Published in final edited form as:

Nature. ; 486(7401): 74–79. doi:10.1038/nature11094.

Cis-regulatory control of corticospinal system development and evolution

Sungbo Shim[1], **Kenneth Y. Kwan**[1], **Mingfeng Li**[1], **Veronique Lefebvre**[2], and **Nenad Šestan**[1]

[1]Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA

[2]Department of Cell Biology, and Orthopaedic and Rheumatologic Research Center, Cleveland Clinic Lerner Research Institute, Cleveland, Ohio 44195, USA

Summary

The co-emergence of a six-layered cerebral neocortex and its corticospinal output system is one of the evolutionary hallmarks of mammals. However, the genetic programs that underlie their development and evolution remain poorly understood. Here we identify a conserved non-exonic element (E4) that acts as a cortex-specific enhancer for the nearby *Fezf2*, which is required for the specification of corticospinal neuron identity and connectivity. We find that SOX4 and SOX11 functionally compete with the repressor SOX5 in the trans-activation of E4. Cortex-specific double deletion of *Sox4* and *Sox11* leads to the loss of *Fezf2* expression and failed specification of corticospinal neurons and, independent of *Fezf2*, a *reeler*-like inversion of layers. We show evidence supporting the emergence of functional SOX binding sites in E4 during tetrapod evolution and their subsequent stabilization in mammals and possibly amniotes. These findings reveal that SOX transcription factors converge onto a *cis*-acting element of *Fezf2* and form critical components of a regulatory network controlling the identity and connectivity of corticospinal neurons.

The emergence and expansion of the neocortex in mammals has been crucial to the evolution of complex perceptual, cognitive, emotional and motor abilities[1–3]. The neocortex is organized into six layers based largely on the distinct subtypes of excitatory projection (or pyramidal) neurons and their patterns of connectivity[4–9]. Upper layer (L2-4) projection neurons form synaptic connections solely with other cortical neurons. In contrast, the majority of neurons in the deeper layers (L5 and L6) project to subcortical regions. Studies

Correspondence: nenad.sestan@yale.edu.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions

S.S., K.Y.K., and N.S. designed research; S.S. performed experiments; S.S., and K.Y.K. performed confocal imaging, M.L. analyzed co-expression and deep sequencing data, S.S., K.Y.K., M.L., and N.S. analyze the data, V.L. generated floxed *Sox4* and *Sox11* allele mice; N.S. conceived the study, S.S., K.Y.K. and N.S. wrote the manuscript. All authors discussed the results and implications and commented on the manuscript at all stages.

Competing Financial Interests

The authors declare no competing financial interests.

of laminar inversion in mice lacking the reelin (RELN) protein[10–14] have shown that neuron identity and connectivity are determined by birth order rather than laminar position, suggesting that neuronal specification and positioning are separately encoded.

The layer-specific pattern of connectivity is dependent on cortical areas. The long-range projections of L5 neurons in somatosensory-motor areas form the corticospinal (CS) system that directly connects the neocortex (Ncx) with various subcortical regions[15–22]. A major component of the system, the CS (or pyramidal) tract, descends through the brainstem and into the spinal cord to provide a high degree of direct control over the precise motor functions affected in many clinical conditions[21,22]. Despite these important functional implications, the genetic programs controlling CS system development and evolution remain unclear.

Phenotypic specification and evolution of neuronal circuits depend on precise regulation of the timing, location, and level of gene expression[23,24]. Transcriptional control via *cis*-regulatory elements has emerged as a crucial mechanism[25–29]. The *cis*-regulatory mechanisms underlying the specification of distinct neuronal cell types and circuits, however, remain poorly understood. Specification of CS neurons and the formation of the CS tract critically depend on *Fezf2* (or *Fezl*, *Zfp312*), which encodes a zinc-finger transcription factor (TF) highly enriched in early cortical progenitor cells and their deeplayer neuron progenies[30–35]. Inactivation of *Fezf2* disrupts the molecular specification of deep-layer neurons and the formation of corticofugal projections, including the CS tract, without affecting the inside-out pattern of neurogenesis and lamination[33–35]. Misexpression of *Fezf2* in L2-3 cortico-cortical projection neurons[35,36] or striatal interneurons[37] alters their molecular profile and induces ectopic subcerebral projections. These findings indicate that the tightly regulated transcription of *Fezf2* is likely critical to the proper specification of distinct types of cortical projection neurons.

In this study, we used bacterial artificial chromosome (BAC) engineering and genetic inactivation in mice to identify and characterize a cortex-specific *Fezf2* enhancer and its *trans*-regulators. We show that three SOX TFs converge onto the *Fezf2* enhancer to control CS system development via functional binding sites that emerged in tetrapods. We also found *Sox4* and *Sox11* to be required for *Fezf2*-independent regulation of cortical RELN expression and laminar organization. Thus, these findings reveal novel developmental genetic programs that control layer formation and CS neuron identity, and the regulatory mechanisms by which they may have evolved.

E4 enhancer controls cortical *Fezf2* expression

Previously, it has been shown that the BAC transgenic mouse harboring 200 kb of the mouse *Fezf2* locus and the *Gfp* reporter gene (*Fezf2-Gfp*[38]) recapitulates the spatio-temporal expression of endogenous *Fezf2* [39,40]. This indicated that the *cis*-regulatory elements required for cortical *Fezf2* expression are located within the BAC sequence. Based on the remarkable similarity in cortical expression pattern of *Fezf2* between mouse and human[39,41], we hypothesized that the regulatory elements are also highly evolutionarily conserved. Comparative sequence analysis revealed several conserved non-exonic elements (CNEEs)

within the *Fezf2-Gfp* BAC (Fig. 1a). To test whether the selected CNEEs regulate *Fezf2-Gfp* expression, we generated multiple lines of BAC transgenic mice in which one of the CNEEs within the *Fezf2-Gfp* BAC was deleted (ΔE1-4) (Fig. 1b; Methods). We analyzed GFP expression in whole mount brain preparations and tissue sections of transgenic mice at embryonic and postnatal ages (Fig. 1c-g, and Supplementary Fig. 1, 2). We found no change in ΔE1-3 mutants, which expressed GFP in Ncx and pontine CS axons similar to wildtype *Fezf2-GFP* mice (Fig. 1d-f). In contrast, the deletion of CNEE E4 (ΔE4), located 7.3 kbp downstream of the *Fezf2* transcription start site, resulted in a dramatic loss of GFP expression in the cortex, but not in the hypothalamus (Hyp) or olfactory bulb (OB) (Fig. 1g). Moreover, in ΔE4 mice, GFP-positive CS axon fascicles were lost from the ventral surface of the pons (Fig. 1g). Consistent with this finding, E4 was identified as an *in vivo* binding site of the enhancer-associated protein p300 (or EP300) in mouse embryonic forebrain[27]. Together, these results demonstrate that E4 is a *cis*-regulatory element acting as a cortex-specific enhancer of *Fezf2*.

Loss of *Fezf2* expression and CS axons in E4$^{-/-}$ mice

To directly test the function of the E4 enhancer, we deleted the E4 region, while leaving intact the entire *Fezf2* coding and proximal promoter regions, using homologous recombination in embryonic stem cells (Fig. 2a, b). Mice lacking E4 (E4$^{-/-}$) were viable, fertile, and without overt behavioral or motor phenotypes. Quantitative (q) RT-PCR analysis revealed that *Fezf2* expression is dramatically down-regulated in Ncx of these mice (Fig. 2c). Immunostaining of tissue sections of E4-null mice for PRKCG (or PKCγ) and L1CAM (or L1), two proteins expressed by CS neurons and their axons[11], revealed that CS axons were absent from the pons (Fig. 2g), similar to mutant mice with a cortex-specific deletion of *Fezf2* (*Fezf2$^{fl/fl}$;Emx1-Cre*) (Fig. 2f). Furthermore, the expression of *Bcl11b* (or *Ctip2*), a gene functioning downstream of *Fezf2*[33,36], was also down-regulated in the E4-null Ncx (Supplementary Fig. 3). Taken together, these results indicate that the E4 enhancer is required for neocortical *Fezf2* expression, molecular specification of L5/6 neurons, and the formation of CS tract.

SOX4 and SOX11 bind and activate the E4 enhancer

We have previously shown that SOX5, a SOXD member of the large family of SOX TFs, binds and represses the transcriptional activity of the E4 enhancer[39]. Although somewhat surprising, SOX5 itself is also required for CS tract formation independent of its repression of *Fezf2*[39]. Because different SOX TFs are known to compete for a common motif to mediate both activation and repression of regulatory elements[42–44], we hypothesized that other SOX members may activate the E4 enhancer, perhaps acting competitively against SOX5-mediated repression. Consistent with this hypothesis, our sequence analysis using MatInspector (Genomatix) revealed eight putative SOX binding sites in E4. To prioritize which SOX TFs may be good candidates as potential *trans*-regulators of the E4 enhancer, we searched for those that are most highly correlated in their spatio-temporal expression pattern with *FEZF2* using the www.humanbraintranscriptome.org database[45]. The highest correlated *SOX* genes were three members of the *SOXC* group (*SOX4*, *SOX11*, and *SOX12*) and *SOX5* (Supplementary Fig. 4). These TFs play crucial roles in regulating cell fate and

differentiation[42–44, 46], and a *de novo* deletion of *SOX11* was described in a patient with autism and intellectual disability[47]. Moreover, *Sox4* and *Sox11* act as transcriptional activators[45] and their expression patterns overlap with that of *Fezf2* in developing cortex (Supplementary Fig. 5).

To determine if SOX4 or SOX11 binds E4, we performed chromatin immunoprecipitation (ChIP)-PCR assays in Neuro-2a cells transiently expressing V5-tagged SOX4 or SOX11 (Methods). Anti-V5 antibodies precipitated E4, but not E1-3, DNA, confirming binding of SOX4 and SOX11 to the E4 enhancer (Fig. 3a). Moreover, recruitment of RNA polymerase II (Pol II) to E4 occurred in the presence of SOX4 or SOX11, which is consistent with increased transcriptional activity. To test the functional consequence of *SoxC* expression on E4, we expressed a luciferase reporter under the control of E4 (pGL4-E4) in Neuro-2a cells (Fig. 3b). Luciferase activity driven by the E4 enhancer was significantly increased by co-transfected *Sox4* or *Sox11*, but not *Sox12*.

To dissect which sequences within the E4 drive *Fezf2* expression, we generated four truncated versions of the E4 sequence (E4F1-4) (Fig. 3c). The luciferase activity of the E4F2 fragment was increased by SOX4 and SOX11, but not SOX12 (Fig. 3d). Using MatInspector, six potential SOX binding sites within the E4F2 sequence were predicted. To better define the precise basepairs with which SOXC proteins interact, we first tested whether the zebrafish E4F2 sequence, which exhibits 74.2% sequence identity with the mouse E4F2, is activated by SOXC proteins. Remarkably, the zebrafish E4F2 sequence was not activated by SOX11 (Supplementary Fig. 6), indicating that the SOXC-interacting sequences of mouse E4F2 lie within the sequences that are divergent in zebrafish. Of the predicted SOX binding sites in mouse E4F2, three putative SOX binding sites (SB1-3) are absent from the zebrafish E4F2. To test the function and specificity of these sites, we generated mutant versions of E4F2 by substituting SB1, SB2, or SB3 with the zebrafish sequence (Fig. 3e). Only mutations of the mouse SB2 site significantly attenuated the ability of SOX4 and SOX11 to activate the luciferase reporter gene, indicating that this site is crucial to species differences in SOXC-mediated transactivation of E4. Next, we assessed the ability of SOX4 and SOX11 to bind the SB2 site *in vitro* using an electrophoretic mobility shift assay (EMSA). We synthesized V5-tagged SOX4 and a higher-affinity, but equally specific, truncated form of SOX11 (SOX11(1-276aa)), since the binding affinity of native SOX11 is weak[43]. Biotinylated SB2 DNA was shifted in the presence of SOX4 or SOX11(1-276aa) and supershifted by an anti-V5 antibody (Fig. 3f), but not when excess unlabeled or mutated SB2 DNA was used. Therefore, SOX4 and SOX11 directly bind to and activate the transcription activity of E4 via SB2.

To test whether SOX4 and SOX11 functionally compete with the repressor SOX5 in the trans-activation of E4, we used the pGL4-E4 luciferase assay. Luciferase activity was significantly increased with increasing amounts of co-transfected *Sox11* and to a lesser extent, *Sox4*, whereas increased amounts of *Sox5* significantly decreased luciferase activity (Fig. 3h, i). Taken together, our sequence analysis and assays, both *in vitro* and *in vivo*, demonstrate that SOX4 and SOX11 functionally compete with SOX5 repression to activate *Fezf2* transcription via direct binding to sites within E4.

Loss of *Fezf2* expression and CS axons in *Sox4*$^{-/-}$;*Sox11*$^{-/-}$ mice

To test the phenotypic consequences of *Sox4* and *Sox11* inactivation, we generated cortex-specific deletions of *Sox4* and *Sox11* using the *Emx1-Cre* line (Methods) to circumvent the prenatal lethality associated with constitutive deletions of the two genes[44]. Single conditional knockout mice (cKO: *Sox4*$^{fl/fl}$;*Sox11*$^{fl/+}$;*Emx1-cre* and *Sox4*$^{fl/+}$;*Sox11*$^{fl/fl}$;*Emx1-cre*) were both viable and fertile, whereas the conditional double knockout mice (cdKO: *Sox4*$^{fl/fl}$;*Sox11* $^{fl/fl}$;*Emx1-cre*) died within the first postnatal week (data not shown). Inspection of the cerebrum during the first postnatal week revealed no obvious gross defects in single *Sox4* or *Sox11* cKO animals (Supplementary Fig. 7). However, in the absence of both *Sox4* and *Sox11*, we observed a reduction in the size of the cerebral hemispheres and OB.

To assess the requirement of *Sox4* and *Sox11* in E4-mediated activity, we transfected primary cortical neurons cultured from heterozygous littermate control or *Sox4; Sox11* cdKO embryos with control or E4-containing luciferase constructs (Fig. 4a). In control neurons, the presence of E4 increased luciferase activity by 2.5±0.3 fold (P=1.9×10^{-4}, one-tailed Student's *t*-test). This increase was abolished in neurons culture from *Sox4; Sox11* cdKO (1.1±0.1 fold; P=0.866), indicating that the two SOXC TFs are major activators of the enhancer in cortical neurons. Next, using qRT-PCR, we found a significant reduction in neocortical *Fezf2* expression in cdKO compared to control or single cKO littermates at postnatal day 0 (P0) (Fig. 4b). In addition, immunostaining for L1CAM and PRKCG revealed a dramatic loss of CS axons in *Sox4; Sox11* cdKO mice but not single cKO littermates at P0 and P6 (Fig. 4c, d, Supplementary Fig. 8), whereas the organization of other brainstem tracts was not affected (Fig. 3c, d). Because the absence of L1CAM and PRKCG staining could reflect their down-regulated expression in the cdKO, as opposed to loss of CS axons, we confirmed the absence of the CS tract using the CRE-responsive *CAG-Cat-Gfp* transgenic line to express GFP in all cortical projection neurons and their axons[39,40] (Supplementary Fig. 9). Two additional observations indicate that the loss of *Fezf2* and CS axons in these mice was not due to an absence of L5 neurons. First, only a moderate increase in cell death was detected in the motor-sensory areas, from which CS axons originate (Supplementary Fig. 10). Second, many BCL11B-immunopositive L5 neurons were present in the cdKO motor-sensory areas, albeit lightly immunostained (Supplementary Fig. 11), indicating that neurons that would normally give rise to the CS tract were present. Analysis of additional laminar markers further revealed an inversion of cortical layers similar to the *reeler* mutant mouse[10–12] (Supplementary Fig. 11), indicating that *Sox4* and *Sox11* are also required for proper laminar positioning of neurons. Previous studies have shown that the CS tract is present in *reeler* mice[12], indicating that this laminar phenotype occurs independently of *Fezf2* and probably in response to defects in the RELN signaling pathway[13,14]. In support of this possibility, RELN was absent from cdKO Ncx (Supplementary Fig. 12). Thus, the combined deletion of *Sox4* and *Sox11* leads to defects in laminar position of neurons and molecular specification and connectivity of CS neurons.

Functional and evolutionary implications of E4 sequence variations

The differences in SOXC-mediated transcription between the mouse and zebrafish E4F2 suggest that species differences in the E4F2 sequence have functional implications. Analysis of SB1-3 sites within the E4F2 sequence in 23 species revealed that SB2 is conserved in all analyzed mammals and the two available non-mammalian amniotes (chicken and lizard) (Supplementary Fig. 13). To investigate the functional consequences, we used luciferase assays in Neuro-2a cells to analyze the activity of E4F2 sequence from different vertebrates (Fig. 5a, b) in response to SOX11, a more powerful *trans*-activator than SOX4 (Fig. 3). Of the constructs containing an E4F2 sequence of a mammal (human, chimpanzee, macaque, or mouse) or non-mammalian amniote (chicken), which have conserved SB1-3 sequences, the luciferase reporter activity was strongly increased following co-transfection with *Sox11*. The reporter activity of a non-amniote tetrapod (*Xenopous*) E4F2 construct was moderately increased compared to the empty or zebrafish luciferase plasmid, but not as high as the increase with mammalian constructs ($P = 3.9\times10^{-16}$, one-tailed Student's *t*-test) (Fig. 5b). In contrast, the reporter construct containing the zebrafish E4F2 sequence, which has a highly divergent SB2 sequence from mammals, exhibited only basal level activity similar to the empty control luciferase plasmid. To confirm that this is dependent on sequence variations between species, we mutated zebrafish E4F2 (ZeE4F2-m2) by "murinizing" its SB1, SB2, or SB3 sequence. Reporter activity of the murinized zebrafish SB2 (ZeE4F2-m2), but not SB1 or SB3, was robustly induced by SOX11 (Fig. 5c). ZeE4F2-m2, however, was insufficient to restore the level of expression observed in the wildtype mouse sequence, suggesting the potential contribution of additional sequences. Thus, evolutionary differences in the E4F2 sequence, especially within SB2, are directly related to functional differences in the ability of SOX11 to activate this regulatory element.

Next, to investigate the functional consequences of species differences in the FEZF2 coding sequence, we tested the ability of the zebrafish *Fezf2* to rescue the effects of mouse *Fezf2* deficiency. We electroporated *in utero* the neocortical wall of the floxed *Fezf2* (*Fezf2*$^{fl/fl}$) mice[40] with *Cre* and CRE-responsive *Gfp*-expressing constructs with or without zebrafish *Fezf2* at embryonic day (E) 12.5 (Fig. 5d). Analysis of electroporated mice at P0 revealed that the zebrafish *Fezf2* was sufficient to rescue the formation of the CS tract in *Fezf2*-null mouse neurons (Fig. 5e). Taken together, these results suggest that sequence variations in the E4 enhancer, but not the *Fezf2* coding region, imparted an essential role of *Fezf2* in CS tract evolution.

Discussion

Our data provide critical insight into the genes and regulatory components controlling CS system development, centering on the *cis* regulation of *Fezf2*. We show that genetic inactivation of E4 results in compartmentalized phenotypic effects largely limited to the loss of cortical *Fezf2* expression and CS system. The spatio-temporal dynamics of cortical *Fezf2* is further controlled by SOX5 and TBR1, two TFs expressed post-mitotically in projection neurons[39,40,46, 48, 49]. Upon its initial activation in early cortical progenitors, *Fezf2* is highly expressed in deep-layer neurons during early corticogenesis but subsequently repressed in L6 neurons by SOX5 and TBR1 to create a postnatal L5-enriched pattern. The earlier

expression of SOX4 and SOX11 is consistent with their role in *Fezf2* activation prior to the L6 upregulation of SOX5, a functional competitor. Despite what is known about the function and regulation of *Fezf2*, a number of key questions have yet to be addressed. First, additional, as yet untested, transactivators and regulatory elements likely contribute to *Fezf2* regulation in a context-dependent manner. These include the mediators of *Fezf2* repression in L2-4 projection neurons. Second, the distinct, and seemingly incongruous, effects of SOX5 in repressing *Fezf2* but being independently required for CS tract formation[39] have not been resolved. Third, the direct transcriptional targets of FEZF2 in CS neurons remain largely unknown.

We also show that SOX4 and SOX11 play a crucial role in regulating RELN expression and the inside-out pattern of cortical layer formation, independent of E4 or *Fezf2* and likely involving interactions with distinct regulatory elements. Moreover, SOX4 and SOX11 have additional roles, as mice lacking both genes exhibit smaller cortex and OB, and increased cell death. Thus, SOX4 and SOX11 have pleiotropic functions, which are likely mediated by distinct regulatory elements and downstream target genes that are involved in multiple developmental processes.

Our results indicate that, following their emergence in tetrapods, functional SOX binding sites have retained high conservation via purifying selection in mammals and some amniotes, thus directly linking species variations in regulatory sequences to functional outcomes. E4 sequence substitutions in SB2 may constitute an evolutionary turning point for *Fezf2* function during forebrain development, possibly facilitating the formation of descending telencephalic pathways including the CS system. Whereas minor projections from the ventral (subpallial) telencephalon to the spinal cord are present in amphibians[2], direct dorsal telencephalo-spinal projections resembling the CS tract have only been reported in mammals and some birds[50]. We propose a scenario in which the concurrent emergence of the described regulatory mechanisms and direct telencephalo-spinal projections in early amniotes, together with subsequent changes in genetic programs driving the patterning and expansion of a six-layered dorsal pallium, made possible the evolution of the CS system in mammals.

METHODS SUMMARY

Selected CNEE sequences were deleted from a *Fezf2-Gfp* BAC by recombineering and GFP expression from modified BACs was analyzed by transgenesis. To confirm the requirement of E4 enhancer for *Fezf2* expression, a germ-line E4 deletion mutant was generated. Putative E4 *trans*-activators SOX4 and SOX11, identified by E4 sequence and co-expression analyses, were tested using ChIP and luciferase reporter assays. To analyze the requirement of *Sox4* and *Sox11* for *Fezf2* expression and cortical development, we generated cortex-specific single and double *Sox4* and *Sox11* null mice. Complete materials and methods are described in supplemental information.

Full Methods and associated references are available in the online version of the paper at www.nature.com/nature.

METHODS

Animals

All experiments were performed in accordance with a protocol approved by Yale University's Committee on Animal Research. The generation of mice with the floxed *Sox4* and *Sox11* alleles (*Sox4*$^{fl/fl}$ and *Sox11*$^{fl/fl}$, respectively) was described elsewhere[44,51]. The *CAG-Cat-Gfp* and *Emx1-Cre* P1 artificial chromosome transgenic mice were a generous gift from Melissa Colbert[52] and Takuji Iwasato[53], respectively. The *Fezf2-Gfp* BAC transgenic mouse was generated by the GENSAT project [38].

Generation of BAC transgenic mouse

To generate BAC reporter constructs harboring deletions of putative enhancer elements (E1-4) in the mouse *Fezf2-Gfp* BAC (RP23-141E17-*Gfp*; the GENSAT project [38]), homology arms A and B flanking the sequences targeted for deletion (E1, Chr14: 13204475-13204944, 470bp; E2, Chr14: 13183314-13183991, 678bp; E3, Chr14: 13180797-13181283, 487bp; E4, Chr14: 13170100-13170960, 861bp) were amplified by PCR, and cloned into the PL451 shuttle vector. SW102 *E. coli* transformed with the *Fezf2-Gfp* BAC was induced for recombineering at 42 °C and electroporated with the linearized shuttle vector. After neomycin selection, *Flpase* expression was induced by arabinose to excise the neomycin cassette[54]. Founder lines were confirmed by PCR using a primers corresponding to sequences from E1-4 (Supplementary Table 1). *Fezf2-Gfp* lines were maintained on a C57BL6/J background and all animals were housed under identical conditions. The intactness of the integrated BACs was confirmed by PCR and deep sequencing of genomic DNA.

Generation of E4-null mice

The targeting vector for the deletion of the E4 enhancer was constructed by a recombineering-based method with an unmodified BAC clone (RP23-141E17). A 6627bp fragment containing the E4 enhancer was retrieved from the BAC and inserted into the PL253 vector by recombination in SW102 bacteria. The removal of the E4 sequence was achieved using the strategy described above for the generation of the ΔE4 BAC transgenic mouse. ES cells (C57BL6/J) were electroporated with the resulting targeting construct and expanded under positive (pGK-neo, from PL451 vector) and negative (*MC1*-TK, from the PL253 vector) selection. Correctly recombined clones were identified by long distance-PCR, confirmed by sequencing, and used for blastocyst injection and subsequent generation of mice with the targeted allele. The neomycin cassette was subsequently removed by flippase-mediated recombination by breeding with *Flp* mice. The mice were genotyped using the P1/P2 primer set (400bp) for the wildtype allele and the P1/P3 primer set (205bp) the E4 deletion allele (Supplementary Table 1).

Immunohistochemistry

Postnatal day 0 (P0) and P6 brains were fixed by immersion in 4% paraformaldehyde (PFA) overnight at 4 °C and sectioned using a vibratome (Leica). Immunostaining was performed as described[20]. The following primary antibodies were used: anti-L1CAM (rat, 1:300,

Millipore), anti-PRKCG (rabbit, 1:300, Santa Cruz), anti-GFP (chicken, 1:3000, Abcam), anti-CUX1 (rabbit, 1:150, Santa Cruz), anti-SATB2 (mouse, 1:200, Genway), anti-BCL11b (rat, 1:250, Santa Cruz), anti-ZFPM2 (rabbit, 1:250, Santa Cruz), anti-RELN (mouse, 1:300, Millipore) and anti-CASP3, active (rabbit, 1:200, Millipore). We tested a total of 12 anti-SOX4 (Abcam, ab52043, ab86809, ab90696; Abgent, AP2045a, AP2045c; LS Bioscience, LS-B3520; Millipore, AB5803, AB10537; Pierce, PA1-38638, PA1-38639; Santa Cruz Biotechnology, sc-17326; Sigma, HPA029901) and 10 anti-SOX11 (Abcam;ab42853; Aviva Systems Biology, ARP33328, ARP38235; LS Bioscience, LS-B1567, LS-C10306; Millipore; AB9090; Pierce, PA5-19728; Santa Cruz Biotechnology, sc-20096, sc-17347; Sigma, HPA000536) commercially available antibodies on western blots and immunostaining. Three antibodies for each SOX4 and SOX11 recognized a single band of the expected size on western blot but none of them were suitable for immunostaining or ChIP assay. In addition, a few antibodies exhibited nuclear immunostaining, but also had strong background staining of tissue sections of the knockout brain, so we concluded that they are not specific.

Plasmids

For expression of *SoxCs* and *Fezf2*, full-length cDNAs (mouse *Sox4*, BC052736; mouse *Sox11*, BC062095; mouse *Sox12*, BC067019, zebrafish *Fezf2*, BC085677) were inserted into pCAGEN. For ChIP and EMSA, PCR amplified products (for mouse *Sox4*, *Sox11* full-length cDNA and *Sox11*(aa 1-276)) were inserted into pcDNA3.1/V5-His TOPO vector (Invitrogen). For luciferase reporter plasmids, PCR amplified products (for E1-4, E4F1-4, and E4F2 of six species) and annealed 190bp complementary synthetic oligonucleotides (for MoE4F2-m1-3, ZeE4F2-m1-3) were inserted into pGL4 (Promega). The sequences of the PCR primers and synthetic oligonucleotides are listed in Supplementary Table 1.

Gene expression analysis

To show which genes in the SOX family are highly correlated with human *FEZF2*, we used the data generated by Kang et al., 2011 (ref. 45) for pairwise Pearson comparisons. This dataset is generated by exon arrays and is available from the www.humanbraintranscriptome.org website or NCBI GEO with the accession number GSE25219. It covers 16 brain and cortical regions over 15 periods, ranging in age from embryonic development to late adulthood. The evaluation of gene expression in each region and each period was detailed in the study [45]. In this study, we averaged gene expression values for each of 11 neocortical regions/area and 15 developmental periods represented in the dataset, and then performed pairwise Pearson correlation analysis for *FEZF2* and the members of the SOX family. Totally, 18 *SOX* genes were represented in the database. The correlation analysis was shown by heatmap and the correlation coefficients were indicated in each cell (Supplementary Fig. 4).

Mouse *Fezf2, Sox4, Sox5, Sox11, and Sox12 in situ* hybridization images were obtained from the Genepaint database (www.genepaint.org)[55].

Chromatin immunoprecipitation

After testing available anti-SOX4 and anti-SOX11 antibodies and finding none to be specific and suitable for chromatin immunoprecipitation (ChIP) assay, we performed ChIP with Neuro-2a cells were transfected with *pcDNA3-Sox4-V5* or *pcDNA3-Sox11-V5* plasmid using Lipofectamine 2000 (Invitrogen). At 36 hrs after transfection, cells were cross-linked in 1% formaldehyde for 10 min at 37 °C, and processed for ChIP assay using the EZ-ChIP kit (Millipore) according to manufacturer's instructions. Capture of the DNA fragments was tested by PCR using primers E1-F, 5'- TGGAGAGAAGGCCAACAAAC -3', E1-R, 5'- GCTGGGGATGGAGAAGAATA -3'; E2-F, 5'- TCACCAAAGCGCCTTTTTAT -3', E2-R, 5'- GTAAGCGGACATGCCATTTT -3'; E3-F, 5'- TGACTTTCCCCAGCCTTCTA -3', E3-R, 5'- CAACAGCTCACCCACACAAT -3'; E4-F, 5'- ATGCCTAGCCCCAAAGAAAT -3', and E4-R, 5'- TTAACTCCCCCTTTGGCTCT -3'.

Electrophoretic Mobility Shift Assay

Double strand DNA probes were generated by annealing with complementary single strand oliogonucleotides. Each DNA probe was biotin-end labeled with Klenow enzyme and then purified using the G-50 Sephadex Column (Roche). The labeled and un-labeled (cold probe; 100 fold) probes were incubated with SOX4-V5, SOX11(1-276aa)-V5 which produced using TNT Quick coupled transcription/translation kit (Promega) in a DNA-binding buffer (75mM NaCl, 1mM EDTA, 1mM DTT, 10mM Tris-HCl(pH 7.5), 6% Glycerol, 2ug BSA, 16ng poly (dI-dC), 0.1ug Salmon sperm DNA) at room temperature for 30 min. For supershift assay, anti-V5 antibody (Invitrogen) were added into the binding reaction and incubated for additional 10 min. Shift assay was carried out in 6% polyacrylamide gel in 0.5x TBE buffer. Subsequently, the labeled DNA was transferred on a Nylon membrane (Amersham Biosciences). For detection, we using the Chemiluminescent Nucleic Acid Detection Module (Thermo Scientific) carried out according to manufacturers' recommendations. All primer sequences were described in the Supplementary Table1.

Luciferase assays

Neuro-2a cells were transfected using Lipofectamine 2000 (Invitrogen) with one of pCAG-*Sox5*, pCAG-*Sox4*, pCAG-*Sox11*, pCAG-*Sox12* or empty pCAGEN, together with one of the pGL4 (Promega) luciferase vectors generated with enhancer sequences as described above. A renilla luciferase plasmid (pRL, Promega) was co-transfected to control for transfection efficiency. The luciferase assays were performed 48hrs after transfection using the dual-luciferase kit (Promega) according to manufacturer's instructions. Primary cortical neurons were prepared from E14.5 heterozygous littermate control (*Sox4*$^{fl/+}$*;Sox11*$^{fl/+}$*;Emx1-cre* and *Sox4*$^{fl/+}$*;Sox11*$^{fl/+}$) *and Sox4;Sox11* cdKO (*Sox4*$^{fl/fl}$*;Sox11*$^{fl/fl}$*;Emx1-cre*) cortices and transfected with pGL4 or pGL4-E4 with pRL using the AMAXA Mouse Neuron Nucleofector Kit (VPG-1001; Lonza). At 48 hr after transfection, the luciferase assays were performed as described above.

In utero electroporation

Plasmid DNA (4 μg/μl) mixture (pCAGEN-*Cre*; pCALNL-*Gfp* or pCAGEN-*Cre*; pCALNL-*Gfp*; pCAGEN-*ZeFezf2*) was injected into the lateral ventricles of embryonic mice at

embryonic day 12.5 (E12.5) and transferred into the cells of ventricular zone by electroporation (five 50ms pulses of 40V at 950ms intervals) as described elsewhere[35, 39]. Brains and tissue sections of electroporated animals were analyzed for GFP expression after fixation with 4% PFA at postnatal day 0 (P0).

Quantitative RT-PCR

Neocortex was dissected from P0 brain and RNA was isolated using the RNeasy kit (Qiagen). QRT-PCR was performed using primers listed in Supplementary Table 1 and the SYBR FAST qPCR kit (KAPA Biosystems) according to the manufacturer's instruction. Gene expression levels were normalized to *Gapdh* expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank W. Han, Y. Imamura Kawasawa, D. Liu, and T. Nottoli for technical help;, A. Giraldez and A.M.M. Sousa for reagents; and the Sestan laboratory for discussions. This work was supported by the National Institutes of Health (NS054273, MH081896), March of Dimes, and McDonnell Scholar Award (N.S.).

References

1. Northcutt RG, Kaas JH. The emergence and evolution of mammalian neocortex. Trends Neurosci. 1995; 18:373–379. [PubMed: 7482801]
2. Nieuwenhuys, R.; ten Donkelaar, HJ.; Nicholson, C. The Central Nervous System of Vertebrates. Springer-Verlag; 1997.
3. Butler AB, Reiner A, Karten HJ. Evolution of the amniote pallium and the origins of mammalian neocortex. Ann. NY. Acad. Sci. 2011; 1225:14–27. [PubMed: 21534989]
4. O'Leary DDM, Koester SE. Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. Neuron. 1993; 10:991–1006. [PubMed: 8318235]
5. Rash BG, Grove EA. Area and layer patterning in the developing cerebral cortex. Curr. Opin. Neurobiol. 2006; 16:25–34. [PubMed: 16426837]
6. Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD. Neuronal subtype specification in the cerebral cortex. Nature Rev. Neurosci. 2007; 8:427–437. [PubMed: 17514196]
7. Leone DP, Srinivasan K, Chen B, Alcamo E, McConnell SK. The determination of projection neuron identity in the developing cerebral cortex. Curr. Opin. Neurobiol. 2008; 18:28–35. [PubMed: 18508260]
8. Hansen DV, Rubenstein JL, Kriegstein AR. Deriving excitatory neurons of the neocortex from pluripotent stem cells. Neuron. 2011; 70:645–660. [PubMed: 21609822]
9. Kwan KY, Sestan N, Anton ES. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. Development. 2012; 139 in press.
10. Caviness VS, Sidman RL. Time of origin of corresponding cell classes in cerebral-cortex of normal and *reeler* mutant mice – Autoradiographic analysis. J. Comp. Neurol. 1973; 148:141–151. [PubMed: 4700506]
11. Steindler DA, Colwell SA. *Reeler* mutant mouse: maintenance of appropriate and reciprocal connections in the cerebral cortex and thalamus. Brain Res. 1976; 113:386–393. [PubMed: 953743]
12. Terashima T. Anatomy, development and lesion-induced plasticity of rodent corticospinal tract. Neurosci. Res. 1995; 22:139–161. [PubMed: 7566696]

13. Bar I, de Rouvroit CL, Goffinet AM. The evolution of cortical development. An hypothesis based on the role of the Reelin signaling pathway. Trends Neurosci. 2000; 23:633–638. [PubMed: 11137154]
14. Rice DS, Curran T. Role of the Reelin signaling pathway in central nervous system development. Annu. Rev. Neurosci. 2001; 24:1005–1039. [PubMed: 11520926]
15. Joosten EAJ, Bar DPR. Axon guidance of outgrowing corticospinal fibres in the rat. J. Anat. 1999; 194:15–32. [PubMed: 10227663]
16. Martin JH. The corticospinal system: From development to motor control. Neuroscientist. 2005; 11:161–173. [PubMed: 15746384]
17. Canty AJ, Murphy M. Molecular mechanisms of axon guidance in the developing corticospinal tract. Prog. Neurobiol. 2008; 85:214–235. [PubMed: 18378059]
18. Lemon RN. Descending pathways in motor control. Annual Review of Neuroscience. 2008; 31:195–218.
19. Rathelot J-A, Strick PL. Subdivisions of primary motor cortex based on cortico-motoneuronal cells. Proc. Natl Acad. Sci. USA. 2009; 106:918–923. [PubMed: 19139417]
20. Nudo RJ, Masterton RB. Descending pathways to the spinal-cord. IV. Some factors related to the amount of cortex devoted to the corticospinal tract. J. Comp. Neurol. 1990; 296:584–597. [PubMed: 2113541]
21. ten Donkelaar HJ, et al. Development and malformations of the human pyramidal tract. J Neurol. 2004; 251:1429–1442. [PubMed: 15645341]
22. Eyre JA. Corticospinal tract development and its plasticity after perinatal injury. Neurosci. Biobehavi. Revi. 2007; 31:1136–1149.
23. Jessell TM. Neuronal specification in the spinal cord: Inductive signals and transcriptional codes. Nature Rev. Genet. 2000; 1:20–29. [PubMed: 11262869]
24. Hobert O, Carrera I, Stefanakis N. The molecular and gene regulatory signature of a neuron. Trends Neurosci. 2010; 33:435–445. [PubMed: 20663572]
25. Wray GA. The evolutionary significance of *cis*-regulatory mutations. Nature Rev. Genet. 2007; 8:206–216. [PubMed: 17304246]
26. Carroll SB. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell. 2008; 134:25–36. [PubMed: 18614008]
27. Visel A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457:854–858. [PubMed: 19212405]
28. Davidson EH. Emerging properties of animal gene regulatory networks. Nature. 2010; 468:911–920. [PubMed: 21164479]
29. Williamson I, Hill RE, Bickmore WA. Enhancers: From developmental genetics to the genetics of common human disease. Dev. Cell. 2011; 21:17–19. [PubMed: 21763601]
30. Hashimoto H, et al. Expression of the zinc finger gene fez-like in zebrafish forebrain. Mech. Dev. 2000; 97:191–195. [PubMed: 11025224]
31. Matsuo-Takasaki M, Lim JH, Beanan MJ, Sato SM, Sargent TD. Cloning and expression of a novel zinc finger gene *Fez* transcribed in the forebrain of Xenopus and mouse embryos. Mech. Dev. 2000; 93:201–204. [PubMed: 10781957]
32. Inoue K, Terashima T, Nishikawa T, Takumi T. *Fez1* is layer-specifically expressed in the adult mouse neocortex. Eur. J. Neurosci. 2004; 20:2909–2916. [PubMed: 15579145]
33. Molyneaux BJ, Arlotta P, Hirata T, Hibi M, Macklis JD. *Fez1* is required for the birth and specification of corticospinal motor neurons. Neuron. 2005; 47:817–831. [PubMed: 16157277]
34. Chen B, Schaevitz LR, McConnell SK. *Fez1* regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc. Natl Acad. Sci. USA. 2005; 102:17184–17189. [PubMed: 16284245]
35. Chen JG, Rasin MR, Kwan KY, Sestan N. *Zfp312* is required for suhcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. Proc. Natl Acad. Sci. USA. 2005; 102:17792–17797. [PubMed: 16314561]

36. Chen B, et al. The *Fezf2-Ctip2* genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. Proc. Natl Acad. Sci. USA. 2008; 105:11382–11387. [PubMed: 18678899]
37. Rouaux C, Arlotta P. *Fezf2* directs the differentiation of corticofugal neurons from striatal progenitors in vivo. Nature Neurosci. 2010; 13:1345–1347. [PubMed: 20953195]
38. Gong SC, et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature. 2003; 425:917–925. [PubMed: 14586460]
39. Kwan KY, et al. SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. Proc. Natl Acad. Sci. USA. 2008; 105:16021–16026. [PubMed: 18840685]
40. Han WQ, et al. TBR1 directly represses *Fezf2* to control the laminar origin and development of the corticospinal tract. Proc. Natl Acad. Sci. USA. 2011; 108:3041–3046. [PubMed: 21285371]
41. Fertuzinhos S, et al. Selective depletion of molecularly defined cortical interneurons in human holoprosencephaly with severe striatal hypoplasia. Cereb. Cortex. 2009; 19:2196–2207. [PubMed: 19234067]
42. Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J. The establishment of neuronal properties is controlled by Sox4 and Sox11. Genes Dev. 2006; 20:3475–3486. [PubMed: 17182872]
43. Dy P, et al. The three SoxC proteins-Sox4, Sox11 and Sox12-exhibit overlapping expression patterns and molecular properties. Nucleic Acids Res. 2008; 36:3101–3117. [PubMed: 18403418]
44. Bhattaram P, et al. Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. Nat. Commun. 2010; 1
45. Kang HJ, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011; 478:483–489. [PubMed: 22031440]
46. Lai T, et al. SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. Neuron. 2008; 57:232–247. [PubMed: 18215621]
47. Lo-Castro A, et al. Deletion 2p25.2: A cryptic chromosome abnormality in a patient with autism and mental retardation detected using aCGH. Eur. J. Med. Genet. 2009; 52:67–70. [PubMed: 18992374]
48. Bedogni F, et al. *Tbr1* regulates regional and laminar identity of postmitotic neurons in developing neocortex. Proc. Natl Acad. Sci. USA. 2010; 107:13129–13134. [PubMed: 20615956]
49. McKenna WL, et al. *Tbr1* and *Fezf2* regulate alternate corticofugal neuronal identities during neocortical development. J. Neurosci. 2011; 31:549–564. [PubMed: 21228164]
50. Wild JM, Williams MN. Rostral wulst in passerine birds. I. Origin, course, and terminations of an avian pyramidal tract. J. Comp. Neurol. 2000; 416:429–450. [PubMed: 10660876]

References

51. Penzo-Mendez A, Dy P, Pallavi B, Lefebvre V. Generation of mice harboring a Sox4 conditional null allele. Genesis. 2007; 45:776–780. [PubMed: 18064674]
52. Kawamoto S, et al. A novel reporter mouse strain that expresses enhanced green fluorescent protein upon Cre-mediated recombination. FEBS Lett. 2000; 470:263–268. [PubMed: 10745079]
53. Iwasato T, et al. Dorsal telencephalon-specific expression of Cre recombinase in PAC transgenic mice. Genesis. 2004; 38:130–138. [PubMed: 15048810]
54. Liu PT, Jenkins NA, Copeland NG. A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res. 2003; 13:476–484. [PubMed: 12618378]
55. Visel A, Thaller C, Eichele G. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res. 2004; 32:D552–D556. [PubMed: 14681479]

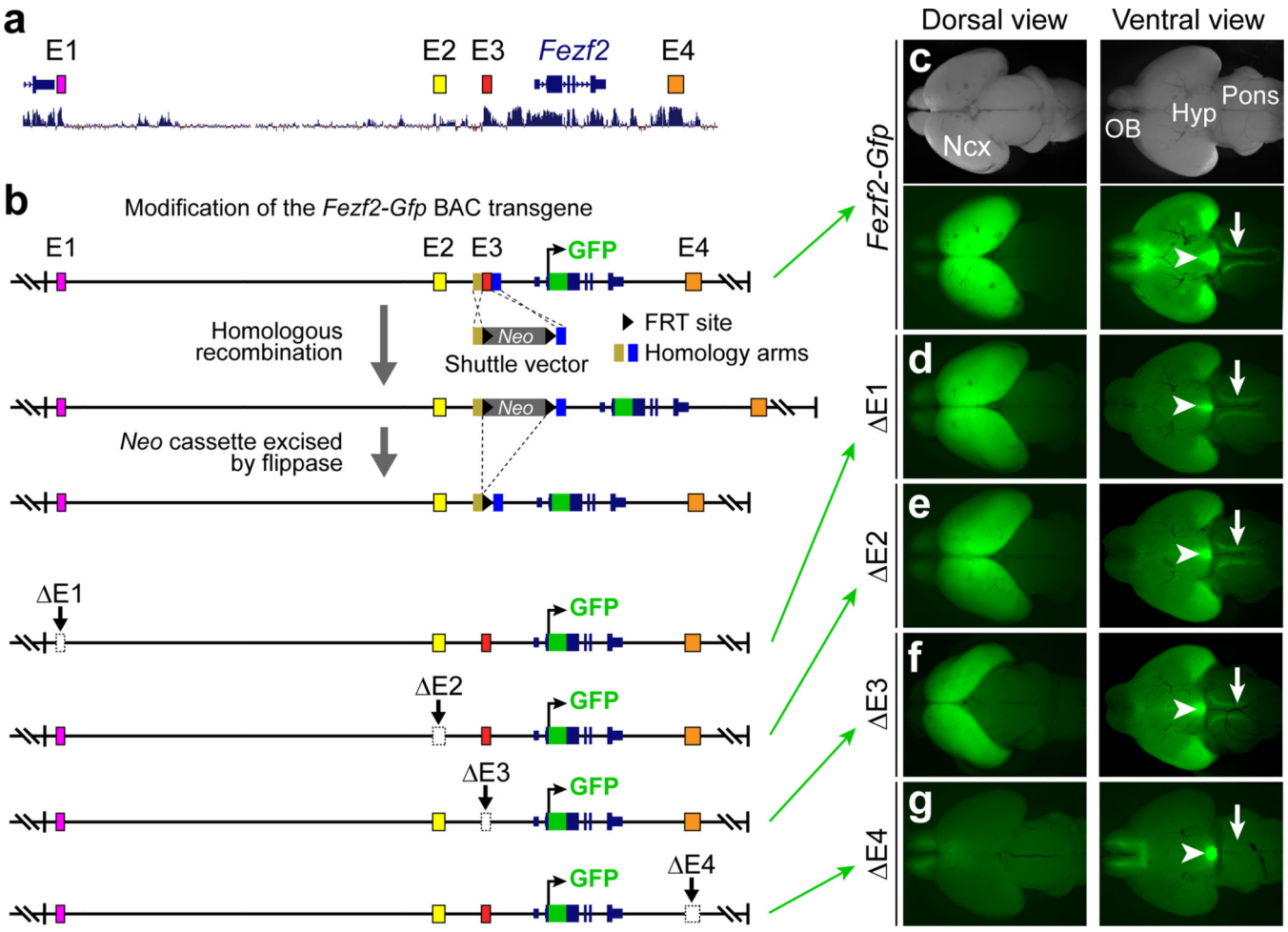

Fig. 1. Identification of a cortex-specific *Fezf2* enhancer

a, The locations of CNNEs (E1-4) analyzed in this study are indicated. **b,** Deletion of each CNEE from the *Fezf2-Gfp* BAC and transgenesis. A positive selection neomycin-cassette (*Neo*) flanked by homology arms was inserted by homologous recombination, resulting in the deletion of each CNNE. After the removal of *Neo* by flippase, the modified BACs (ΔE1-4) were used for transgenesis. **c-g,** Whole-mount brains of P0 control *Fezf2-Gfp* (**c**) and founder mutant (**d-g**) transgenic mice. $n \geq 3$ founders per mutant line. In *Fezf2-Gfp* mice (**c**), GFP was expressed in the neocortex (Ncx), olfactory bulb (OB), hypothalamus (Hyp) (arrowhead), and CS axons in the pons (arrow). The deletion of E4 (**g**), but not E1-3 (**d-f**), led to a specific loss of GFP expression in the Ncx and CS axons (arrow), but not the OB or Hyp.

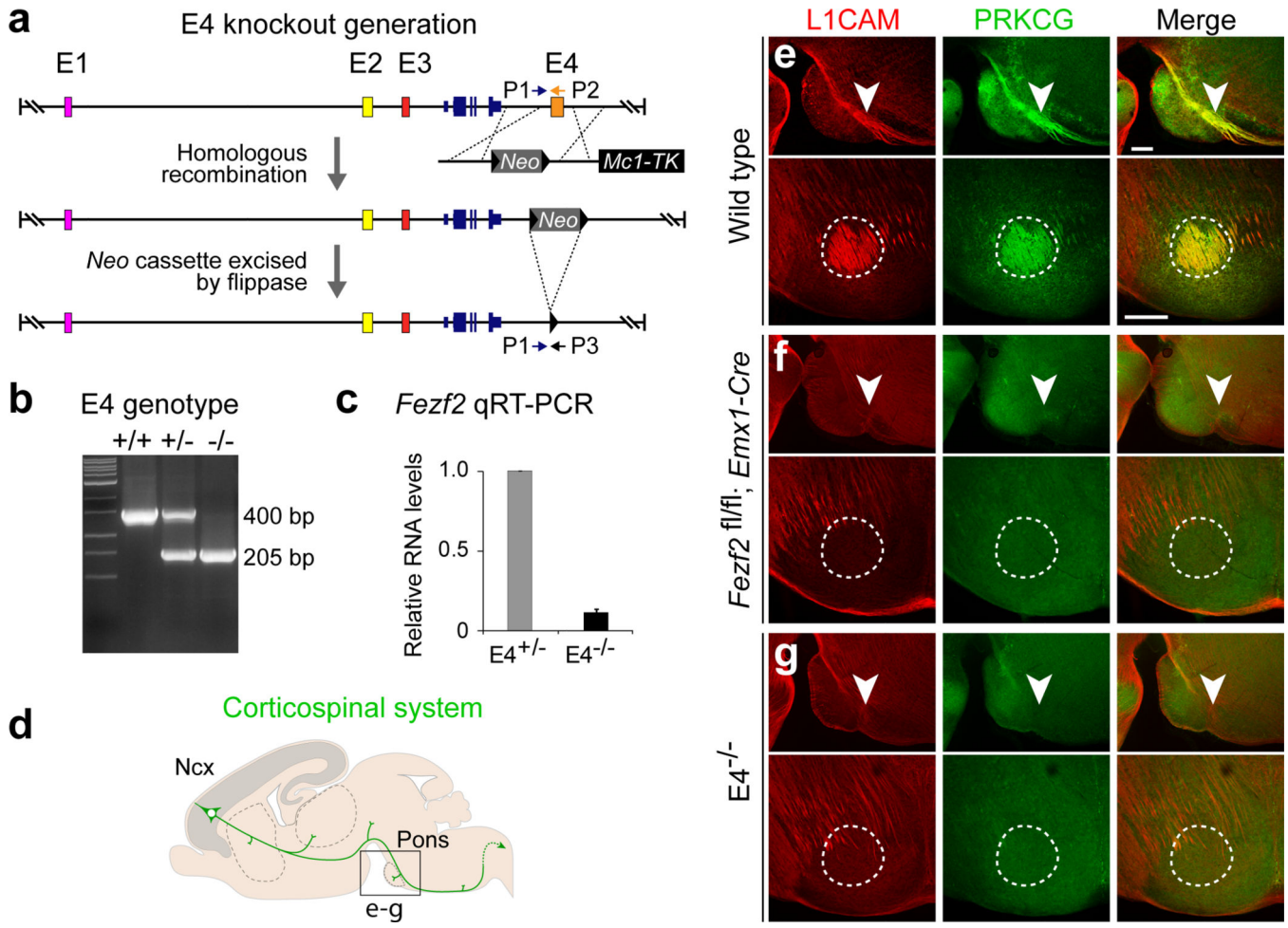

Fig. 2. Loss of neocortical *Fezf2* expression and CS axons in E4 knockout mice

a, Generation of E4 knockout mice ($E4^{-/-}$). **b,** Duplex PCR genotyping of wildtype and mutant alleles using primers P1, P2, and P3. **c,** Analysis of neocortical *Fezf2* expression by quantitative (q) RT-PCR. Normalized to *Gapdh*, *Fezf2* mRNA levels were significantly reduced in the $E4^{-/-}$ compared to heterozygous littermate controls ($E4^{+/-}$). ($P = 2.1\times10^{-6}$, one-tailed Student's *t*-test, $n = 3$ per genotype). Error bars represent s.e.m. **d,** A schematic depiction of CS system. **e-g,** Sagittal (top row) and coronal (bottom row) sections of the pons from wildtype (**e**), cortex-specific *Fezf2* knockout (**f**), and $E4^{-/-}$ (**g**) mice immunostained for axon marker L1CAM (red) and PRKCG (green). The near complete loss of CS axons (arrowheads and dashed outlines) in the $E4^{-/-}$ is a phenocopy of cortex-specific *Fezf2* deletion. Scale bars represent 200 μm.

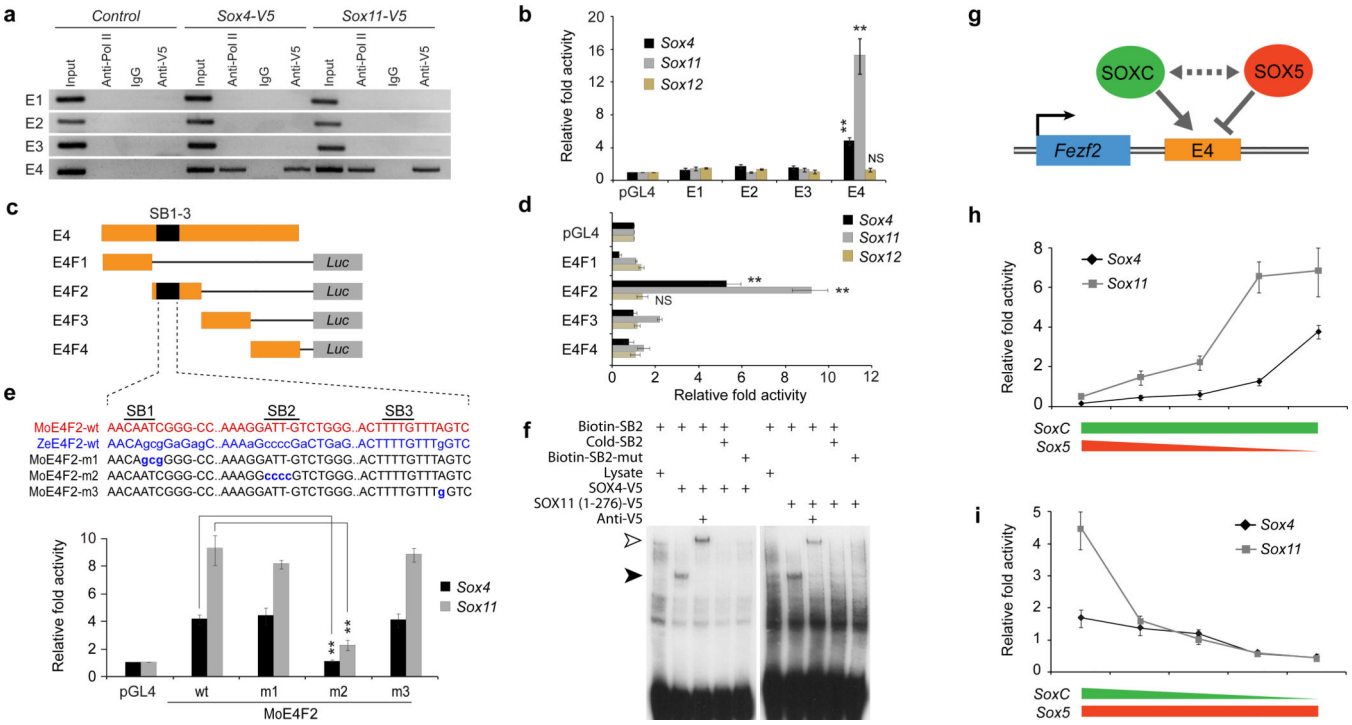

Fig. 3. SOX4 and SOX11 bind to and activate E4 via competition with SOX5

a, Chromatin immunoprecipitation (ChIP) from Neuro-2a cells expressing V5-tagged SOX4 and SOX11. Captured DNA was analyzed by PCR using primers specific for E1-4. SOX4 and SOX11 bound and recruited Pol II to E4 but not E1-3. **b**, Analysis of SOXC transactivation using empty (pGL4) or E1-4 containing luciferase vectors. The activity of E4, but not E1-3, was significantly increased by *Sox4* (≥ 4 fold) or *Sox11* (≥ 13 fold), but not *Sox12* (≤ 1.5 fold). **c-d**, Analysis of functional elements within E4 using luciferase vectors containing deletion fragments of E4 (E4F1-4) (**c**). E4F2, but not the other fragments, was significantly activated by co-transfection of *Sox4* (≥ 5 fold) or *Sox11* (≥ 9 fold), but not *Sox12* (≤1.5 fold) (**d**). **e**, Analysis of putative SOX binding sites (SB1-3) using mouse (Mo) E4F2 luciferase vectors mutagenized by substitution (blue lowercase nucleotides) with zebrafish (Ze) sequence. Targeted mutation of SB2 (MoE4F2-m2) significantly diminished the *trans*-activating ability of SOX4 or SOX11. **f**, Electrophoretic mobility shift assay (EMSA) with biotin-labeled SB2 DNA. SOX4-V5 and SOX11(1-276)-V5 shifted wildtype (arrowhead) but not mutated SB2 DNA. SOXC-SB2 complexes were super-shifted (open arrowhead) by an anti-V5 antibody. **g**, A schematic model of E4 regulation by SOX5 and SOXC. **h-i,** Analysis of competition between SOX5 and SOXCs using the E4-containing luciferase vector. Decreasing concentrations of co-transfected *Sox5* led to a dose-dependent increase in E4 activation in response to *Sox4* or *Sox11* (**h**), whereas decreasing concentrations of co-transfected *Sox4* or *Sox11* led to a dose-dependent increase in *Sox5* repression of E4 (**i**). One-tailed Student's t-test; ***P*<0.01, ****P*<0.001. n = 4 per condition. Error bars represent s.e.m.

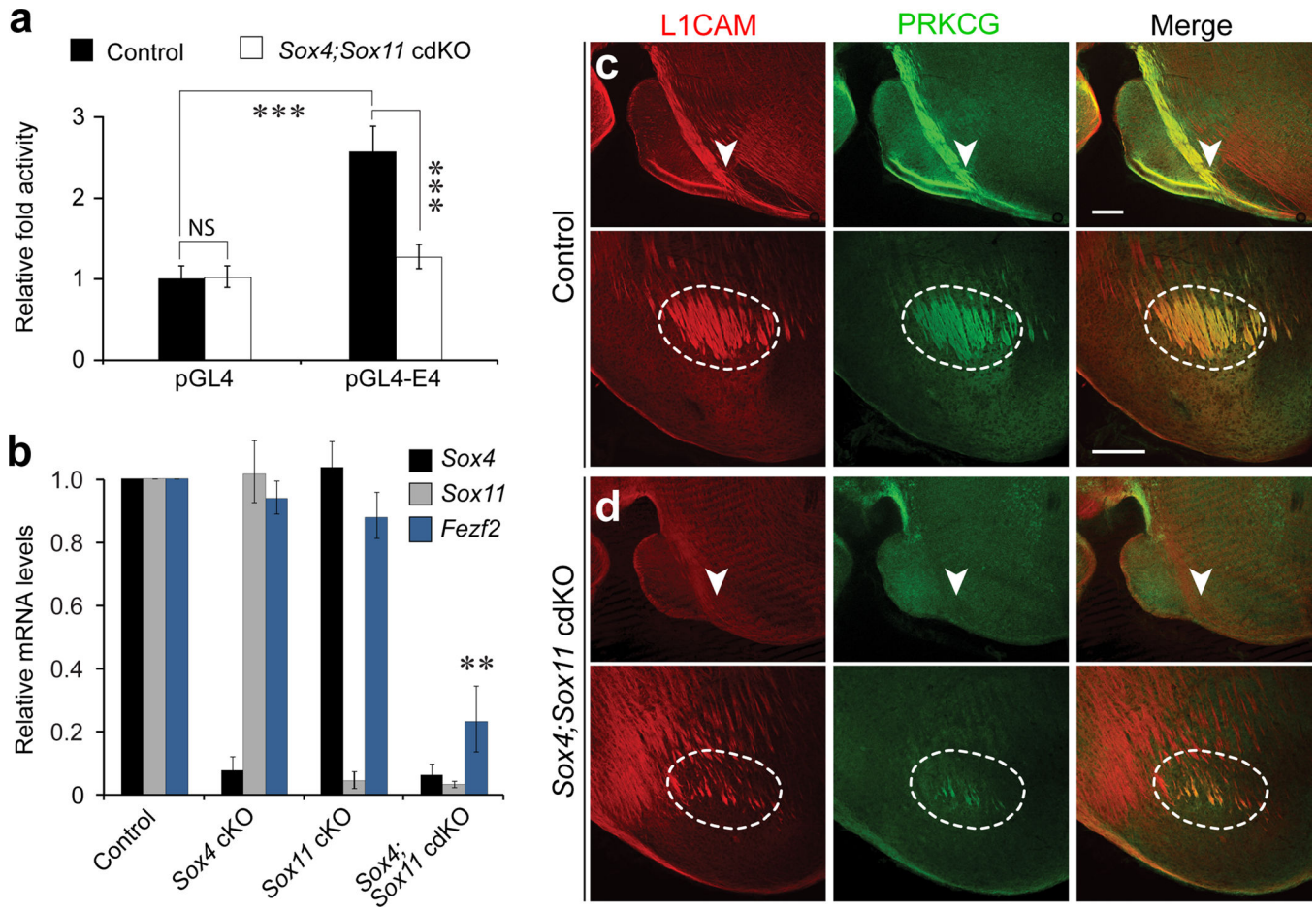

Fig. 4. *Sox4* and *Sox11* are required for *Fezf2* expression and CS tract formation

a, The requirement of SOX4 and SOX11 for E4 transactivation was determined in neurons cultured from E14.5 heterozygous littermate control and *Sox4;Sox11* cdKO mice and transfected with the E4 luciferase construct. The activity of E4 in double knockout neurons was not significantly above background (1.1 fold. **b**, Analysis of neocortical *Fezf2* expression by qRT-PCR in cortex-specific *Sox4* and/or *Sox11* mutants. Normalized to *Gapdh*, *Fezf2* mRNA levels were dramatically reduced in cdKO but not *Sox4* or *Sox11* mutants. **c-d,** Sagittal (top row) and coronal (bottom row) sections of the control and cdKO pons immunostained for L1CAM (red) and PRKCG (green). The dramatic loss of L1CAM and PRKCG-positive CS tract axons in the cdKO (arrowhead and dashed outline) is similar to that of the cortex-specific *Fezf2* mutant. Errors bars represent s.e.m. One-tailed Student's t-test; NS=Not significant, ***P*<0.01, ****P*<0.001. n = 4 per genotype. Scale bars represent 200 μm.

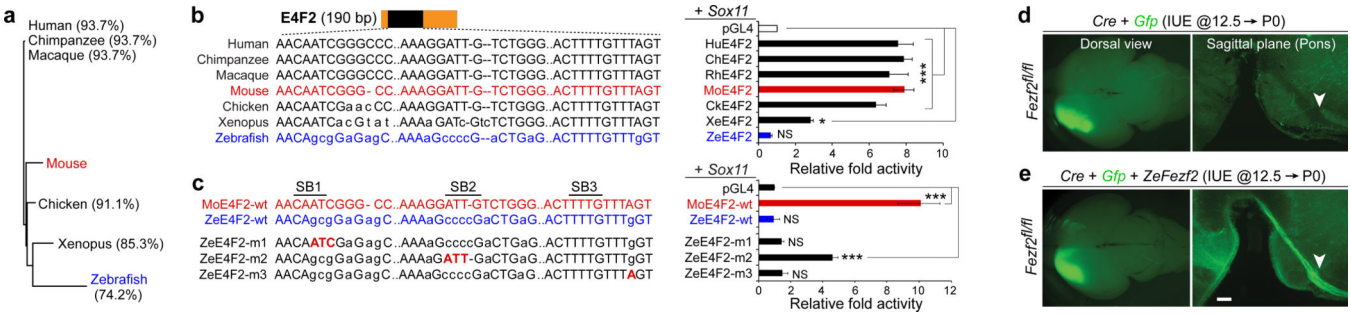

Fig. 5. Functional analysis of species differences in E4 sequence

a, Hierarchical clustering of E4F2 sequences from seven vertebrates. Percentage nucleotide identity relative to mouse E4F2 is indicated in parentheses. **b,** Analysis of species differences in *Sox11* activation of E4F2. *Sox11 trans*-activated the E4F2 sequence of five vertebrates (≥ 7.3 fold) and *Xenopus* (2.8 fold). In contrast, the activity of zebrafish E4F2, which is divergent in SB1 and SB2, was not activated by *Sox11*. **c,** Murinization (red uppercase nucleotides) of zebrafish SB2 (ZeE4F2-m2) partially rescued the loss of transactivation of zE4F2 by *Sox11*. **d-e,** Cell-autonomous rescue of mouse *Fezf2* loss-of-function by zebrafish *Fezf2* (*ZeFezf2*). *In utero* electroporation (IUE) of $Fezf2^{fl/fl}$ neocortical wall at E12.5 with *Cre* and CRE-responsive *Gfp* plasmids. *Fezf2*-deficient L5 neurons do not form CS tract at P0 (**d**). Co-electroporation of a *ZeFezf2* plasmid cell-autonomously rescued the formation of CS tract by *Fezf2*-deficient neurons (**e**). Errors bars represent s.e.m. One-tailed Student's t-test; **P*<0.05, ***P*<0.01, ****P*<0.001. n = 4 per condition. Scale bar represents 200 μm.