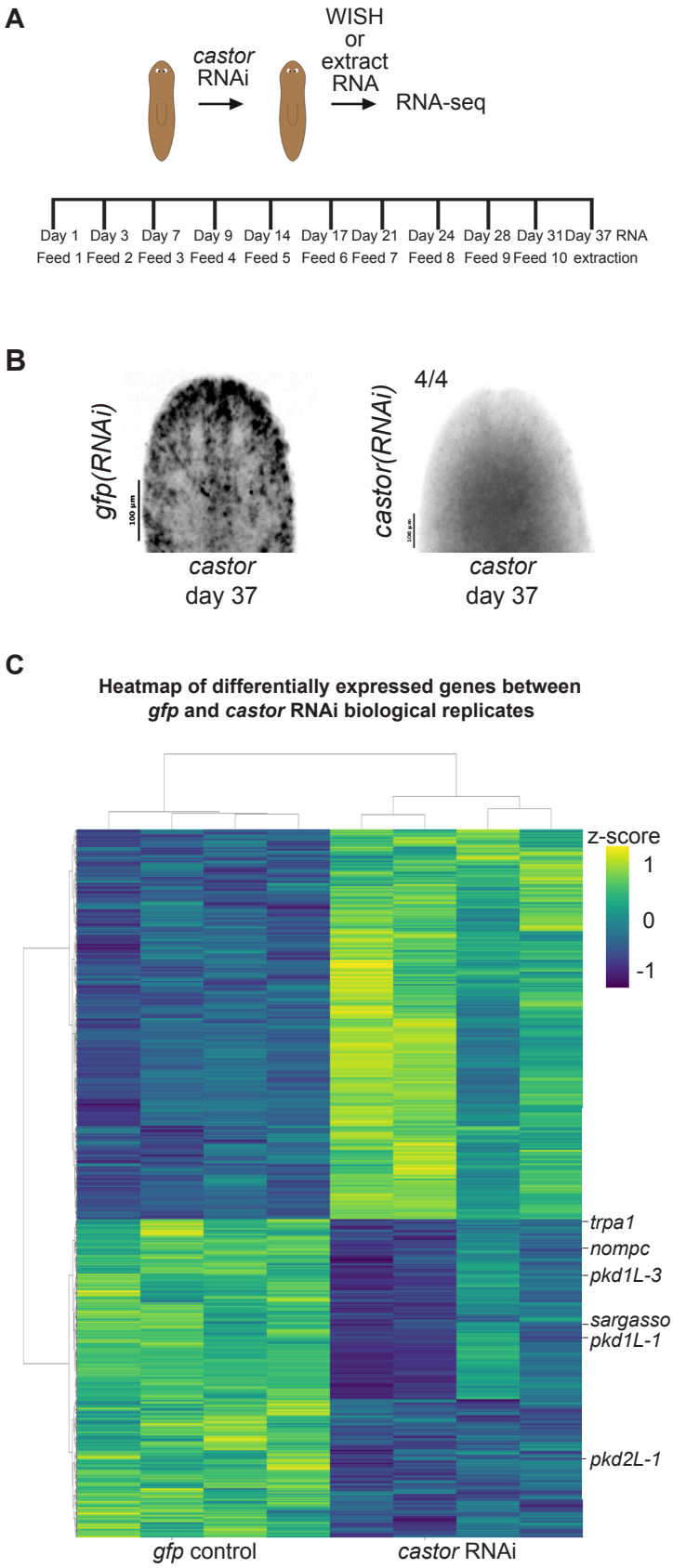
SUPPLEMENTAL FIGURE 7



SUPPLEMENTAL FIGURE 8



SUPPLEMENTAL FIGURE 9



SUPPLEMENTAL FIGURE 10