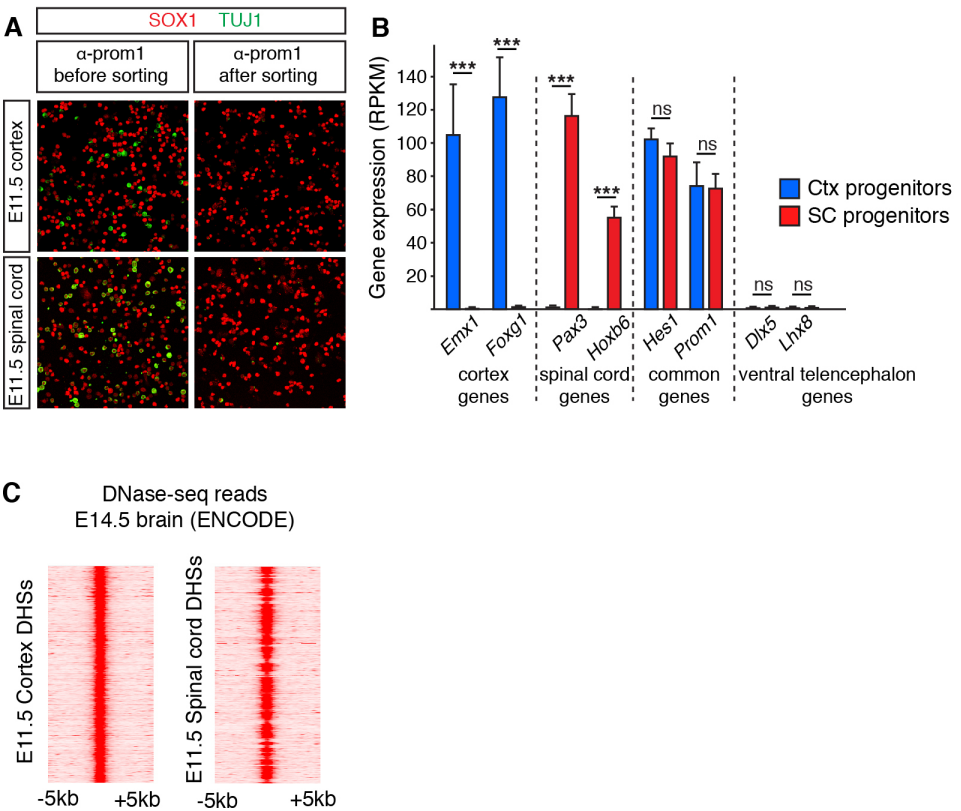
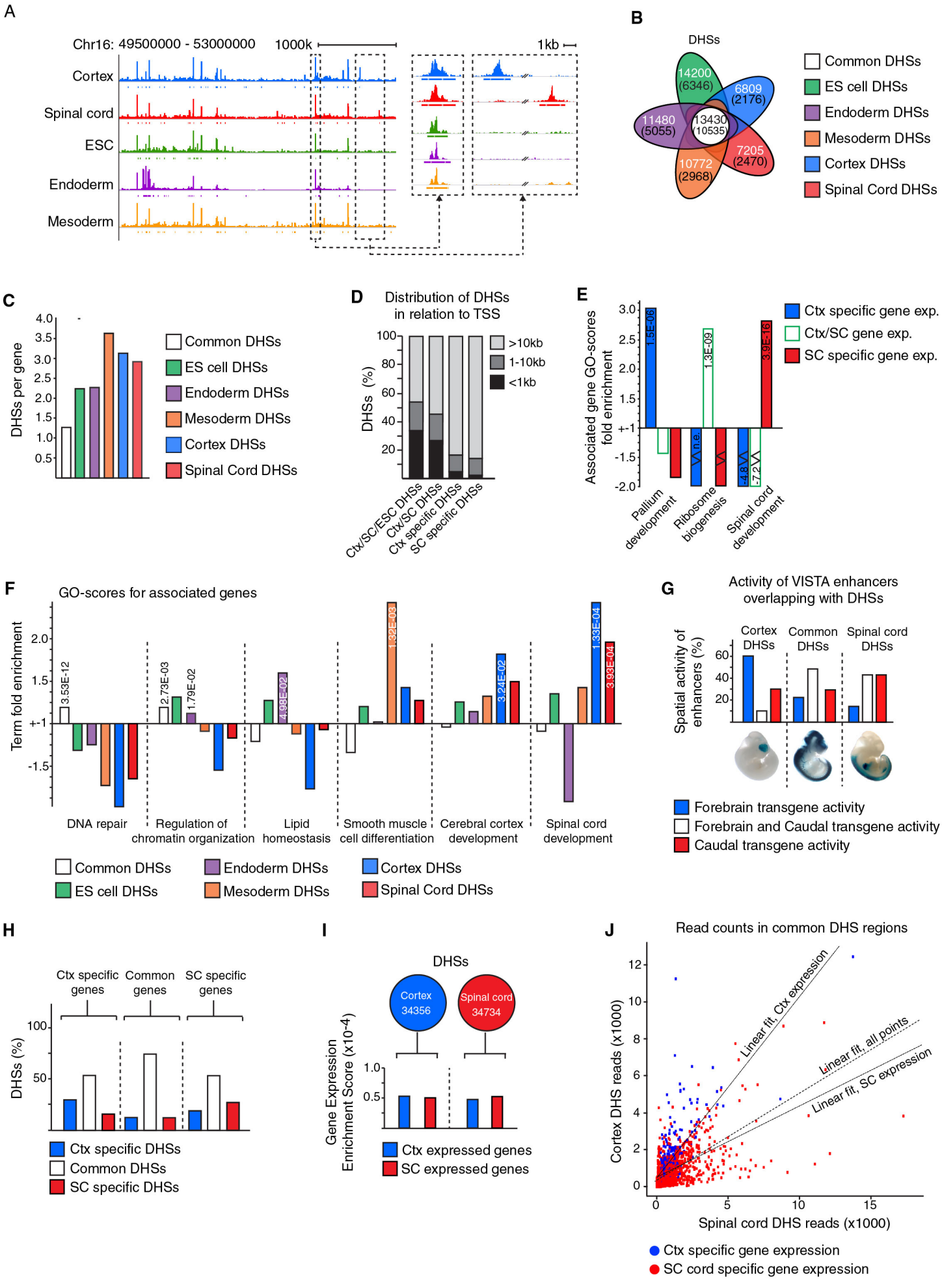
Supplemental Fig S1



Supplemental Fig S1 (A) Sorting of E11.5 mouse cortex and spinal cord neural cells using MACS with anti-PROMININ-1 microbeads enriches for NSCs as revealed by the expression of the progenitor marker SOX1 and the depletion of TUJ1^{+} cells. (*B*) RNA-seq analysis of sorted NSCs reveals that cells isolated from the cortex and the spinal cord express marker genes consistent with their region of origin. (*C*) Heat map comparisons reveal that the vast majority of the DHS identified in E11.5 mouse cortical NSCs, and many of the open regions identified in E11.5 mouse spinal cord NSCs, are represented in the E14.5 mouse brain tissue (ENCODE).

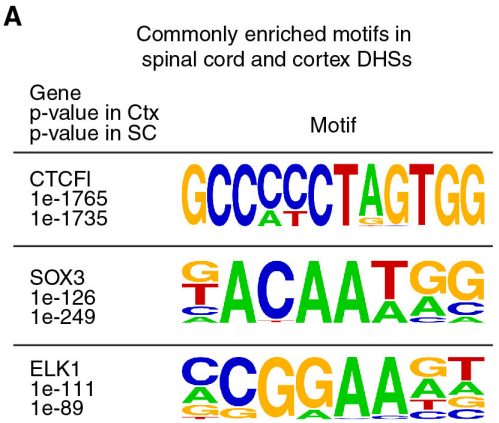
Supplemental Fig S2

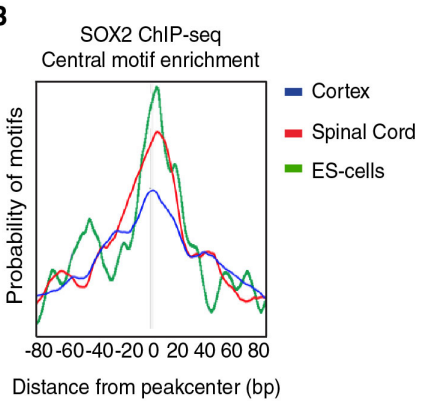


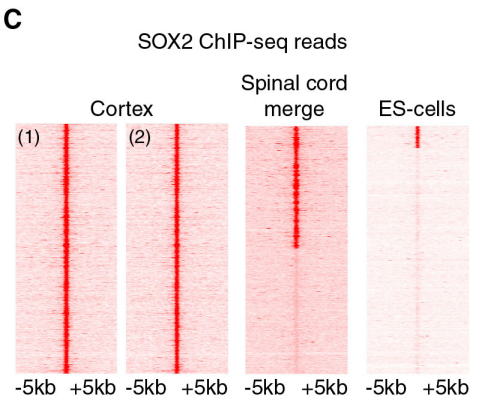
Supplemental Fig S2 (*A*) Dnase I cleavage density profiles for cortex (blue), spinal cord (red), ESCs (green), endoderm (purple) and mesoderm (yellow) aligned to a region of mouse chromosome 16 illustrates specific and common DHS regions. (*B*) Venn diagram showing specific and overlapping (white circle) DHSs, with inset values showing the number of DHSs in each group (white) and number of associated genes (black within brackets). (*C*) Graph showing the average number of DHSs per gene in each group. (*D*) Distance to nearest TSS of specific and common groups of DHSs. Bar sections show the percent of DHSs less than 1kb (black), 1-10kb (dark grey) and >10kb (light grey) from TSS. (*E*) Relative GO-term enrichments for the genes with a specific or common expression pattern in RNA-seqs from spinal cord and cortex tissue. (*F*) Relative enrichment scores for various GO-terms significant for DHS associated genes in each group, with p-values displayed for significant enrichments. (*G*) Graph depicts spatial activity in E11.5 embryos of transgenic enhancers from the VISTA Enhancer Browser that overlap with DHSs indentified in cortical and spinal cord NSCs. LacZ stained E11.5 embryos show regional activity of three representative transgenic enhancers. (*H*) Percentage of specific or common DHSs associated with genes with a specific and common expression pattern. (*I*) Specific expression pattern for genes associated with DHSs found in cortex or spinal cord. (*J*) Scattered plot showing the peak size of common DHSs in cortical and spinal cord NSCs, depending on whether they are associated with genes primarily expressed in the cortex (blue dots) or spinal cord (red dots). The specific relationship between chromatin accessibility and gene expression is reflected by differences in the angle of the group specific regression lines. P-values for both the cortex and spinal cord specific regression lines, assuming a null hypothesis where the cortex specific or spinal cord specific data

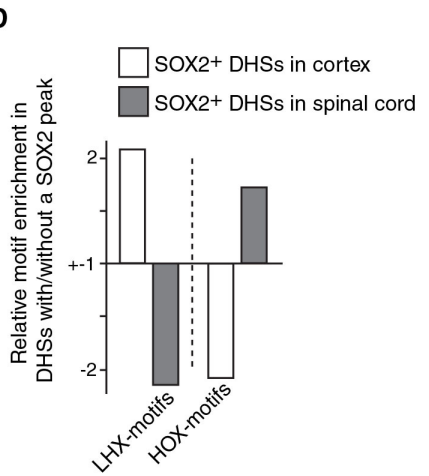
come from the same distribution, are <2.2e-16. Gene expression was defined as cortex or spinal cord specific if the expression levels differed more than 3-fold between the spinal cord and cortex.

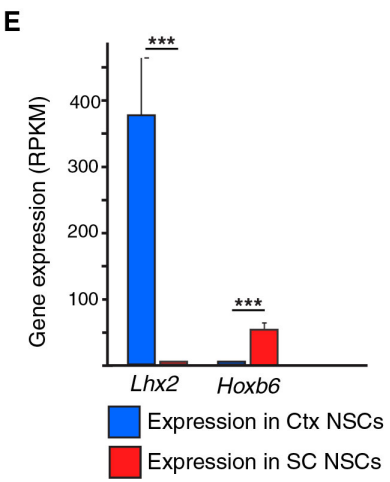
Supplemental Fig S3

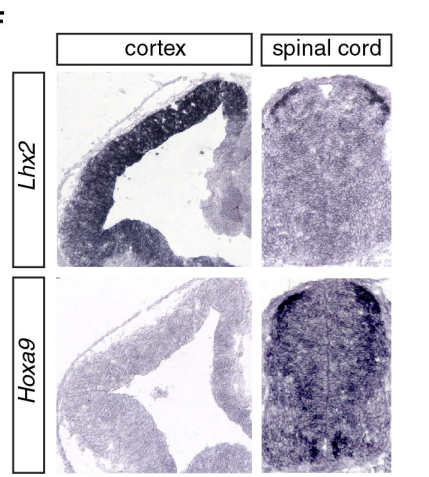








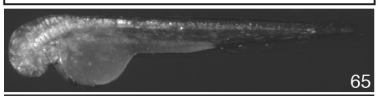
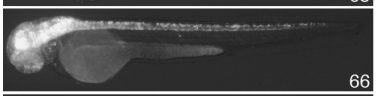
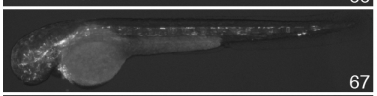
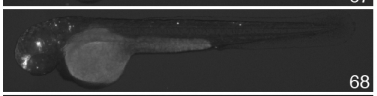
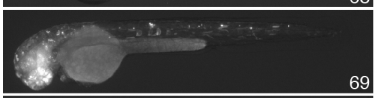
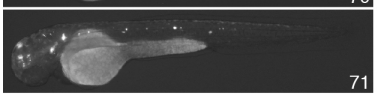


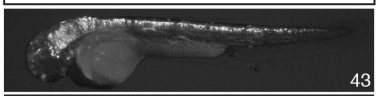
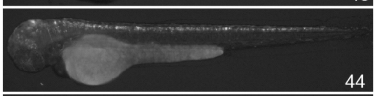
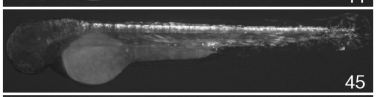
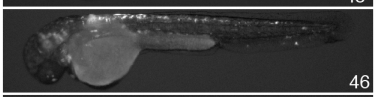
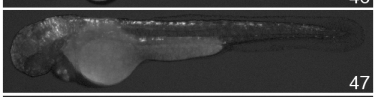
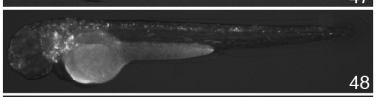
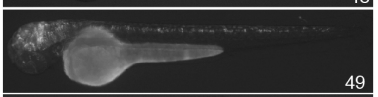
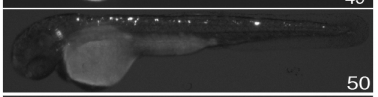
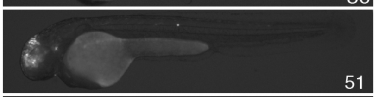
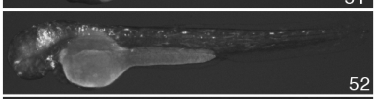
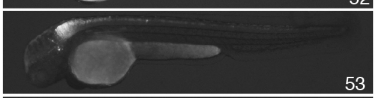
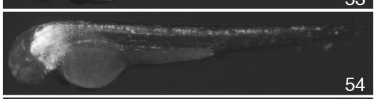
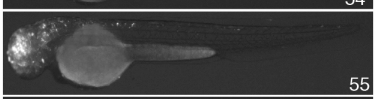
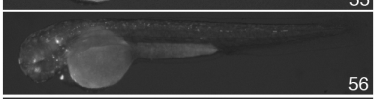
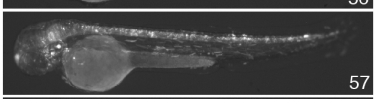
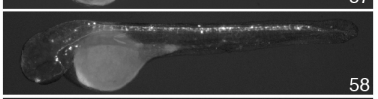
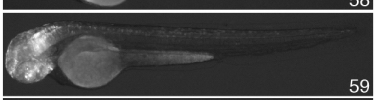
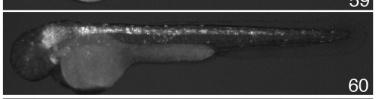
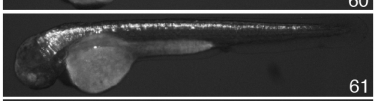
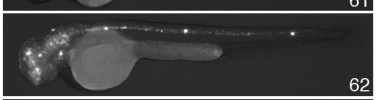
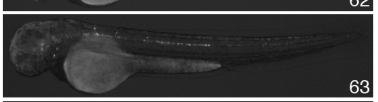
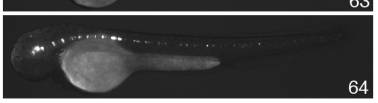


Supplemental Fig S3 (A) Top three binding motifs commonly enriched above a random background in cortical and spinal cord DHSs. P-values for motifs in cortical and spinal cord DHSs are shown. (B) Centrally enriched SOX motifs in ChIP-seq peaks. Probabilities of finding motifs at distances from peak centres are shown. (C) Heatmaps of reads from replicate cortex SOX2 ChIP-seqs and merged SOX2 ChIP-seqs in the spinal cord and ES-cells. (D) Relative enrichment of LHX and HOX motifs in DHSs with a SOX2 peak in cortex (white) or spinal cord (grey) over DHSs from the same tissue without a SOX2 peak. (D) Raw rpkm values for LHX2 and HOXA9 in cortex (blue) and spinal cord (red) progenitors, with error bars representing s.d. (F) In situ hybridization analysis shows that LHX2 is mainly expressed in NSCs of the cortex, whereas HOXA9 is mainly expressed in NSCs of the spinal cord. Stars represent padj- values of $0.05>p>0.01$ (*), $0.01>p>0.001$ (**) or $p<0.001$ (***).

Supplemental Fig S4

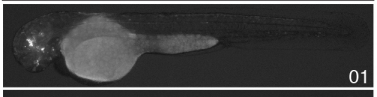
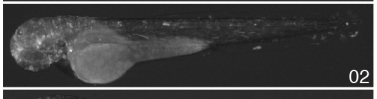
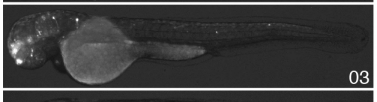
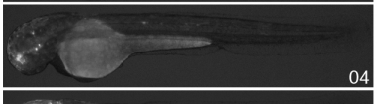
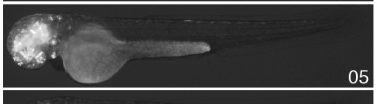
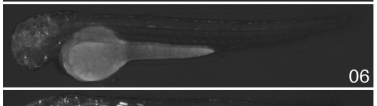
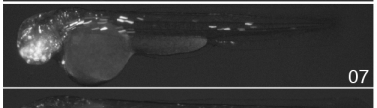
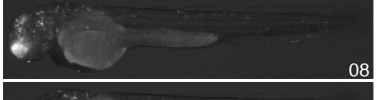
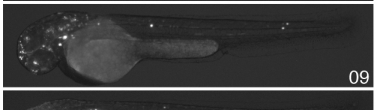
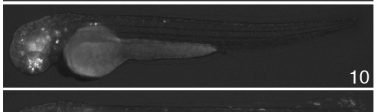
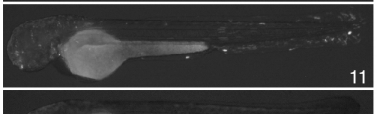
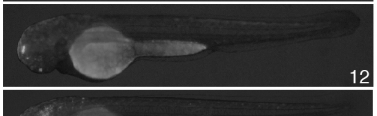
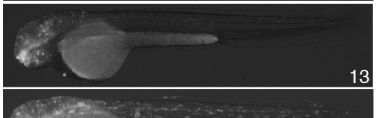
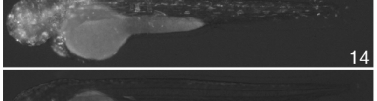
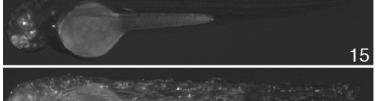
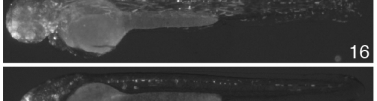
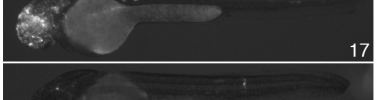
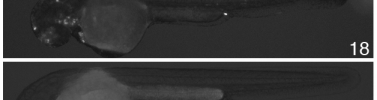
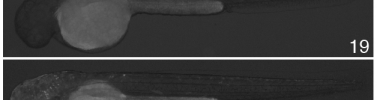
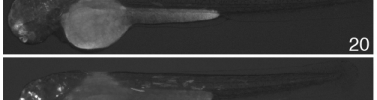
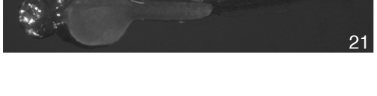
A

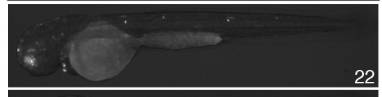
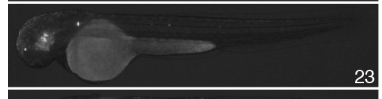
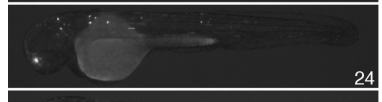
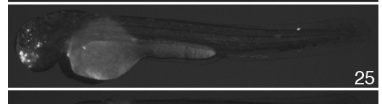
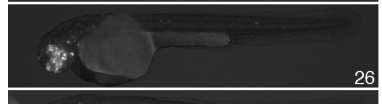
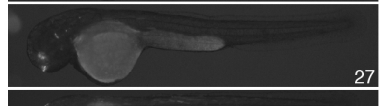
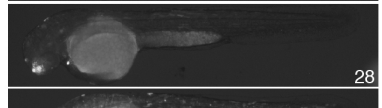
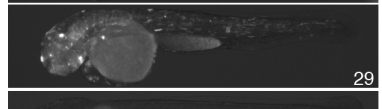
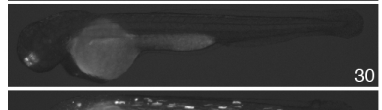
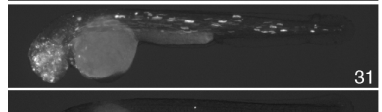
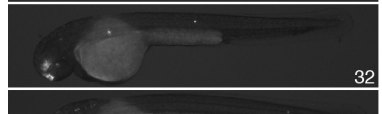
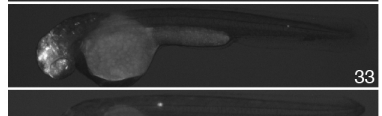
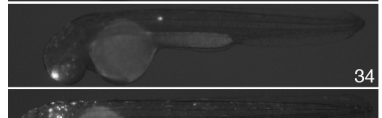
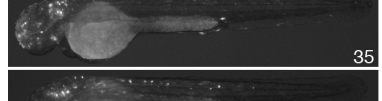
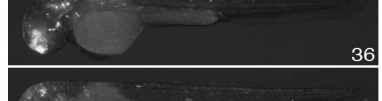
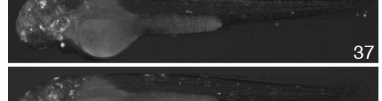
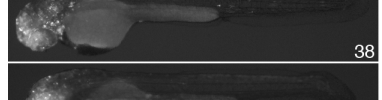
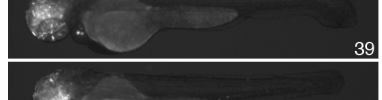
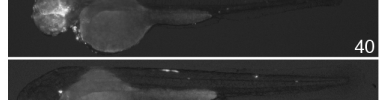
common SOX2 bound regions	genomic coordinates percentage of GFP+ embryos
	chr13:39,973,048-39,973,577 79% (94/107)
	chr13:44,847,180-44,847,446 94% (93/99)
	chr1:130,301,483-130,301,941 36% (25/69)
	chr18:69,659,847-69,660,238 32% (38/120)
	chr9:61,316,198-61,316,636 79% (54/68)
	chr12:105,969,878-105,970,338 83% (77/93)
	chr3:127,328,603-127,329,076 95% (39/41)

spinal cord SOX2 bound regions	genomic coordinates percentage of GFP+ embryos
	chr1:78,136,053-78,136,291 71% (82/115)
	chr1:78,187,168-78,187,559 76% (26/34)
	chr1:78,194,457-78,194,938 89% (39/44)
	chr10:25,628,461-25,628,721 9% (6/68)
	chr10:84,173,587-84,173,971 51% (72/140)
	chr13:39,615,809-39,616,037 58% (48/83)
	chr13:73,171,995-73,172,642 83% (25/30)
	chr16:7,293,439-7,293,797 6% (5/83)
	chr3:5,237,364-5,237,841 10% (7/70)
	chr3:34,457,911-34,458,297 31% (22/71)
	chr4:21,613,032-21,613,512 81% (78/96)
	chr4:153,877,337-153,877,560 84% (91/108)
	chr4:153,978,185-153,978,553 59% (81/137)
	chr5:7,5408,870-75,409,344 25% (24/96)
	chr5:101,950,246-101,950,643 91% (64/70)
	chr6:144,103,646-144,103,832 94% (68/72)
	chr6:14,4261,474-14,4261,772 61% (72/118)
	chr7:77,992,859-77,993,264 42% (92/219)
	chr7:146,716,432-146,716,915 48% (72/150)
	chr8:105,250,588-105,250,874 47% (26/55)
	chr9:90,595,429-90,595,946 49% (74/151)
	chr9:91,282,567-91,282,882 51% (90/178)

Supplemental Fig S4

B

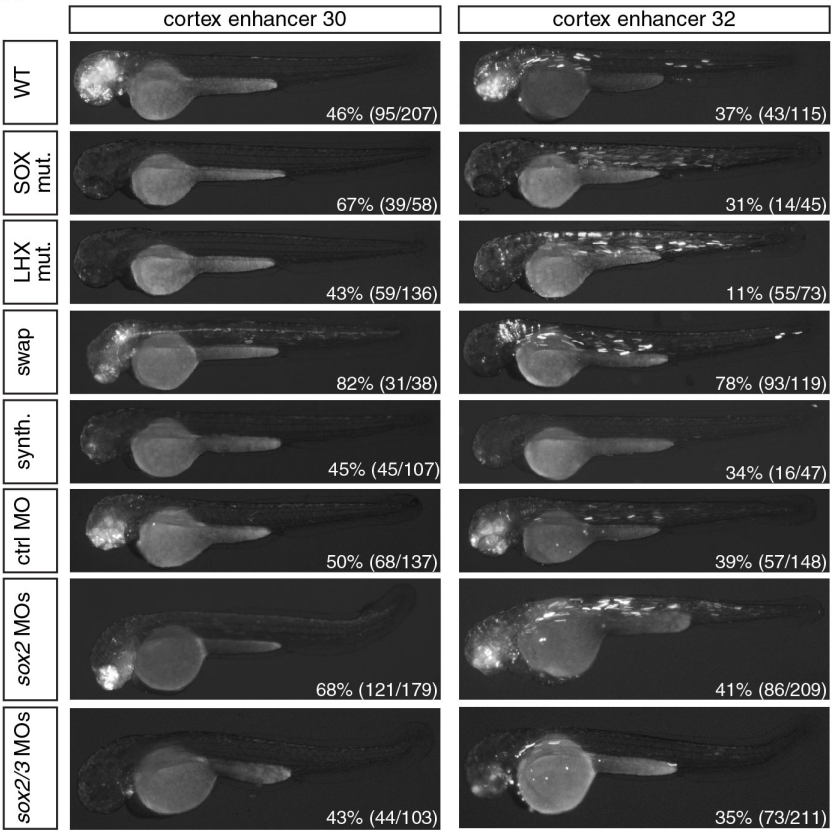
cortex SOX2 bound regions	genomic coordinates percentage of GFP+ embryos
 01	chr11:18,993,601-18,993,829 53% (68/128)
 02	chr11:21,902,978-21,903,327 25% (23/91)
 03	chr11:21,917,254-21,917,666 2% (2/82)
 04	chr12:51,223,115-51,223,480 18% (31/171)
 05	chr12:85,882,495-85,882,801 46% (95/207)
 06	chr13:79,570,188-79,570,607 23% (17/74)
 07	chr14:13,170,295-13,170,943 37% (43/115)
 08	chr14:98,472,038-98,472,941 51% (87/170)
 09	chr14:104,696,957-104,697,591 2% (3/156)
 10	chr14:118,765,732-118,766,098 15% (17/112)
 11	chr14:123,342,366-123,342,632 40% (34/85)
 12	chr16:6,962,187-6,962,478 7% (8/111)
 13	chr16:91,140,746-91,140,962 45% (63/139)
 14	chr17:48,051,683-48,052,032 78% (127/162)
 15	chr18:661,82,322-66,182,620 52% (13/25)
 16	chr19:25,682,374-25,682,748 82% (68/83)
 17	chr19:59,541,384-59,541,943 76% (58/76)
 18	chr1:41,975,469-41,976,004 58% (42/73)
 19	chr1:170,838,880-170,839,217 4% (5/113)
 20	chr2:57,230,684-57,231,160 10% (15/153)
 21	chr2:70,312,746-7,0313,241 75% (79/106)

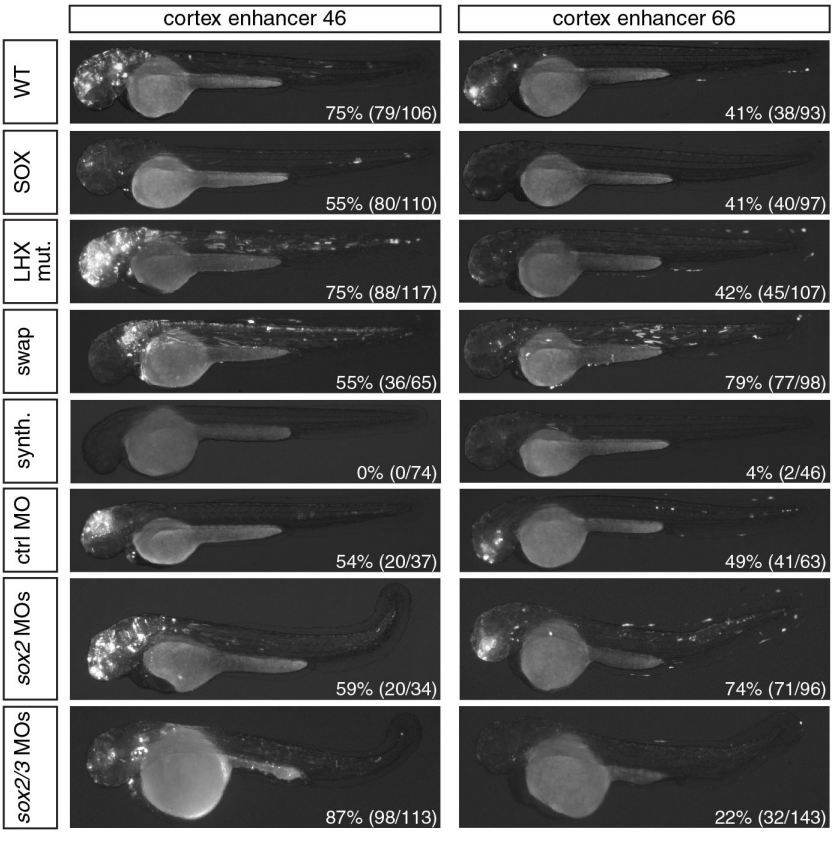
cortex SOX2 bound regions	genomic coordinates percentage of GFP+ embryos
 22	chr2:71,256,752-71,257,464 22% (32/143)
 23	chr2:71,258,885-71,259,200 37% (22/59)
 24	chr2:105,559,807-105,560,544 17% (7/42)
 25	chr2:106,352,960-106,353,309 9% (8/91)
 26	chr3:34,403,059-34,403,528 68% (99/145)
 27	chr3:66,550,376-66,551,046 22% (10/46)
 28	chr3:143,837,181-143,837,495 27% (48/179)
 29	chr3:143,861,908-143,862,219 60% (60/100)
 30	chr4:6,840,457-6,841,053 18% (16/91)
 31	chr4:109,435,378-109,435,974 59% (57/96)
 32	chr4:124,072,787-124,073,110 54% (55/101)
 33	chr6:15,589,441-15,589,801 37% (39/106)
 34	chr6:16,892,534-16,892,904 14% (6/42)
 35	chr6:98,717,774-98,718,235 52% (109/210)
 36	chr6:143,943,612-143,943,949 31% (30/97)
 37	chr6:144,629,708-144,629,984 28% (67/241)
 38	chr7:76,723,855-76,724,132 34% (66/193)
 39	chr7:116,334,831-116,335,210 41% (24/59)
 40	chr8:94,594,789-94,595,003 34% (22/65)
41	chr9:40,735,460-40,735,822 41% (38/93)
42	chrX:90,547,258-90,547,673 43% (28/65)

Supplemental Fig S4 (A, B) Representative examples of resulting transgenic zebrafish embryos after injection with GFP-reporters containing enhancer elements bound commonly (A) or specifically by SOX2 in the mouse spinal cord (A) or cortex (B). Genomic coordinates of mouse enhancers are shown. Numbers within brackets represent fraction of zebrafish embryos with appropriate GFP-expression over the number of live embryos, 50-55 hours after the injection of transgenic reporter constructs.

Supplemental Fig S5

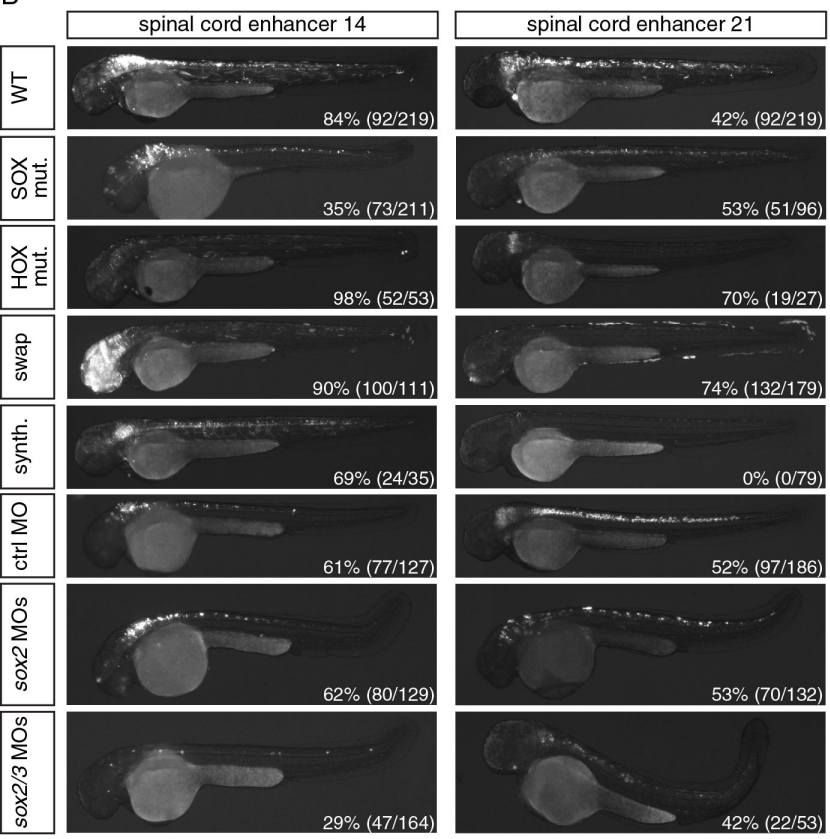
A

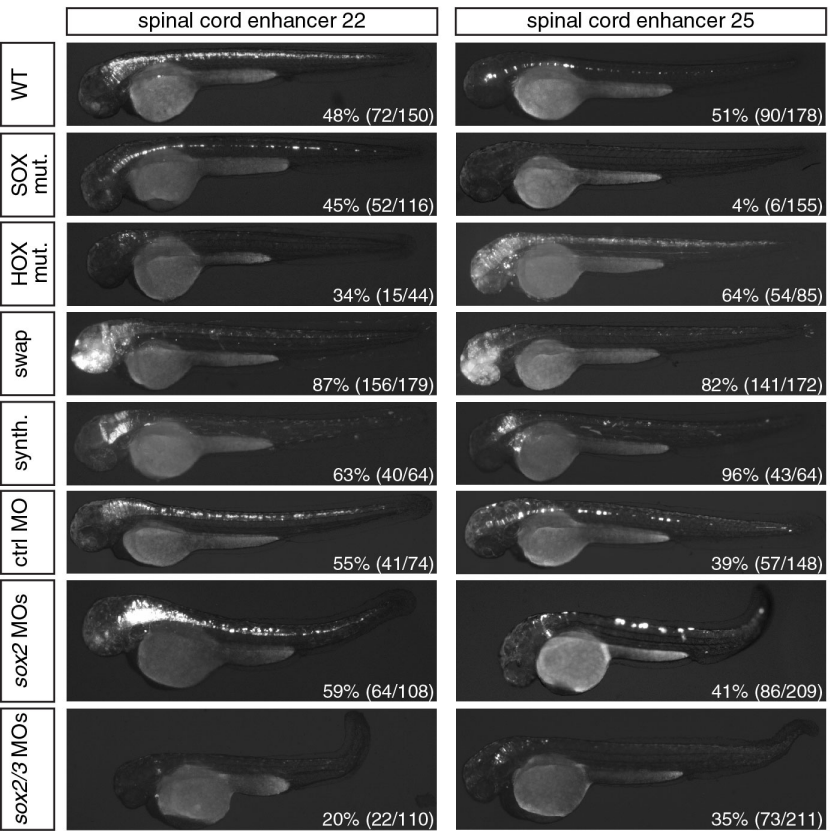




Supplemental Fig S5

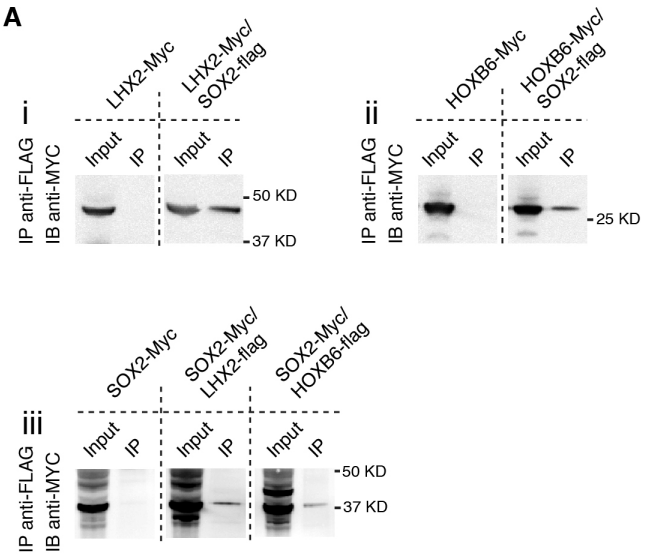
B





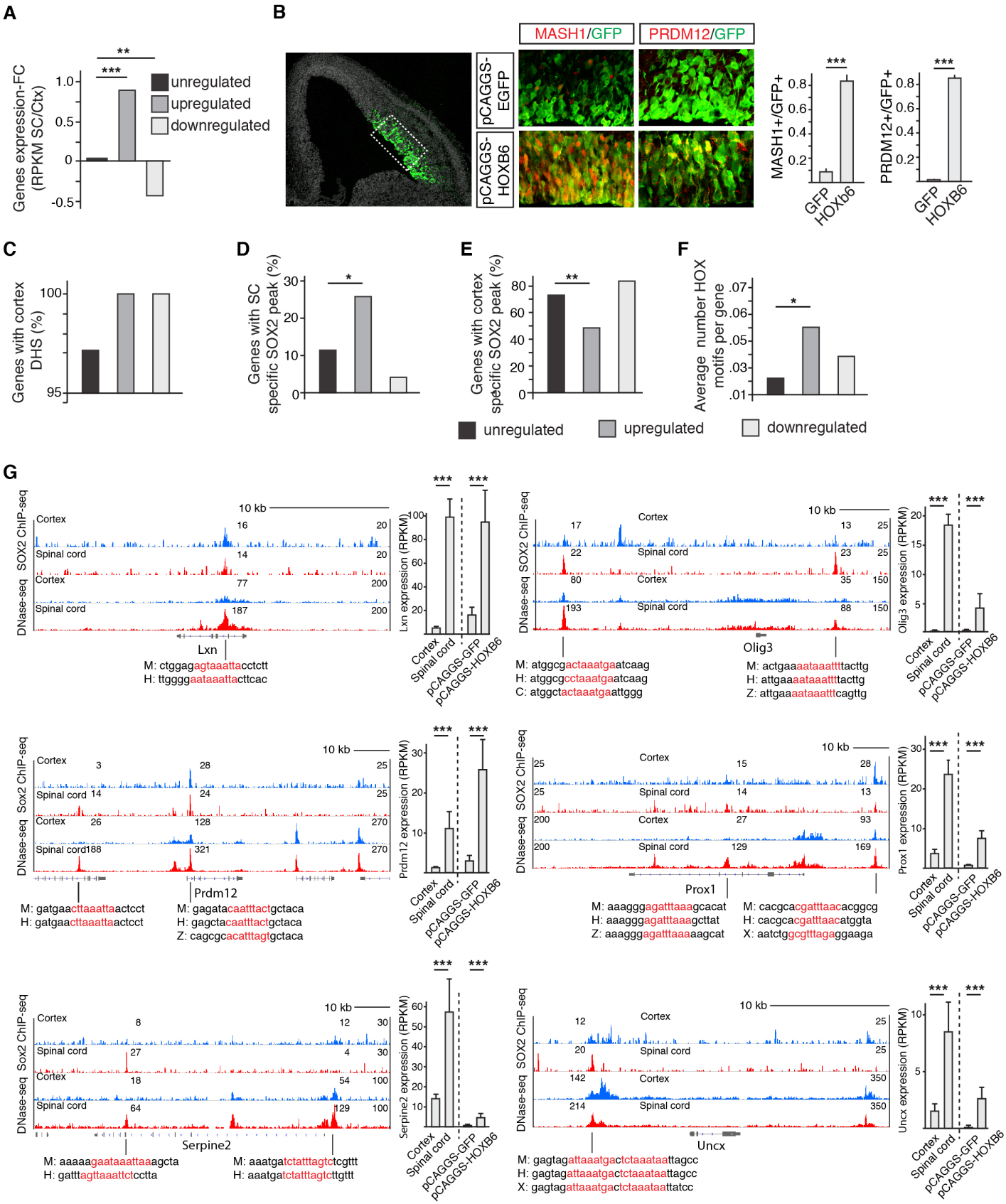
Supplemental Fig S5 (A,B) Activity of transgenic wild type enhancers bound by SOX2 in the mouse cortex (A) or spinal cord (B), compared to activity of enhancers upon mutation of SOX-motifs (SOX mut.) and LHX-motifs (LHX-mut.) or HOX-, PBX-, and MEIS-motifs (HOX mut.). Enhancer versions in which the nucleotide sequences have been randomized (synthetic) apart from intact LHX- and SOX-motifs or intact HOX-, PBX-, MEIS- and SOX-motifs. Swap represents GFP-reporter variants containing enhancer versions in which the LHX-motifs were swapped for HOX-, PBX-, and MEIS-motifs or in which HOX-, PBX- and MEIS-motifs were swapped for LHX-motifs. Ctrl MOs, *sox2* MOs and *sox2/3* MOs show activity of GFP-reporters upon co-injection with control morpholinos or morpholinos targeting *sox2* or *sox2/sox3*. Fraction of embryos with representative reporter activity over live embryos 50-55 hours after injection is shown as absolute numbers (within brackets) or as percentage.

Supplemental Fig S6



Supplemental Fig S6 (*A*) Co-immunoprecipitation experiments in transfected HEK293 cells reveal that MYC-tagged versions of LHX2 and HOXB6 can interact with FLAG-tagged SOX2, and vice versa.

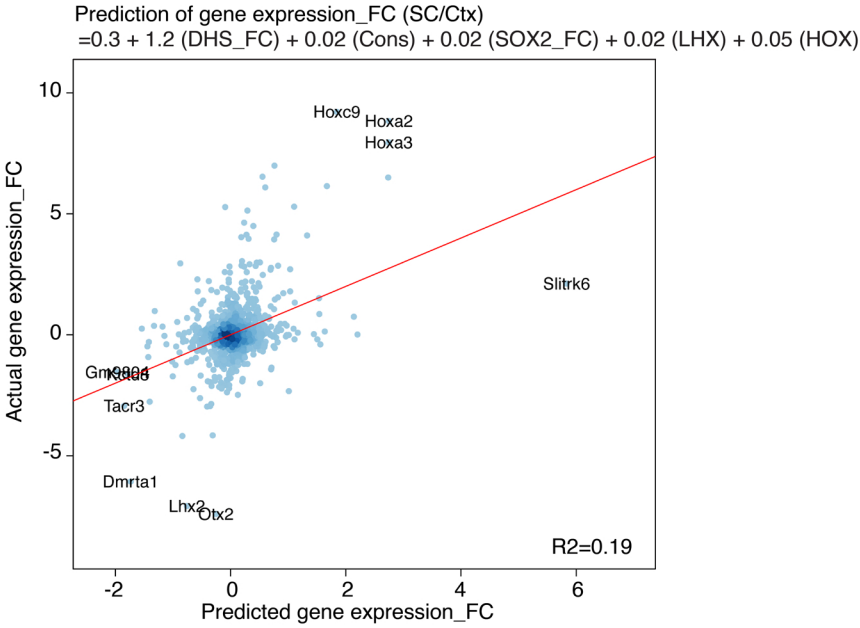
Supplemental Fig S7



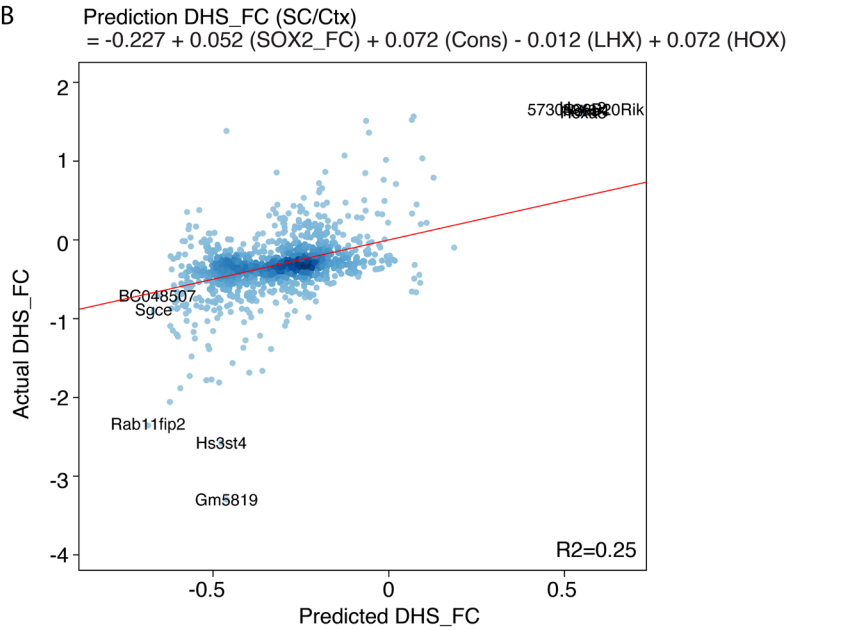
Supplemental Fig S7 (*A*) RPKM fold change (FC) values in the spinal cord versus the cortex of genes deregulated upon HOXB6 misexpression in E13.5 cortices. (*B*) Immunohistochemical analysis of HOXB6 electroporated cortices demonstrate a broad upregulation of MASH1 and PRDM12, which are normally expressed predominately in the spinal cord compared to the cortex. (*C-F*) DHSs (C), SOX2 binding in spinal cord (D) or cortex (E) and HOX motif enrichment (F) within 50 kb of TSSs of genes up- (dark grey), down- (light grey) or unregulated (black) in cortical NSCs following HOXB6 misexpression. (*G*) DNAse I- and SOX2 ChIP-seq reads from spinal cord (red) and cortex (blue) aligned to regions around spinal cord specific genes that are upregulated by HOXB6 misexpression in cortex. Inset values to the right show read-scale and peak read stack height (internal) for regulatory regions with conserved HOX motifs (shown below). Bar graph show raw rpkms for each HOXB6 upregulated gene in cortex or spinal cord NSCs, as well as, in GFP- and HOXB6 electroporated cortical NSCs. Error bars represents s.d. Stars represent p- or padj-values of $0.05>p>0.01$ (*), $0.01>p>0.001$ (**) or $p<0.001$ (***).

Supplemental Fig S8

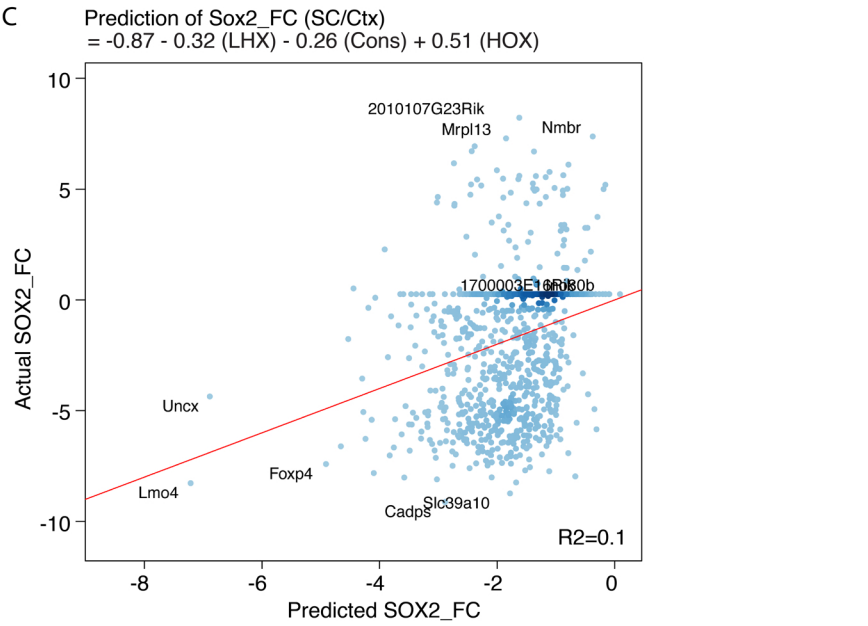
A



B



C



Supplemental Fig S8 (*A-C*) Formulas for predicting gene expression, DHS and SOX2 binding fold change (FC) values in cortical versus spinal cord NSCs. Scatter plot shows actual versus predicted relative gene expression values in cortical and spinal cord NSCs. Analysis is based on a test gene set of SOX2 bound genes, different from the training set used to generate presented formula.