



# Co-occupancy identifies transcription factor co-operation for axon growth

Ishwariya Venkatesh <sup>1,2</sup>✉, Vatsal Mehra<sup>1</sup>, Zimei Wang<sup>1</sup>, Matthew T. Simpson<sup>1</sup>, Erik Eastwood<sup>1</sup>, Advaita Chakraborty<sup>1</sup>, Zac Beine<sup>1</sup>, Derek Gross<sup>1</sup>, Michael Cabahug<sup>1</sup>, Greta Olson<sup>1</sup> & Murray G. Blackmore <sup>1,2</sup>✉

Transcription factors (TFs) act as powerful levers to regulate neural physiology and can be targeted to improve cellular responses to injury or disease. Because TFs often depend on cooperative activity, a major challenge is to identify and deploy optimal sets. Here we developed a bioinformatics pipeline, centered on TF co-occupancy of regulatory DNA, and used it to predict factors that potentiate the effects of pro-regenerative Klf6 in vitro. High content screens of neurite outgrowth identified cooperative activity by 12 candidates, and systematic testing in a mouse model of corticospinal tract (CST) damage substantiated three novel instances of pairwise cooperation. Combined Klf6 and Nr5a2 drove the strongest growth, and transcriptional profiling of CST neurons identified Klf6/Nr5a2-responsive gene networks involved in macromolecule biosynthesis and DNA repair. These data identify TF combinations that promote enhanced CST growth, clarify the transcriptional correlates, and provide a bioinformatics approach to detect TF cooperation.

<sup>1</sup>Department of Biomedical Sciences, Marquette University, Milwaukee, WI, USA. <sup>2</sup>These authors contributed equally: Ishwariya Venkatesh, Murray G. Blackmore. ✉email: [ishwariya.venkatesh@marquette.edu](mailto:ishwariya.venkatesh@marquette.edu); [murray.blackmore@marquette.edu](mailto:murray.blackmore@marquette.edu)

As they mature, neurons in the central nervous system (CNS) decline in their capacity for robust axon growth, which broadly limits recovery from injury<sup>1–5</sup>. Axon growth depends on the transcription of large networks of regeneration-associated genes (RAGs) and coaxing activation of these networks in adult CNS neurons is a major unmet goal<sup>2,4</sup>. During periods of developmental axon growth, RAG expression is supported by pro-growth transcription factors (TFs) that bind to relevant promoter and enhancer regions, to activate transcription<sup>3,6</sup>. One factor that limits RAG expression in mature neurons is the developmental downregulation of these pro-growth TFs<sup>2,4</sup>. Thus, identifying TFs that act developmentally to enable axon growth and supplying them to mature neurons is a promising approach to improve regenerative outcomes in the injured nervous system.

An ongoing challenge, however, is to decode the optimal set of factors for axon growth. To date, progress has centered on identifying individual TFs whose ectopic expression in mature neurons leads to improved axon growth. For example, we showed previously that the TF Klf6 is developmentally downregulated, and that viral re-expression drives improved axon growth in mature corticospinal neurons, which are critical mediators of fine movement<sup>7</sup>. The restoration of axon growth from this and other single-factor studies remains partial, however, indicating the need for additional intervention<sup>7–12</sup>. One likely possibility is that, because multiple TFs generally act in a coordinated manner to regulate transcription, a more complete restoration of axon growth will depend on multiple TFs. Consistent with this, regeneration-competent neurons recruit groups of interacting TFs in the hours to days post injury, which results in transcriptional remodeling, leading to reactivation of growth gene networks and functional recovery<sup>13–15</sup>. The need for multi-TF interventions in CNS neurons is recognized conceptually<sup>4,16</sup>, but a major challenge has been to develop a systematic pipeline aimed at discovering growth-relevant TF combinations.

Here we developed a bioinformatics framework to detect cooperative TF promotion of axon growth. The framework centers on the concept of TF co-occupancy, in which functional interactions between TFs are detected by virtue of shared binding to common sets of regulatory DNA. We deployed the pipeline to predict TFs that potentially synergize with Klf6 to drive enhanced axon outgrowth. In vitro phenotypic screening confirmed cooperative TF activity for 12 candidates and 2 independent bioinformatic analyses converged to prioritize interest in a core of 3 factors, nuclear receptor subfamily 5, group A, member 2 (Nr5a2), retinoic acid receptor  $\beta$  (Rarb), and Eomes. Systematic tests of forced TF co-expression in vivo using a pyramidotomy model of axon injury revealed strong and consistent promotion of corticospinal tract (CST) axon growth by combined expression of Klf6 and Nr5a2. Finally, transcriptional profiling of purified CST neurons revealed transcriptional correlates to the evoked growth, notably gene modules related to biosynthesis and DNA repair. Overall, we have identified novel combinations of TFs that drive enhanced axon growth in mature CNS neurons and provide a generalized computational roadmap for the discovery of cooperative activity between TFs with potential application to a wide range of neural activities.

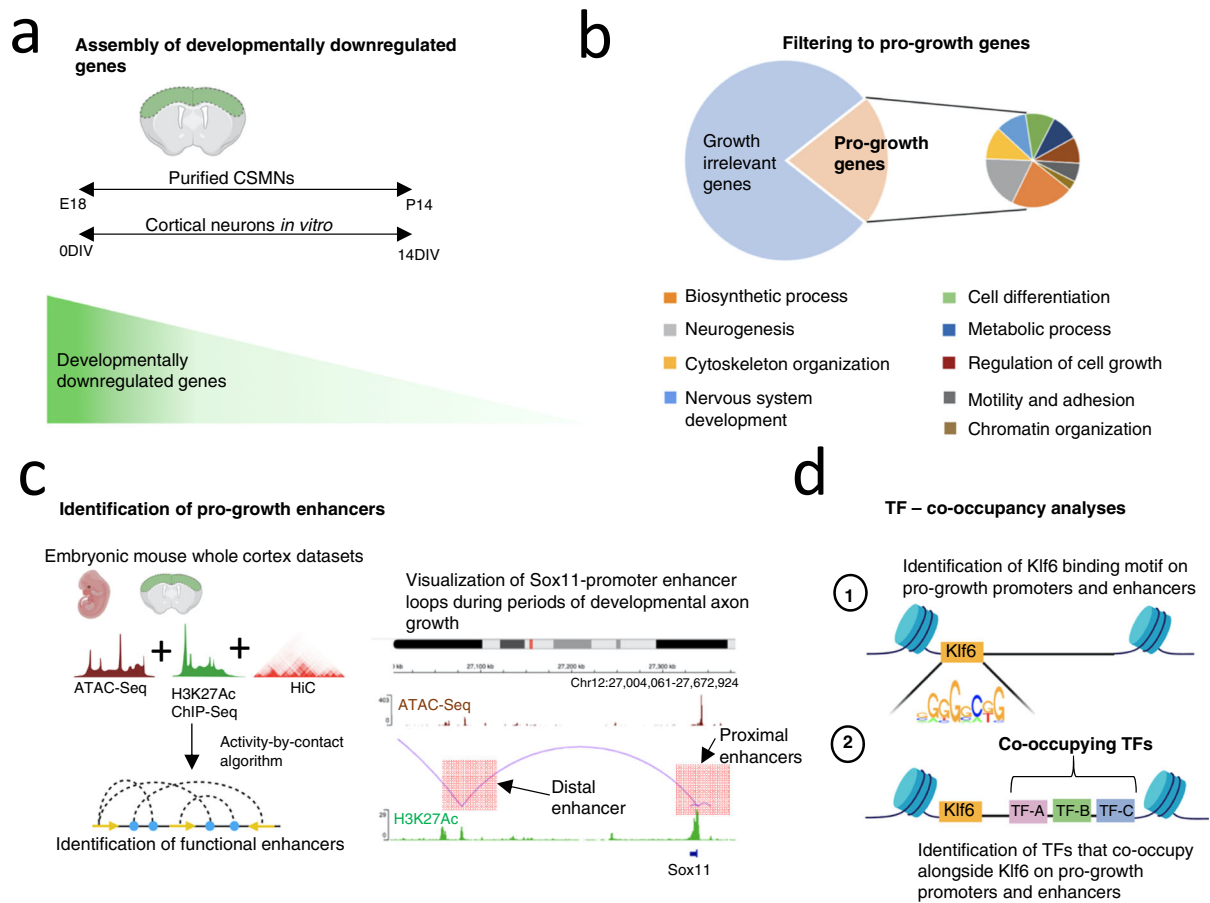
## Results

**In vitro screening identifies TFs that synergize with Klf6 to drive increased neurite outgrowth.** TFs that functionally synergize often do so by binding DNA in close proximity and initiating complementary transcriptional mechanisms<sup>17–20</sup>. Thus, analysis of TF co-occupancy can reveal instances of cooperation between TFs<sup>21,22</sup>. It was shown previously that the TF Klf6 is

expressed in CST neurons during developmental periods of axon growth, is downregulated during postnatal maturation, and that forced re-expression in adult neurons modestly enhances axon growth<sup>7</sup>. We therefore reasoned that during development, additional TFs likely co-occupy regulatory elements alongside Klf6 and cooperate to regulate transcription. We further hypothesized that identifying these co-occupiers and supplying them to adult neurons along with Klf6 may lead to further enhancement of axon growth after injury.

To identify potential Klf6 co-occupiers, we first assembled a list of genes that decline in expression as cortical neurons mature, using data from corticospinal motor neurons (CSMN)s in vivo and cortical neurons ages in vitro<sup>23,24</sup> (Fig. 1a). Genes common to both datasets were selected for further analysis. To tighten the focus on genes that contribute to axon growth, we selected those genes associated with Gene Ontology (GO) terms linked to axon growth, including cytoskeleton organization, microtubule-based transport, nervous system development, and regulation of growth (Fig. 1b). Supplementary Data 1 summarizes the final list of 308 developmentally downregulated pro-growth genes and their associated GO terms. For each gene, promoters were assigned as 1500 bp upstream/300 bp downstream of the transcription start site (TSS). To assign enhancers, we took advantage of recent advances in machine learning-based algorithms, deploying the activity-by-contact (ABC) pipeline to reconstruct functional enhancer–gene pairs during active periods of developmental axon growth<sup>25</sup>. Utilizing open-access ENCODE datasets of chromatin accessibility, H3K27Ac enrichment, HiC, and expression datasets from embryonic forebrain, ABC identified a total of 1230 enhancers linked to the selected growth-relevant genes<sup>26</sup> (Fig. 1c). Each gene was associated with those between 1 and 13 enhancers (mean  $6 \pm 3$  SEM), 80% of which were located 10–1500 kb from the TSS. The full list of enhancer–gene pairs with genomic coordinates is summarized in Supplementary Data 1. Finally, to predict the binding of Klf6 and potential co-occupying TFs, we examined promoters and enhancers using an anchored TF motif analysis algorithm. This algorithm scans sequences for the canonical Klf6-binding motif and then identifies TF-binding motifs that are over-represented in nearby DNA<sup>27</sup>. The result was 62 candidate TFs that were predicted to frequently co-occur with Klf6 in pro-growth promoters, pro-growth enhancers, or both (Fig. 1d).

We next used high content screening to test the prediction that these candidate TFs functionally synergize with Klf6 to enhance neurite outgrowth (Fig. 2a). Using a well-established screening platform, postnatal cortical neurons received candidate genes by plasmid electroporation and were cultured at low density on laminin substrates, followed 2 days later by automated tracing to quantify neurite outgrowth<sup>7,12,24,28–31</sup>. Nuclear-localized enhanced green fluorescent protein (EGFP) served to mark transfected neurons and  $\beta$ III tubulin immunohistochemistry labeled neuronal processes for automated tracing (Fig. 2b). All screening plates included wells with three standard treatments as follows: (1) *EBFP* plasmid alone, (2) *Klf6* mixed with mCherry control, and (3) *Klf6* mixed with *VP16-Stat3*, a combination that was shown previously to elevate neurite length above *Klf6* alone<sup>7</sup>. This design accounts for inter-plate variability by normalizing all lengths to *EBFP*, while establishing on each plate both the *Klf6*-only effect size and the sensitivity to potential increase above that level. As expected, across the screen forced expression of *Klf6* alone increased neurite outgrowth by 28% ( $\pm 4.37\%$  SEM) compared to *EBFP* control, whereas combination with *VP16-Stat3* further elevated lengths to 61% ( $\pm 5.59\%$ ) of *EBFP* ( $p$ -value  $< 0.01$ , analysis of variance (ANOVA) with post hoc Fisher's least significant difference (LSD)). The remainder of each plate was devoted to *Klf6* mixed with candidate TFs, which were delivered



**Fig. 1 Integrated bioinformatics pipeline identifies KLF6 partner TFs likely involved in the regulation of axon growth.** **a** Time-course gene expression datasets from mouse cortical neurons were integrated to isolate genes that are downregulated across cortical maturation in vivo and in vitro (CSMNs, corticospinal motor neurons). **b** Gene Ontology analyses delineated developmentally downregulated genes that are likely growth relevant, based on functional enrichment in GO categories pertinent to axon growth. **c** Activity-by-contact (ABC), an enhancer–target gene pairing algorithm, defined specific enhancer regions genomewide, which modulate the expression of growth-relevant transcripts. Genome tracks are loaded to visualize the proximal and distal enhancer–promoter loops involved in the regulation of the pro-growth gene *Sox11* during periods of developmental axon growth. ATAC-Seq profiles are shown in magenta and H3K27Ac profiles are shown in green. **d** TF co-occupancy analyses predicted TFs that co-occupy regulatory DNA alongside Klf6 to regulate the expression of growth-relevant transcripts.

in their native form, not modified with VP16, to probe endogenous activity. Each combination was tested in three separate experiments and a minimum of 150 individual cells quantified per experiment. Notably, of the 62 candidate TFs, 12 significantly elevated neurite length above the level of Klf6 alone (“hard hits”) ( $p$ -value < 0.05, ANOVA with post hoc Fisher’s LSD) and an additional five missed statistical significance but improved growth in two of three replicate experiments (“soft hits”) (Fig. 2a). Detailed screening data are available in Supplementary Data 2 and Supplementary Fig. 1 provides data for the effects of TFs on additional morphological parameters. Interestingly, all 12 TFs that elevated neurite length came from the one-third of candidates that were predicted to co-occupy with Klf6 in both promoters and enhancers, and no hits came from TFs predicted to co-occupy solely in promoters or in enhancers (Fig. 2c). These data support the ability of co-occupancy analysis to predict functional interactions between TFs and emphasize the importance of scanning both the promoter and enhancer elements.

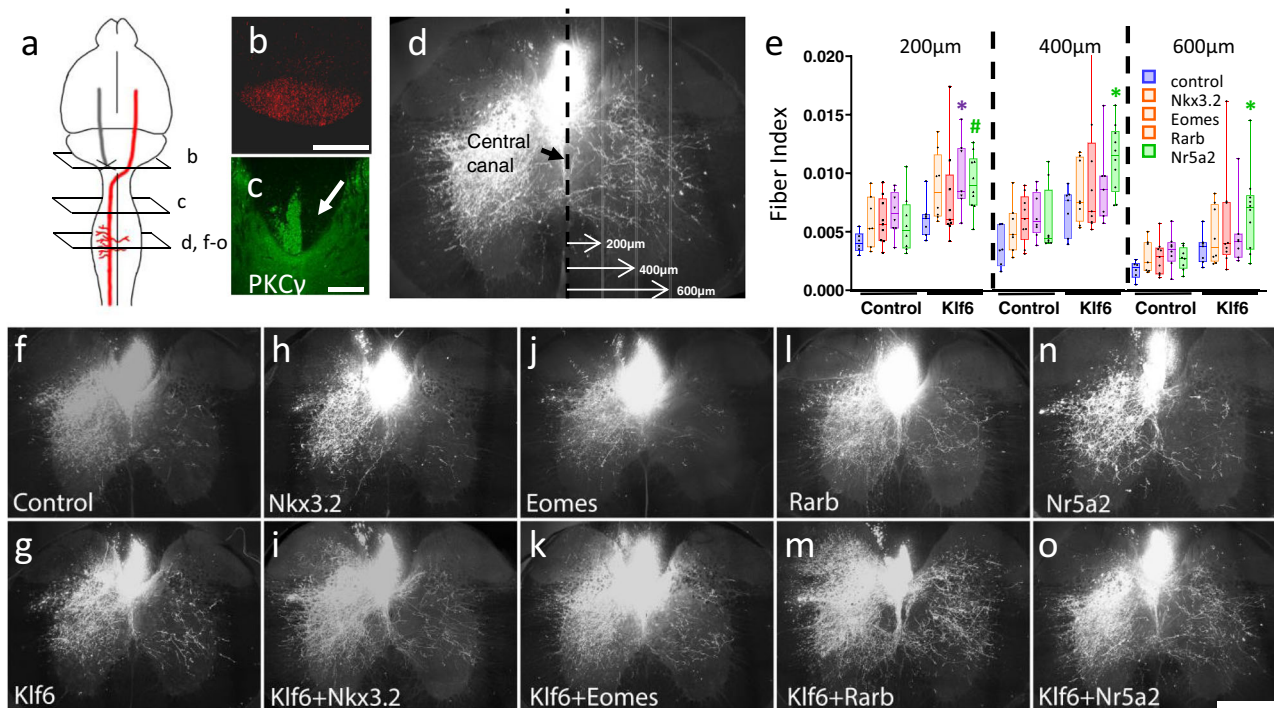
We then used network analyses to prioritize interest in the 12 “hard hit” TFs. First, using open-access TF-binding data, top transcriptional targets of the 12 hit TFs were used to generate individual TF target gene networks. These networks were then

merged, expanded to bring in immediately connected genes based on known interactions, and restructured to move genes with maximal connections to the center and minimal connections to the periphery. This resulted in a unified network with a core of TFs predicted to exert maximal influence, surrounded by shells of TFs with progressively lower connectivity (Fig. 2c). Notably, during the expansion step, Klf6 itself was among the genes that were spontaneously added to the network. Klf6 was the only added TF to be placed in the core, even as additional pro-growth TFs including Sox11, Stat3, Myc, and Jun appeared in inner shells. The appearance of these TFs and their central positions in the network substantiates the premise that the constructed network is Klf6-centered and relevant to axon growth<sup>2,7–11,32–34</sup>. Moreover, all five of the “soft hit” TFs also appeared but were placed in outer shells of the network, hinting at interactions near the threshold of screening sensitivity. Most importantly, three hit TFs (Nr5a2, Eomes, and Rarb) occupied the core with Klf6, prioritizing subsequent interest. Thus, an integrated pipeline of co-occupancy analyses, high content screening, and network analyses point toward three TFs as potentially cooperative with Klf6.

For independent validation, we performed a second analysis, this time starting from a previously published set of genes that are upregulated in cortical neurons in vitro upon forced expression





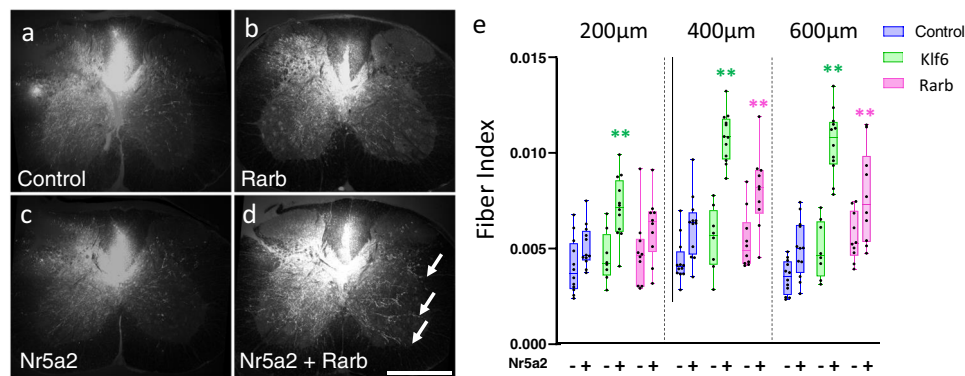


**Fig. 3 Nr5a2 synergizes with Klf6 to promote CST growth.** **a** Adult mice received cortical AAV delivery of single or combinatorial TFs and contralateral pyramidotomy. **b** Co-injected tdTomato-labeled CST axons, visible in cross-section of the medullary pyramids. **c** PKC $\gamma$  immunohistochemistry confirmed unilateral ablation of the CST (arrow). **d**, **e** Cross-midline growth of CST axons was quantified in transverse sections of cervical spinal cord by counting intersections between labeled axons and virtual lines at 200, 400, and 600  $\mu$ m from the midline, normalized to total labeled axons counted in the medulla (Fiber Index). Candidate TFs administered singly did not significantly elevate axon growth, whereas Nr5a2 and Rarb elevated growth above the level of Klf6 alone. [ $\#p = 0.0502$ ,  $*p < 0.05$  (K + R 200  $\mu$ m–0.0284, K + N 400  $\mu$ m–0.0129, K + N 600  $\mu$ m–0.0367) two-way ANOVA, post hoc Dunnett’s]. **f–o** Representative images of cervical spinal cord showing elevated cross-midline growth in animals treated with Klf6 and Nr5a2. Scale bar is 500  $\mu$ m. *n* is between six and ten animals for all groups; PKC $\gamma$ , medulla, and spinal images from all animals are in Supplementary Figs. 2–5. Box plots represent median and IQR (25th to 75th percentile), and whiskers extends to the maximum and minimum values. Dots in **e** represent values of single animals. Source data are provided as a Source data file. Scale bars are 200  $\mu$ m (**b**, **c**) and 500  $\mu$ m (**f–o**).

**Nr5a2 synergizes with Klf6 and Rarb to drive enhanced CST sprouting following pyramidotomy injury.** We next asked whether forced expression of the three hit TFs, singly or in conjunction with Klf6, can promote the growth of CST axons in vivo. Based on serotype availability, we first verified the ability of AAV2-Retro vectors, previously shown to effectively transduce CST neurons in a retrograde manner<sup>35</sup>, to also transduce neurons when applied directly to cell bodies. We delivered a titer-matched mixture of AAV-H2B-mEGFP and AAV-H2B-mScarlet to the cortex of adult mice, followed by retrograde labeling of CST neurons by cervical injection of CTB-647. Two weeks later, 3.6% ( $\pm 1.5\%$ ) of transfected CST neurons expressed only EGFP, 1.0% ( $\pm 0.3\%$ ) expressed only mScarlet, and 95.4% ( $\pm 1.5\%$ ) expressed both fluorophores (Supplementary Fig. 2). Accordingly, TFs were delivered to adult mice by cortical injection of Retro-AAV (hereafter shortened to AAV), with AAV-Cre acting as control and also included in single-gene treatments to equalize the viral load in all animals (see Supplementary Data 3). tdTomato tracer was co-injected in serotype AAV9, avoiding potential complications of retrograde spread of the tracer itself. Mice received unilateral pyramidotomy, followed 8 weeks later by quantification of cross-midline sprouting of corticospinal (CST) axons in the cervical spinal cord, measured at 200, 400, and 600  $\mu$ m from the midline (Fig. 3a, d). Axon counts were normalized to tdTomato+ axons in the medullary pyramids (Fig. 3b) and injuries were confirmed by unilateral ablation of PKC $\gamma$  in spinal sections (Fig. 3c). Raw images of all animals are provided in Supplementary Figs. 3–5. In addition to Cre, we also included Nkx3.2, a

non-hit/non-core TF, as an additional negative control; RNA-scope or immunohistochemistry confirmed overexpression of all TFs (Supplementary Fig. 6a–e). As expected, Klf6 expression significantly increased normalized axon counts at all distances across the midline ( $p < 0.01$ , two-way ANOVA) (Fig. 3e, g). When expressed singly, none of the candidate TFs significantly elevated axon growth above the level of control (Fig. 3e, h, j, l, n). When combined with Klf6, however, significant effects emerged (Fig. 3e–o). Most notably, co-expression of Nr5a2 with Klf6 consistently elevated axon counts above that of Klf6 alone and above the maximum growth in Cre controls in all animals. Co-expression of Rarb with Klf6 drove growth that exceeded Klf6 alone at 200  $\mu$ m from the midline (Fig. 3e). In three of nine animals, combined Eomes/Klf6 expression yielded high axon counts, but this effect was inconsistent and did not reach statistical significance (Fig. 3e). Finally, as expected, Nkx3.2 produced no significant elevation in axon growth above the level of Klf6 alone (Fig. 3e, i). In summary, these data identify robust synergy between Klf6 and candidate TFs, most notably Nr5a2.

We then performed a second in vivo pyramidotomy experiment to confirm Nr5a2’s synergy with Klf6 and to determine whether Nr5a2 might similarly potentiate the growth of other pro-growth TFs. AAV delivery, injury, and axon quantification were identical to the first experiment and detailed histology is available in Supplementary Figs. 3–5. In this experiment, Nr5a2 acted as the base gene, to which control, Klf6, or Rarb were added. Consistent with the prior experiment, combined Nr5a2/Klf6 expression elevated CST growth above the level of either



**Fig. 4 Nr5a2 synergizes with Rarb to promote CST growth. a–d** Adult mice received cortical AAV delivery of single or combinatorial TFs and contralateral pyramidotomy. Co-injected AAV-tdTomato-labeled CST axons. Elevated cross-midline sprouting of CST fiber is evident in animals treated with combined Nr5a2 and Rarb (arrows, **d**). **e** Cross-midline growth of CST axons was quantified in transverse spinal sections. Nr5a2 administered alone had no significant effect on axon growth (blue bars), but increased axon growth when combined with either Klf6 or Rarb, significantly above the level of either alone (green and pink bars, respectively). [ $**p < 0.01$  (K + N 200  $\mu\text{m} = 0.0012$ , K + N 400  $\mu\text{m} = 0.0001$ , K + N 600  $\mu\text{m} = 0.0001$ , K + R 400  $\mu\text{m} = 0.0007$ , K + R 600  $\mu\text{m} = 0.0136$ , two-way ANOVA, post hoc Dunnett's)]. Scale bar is 500  $\mu\text{m}$ .  $n$  is between 10 and 14 animals per group. PKC $\gamma$ , medulla, and spinal images from all animals are in Supplementary Figs. 2–5. Box plots represent median and IQR (25th to 75th percentile), and whiskers extends to the maximum and minimum values. Dots in **e** represent values of single animals. Source data are provided as a Source data file.

alone (Fig. 4e). Moreover, axon growth in the lowest Klf6/Nr5a2 animal still exceeded that in the maximum control animal, illustrating the strength and consistency of the effect. Interestingly, combined Nr5a2 and Rarb also elevated CST growth above the level of either alone (Fig. 4a–e). This effect did not appear as striking as that of Nr5a2/Klf6, but nevertheless substantiates the ability of network analysis to predict functional TF cooperation. Thus, combinations of TFs, rationally selected based on the basis of predicted co-occupancy of pro-growth regulatory DNA, can enhance CST axon growth in vivo.

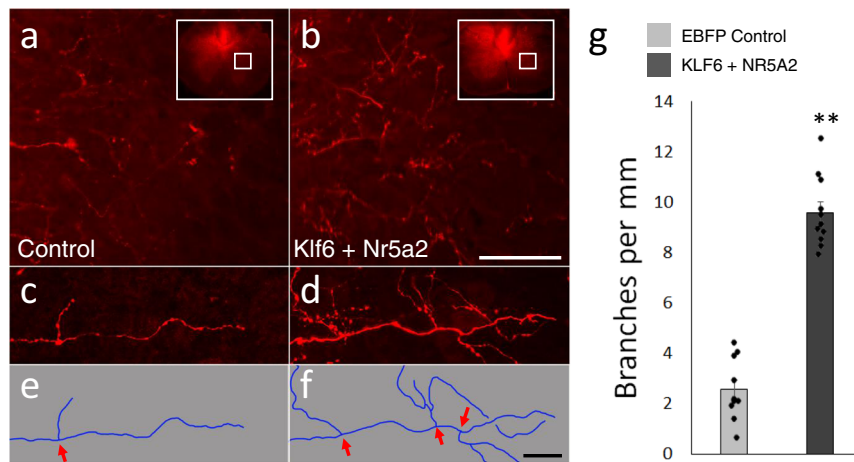
Next, focusing on the Klf6 and Nr5a2 combination, we performed additional experiments to substantiate and clarify the phenotypic effects in vivo. First, to further confirm successful co-expression, we performed additional cortical injections, using the same viral loads as the pyramidotomy experiments, followed by RNAscope-based co-detection of Klf6 and Nr5a2. In animals that received AAV-tdTomato with AAV-Klf6 and AAV-Nr5a2, >90% of tdTomato-positive cells showed a strong label with both transgenes (Supplementary Fig. 6). Thus, tdTomato-labeled axons in the spinal cord in Nr5a2/Klf6 animals largely arose from CST neurons expressing both TFs. We next compared inflammation, gliosis, and cell death in the cortex of animals that received control vs. Klf6/Nr5a2. Three days after injection, CD11b, Glial Acidic Fibrillary Protein (GFAP), and Terminal Deoxynucleotidyl Transferase dUTP Nick End Labeling (TUNEL) reactivity near the needle tracks were similar across treatments (Supplementary Fig. 7a–d, i, j). We also examined cortices in which CST neurons were identified by retrograde labeling and again detected similarly low levels of GFAP, CD11b, and TUNEL reactivity (Supplementary Fig. 7e–h, i, j). These data support a model in which Klf6/Nr5a2 act directly in CST neurons to influence axon growth, as opposed to acting indirectly by influencing gliosis, inflammation, or cell death.

We next re-examined the morphology of Klf6/Nr5a2-evoked growth by quantifying the frequency of branch points in CST axons. Starting with tissue sections from the second pyramidotomy experiment, axons that sprouted across the cervical midline were visualized at high magnification by confocal microscopy (Fig. 5a–d). Segments of axons that remained continuously within the plane of visualization for a minimum of 100  $\mu\text{m}$  were identified and traced, and the number of definitive branches along the segment was normalized to the traced length. Interestingly,

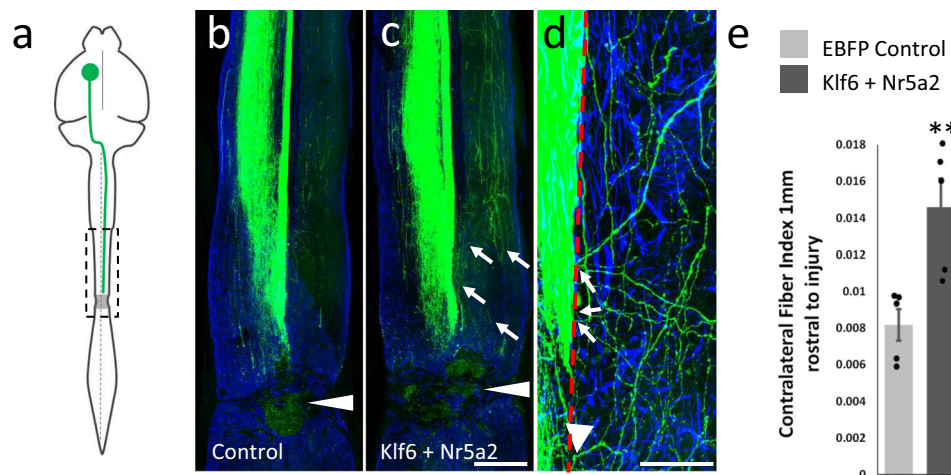
CST axons in Klf6/Nr5a2-treated animals displayed significant elevation of the frequency of branch formation (Fig. 5e–g control:  $2.5 \pm 0.4$  SEM branches/mm, Klf6/Nr5a2:  $9.6 \pm 0.5$  SEM branches/mm). This finding is consistent with the in vitro screening data, which showed that Klf6/Nr5a2 expression increased the length of the longest neurite when branches were included in the tracing, but not when only the longest unbranched path was analyzed (Supplementary Fig. 1d, e). Combined, these data indicate that combined Klf6/Nr5a2 expression in cortical neurons can act to increase branch formation and growth.

Finally, we tested the effects of forced expression of Klf6/Nr5a2 in a model of direct axon injury. Adult mice received cortical injections of AAV-EGFP tracer and either AAV-Cre control or combined AAV-Klf6 and AAV-Nr5a2. Animals received a severe crush injury to the thoracic spinal cord, followed 4 weeks later by perfusion and visualization of axons in horizontal spinal sections (Fig. 6a–d). GFAP reactivity defined the injury site, which, as expected, spanned the entirety of the cord in all animals. In control and Klf6/Nr5a2-treated animals alike, CST axons were completely interrupted by the injury and in no animals did we observe an extension of axons from the severed ends into or beyond the injury site. The distance between the lesion edge and CST axons was similar between groups, suggesting similar retraction behavior. Notably, however, axons treated with Klf6/Nr5a2 displayed a strong sprouting response into spinal tissue rostral to the injury, including robust extension across the spinal midline. Individual axons were clearly visualized crossing the midline and Klf6/Nr5a2 animals showed a significant elevation of axon density in contralateral spinal cord, normalized to axon counts in the medulla (Fiber Index, Fig. 6e). Thus, although Klf6/Nr5a2 does not confer to CST axons an ability to traverse the strongly inhibitory environment of a spinal lesion, it does act in injured axons to increase the propensity for growth by axon collaterals.

**Combined Klf6-Nr5a2 treatment leads to the upregulation of gene modules related to macromolecule biosynthesis and DNA repair.** To probe the underlying mechanisms of growth promotion, we profiled the transcriptional consequences of single and combined expression of Klf6 and Nr5a2 in CST neurons. To



**Fig. 5 Combined Klf6/Nr5a2 treatment increases the frequency of CST axon branching.** **a, b** CST axons transduced with control (**a**) or Klf6 and Nr5a2 (**b**), which have extended across the midline into contralateral cervical spinal cord. **c–f** At higher magnification, instances of branching can be identified (arrows). **g** Animals that received combined Klf6 and Nr5a2 showed a significant increase in the frequency of branch formation above control animals ( $p = 1.89 \times 10^{-10}$ , two-tailed  $t$ -test). A minimum of 5 mm of growth, from three separate sections, was quantified for each of ten animals in both groups. Scale bars are 100  $\mu\text{m}$  (**a, b**) and 20  $\mu\text{m}$  (**c–f**). Dots in **g** represent values in single animals. Source data are provided as a Source data file.

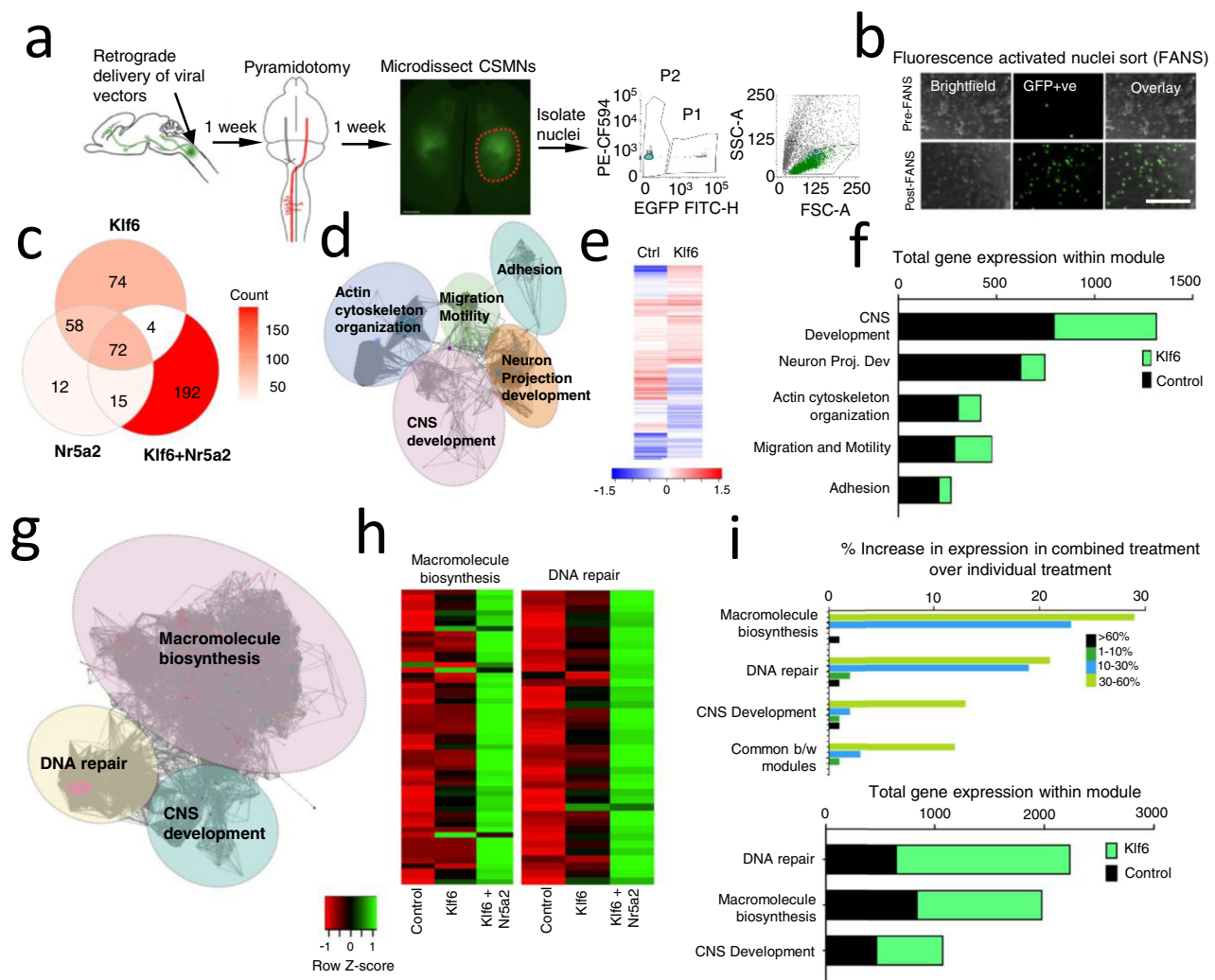


**Fig. 6 Combined expression of Klf6 and Nr5a2 in corticospinal neurons increases collateral sprouting but not extension through sites of spinal lesion.** **a** Adult mice received cortical injection of AAV-EGFP tracer with AAV-Cre control or combined AAV-Klf6 and AAV-Nr5a2, and T10/11 spinal crush injuries. **b, c** Horizontal spinal sections with examples of CST axons (green), spinal injuries (arrowheads), and reactive astrocytes (GFAP, blue). Neither EBFP control nor KLF6/Nr5a2-treated axons traverse the injury, but Klf6/Nr5a2-treated axons display cross-midline sprouting rostral to the injury (arrows). **d** CST collaterals crossing the midline (dotted line) and extending into contralateral cord. **e** Quantification of CST axons that extend 400  $\mu\text{m}$  from the midline, normalized to total CST axons counted in the medullary pyramids, shows a significant elevation in cross-midline growth (\*\* $p = 0.0069$ , two-tailed  $t$ -test,  $n = 5$  control, 5 Klf6/Nr5a2). Scale bars are 1 mm (**b, c**) and 100  $\mu\text{m}$  (**d**). Dots in **e** represent values of single animals. Source data are provided as a Source data file.

isolate CST neurons, nuclear-localized fluorophores and TFs were expressed by cervical injection of Retro-AAV2, which results in highly efficient and selective transgene expression by retrograde transduction<sup>35</sup> (Fig. 7a, b). Animals received pyramidotomy injury, followed 1 week later by purification of labeled nuclei of CST neurons from the spared hemisphere by flow cytometry, isolation of messenger RNA, and construction and sequencing of RNA libraries. Differential gene expression was determined using EdgeR software<sup>36</sup>. Compared to tdTomato control, single expression of Klf6 resulted in the upregulation of 208 transcripts (transcripts with fragments per kilobase of transcript per million mapped reads (FPKM) > 5 and  $\log_2$ fold change > 1;  $p$ -value < 0.05, false discovery rate (FDR) < 0.05) (Fig. 7c). Reminiscent of prior findings in cultured neurons<sup>7</sup>, network analysis of Klf6-responsive genes revealed subnetworks with functions highly relevant to axon

growth, including CNS development, neuron projection development, migration, adhesion, and cytoskeleton organization (Fig. 7d–f). Indeed, neuron projection and cytoskeleton networks comprised >60% of Klf6-responsive genes, consistent with the effects of Klf6 overexpression on axon growth (Fig. 7d–f). Single overexpression of Nr5a2 had little unique effects on gene expression, with only 12 transcripts identified as upregulated only in the Nr5a2 group and 130 transcripts shared with the Klf6 treatment group (Fig. 7c). Dual Klf6/Nr5a2 expression, however, drove upregulation of 192 transcripts that were not significantly increased by either alone (Supplementary Data 4 and Fig. 7c). Interestingly, network analysis of these transcripts revealed modules involved in macromolecule biosynthesis, DNA repair, and development, all of which showed substantial increases in the net expression above control levels (Fig. 7g–i).





**Fig. 7 Combined Klf6, Nr5a2 treatment induces expression of gene modules involved in macromolecule biosynthesis and DNA repair.** **a** Overview of sample collection for RNA-Seq analysis. (CSMNs, corticospinal motor neurons). **b** Nuclei before and after FANS, confirming successful purification of mNeogreen+ nuclei. **c** Venn diagram showing transcript comparison across the three gene treatments. **d** Regulatory network analysis of genes upregulated after Klf6 overexpression revealed subnetworks enriched for distinct functional categories highly relevant to axon growth. Nodes correspond to target genes and edges to multiple modes of interaction (physical, shared upstream regulators, shared signaling pathways, and inter-regulation). Only significantly enriched GO categories were included in the network analysis ( $p < 0.05$ , right-sided hypergeometric test with Bonferroni correction, Cytoscape). **e** Heatmap showing selection of differentially expressed genes between control and Klf6 treatment. **f** The number of genes within each functional subnetwork (top) and the total FPKM values in each in control and Klf6 overexpression groups. **g** Regulatory network analysis of genes uniquely upregulated after combined Klf6 and Nr5a2 overexpression (not upregulated in single treatments) revealed subnetworks enriched for functions in macromolecule biosynthesis and DNA repair ( $p < 0.05$ , right-sided hypergeometric test with Bonferroni correction, Cytoscape). **h** Heatmap showing selection of differentially expressed genes in DNA repair and macromolecule biosynthesis modules. Genes that are upregulated (green) and downregulated (red) relative to control are shown. Color values indicate average FPKM expression values as shown in the legend. Columns correspond to treatment groups and rows to genes. **i** The percent change in FPKM values within each module, revealing that combined Klf6 and Nr5a2 treatment most affect genes in DNA repair and macromolecule biosynthesis modules (top). Graph shows the number of genes within functional subnetworks (bottom).  $n = 3$  animals pooled per replicate, 2 replicates/gene treatment with similar results (Supplementary Data 4). Scale bar is 100  $\mu\text{m}$ . Source data are provided as a Source Data file.

To further validate the findings of this bulk RNA sequencing (RNA-Seq) approach, we performed a replicate study, this time employing a single-nuclei RNA-Seq approach. Nuclear-localized fluorophores and TFs (Control or Klf6 + Nr5a2) were expressed by cervical injection of Retro-AAV2 as described above (Supplementary Fig. 8a). Animals received pyramidotomy injury, followed 1 week later by purification of labeled nuclei of CST neurons by flow cytometry, construction, and sequencing of single-nuclei libraries. Using this approach, we sequenced and analyzed ~3000 control nuclei and Klf6/Nr5a2-treated nuclei (Supplementary Fig. 8b). The majority of nuclei partitioned into

clusters highly enriched for established markers of CST neurons, including Bcl11b, Cdh13, Tmem163, Crim1, Cntn6, and Slco2a1, indicating successful purification of CST neurons<sup>23,37,38</sup> (Supplementary Fig. 8c). We integrated control and Klf6/Nr5a2 single-nuclei datasets using SEURAT<sup>39</sup>, and carried out differential gene expression analyses to identify Klf6/Nr5a2-responsive target genes, specifically in the high-confidence CST cluster (Supplementary Data 4). As expected, Klf6 and Nr5a2 themselves were detected as strongly upregulated. In addition, when comparing the output of the initial bulk RNA-Seq and single-nuclei RNA-Seq, 144 genes were commonly called as significantly regulated by



Klf6/Nr5a2 in both the datasets, 101 of which were shared with the 192 transcripts called as Klf6/Nr5a2-specific in the bulk RNA-Seq experiment (>50% overlap) (Supplementary Fig. 8d). Importantly, all shared target genes were located in clusters functionally linked to roles in Macromolecule Biosynthesis, DNA repair, and CNS development, providing independent support for the regulation of these processes by combined expression of Klf6 and Nr5a2 (Supplementary Fig. 8e–g). Intriguingly, recent analyses of regeneration-competent peripheral neurons also showed that genes involved in macromolecule biosynthesis and DNA repair are upregulated in the course of successful regeneration<sup>14</sup>, supporting their potential involvement in axon growth. In aggregate, these data substantiate the ability of forced Klf6 expression to activate gene networks relevant to axon growth and point toward the ability of dual Klf6/Nr5a2 expression to activate gene networks involved in macromolecule synthesis and DNA repair as a likely explanation for their combined enhancement of axon growth.

## Discussion

Using interlocking bioinformatics approaches, high content screening, and *in vivo* testing, we have identified TF combinations that synergize to enhance CST axon growth. In addition, transcriptional profiling and network analyses provide molecular correlates to the evoked responses. These findings clarify cellular functions that drive axon extension and serve as a roadmap for ongoing efforts to discover combinatorial gene treatments to promote regenerative axon growth.

Klf6 belongs to the KLF family of TFs, which have well-studied roles in regulating axon growth<sup>7,12,28,40</sup>. Both Rarb and Nr5a2/LRH-1 belong to the nuclear receptor family of TFs, with diverse functions in development and cellular metabolism<sup>41</sup>. Rarb has previously been linked to axon growth and a Rarb agonist is under evaluation for clinical efficacy in a brachial plexus avulsion model of injury<sup>42</sup>. In contrast, Nr5a2 is unstudied in the context of axon extension. It has, however, been linked to the differentiation of neural stem cells and to cellular reprogramming, and can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells<sup>43,44</sup>. Our present work adds to the growing body of evidence that TFs involved in neuronal differentiation may positively influence axon growth<sup>9,13</sup>.

An important finding is that although Nr5a2 does not improve axon growth when expressed individually, it strongly potentiates the effect of overexpressed Klf6. The importance of combinatorial gene activity for axon growth, as opposed to single-factor effects, is increasingly recognized<sup>2,4,45</sup>. Regeneration-competent cell types such as zebrafish retinal ganglion cells and peripherally injured sensory neurons respond to axotomy by upregulating groups of TFs that likely synergize to drive regenerative axon growth<sup>14,15,46</sup>. An unmet challenge, however, has been to turn the conceptual appreciation for the role of TF synergy into an operational workflow to discover effective gene combinations for axon extension in the injured nervous system. Whereas prior efforts have focused on direct protein–protein interactions to predict TF synergy<sup>46</sup>, here we exploited the concept of TF co-occupancy. A search for factors with recognition motifs in proximity to those of Klf6 yielded a set of 62 candidate TFs, which was further narrowed by phenotypic screening. A key observation from the cell culture experiments is that none of the hit TFs were predicted to co-occupy solely in promoter regions; rather, all hit TFs were characterized by predicted co-occupancy in both promoters and enhancers. This suggests a key role for enhancer elements in regulating pro-growth gene transcription, in line with the recent finding in Dorsal Root Ganglion (DRG) neurons that modulation of enhancers may be critical for transcriptional activation of growth genes<sup>13,47</sup>.

Finally, guided by network analysis of the screening results, a systematic campaign of combinatorial testing *in vivo* identified highly consistent enhancement of CST axon growth by combined expression of Klf6 and Nr5a2. Importantly, in a model of pyramidalotomy injury, delivery of Klf6 and Nr5a2 did not affect cell survival but did affect both the overall density and frequency of branching of collateral axons in contralateral spinal cord. Moreover, in axotomized CST axons, forced expression also increased the growth of collateral branches, although not regenerative growth across a spinal lesion. Importantly, in contrast to prior reports of CST regeneration through crush sites, injuries here were performed using double crushes with wider forceps, thus producing complete lesions without astrocytic bridges<sup>48,49</sup>. The absence of regenerative advance in this model indicates that treated axons remained sensitive to environmental inhibition at the lesion, but do not rule out the possibility that Klf6 and Nr5a2 treatment could produce regenerative growth into more favorable environments, e.g., progenitor-derived tissue grafts<sup>50</sup>. In addition, the strong effects of combined Klf6 and Nr5a2 on collateral growth in both spared and injured CST axons suggests that this treatment could be highly beneficial in situations of partial injury by enhancing increased branch elaboration by spared axons and potentially fostering relay circuitry by injured axons<sup>51,52</sup>. A critical question for future research is whether the stimulated branches succeed in forming functional synapses on target cells in the spinal cord. We believe this evidence of TF synergy in driving growth of axons following CNS injury *in vivo* is a critical advance for the field. Although these experiments were anchored on Klf6, these data show a pipeline that can be re-deployed to identify additional TF combinations that synergize to drive growth, or potentially other cellular functions.

A key insight from the current work is to clarify *in vivo*, in purified CST neurons, the transcriptional events that are triggered by Klf6, Nr5a2, and the two together. Forced expression of Klf6 activates highly interconnected groups of genes with functions in various aspects of axon growth spanning terms such as Neuron Projection Development, Cytoskeleton organization, Motility, and Adhesion. We have previously shown that gene modules involved in cytoskeleton remodeling and motility are evoked in response to Klf6 overexpression in cortical neurons *in vitro*<sup>7</sup>, consistent with our *in vivo* findings here. This overlap underscores the utility of *in vitro* screening approaches in delineating transcriptional networks relevant to axon growth. Importantly, this finding illustrates the potential of TF-based interventions to trigger broad changes in gene expression; it is most likely that improvements in axon growth reflect the net change in a wide set of transcripts, as opposed to acting through any single target. Combined Klf6/Nr5a2 expression, which led to large and consistent increases in axon growth above those triggered by Klf6 alone, also caused the upregulation of unique modules of genes with functions in macromolecule biosynthesis and DNA repair. These modules were activated by neither Klf6 nor Nr5a2 alone, suggesting that enhanced biosynthesis and/or DNA repair contribute to the enhanced growth in Klf6/Nr5a2-treated animals. A role for the biosynthesis of macromolecules is highly plausible, as active growth depends on re-synthesis of cellular material. Indeed, regeneration-competent zebrafish retinal ganglion cells and mammalian DRGs also respond to injury by upregulating genes involved in macromolecule biosynthesis<sup>13–15</sup>. Moreover, a very recent study in DRG neurons found that peripheral injury, which triggers axon growth, resulted in a sustained upregulation of genes involved in biosynthesis. In contrast, a central injury that does not trigger regeneration led to eventual downregulation, suggesting that a reduction in biosynthesis pathways may restrict axon growth after spinal injury<sup>53</sup>. Our findings regarding the effects of Klf6/Nr5a2 add to multiple lines of evidence that the

upregulation of genes involved in biosynthesis may be an evolutionarily conserved molecular signature of neurons mounting a successful growth response.

Regarding DNA repair pathways, it is interesting to note that increased cellular metabolism and transcription frequently lead to DNA damage<sup>54,55</sup>. Moreover, disruption of DNA repair machinery in mouse retinal progenitors or cortical progenitors leads to reduced, disturbed trajectories of axon growth and guidance<sup>56,57</sup>. Thus, during development, proper axon growth may require efficient DNA repair machinery. Similarly, peripherally injured DRG neurons upregulate several DNA damage response marker genes and blocking this response leads to reduced neurite outgrowth in vitro and impaired regeneration in vivo<sup>58</sup>. The finding that combined *Klf6*/*Nr5a2* expression activates genes involved in DNA repair suggests the intriguing hypothesis that this repair machinery may aid successful axon regrowth in CNS neurons; future work can address this notion.

In summary, a bioinformatic and screening platform, centered on the concept of TF co-occupancy, has revealed synergy between TFs *Klf6* and *Nr5a2*, which leads to improved axon growth after CNS injury.

## Methods

**Mice and husbandry conditions.** All animal testing and research was carried out in compliance with ethical regulations laid out by the National Institutes of Health (NIH) guide for the care and use of animals, and all experimental protocols involving animals were approved by the Institutional Animal Care and Use Safety committee at Marquette University (protocol number AR-309, AR-314). Mice were bred and raised under a 24 h light–dark cycle with 12 h of light and 12 h of darkness. Ambient temperature was maintained at 22 °C ± 2 °C and humidity between 40 and 60%.

### Identification of developmentally downregulated, growth-relevant transcripts.

All expression datasets used to identify developmentally downregulated genes are deposited at NCBI GEO (In vivo CSMNs data by ref. <sup>23</sup> - GSE2039; In vitro primary neurons aged in culture by ref. <sup>24</sup>, SRP151916). In vitro RNA-Seq datasets were processed using the updated Tuxedo suite of tools<sup>24</sup>. Trimmed reads were mapped to the rat reference genome [UCSC, Rat genome assembly: Rnor\_6.0] using HISAT2 aligner software (unspliced mode along with –qc filter argument to remove low-quality reads prior to transcript assembly)<sup>59</sup>. Transcript assembly was performed using Stringtie<sup>60</sup> and assessed through visualization on UCSC Genome Browser. Differential gene expression analysis was performed using Ballgown software (Default parameters)<sup>60</sup>. Transcripts were considered significant if they met the statistical requirement of having a corrected *p*-value of <0.05, FDR < 0.05.

For the in vivo CSMN microarray datasets, weighted correlation network analysis was performed<sup>61</sup>. First, relevant microarray datasets (GSE2039) were loaded into R and basic pre-processing of data was performed to remove outlier data and handle probes with missed data. Outlier data were identified by performing hierarchical clustering to detect array outliers. Numbers of missing samples in each probe profile were counted and probes with extensive numbers of missing samples were removed. Next, a weighted correlation network was constructed by creating a pairwise Pearson correlation matrix, which was transformed into an adjacency matrix using a power of 10. Then, topological overlap was calculated to measure network interconnectedness and average linkage hierarchical clustering was done to group genes on the basis of the topological overlap dissimilarity measure. Finally, using a dynamic tree-cutting algorithm, we identified 20 gene modules. Gene modules with developmental downregulation in expression of twofold and above were isolated for ontology analyses. GO analyses were performed using Database for Annotation, Visualization and Integrated Discovery Bioinformatics Resource<sup>62</sup> and genes with terms relevant to axon growth (FDR < 0.05) based on literature review were user-selected for further analyses.

**Enhancer identification.** Enhancers relevant to pro-growth genes were identified by running the ABC algorithm<sup>25</sup> using code described here—<https://github.com/broadinstitute/ABC-Enhancer-Gene-Prediction>. ENCODE datasets used for the analysis are as follows: ATAC-Seq (ENCSR310MLB, ENCSR836PUC), H3K27Ac histone chromatin immunoprecipitation sequencing (ChIP-seq) (ENCSR094TTT, ENCSR428OEK), RNA-seq (ENCSR362AIZ, ENCSR080EVZ), and HiC (GSE96107). The algorithm has three sequential steps as follows: definition of candidate enhancers, quantification of enhancer activity, and calculation of ABC scores. For definition of candidate enhancers, indexed and sorted bam files of ATAC-Seq and H3K27Ac ChIP-seq datasets were supplied and peaks were called using MACS2, with the argument –nStrongestPeaks 15000. For quantification of enhancer activity, reads counts were calculated using peak files processed in the previous step, to yield a final list of candidate enhancer regions with ATAC-seq and

H3K27ac ChIP-seq read counts within gene bodies and promoters. Finally, ABC scores were calculated using arguments –hic\_resolution 5000, –scale\_hic\_using\_powerlaw, and a threshold of 0.01. The default threshold of 0.01 corresponds to ~70% recall and 60% precision. All enhancer–gene pairs that passed the above threshold were used for subsequent analyses.

**Co-occupancy motif analyses.** TF-binding site/motif analysis on pro-growth gene promoters and enhancers was performed using opomsum v3.0 software<sup>27</sup>. For promoter analyses, the list of pro-growth genes was supplied along with promoter coordinates to be used for scanning (upstream/downstream of TSS—1000/300 bps). For enhancer motif analyses, a unified file containing a list of FASTA formatted sequences corresponding to pro-growth enhancer regions was supplied.

Mouse anchored TF-cluster analyses were performed using search parameters—JASPAR CORE profiles that scan all vertebrate profiles with a minimum specificity of 8 bits, conservation cutoff of 0.40, matrix score threshold of 90%, and results sorted by Z-score ≥ 10. TFs were sorted into three categories as follows: TFs with over-represented sites within pro-growth promoters, TFs with over-represented sites within pro-growth enhancers, and TFs with over-represented binding sites within both pro-growth promoters and enhancers.

**Network analyses.** Hit TF target gene networks were constructed using ENCODE ChIP-Seq data for hit TFs using the Harmonizome database<sup>63</sup>. For each of the 12 hit TFs, TF target gene lists were retrieved from Harmonizome individually and merged before visualization on the Cytoscape platform (v3.7.1). Cytoscape plug-ins ClueGO and CluePedia were used for network visualization. iRegulon plug-in on the Cytoscape platform was used to expand the network one level and bring in additional regulators. Next, we calculated connectivity scores to assign centrality using the Network Analyzer option within Cytoscape 3.7.1. TFs were ranked in descending order of connectivity scores to assign centrality. Core TFs were those TFs that had the highest connectivity scores and were centrally located in the network, followed by TFs with decreasing connectivity scores that occupied peripheral positions in the network. Upstream regulator analysis on *Klf6* target genes was run on differentially expressed genes listed in ref. <sup>7</sup> using motif analyses described above. Motif analyses were done in batches such that each batch had genes belonging to one functional subnetwork to identify *Klf6* synergizers by functional category.

**Cloning and virus production.** Constructs for candidate genes were purchased from Dharmacon or Origene, and relevant accession numbers for all 62 candidate TFs are summarized in Supplementary Data 2. mScarlet was a gift from Dorus Gadella (RRID:Addgene\_85044), pAAV-CAG-tdTomato (codon diversified) was a gift from Edward Boyden (RRID:Addgene\_59462), and mNeonGreen sequences were synthesized by Genscript, based on the amino acid sequence of ref. <sup>64</sup>. For viral production, genes were cloned into an AAV-CAG backbone (Addgene Plasmid #59462) using standard PCR amplification<sup>12</sup>. Maxipreps were prepared by Qiagen Endo-free kits and fully sequenced, and AAV9-tdTomato (Addgene Plasmid #59462) and AAV2-Retro of all other constructs were produced at the University of North Carolina Viral Vector Core and brought to 1 × 10<sup>13</sup> particles per ml in sterile saline prior to injection. Viruses were mixed at a 3:3:2 ratio of combinatorial test genes and tracer; viral treatments are detailed in Supplementary Data 3.

**Cortical cell culture and analysis of neurite outgrowth.** All animal procedures were approved by the Marquette University Institutional Animal Care and Use Committee. Cortical neurons were prepared from early postnatal (P5–P7) Sprague–Dawley rat pups (Harlan), with each experimental replicate derived from a separate litter. Procedures for dissociation, transfection, immunohistochemistry, imaging, and neurite outgrowth analysis were performed as in ref. <sup>7</sup>. Plasmid expressing nuclear-localized EGFP, mixed with text plasmids at a 1:3 ratio, served to identify transfected cells. Each 24-well culture plate included three standard treatments: enhanced blue fluorescent protein (EBFP) control, *Klf6* plasmid mixed in equal parts with mCherry control, and *Klf6* mixed in equal parts with VP16-Stat3. Remaining wells contained *Klf6* mixed with candidate TFs. All morphological values were normalized to within-plate EBFP control values, the *Klf6*-only treatment served to set the level of *Klf6*'s individual effect, and the *Klf6*/VP16-Stat3 combination confirmed sensitivity of the assay plate to TF synergy<sup>7</sup>. For all cell culture experiments, neurite length from a minimum of 150 cells per treatment was averaged and each experiment was repeated a minimum of three times on separate days. These averaged values were the basis for ANOVA with Fisher's multiple comparisons. Neurite outgrowth values, number of cells analyzed for every screening experiment are summarized in Supplementary Data 2.

**Viral delivery to cortical neurons and pyramidotomy injuries.** In vivo experiments were performed in a double-blind manner, with non-involved lab personnel maintaining blinding keys. Experiments used adult C57BL/6 mice, 8–10 weeks of age at the start of the experiment. The first pyramidotomy experiment used female mice and the second pyramidotomy and spinal crush experiment used mixed sex at approximately equal numbers; sex had no significant effect on axon growth in either experiment (*p* > 0.05, two-way ANOVA). Animals were randomized prior to

viral treatment, with each surgical day including equal numbers from each group. Cortical neurons were transduced using intracerebral microinjection<sup>7,8</sup>. Briefly, mice were anesthetized with ketamine/xylazine (100/10 mg/kg, intraperitoneally), mounted in a stereotaxic frame, and skull exposed and scraped away with a scalpel blade. Virus particles (0.5  $\mu$ l) were delivered at two sites located 0 mm/1.3 mm and 0.5 mm/1.3 mm (anterior/lateral from Bregma) in pyramidotomy experiments, and -0.6 mm/1.3 mm and -1.2 mm/1.3 mm (anterior/lateral from Bregma) in spinal crush experiments, at a depth of 0.55 mm and at a rate of 0.05  $\mu$ l/min using a pulled glass micropipette connected to a 10  $\mu$ l Hamilton syringe driven by a programmable pump (Stoelting QSI), with 1 min dwell time. For spinal injections, mice were mounted in a custom spine stabilizer and viral particles or Cholera Toxin Subunit B conjugated to Alexa Fluor 647 (CTB-647) in sterile 0.9% NaCl (C22841-ThermoFisher, Waltham, MA, final concentration of 2%) were injected to the spinal cord through a pulled glass micropipette fitted to a 10  $\mu$ l Hamilton syringe driven by a Stoelting QSI pump (catalog # 53311) (pumping rate: 0.04  $\mu$ l/min) between C4 and C5 vertebrae, 0.35 mm lateral to the midline, and to depths of 0.6 mm and 0.8 mm, 0.5  $\mu$ l at each depth. Unilateral pyramidotomy was performed<sup>8,12</sup>. Briefly, a ventral midline incision was made to expose the occipital bone, the ventrocaudal part of which was removed using fine rongeurs. The dura was punctured and the right pyramid cut completely using a micro feather scalpel. For spinal crush injuries, mice were anesthetized and mounted in a spine stabilization device, a laminectomy performed at T12 vertebrae/T10/11 spinal cord, and forceps of width 250  $\mu$ m used to crush the cord for 10 s, then grip-reversed and repeated for another 10 s.

**RNAScope.** RNAScope kits are commercially available from Advanced Cell Diagnostics (ACD). We utilized the Multiplex v2 system for all RNAScope experiments (Cat-323100, ACD). Mouse brains were snap-frozen on dry ice and cryosectioned (25  $\mu$ m). Slides were baked at 60 °C for 40 min followed by RNAScope procedures according to manufacturer's instructions. Briefly, hydrogen peroxide solution was applied to baked slices and incubated for 10 min at room temperature (RT). Next, slices were incubated in boiling antigen retrieval solution (<98 °C) for 5 min. Following retrieval, slices were washed in Nuclease-free water three times, 5 min per wash. Then, brain slices were dehydrated in 100% ethanol briefly followed by treatment with ProteaseIII for 20 min at 40 °C. Following protease incubation, slices were washed in Nuclease-free water three times, 5 min per wash, and probes were applied and allowed to incubate for 2 h at 40 °C. The following probes were purchased off the catalog RNAScope® Probe-Mm-Klf6 (Cat-426901, ACD), RNAScope® Probe-Mm-Nkx3-2 (Cat-526401, ACD), and RNAScope® Probe-Mm-Rarb (Cat-463101, ACD). We designed a custom probe for Nr5a2, targeting a region common to all variants (RNAScope® Probe-Mm-Nr5a2-O1-C2, Cat-547841-C2, ACD). All probes were detected with tyramide signal amplification plus fluorophore used at 1:750 dilution. Before mounting the slices, a brief 5 min incubation with 4',6'-diamidino-2-phenylindole was performed to label the nuclei. All RNAScope experiments were carried out between 1 and 3 days post cryosectioning, to ensure tissue integrity.

**Immunohistochemistry.** Adult animals were perfused with 4% paraformaldehyde (PFA) in 1 $\times$  phosphate-buffered saline (PBS) (15710-Electron Microscopy Sciences, Hatfield, PA), brains, and spinal cords removed, and post-fixed overnight in 4% PFA. Transverse sections of the spinal cord or cortex were embedded in 12% gelatin in 1 $\times$  PBS (G2500-Sigma Aldrich, St.Louis, MO) and cut via Vibratome to yield 100  $\mu$ m sections. Sections were incubated overnight with primary antibodies PKC $\gamma$  (SC C-19, Santa Cruz, Dallas, TX, 1:500, RRID: AB\_632234), GFAP (DAKO, Z0334 1:500, RRID:AB\_10013482), or Cd11b (Invitrogen 14-01120-82 1:500, RRID:AB\_2536484), or Eomes (ab23345, RRID:AB\_778267, 1:500), and rinsed and then incubated for 2 h with appropriate Alexa Fluor-conjugated secondary antibodies (R37116/7, ThermoFisher, Waltham, MA, 1:500). For TUNEL staining (In Situ Cell Death Kit, Roche), fresh-frozen transverse cryostat sections (30  $\mu$ m) were post-fixed in 4% PFA 15 min, incubated in ethanol/acetic acid for 10 min, 0.4% Triton for 30 min, and with probe mixture for 1 h<sup>8</sup>. Staurosporine control (1  $\mu$ l, 1 mM; Sigma) was cortically injected 2 days prior to killing. Fluorescent images were acquired using Olympus IX81 or Zeiss 880LSM microscopes.

**Quantification of axon growth.** In pyramidotomy experiments, axon growth was quantified from four 100  $\mu$ m transverse sections of the spinal cord of each animal that spanned C2 to C6 spinal cord. Starting from a complete series of transverse sections, the rostral-most was selected from C2 as identified by the morphology of the ventral horns of the spinal cord, and the next three selected at 1.6 mm intervals in the caudal direction. Each section was imaged on Nikon confocal microscope (Nikon AR1+) using a  $\times$ 20 Plan Apochromatic (MRD00205, NA 0.75) objective, gathering 20  $\mu$ m of images at 5  $\mu$ m intervals and then creating a maximum intensity projection. On the resulting image, virtual lines of 10  $\mu$ m width were placed at 200, 400, and 600  $\mu$ m from the midline and intersection of tdTomato+ axons with these lines were quantified. For normalization, 100  $\mu$ m transverse sections of the medullary pyramids were prepared by vibratome sectioning and imaged on a Nikon confocal microscope (Nikon AR1+) using a  $\times$ 60 Apochromatic Oil DIC N2 (MRD71600, NA 1.4) objective. Seven virtual lines of 10  $\mu$ m width were distributed evenly across the medial/lateral axis of the medullary pyramid, the

number of tdTomato+ axons visible in each was determined, and the total number of axons extrapolated based on the area of the entire pyramid and the area of the sampled regions. Fiber index was calculated as the average number of axons at each distance from the midline across the four replicate spinal sections, divided by the calculated number of axons detected in the pyramid. Counting of digital images was performed by three blinded observers, with final values reflecting the average. Exclusion criteria for pyramidotomy experiments were animals with <80% decrease in PKC $\gamma$  in the affected CST. For spinal crush injuries, quantification and normalization of cross-midline sprouting was nearly identical to those for pyramidotomy, with the exception that quantification was performed in four horizontal sections of the spinal cord, with the sampling area set 400  $\mu$ m from the midline. To quantify axon branching, the following procedure was employed. First, in transverse sections of cervical spinal cord, tdTomato+ CST axons were identified in a sampling region between 200 and 600  $\mu$ m from the midline. Next, axon segments traced, with the requirement that sampled segments must remain in a single confocal imaging plane at  $\times$ 60 magnification for a minimum of 100  $\mu$ m traced length. Finally, using Z-stacks to definitively distinguish true branches from near-plane intersections, the number of branches was counted in each sampled segment. A minimum of 5 mm of total length, from three separate spinal sections, were sampled for each animal.

**Fluorescence-activated nuclei sorting.** Equal numbers of male and female mice were used in every experiment in concordance with National Institutes of Health guidelines. Adult mice received retrograde injections of viral vectors for Ctrl treatment, Klf6-alone treatment, or combined Klf6 + Nr5a2 treatment. One week later, animals were challenged with pyramidotomy injuries, as described above. One week post injury, animals were killed and the motor cortices were dissected. Tissue was collected only from the hemisphere contralateral to the pyramidotomy injury, thus isolating spared neurons and not neurons that experienced direct axotomy. Dissected cortices were minced finely using razor blades and transferred to pre-chilled 15 ml Dounce homogenizer filled with 3 ml Nuclear release buffer (320 mM Sucrose, 5 mM CaCl<sub>2</sub>, 3 mM MgCl<sub>2</sub>, 10 mM Tris-HCl, 0.3% Igepal). Tissue was dounced 15 $\times$  while on ice and filtered sequentially via a 50 and 20  $\mu$ m filter, and used as input for flow cytometry. Dissociated nuclei were flow-sorted on a BD FACS Melody using an 80  $\mu$ m nozzle and a sequential gating strategy (Sort type: Purity). Specifically, nuclei were first gated by side scatter height (SSC-H) vs. forward scatter height (FSC-H) and side scatter area SSC-A vs. FSC-A to separate debris vs. intact nuclei. Nuclei that passed these filters were then filtered by SSC-W vs. SSC-H, to eliminate potential doublets. Finally, nuclei were gated by levels of fluorescence marker such that only the brightest nuclei are collected to a goal of ~30,000 events for RNA-Seq and 4000 events for single-nuclei RNA-Seq.

**RNA-Seq data generation and analysis.** Following fluorescence-activated nuclei sorting, nuclei were sorted directly into Trizol, lysed, and RNA was extracted according to the manufacturer's instructions. Only samples with an RNA integrity number score of 8 and above were used for library prep. Total RNA was used for library prep using Takara SMARTer stranded total RNA-Seq kit v2 according to the manufacturer's instructions. Samples were sequenced at the University of Wisconsin-Madison Genomics core to a depth of 25 million paired-end reads on an Illumina HiSeq platform, with two replicates per treatment. Raw FASTQ files were fed into ENCODE consortia RNA-Seq pipeline described here— <https://github.com/ENCODE-DCC/rna-seq-pipeline>. Read alignment was performed using STAR<sup>65</sup> and transcript quantification performed using Kallisto<sup>66</sup>. Transcript counts from Kallisto were used for performing differential gene expression analyses using EdgeR<sup>36</sup>. Transcripts were considered significant if they met the statistical requirement of having a corrected *p*-value of <0.05, FDR <0.05. Differentially expressed genes were further filtered to only include transcripts with FPKM > 5 and log2Fold change > 1. RNA-Seq datasets have been deposited with NCBI GEO (PRJNA630017). Network analysis on significantly differentially expressed genes (upregulated) was performed on the cytoscape platform v3.7.1 using plug-ins clueGo and CluePedia. Functional analysis mode was selected and the following GO categories were selected for analyses—GO Biological Process, Kyoto Encyclopedia of Genes and Genomes pathways, Reactome Pathways, and Wiki pathways. Network specificity was set to medium and GO tree interval was set between 3 and 8. GO term  $\kappa$ -score was set to default values (0.4) and leading group term was set to rank by highest significance. Lastly, GO term fusion was also selected and only subnetworks with significantly enriched GO terms were used to generate network visualizations (*p* < 0.05, right-sided hypergeometric test with Bonferroni correction, Cytoscape). Expression heatmaps were assembled using heatmapr.ca, using average linkage and Euclidean distance parameters for clustering.

**Single-nuclei RNA-Seq data generation and analysis.** Equal numbers of male and female mice were used in every experiment in concordance with NIH guidelines. Adult mice (8 weeks) received retrograde injections of viral vectors for Ctrl treatment or combined Klf6 + Nr5a2 treatment. One week later, animals were challenged with pyramidotomy injuries, as described above. One week post injury, animals were killed and the motor cortices were dissected. Dissected cortices were frozen on dry ice and stored at -80° until library preparation. Frozen tissue was transferred to a pre-chilled 3 ml Kimble dounce homogenizer filled with 4 ml Nuclei



Lysis Buffer (Sigma, NUC101), supplemented with RNAase inhibitors (10 mg/ml) (RNAase OUT, ThermoFisher 10777019). Tissue was homogenized (20 strokes with pestle A and 25 strokes with pestle B) and allowed to incubate on ice for 5 min. Following incubation, nuclei was centrifuged for 5 min at 4° (800 G, low brake and acceleration) and the pellet was resuspended in 4 ml Nuclei lysis buffer. Homogenate was incubated for 5 min on ice. Following incubation, nuclei was centrifuged as described above and resuspended in 0.5 ml 1× PBS with 1% bovine serum albumin. Dissociated nuclei were filtered using a 20 µm filter and flow-sorted on a BD FACS Melody using an 80 µm nozzle to a goal of ~4500–5000 events (gate strategy described above). Nuclei were sorted directly into 10× RT buffer (enzyme added only prior to Gel Bead-in Emulsion (GEM) generation) and library preparation was performed on the 10× chromium platform according to manufacturer's instructions (10× Genomics- Next GEM, Cat # PN1000121). Samples were sequenced at the University of Wisconsin-Madison Genomics core to a depth of ~70,000 reads/nuclei (3000–4000 nuclei per library) on an Illumina NovaSeq platform. Raw FASTQ files were fed into the Cellranger pipeline (default parameter) described here—<https://github.com/10XGenomics/cellranger>. Datasets were then integrated using SEURATv3<sup>39</sup> and differential expression testing was performed according to default conditions (non-parametric Wilcoxon rank-sum test).

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

There are no restrictions on data availability. All the datasets generated and analyzed during the current study are available in the NCBI repository (PRJNA630017). The following publicly available datasets were also analyzed in this study—ATAC-Seq (ENCSR310MLB, ENCSR836PUC), H3K27Ac histone ChIP-seq (ENCSR094TTT; <https://www.encodeproject.org/files/ENCF195BGJ/>, ENCSR428OEK), RNA-seq (ENCSR362AIZ, ENCSR080EVZ), and HiC (GSE96107). Source data are provided with this paper.

### Code availability

There was no custom code development and all software used in data analyses are previously published, open access, and have been cited under the relevant methods section. Links to relevant software repositories/documentation is listed here. WGCNA—<https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/>; RNA-Seq analyses—<https://github.com/ENCODE-DCC/rna-seq-pipeline>; EdgeR—<https://www.bioconductor.org/packages/release/bioc/html/edgeR.html>; Activity-by-contact (ABC)—<https://github.com/broadinstitute/ABC-Enhancer-Gene-Prediction>; oPOSSUM 3.0—<http://opossum.cisreg.ca/oPOSSUM3/>; Cytoscape-ClueGO—<http://www.ici.upmc.fr/cluego/cluegoDocumentation.shtml>; Harmonizome—<https://amp.pharm.mssm.edu/Harmonizome/>; iRegulon—<http://iregulon.aertslab.org/>; Cellranger—<https://github.com/10XGenomics/cellranger>. SEURAT—<https://github.com/satijalab/seurat>; FACSChorus—<https://www.bdbiosciences.com/en-us/instruments/research-instruments/research-software/flow-cytometry-acquisition/facschorus-software>; and DAVID—<https://david.ncifcrf.gov/summary.jsp>.

Received: 15 July 2020; Accepted: 29 March 2021;

Published online: 05 May 2021

### References

- Moore, D. L. & Goldberg, J. L. Multiple transcription factor families regulate axon growth and regeneration. *Dev. Neurobiol.* <https://doi.org/10.1002/dneu.20934> (2011).
- Venkatesh, I. & Blackmore, M. G. Selecting optimal combinations of transcription factors to promote axon regeneration: why mechanisms matter. *Neurosci. Lett.* **652**, 64–73 (2017).
- O'Donovan, K. J. Intrinsic axonal growth and the drive for regeneration. *Front. Neurosci.* **10**, 486 (2016).
- Mahar, M. & Cavalli, V. Intrinsic mechanisms of neuronal axon regeneration. *Nat. Rev. Neurosci.* **19**, 323–337 (2018).
- He, Z. & Jin, Y. Intrinsic control of axon regeneration. *Neuron* **90**, 437–451 (2016).
- Goldberg, J. L. How does an axon grow? *Genes Dev.* **17**, 941–958 (2003).
- Wang, Z. et al. KLF6 and STAT3 co-occupy regulatory DNA and functionally synergize to promote axon growth in CNS neurons. *Sci. Rep.* **8**, 12565 (2018).
- Wang, Z., Reynolds, A., Kirry, A., Nienhaus, C. & Blackmore, M. G. Overexpression of Sox11 promotes corticospinal tract regeneration after spinal injury while interfering with functional recovery. *J. Neurosci.* **35**, 3139–3145 (2015).
- Norsworthy, M. W. et al. Sox11 expression promotes regeneration of some retinal ganglion cell types but kills others. *Neuron* **94**, 1112–1120.e4 (2017).
- Luo, X. et al. Enhanced transcriptional activity and mitochondrial localization of STAT3 co-induce axon regrowth in the adult central nervous system. *Cell Rep.* **15**, 398–410 (2016).
- Belin, S. et al. Injury-induced decline of intrinsic regenerative ability revealed by quantitative proteomics. *Neuron* **86**, 1000–1014 (2015).
- Blackmore, M. G. et al. Kruppel-like Factor 7 engineered for transcriptional activation promotes axon regeneration in the adult corticospinal tract. *Proc. Natl Acad. Sci. USA* **109**, 7517–7522 (2012).
- Palmisano, I. et al. Epigenomic signatures underpin the axonal regenerative ability of dorsal root ganglia sensory neurons. *Nat. Neurosci.* **22**, 1913–1924 (2019).
- Li, S. et al. The transcriptional landscape of dorsal root ganglia after sciatic nerve transection. *Sci. Rep.* **5**, 16888 (2015).
- Dhara, S. P. et al. Cellular reprogramming for successful CNS axon regeneration is driven by a temporally changing cast of transcription factors. *Sci. Rep.* **9**, 14198 (2019).
- Hilton, B. J. & Bradke, F. Can injured adult CNS axons regenerate by recapitulating development? *Development* **144**, 3417–3429 (2017).
- Kidder, B. L., Yang, J. & Palmer, S. Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. *PLoS ONE* **3**, e3932 (2008).
- He, A., Kong, S. W., Ma, Q. & Pu, W. T. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc. Natl Acad. Sci. USA* **108**, 5632–5637 (2011).
- Liu, L. et al. Modeling co-occupancy of transcription factors using chromatin features. *Nucleic Acids Res.* **44**, e49–e49 (2015).
- Lee, Y. & Zhou, Q. Co-regulation in embryonic stem cells via context-dependent binding of transcription factors. *Bioinformatics* **29**, 2162–2168 (2013).
- Wang, J. et al. Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* **41**, D171–D176 (2013).
- Wang, J. et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* **22**, 1798–1812 (2012).
- Arlotta, P. et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* **45**, 207–221 (2005).
- Venkatesh, I., Mehra, V., Wang, Z., Califf, B. & Blackmore, M. G. Developmental chromatin restriction of pro-growth gene networks acts as an epigenetic barrier to axon regeneration in cortical neurons. *Dev. Neurobiol.* **78**, 960–977 (2018).
- Fulco, C. P. et al. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).
- Gerstein, M. B. et al. Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91–100 (2012).
- Kwon, A. T., Arenillas, D. J., Hunt, R. W. & Wasserman, W. W. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. *Genes Genomes Genet.* **2**, 987–1002 (2012).
- Moore, D. L. et al. KLF family members regulate intrinsic axon regeneration ability. *Science* **326**, 298–301 (2009).
- Califf, B. L., Maunze, B., Krueger, N. L., Simpson, M. T. & Blackmore, M. G. The application of CRISPR technology to high content screening in primary neurons. *Mol. Cell. Neurosci.* <https://doi.org/10.1016/j.mcn.2017.01.003> (2017).
- Simpson, M. T. et al. The tumor suppressor HHEX inhibits axon growth when prematurely expressed in developing central nervous system neurons. *Mol. Cell. Neurosci.* **68**, 272–283 (2015).
- Blackmore, M. G. et al. High content screening of cortical neurons identifies novel regulators of axon growth. *Mol. Cell. Neurosci.* **44**, 43–54 (2010).
- Fagoe, N. D., Attwell, C. L., Kouwenhoven, D., Verhaagen, J. & Mason, M. R. J. Overexpression of ATF3 or the combination of ATF3, c-Jun, STAT3 and Smad1 promotes regeneration of the central axon branch of sensory neurons but without synergistic effects. *Hum. Mol. Genet.* **24**, 6788–6800 (2015).
- Lerch, J. K., Martínez-Ondaro, Y. R., Bixby, J. L. & Lemmon, V. P. cJun promotes CNS axon growth. *Mol. Cell. Neurosci.* **59**, 97–105 (2014).
- Mehta, S. T., Luo, X., Park, K. K., Bixby, J. L. & Lemmon, V. P. Hyperactivated Stat3 boosts axon regeneration in the CNS. *Exp. Neurol.* **280**, 115–120 (2016).
- Wang, Z., Maunze, B., Wang, Y., Tsoulfas, P. & Blackmore, M. G. Global connectivity and function of descending spinal input revealed by 3D microscopy and retrograde transduction. *J. Neurosci.* **38**, 10566–10581 (2018).
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- Fink, K. L., López-Giráldez, F., Kim, I.-J., Strittmatter, S. M. & Cafferty, W. B. J. Identification of intrinsic axon growth modulators for intact CNS neurons after injury. *Cell Rep.* **18**, 2687–2701 (2017).
- Economou, M. N. et al. Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**, 79–84 (2018).
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

40. Wang, Z., Winsor, K., Nienhaus, C., Hess, E. & Blackmore, M. G. Combined chondroitinase and KLF7 expression reduce net retraction of sensory and CST axons from sites of spinal injury. *Neurobiol. Dis.* **99**, 24–35 (2017).
41. Robinson-Rechavi, M., Escriba Garcia, H. & Laudet, V. The nuclear receptor superfamily. *J. Cell Sci.* **116**, 585–586 (2003).
42. Yip, P. K. et al. Lentiviral vector expressing retinoic acid receptor  $\beta 2$  promotes recovery of function after corticospinal tract injury in the adult rat spinal cord. *Hum. Mol. Genet.* **15**, 3107–3118 (2006).
43. Hale, M. A. et al. The nuclear hormone receptor family member NR5A2 controls aspects of multipotent progenitor cell formation and acinar differentiation during pancreatic organogenesis. *Development* **141**, 3123–3133 (2014).
44. Stergiopoulos, A. & Politis, P. K. Nuclear receptor NR5A2 controls neural stem cell fate decisions during development. *Nat. Commun.* **7**, 12230 (2016).
45. Liu, Y. et al. A sensitized IGF1 treatment restores corticospinal axon-dependent functions. *Neuron* **95**, 817–833.e4 (2017).
46. Chandran, V. et al. A systems-level analysis of the peripheral nerve intrinsic axonal growth program. *Neuron* **89**, 956–970 (2016).
47. Puttagunta, R. et al. PCAF-dependent epigenetic changes promote axonal regeneration in the central nervous system. *Nat. Commun.* **5**, 3527 (2014).
48. Du, K. et al. Pten deletion promotes regrowth of corticospinal tract axons 1 year after spinal cord injury. *J. Neurosci.* **35**, 9754–9763 (2015).
49. Leibinger, M. et al. Transneuronal delivery of hyper-interleukin-6 enables functional recovery after severe spinal cord injury in mice. *Nat. Commun.* **12**, 391 (2021).
50. Kadoya, K. et al. Spinal cord reconstitution with homologous neural grafts enables robust corticospinal regeneration. *Nat. Med.* **22**, 479–487 (2016).
51. Tuszynski, M. H. & Steward, O. Concepts and methods for the study of axonal regeneration in the CNS. *Neuron* **74**, 777–791 (2012).
52. Chen, M. & Zheng, B. Axon plasticity in the mammalian central nervous system after injury. *Trends Neurosci.* **37**, 583–593 (2014).
53. Ewan, E. E., Carlin, D., Goncalves, T. M., Zhao, G. & Cavalli, V. Ascending dorsal column sensory neurons respond to spinal cord injury and downregulate genes related to lipid metabolism. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-79624-0> (2021).
54. Haffner, M. C., De Marzo, A. M., Meeker, A. K., Nelson, W. G. & Yegnasubramanian, S. Transcription-induced DNA double strand breaks: both oncogenic force and potential therapeutic target? *Clin. Cancer Res.* **17**, 3858–3864 (2011).
55. Sebastian, R. & Oberdoerffer, P. Transcription-associated events affecting genomic integrity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160288 (2017).
56. Onishi, K. et al. Genome stability by DNA polymerase  $\beta$  in neural progenitors contributes to neuronal differentiation in cortical development. *J. Neurosci.* **37**, 8444–8458 (2017).
57. Baleriola, J. et al. Increased neuronal death and disturbed axonal growth in the Polp-deficient mouse embryonic retina. *Sci. Rep.* **6**, 25928 (2016).
58. Krishnan, A. et al. A BRCA1-dependent DNA damage response in the regenerating adult peripheral nerve milieu. *Mol. Neurobiol.* **55**, 4051–4067 (2018).
59. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
60. Perte, M., Kim, D., Perte, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
61. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
62. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
63. Rouillard, A. D. et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* **2016**, baw100 (2016).
64. Shaner, N. C. et al. A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods* **10**, 407–409 (2013).
65. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
66. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

### Acknowledgements

This work was supported by grants from NINDS (5R01NS083983 and R21NS106309), The Craig Neilsen Foundation, The Bryon Riesch Paralysis Foundation, and computational allocations from NSF-XSEDE. We thank Erik Van Newenhizen for technical assistance. We acknowledge ENCODE consortia for generating the developmental time-series NGS datasets used in this study. We acknowledge the use of Biorender (academic license) to generate illustrations and graphics used in the manuscript.

### Author contributions

M.G.B. and I.V. conceived and planned the experiments. I.V., V.M., Z.W., M.T.S., E.E., A.C., Z.B., D.G., M.C., and G.O. carried out the experiments. M.B.M. and I.V. contributed to the interpretation of the results. M.B.M. and I.V. were involved in writing the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information


**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-22828-3>.

**Correspondence** and requests for materials should be addressed to I.V. or M.G.B.

**Peer review information** *Nature Communications* thanks Valeria Cavalli, Xiao-Ming Xu, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021