Supplementary Files

Supplementary methods

Whole genome sequencing and mapping of the *sox1a:eGFP* line

Extraction of gDNA:

About 60 embryos of a *sox1a:eGFP* outcross (with ABO wild type line) were collected at 24 hpf and embryo medium was removed as much as possible. 500 µl extraction buffer without proteinase K (80 mM Tris pH 8.5, 200 mM NaCl, 0.5% SDS, 5 mM EDTA) was added and pipetted up and down to destroy the embryos. 500 µl extraction buffer with proteinase K (final conc. 1 mg/ml) was added. The solution was incubated at 65°C overnight without agitation. Genomic DNA was extracted with phenol/chloroform/isoamylalcohol (IAA, ratio 25/24/1). 500 µl phenol/chloroform/IAA was added and agitated gently for 10 minutes at room temperature (RT) to avoid shearing of genomic DNA. The solution was centrifuged for 10 minutes at 13,000 rpm at RT and the supernatant was transferred into a fresh tube using a cut 1 ml pipette tip. 500 µl of 100% isopropyl alcohol was added and centrifuged for 10 minutes at 13,000 rpm at RT. The supernatant was removed, and the pellet was washed with 75% ethanol by centrifugation for 5 minutes at 13,000 rpm at RT. The pellet was resuspended in 100 µl of deionized water for 10 minutes at 40°C. Genomic DNA was kept at 4°C until further processing. The quality and the concentration of the genomic DNA were assessed with a Nanodrop (Thermo Scientific) and a Qubit (Invitrogen), respectively.

Library preparation and sequencing:

The indexed sequencing library was prepared from 1 µg of genomic DNA, using the Tauseq DNA Nano kit (Illumina). The quality of the library was assessed with a Bioanalyser 2100 (Agilent, DNA-chip). 10 pM of the library was used for the cluster generation with a cbot (Illumina) in one lane of a high-throughput flow cell. The sequencing was done on a Hiseq1500 using SBS v3 kits (Illumina) in paired-end mode. Cluster detection and base calling were performed using RTA (Illumina, version 1.13). The sequencing generated 127 million of paired reads of 50 nucleotides, with 83% having a quality Phred score greater or equal to 30. The quality of the reads was assessed using the FASTX-toolkit (Hannon lab, version 0.0.3). The

alignment of the reads was done using bwa (Li and Durbin, 2009) (version 0.7.15) with default options. For the alignment, a custom-made genome was created. The sequence of the insert was added to the *Danio rerio* reference genome sequence (built GRCz10) as an additional chromosome. The results of the alignment were refined using the tool FixMateInformation of the Picard toolbox (http://broadinstitute.github.io/picard). The localization of the insert was done manually, retrieving the pairs of reads with one mate aligned against the insert and the other mate aligned against a canonical chromosome. The results were manually curated, as well as the precise positions of the insert. The results were visualized using IGV (Robinson et al., 2011).

Data have been deposited in NCBI's Sequence Read Archive (SRA), SRA accession: SRR8237123.

Mouse spinal cord immunofluorescence

Mouse embryo fixation, embedding and cryo-sectioning was previously described, and so were the immuno-labelling protocols (Briscoe et al., 2000). The following primary antibodies were used goat-anti-Sox1 (1/200; R&D) and rabbit anti-glycine (1/1000; ImmunoSolution). Immunofluorescence microscopy was carried out using a Leica TCS SP2 confocal microscope. *In situ* hybridisations were performed as described (Yamada et al., 1993) using a *Gad1* and *Scl6a5* mRNA probe synthetized from PCR-amplified cDNA generated from GD11.5 dissected spinal cord using the primers listed below. Pictures were then taken with an Axio Observer Z1 microscope (Zeiss). All the images were processed with Photoshop 7.0 software (Adobe Systems, San Jose, CA, USA).

Gad1 FW: *ACAGAGACCGACTTCTCCAA*

Gad1 T7-REV: *GGTAATACGACTCACTATAGGGACATGCAGCCAAAGGTTGTA*

Scl6a5 FW: *GTATCCCACGAGATGGATTGTT*

Scl6a5 T7-REV: *GGTAATACGACTCACTATAGGGCCATACAGAACACCCTCACTCA*

Supplementary Figurs

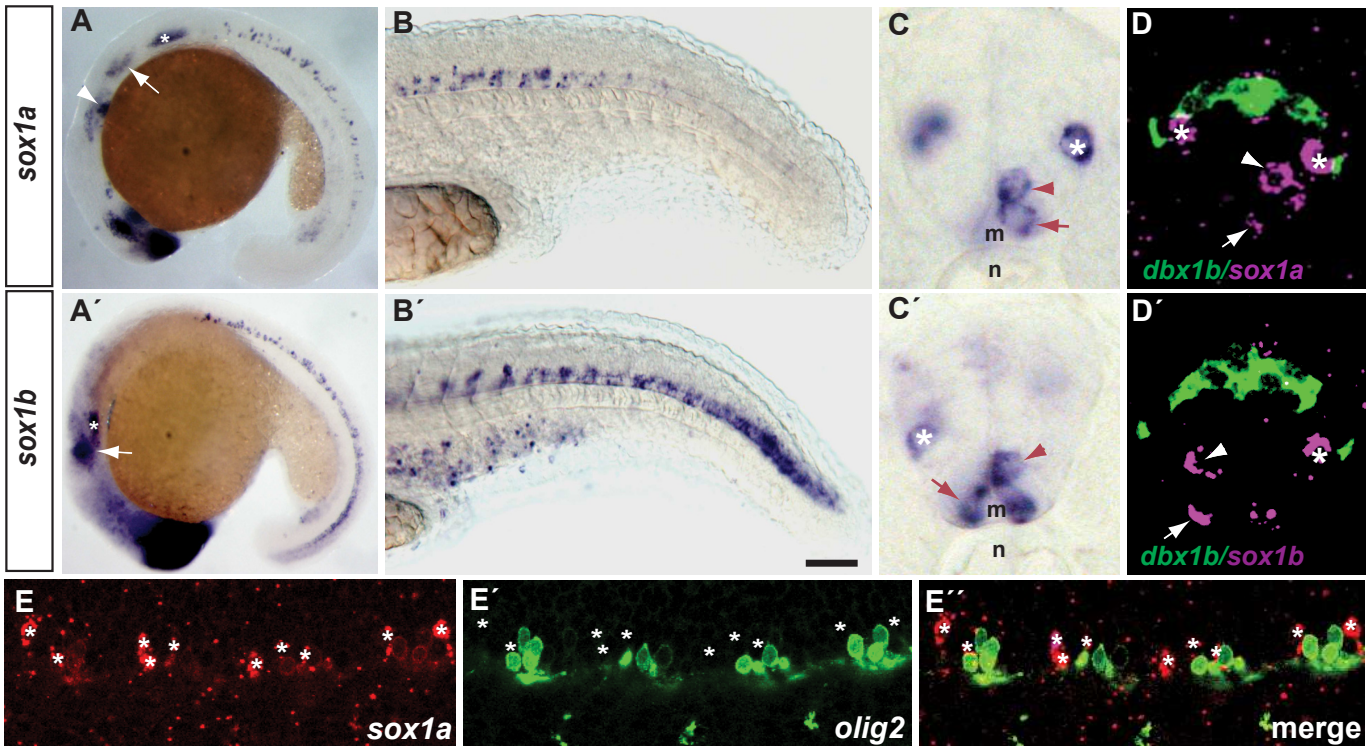


Figure S1

***sox1a* and *sox1b* genes are expressed in overlapping patterns.**

A, A´: Lateral view of whole mount embryos (22-somite stage) hybridized to *sox1a* (A) and *sox1b* (A´) antisense probes. **B, B´**: Lateral views of tail regions of *sox1a* (B) and *sox1b* (B´) antisense RNA labelled embryos at 24 hpf. **C, C´**: Transverse sections at the level between the 8th and 13th somite at 24 hpf labelled with *sox1a* (C) and *sox1b* (C´) antisense probes. *Sox1a* is expressed in the ventral telencephalon, hypothalamus, in single cells in the spinal cord, eye field, trigeminal placode (arrow head) and primordia of otic vesicle (arrow) and lateral line (asterisk) (A, B). *sox1b* is expressed in a similar pattern with exception of the tail region (A´, B´). In the spinal cord, the mRNA of *sox1a* (C) and *sox1b* (C´) is present in the lateral floor plate (C, C´ arrow), cells adjacent to the ventricle (C, C´ arrowhead) and more dorsally located cells at the pial surface of the spinal cord (C, asterisk). **D, D´**: Transverse sections through the spinal cord of a *dbx1b:eGFP* transgenic embryo (24 hpf) expressing GFP in V0, V1 interneurons (green). The transgenic line was co-stained with either *sox1a* (D) or *sox1b* (D´) antisense RNA (purple). *sox1a* and *sox1b* expressing cells are distinct from *dbx1b* expressing interneurons. **E-E´´**: A *olig2:eGFP* expressing transgenic line was co-stained with *sox1a* probe. The *sox1a*+ cells do not co-express *olig2:eGFP*. Anterior left, dorsal up. Transverse sections (C-D´) are at the level of the hindgut extension, dorsal up. Abbreviations: m, medial floor plate; n, notochord. Scale bar: 25 µm.

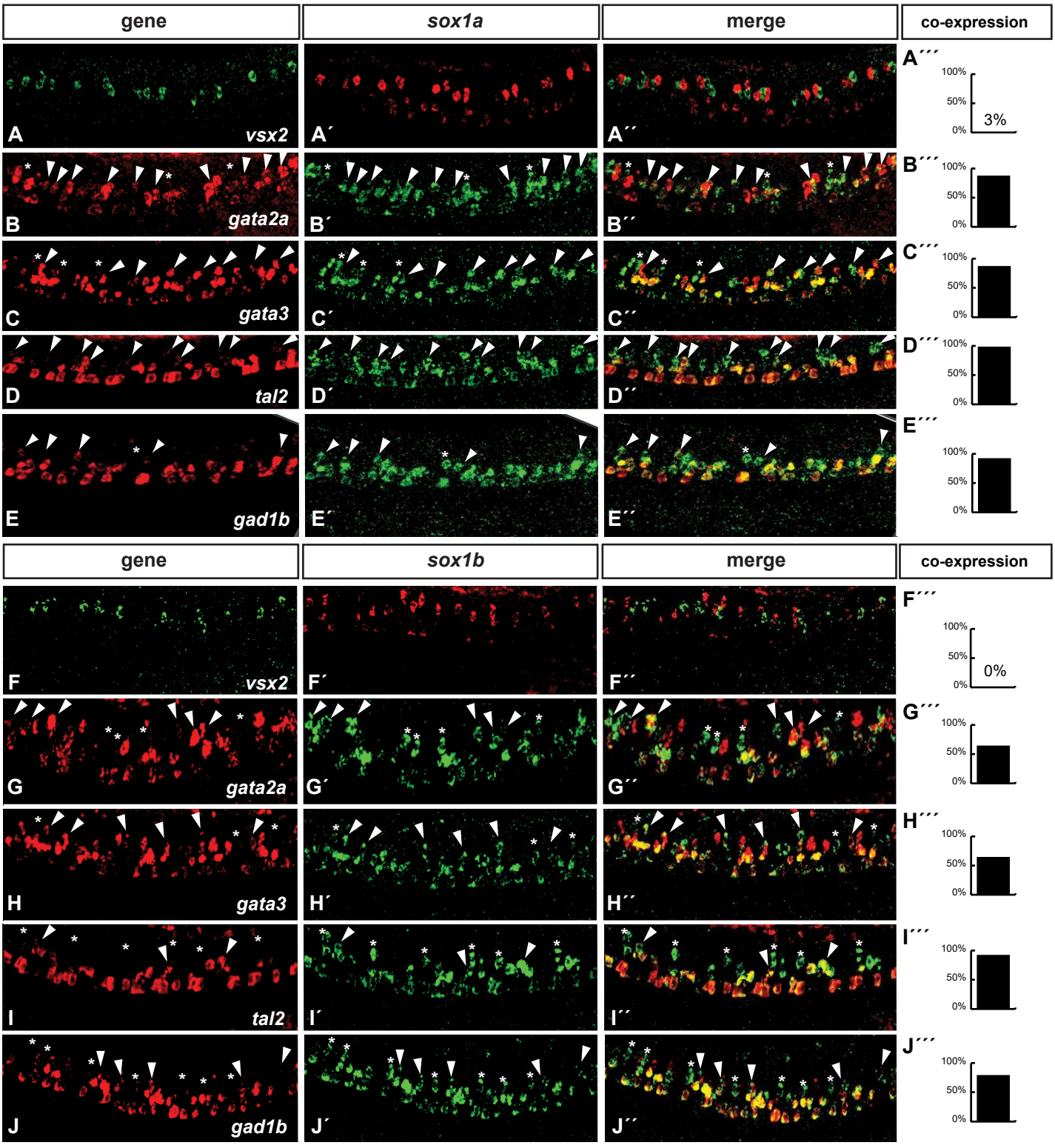
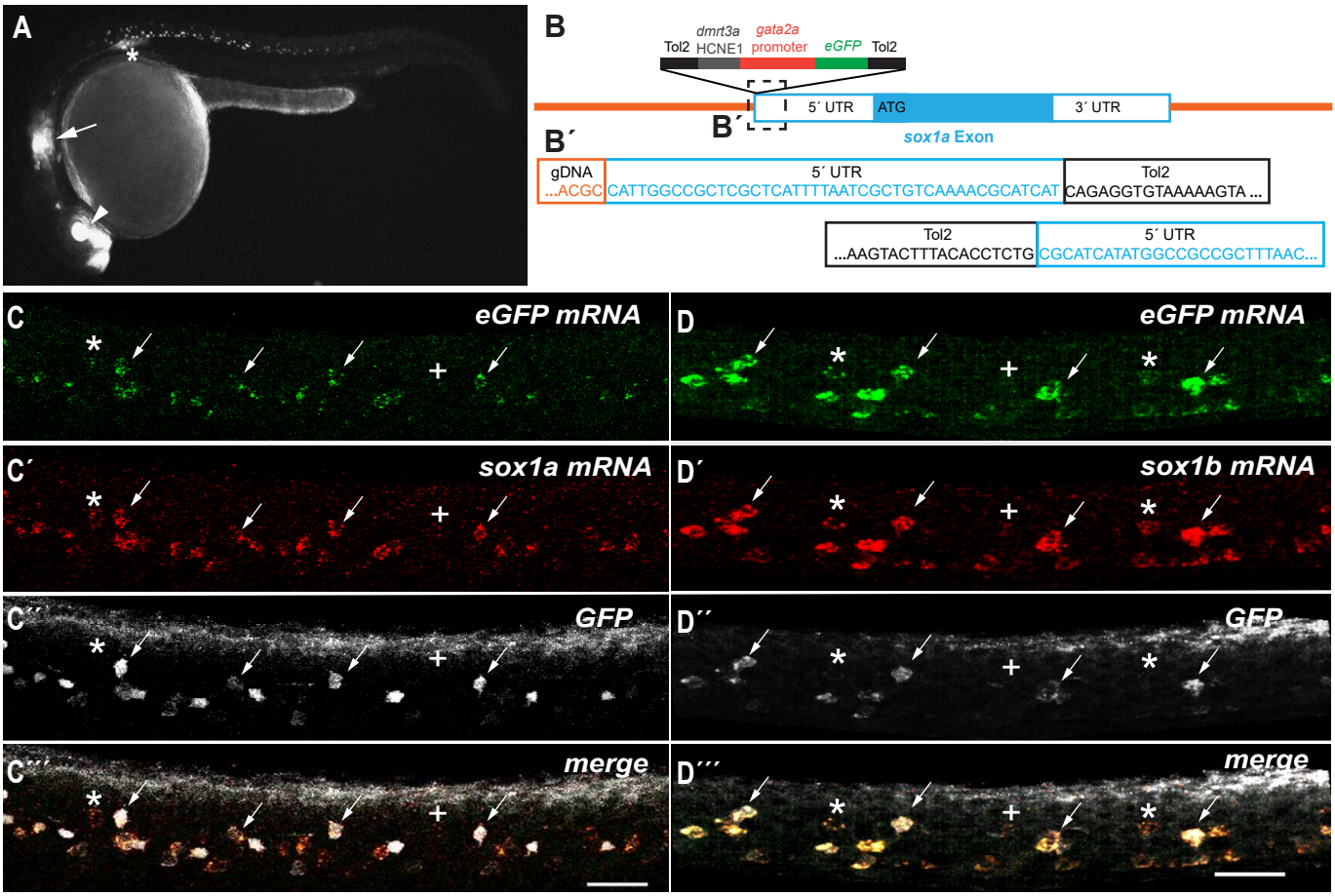


Figure S2

Mapping of *sox1a* and *sox1b* expression relative to genes marking diverse types of interneurons in the spinal cord.

A-A´´: Expression of *vsx2* (A) and *sox1a* (A´) and merged view (A´´). In most cells *vsx2* mRNA does not co-localized with the mRNA of *sox1a* (A´´). Thus, *sox1a* expressing cells are different from *vsx2*+ V2a interneurons (3% *sox1a*+; *vsx2*+, n=115 cells, 7 embryos). **B-B´´**: Expression of *gata2a*, a marker of V2b interneurons, relative to *sox1a* and merged view (B´´). 87% of *gata2a*+ cells in the V2 domain also expressed *sox1a* (n=112 cells). In addition, KA´ (98%, n=63 cells) and KA´´ cells (95%, n=86 cells) expressed *gata2a* and *sox1a* mRNA (5 embryos). **C-C´´**: Cells expressing *gata3*, a marker of V2b interneurons, and *sox1a* and the merged panels (C´´). 87% of *gata3*+ V2 cells also expressed *sox1a* (n=193 cells). Note, 99% of *gata3*+ KA´ (n=195) and 100% of KA´´ cells (n=160 cells) were positive for both *gata3* and *sox1a* mRNA (10 embryos). **D-E´´**: Cells expressing *tal2* (D), *gad1b* (E) and *sox1a* and the merged view (D´´, E´´). 98% of *tal2*+ and 92% of *gad1b*+ V2 domain cells are positive for *sox1a* at 24 hpf (n=48 cells, n=60 cells, respectively). *gad1b* only expressing cells (E-E´´) are intermingled with *sox1a+; gad1b*+ cells (E-E´´). Note, 100% (n=95 cells) of *tal2* expressing KA cells expressed *sox1a*, while 94% (n=100 cells) and 100% (n=84 cells) of *gad1b* expressing KA´´ and KA´ also were positive for *sox1a,* respectively (5 embryos each). **F-J´´**: The expression of *vsx2, gata2, gata3, tal2* and *gad1b* in the spinal cord (F-J), *sox1b* expressing cells and (F´-J´) and merged images (F´´-J´´). Essentially, comparable results were obtained for *sox1b* (F-J´´) as for *sox1a* (A-E´´). *sox1a/b* single labelled cells are indicated by asterisks and co-expressing cells by arrowheads. **A´´´-J´´´**: *sox1a/b* co-expression with the corresponding gene is shown as mean percent. All embryos are 24 hpf. Dorsal up, anterior left. Scale bar: 25µm.



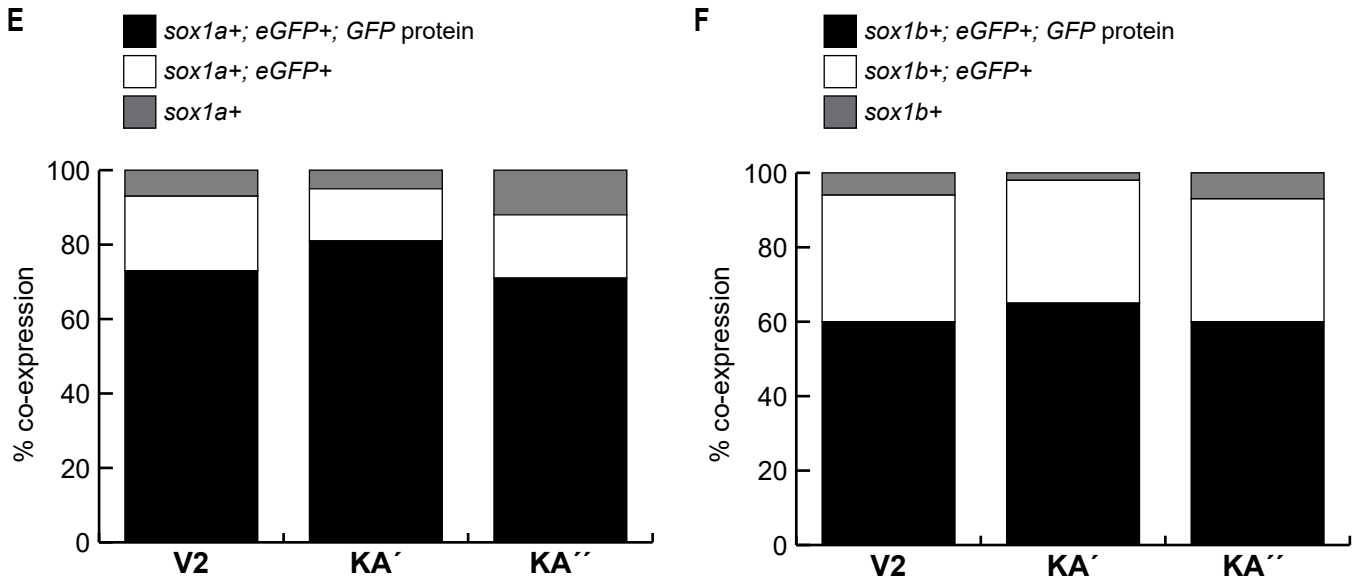


Figure S3

Description and characterization of *sox1a:eGFP* gene trap line.

A: An embryo of the *sox1a:eGFP* gene trap line showing *eGFP* expression in the forebrain, lens (arrowhead), ear (arrow), lateral line (asterisk) and in cells along the spinal cord at 24 hpf. **B**: Schematic drawing showing the integration place of the *sox1a:eGFP* line with close-up (dashed box) of flanking genomic DNA sequence (orange), parts of the *sox1a* 5´ untranslated region (UTR, blue) and the *Tol2* sites (black) of the enhancer test construct (sequences in **B´**). A *Tol2*-based enhancer test construct with a highly conserved non-coding element (HCNE) of the *dmrt3a* gene (grey), followed by a *gata2a* promoter (red) and *eGFP* coding sequence (green) integrated via the *Tol2*-system into the 5´ UTR of the *sox1a* gene (blue). Integration site was verified by whole genome sequencing and subcloning, showing that the construct integrated 43 bp downstream of the 5´ UTR start site in forward direction like the *sox1a* gene. **C-C´´´, E**: Double FISH for *sox1a* mRNA (red) and *eGFP* mRNA (green) with IHC for GFP (white) and quantified as percentage of co-expression in **E** at 24 hpf obtained from 8 embryos. Data are stated as mean percent±s.d.. About 73±3% of *sox1a*+ cells co-express *eGFP* mRNA and are immunoreactive for GFP (arrows, n=124 cells). In addition, 20±3% of cells co-express *sox1a* mRNA and *eGFP* mRNA, but do not express detectable amounts of GFP protein (asterisk) and 7±1% only express *sox1a* mRNA without detectable amounts of *eGFP* mRNA and GFP protein (cross). In the KA´ domain 81±2% of the counted cells (n=170 cells) are *sox1a*+; *eGFP*+ and immunoreactive for GFP, 14±2% co-express solely *sox1a* and *eGFP* mRNA and 5±1% express only *sox1a* mRNA. In the KA´´ domain 71±2% of all counted cells (n=167) express *sox1a* mRNA, *eGFP* mRNA and are immunoreactive for GFP protein. About 17±2% express both *sox1a* and *eGFP* mRNA and 12±2% of cells express solely *sox1a* mRNA. **D-D´´´, F**: Double FISH for *sox1b* mRNA (red) and *eGFP* mRNA (green) with IHC for GFP (white) and quantified as percentage of co-expression in **F** at 24 hpf obtained from 5 embryos. In the V2 domain about 60±3% of all counted cells (n=77) express *sox1b* mRNA, *eGFP* mRNA and GFP protein (arrows).

About 34±4% express *sox1a* and *eGFP* mRNA (asterisks) and 6±1% only express *sox1b* mRNA (cross). In the KA´ domain 65±5% of cells (n=104) are positive for *sox1b* mRNA, *eGFP* mRNA and GFP protein. About 33±5% of cells are positive for *sox1b* and *eGFP* mRNA and only 2±1% express solely *sox1b* mRNA. In the KA´´ domain 60±2% of cells (n=116) express *sox1b* and *eGFP* mRNA and are immunoreactive for GFP protein, about 33±3% of cells co-express *sox1b* and *eGFP* mRNA and 7±2% express solely *sox1b* mRNA. Only one side of the spinal cord is shown. Data are obtained from two independent experiments. Anterior left, dorsal up. Scale bars: 25 µm

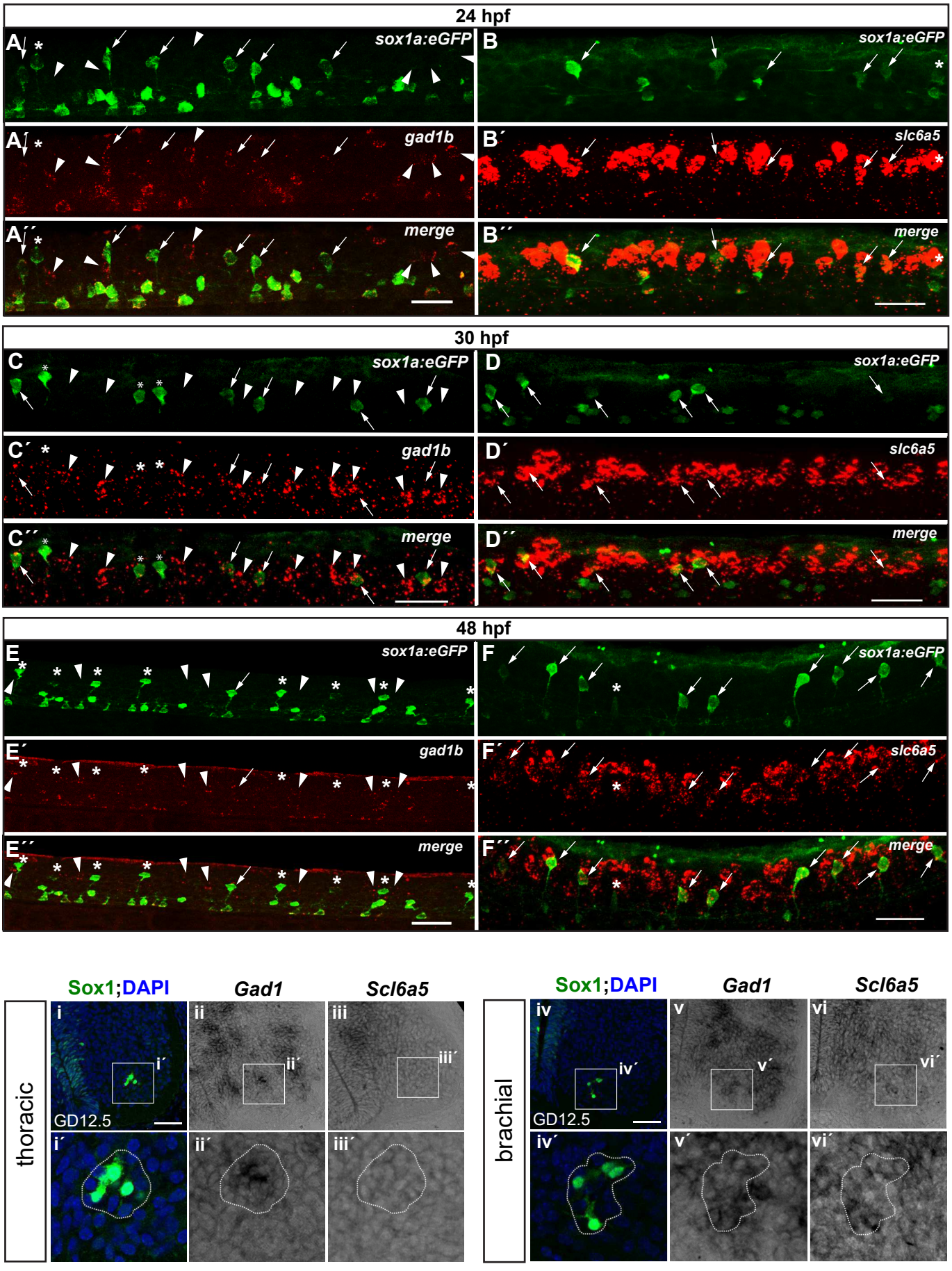


Figure S4

***Sox1a:eGFP+* neurons are initially both, GABA- and glycinergic and later become restricted to a glycinergic neurotransmitter type.**

A-A´´: At 24 hpf in *sox1a:eGFP; gad1b* double labelling experiments, about 78% of cells co-express *sox1a* and *gad1b* (arrows, n=91). We noted also *gad1b* only expressing cells (arrowheads) and a few cells expressing only *sox1a:eGFP* (asterisk). **B-B´´**: In *sox1a:eGFP; slc6a5* double labelling experiments at 24 hpf, about 64% of *sox1a:eGFP+* cells co-express *slc6a5* (arrows, n=34) and 36% are only *sox1a:eGFP+* (asterisk, n=19). **C-C´´**: At 30 hpf, 45% of the *sox1a:eGFP* cells are positive for *gad1b* (arrows, n=51), 55% cells express only *sox1a:eGFP* (asterisk, n=61) and are intermingled with *gad1b+;sox1a:eGFP-* cells (arrowheads). **D-D´´**: At 30 hpf, 70% of *sox1a:eGFP+* cells are glycinergic (arrows, n=56). **E-E´´**: At 48 hpf only 8% of *sox1a:eGFP+* neurons are GABAergic (arrow, n=8) and are intermingled with *sox1a:eGFP* only expressing cells (asterisk) and *gad1b+* cells (arrowheads). **F-F´´**: At 48 hpf, 93% of *sox1a:eGFP+* neurons have become glycinergic (arrows, n=84). A small number of *eGFP+* cells are negative for *slc6a5* (asterisk). **A-F**: All embryos in lateral views of spinal cord over yolk extension. Data are obtained from two independent experiments. Only one side of the spinal cord is shown. Anterior left, dorsal up. **i-iv´:** Immunodetection for Sox1 and DAPI labelling (i, i´,iv, iv´), ISH for *Gad1* (ii, ii´, v, v´) and *Scl6a5* (iii, iii´, vi, vi´) on adjacent transverse sections of GD12.5 spinal cord at thoracic and brachial levels. x´ panels are zoom in on the squared region of interest shown in x panels. Scale bars: 25 µm in A´´, B´´, C´´, D´´, E´´, F´´; 100 µm in i, iv.

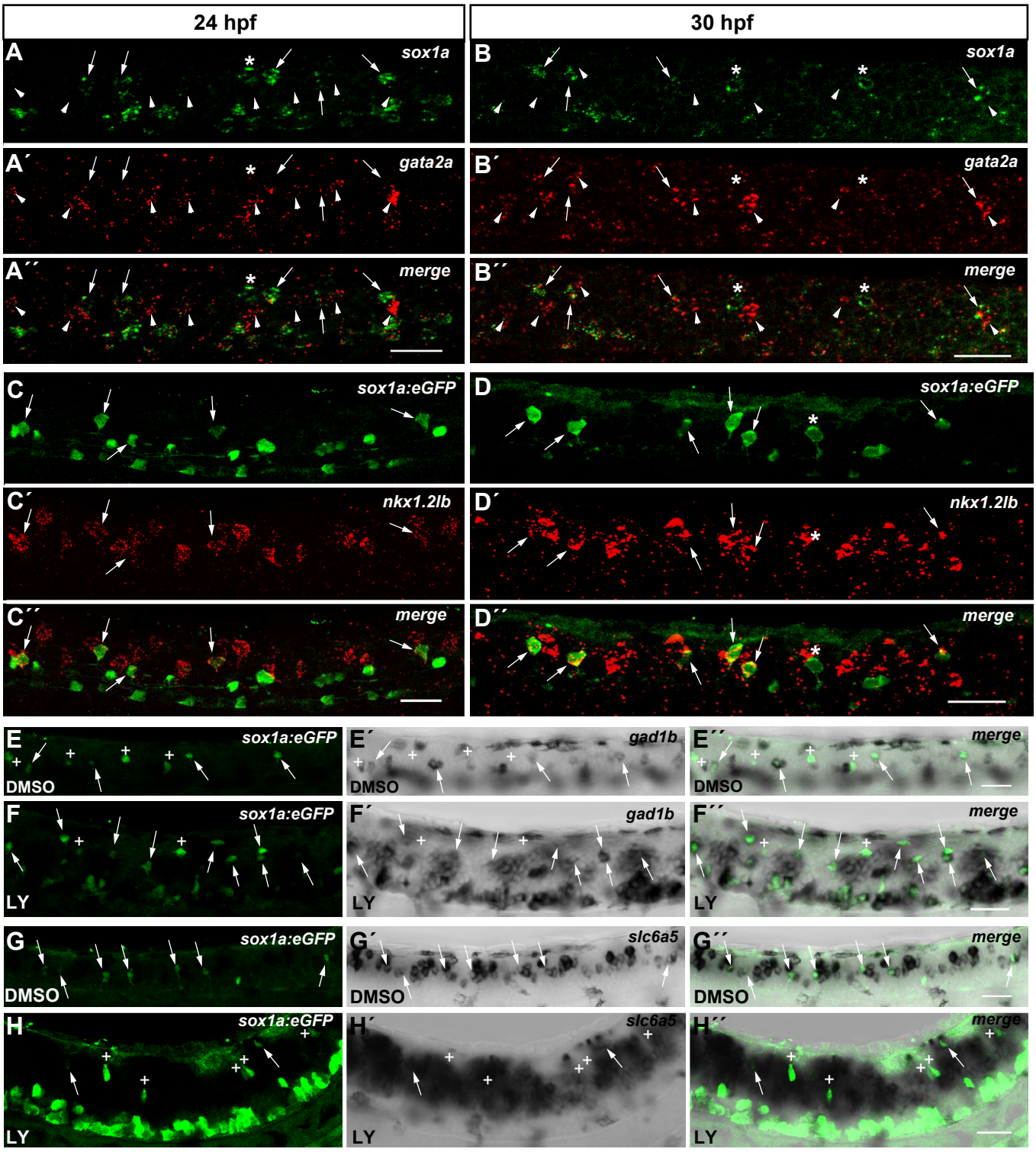


Figure S5

V2s neurons co-express *gata2a* and *nkx1.2lb* at 24 and 30 hpf and are influenced by Notch inhibition.

A-B´´: Double FISH experiments showing expression of *sox1a* (green) and *gata2a* (red) mRNA at 24 (**A-A´´**) and 30 hpf (**B-B´´**) on a 5-somite long stretch of the spinal cord above the yolk extension. At 24 hpf, 65% of *sox1a*+ cells co-express *gata2a* (arrows, n=109 cells) whereas only 32% of cells co-express both genes at 30 hpf (n=107, Fig. 4A). **C-D´´**: IHC for GFP (green) and FISH of *nkx1.2lb* mRNA (red) in *sox1a:eGFP* transgenic embryos at 24 (**C-C´´**) and 30 hpf (**D-D´´**). Noteworthy is that *nkx1.2lb* is only co-expressed in the V2s neurons and not in KA neurons. However, *nkx1.2lb* is also expressed in other *sox1a*- cells from the intermediate to the dorsal region of the spinal cord. At both time points *sox1a:eGFP*+ cells co-express to more than 80% *nkx1.2lb* (arrows, Fig. 4A). **E-H´´**: *sox1a:eGFP* embryos raised in 0.04% DMSO vehicle control (**E-E´´, G-G´´**) or Notch inhibitor (LY 411575 in 0.04% DMSO). *sox1a:eGFP* embryos were treated from 16-24 hpf (**F-F´´, H-H´´**) and analysed at 30 hpf by ISH for *gad1b* or *slc6a5* mRNA (black) and IHC for GFP (green). Single labelled *sox1a:eGFP*+ V2s cells are indicated with crosses. Note, inhibition of Notch signalling leads most likely to the precociously differentiation of other spinal cord progenitor cells increasing the overall number of GABA- and glycinergic cells. **E-G´´**: Inhibition of Notch signalling from 16-24 hpf increased the numbers of gad1b+ V2s cells (arrows) at 30 hpf by 2.3-fold (Fig. 4F). F-H´´: At the same time, Notch inhibition caused a 2.3-fold decrease of slc6a5+ V2s cells (arrows, Fig. 4F). Only one side of the spinal cord is shown. Data are obtained from two independent experiments. Dorsal up, anterior left. Scale bars: 25 µm.

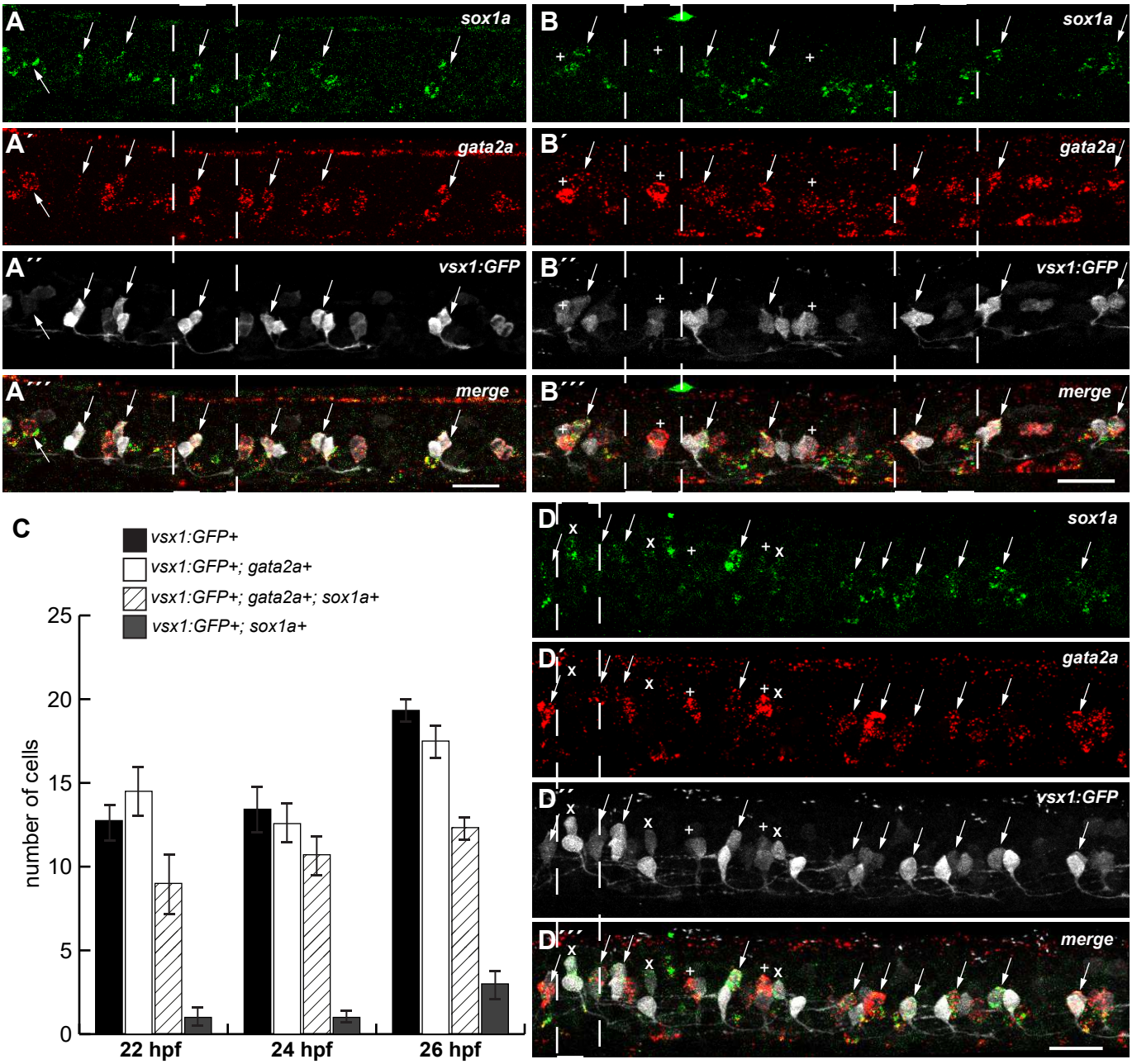


Figure S6

Co-expression analysis of *vsx1:GFP, sox1a* and *gata2a* identifies 4 different cell types in the V2 domain.

A-B´´´, D-D´´´: Double FISH experiments showing expression of *sox1a* (green), *gata2a* (red) and GFP immunohistochemistry (white) in *Tg(vsx1:GFP)* embryos at 22 hpf **(A-A´´´)**, 24 hpf **(B-B´´´)** and 26 hpf **(D-D´´)**. Arrows indicate cells which express *gata2a* and *sox1a* mRNA and are *vsx1:GFP*+. Crosses label *vsx1:GFP*+*; gata2a*+ cells and x indicate *vsx1:GFP*+*; sox1a*+ cells. Cells that express only *vsx1:GFP* are most likely V2a cells. **A-A´´´**: At 22 hpf *sox1a* is co-expressed with *gata2a* mRNA in one cell of a *vsx1:GFP*+ pair (arrows, n=36 cells). **B-B´´´**: At 24 hpf *sox1a* mRNA is still co-expressed with *gata2a* mRNA in one cell of a *vsx1:GFP*+ pair (arrows, n=75 cells) and some *vsx1:GFP*+cells are solely and strongly expressing *gata2a*+ (crosses, n=88). **C-C´´´**: Graph showing quantification (data are presented as mean±s.e.m.) of co-expression analysis of *vsx1:GFP*, *sox1a* and *gata2a* obtained from 4 embryos at 22 hpf, 7 embryos at 24 hpf and 6 embryos at 26 hpf. **D-D´´´**: At 26 hpf a few cells express only *sox1a* mRNA in a *vsx1:GFP*+ neuron (indicated by x, n=18 cells). However, most cells still co-express *gata2a* and *sox1a* mRNA, as seen in a *vsx1:GFP*+ pair (arrows, n=74 cells). Some *vsx1:GFP*+ cells express strongly and only *gata2a* mRNA (crosses, n=105). Dashed boxes are close-ups in Fig. 4B-E´´´. Only one side of the spinal cord is shown as lateral view. Data are obtained from two independent experiments. Dorsal up, anterior left. Scale bars: 25 µm.

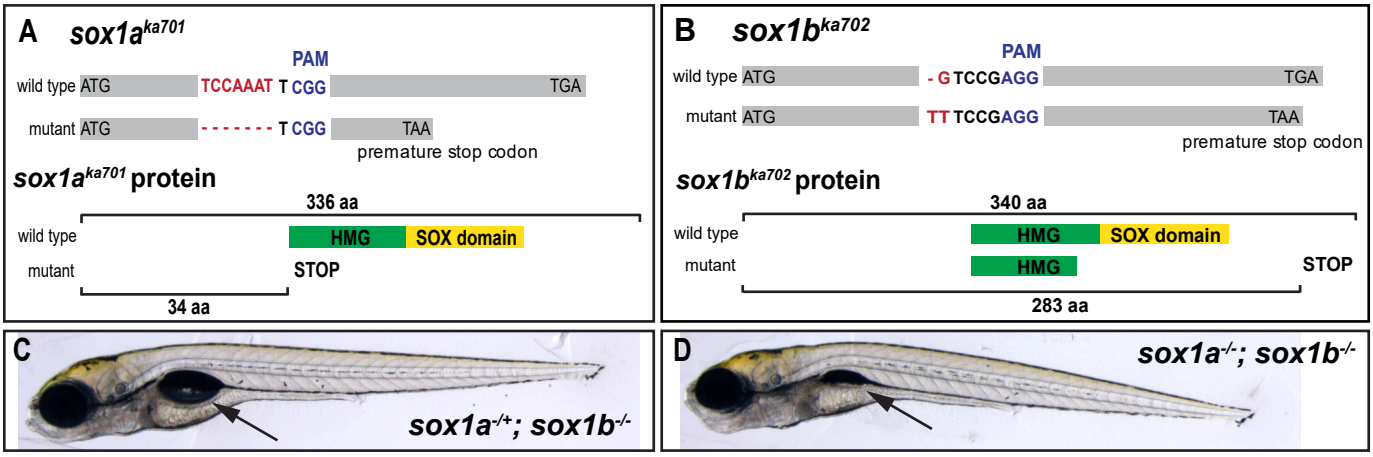


Figure S7

***sox1a/b* mutant allele description and larval phenotypes.**

A-B: Schematic drawing of CRISPR/Cas9 generated *sox1a*^{ka701} and *sox1b*^{ka702} mutant alleles showing induced mutations relative to the *sox1a* and *sox1b* wild type coding sequences and protein structures including the predicted mutant protein structure. **A**: In the *sox1a*^{ka701} mutant allele a stretch of 7 bp was deleted (wild type sequence in red), which leads to a frame shift and a premature stop codon resulting in a putative truncated protein of 34 aa. The truncated protein will lack the DNA-binding HMG domain (green) and the partner-protein-binding SOX transcription factor domain (yellow). In addition, the 9 amino acids before the stop codon are altered (red box in mutant protein). **B**: In the *sox1b*^{ka702} allele, a 2 bp indel mutation was generated: one nucleotide was inserted (thymine) and the guanine at the position four nucleotides before the PAM (blue) was exchanged to thymine, which leads to a frame shift and a premature stop. The putative truncated protein contains parts of the HMG domain. After 77 amino acids, the following 206 amino acids are aberrant (red box in mutant protein), and thus the mutant protein will lack the entire SOX transcription factor domain. **C**: A *sox1a*$^{ka701/+}$; *sox1b*$^{ka702/ka702}$ larvae with a filled swim bladder (arrow) at 5 dpf. **D**: A *sox1a*$^{ka701/ka701}$; *sox1b*$^{ka702/ka702}$ larvae where the swim bladder is not inflated at 5 dpf.

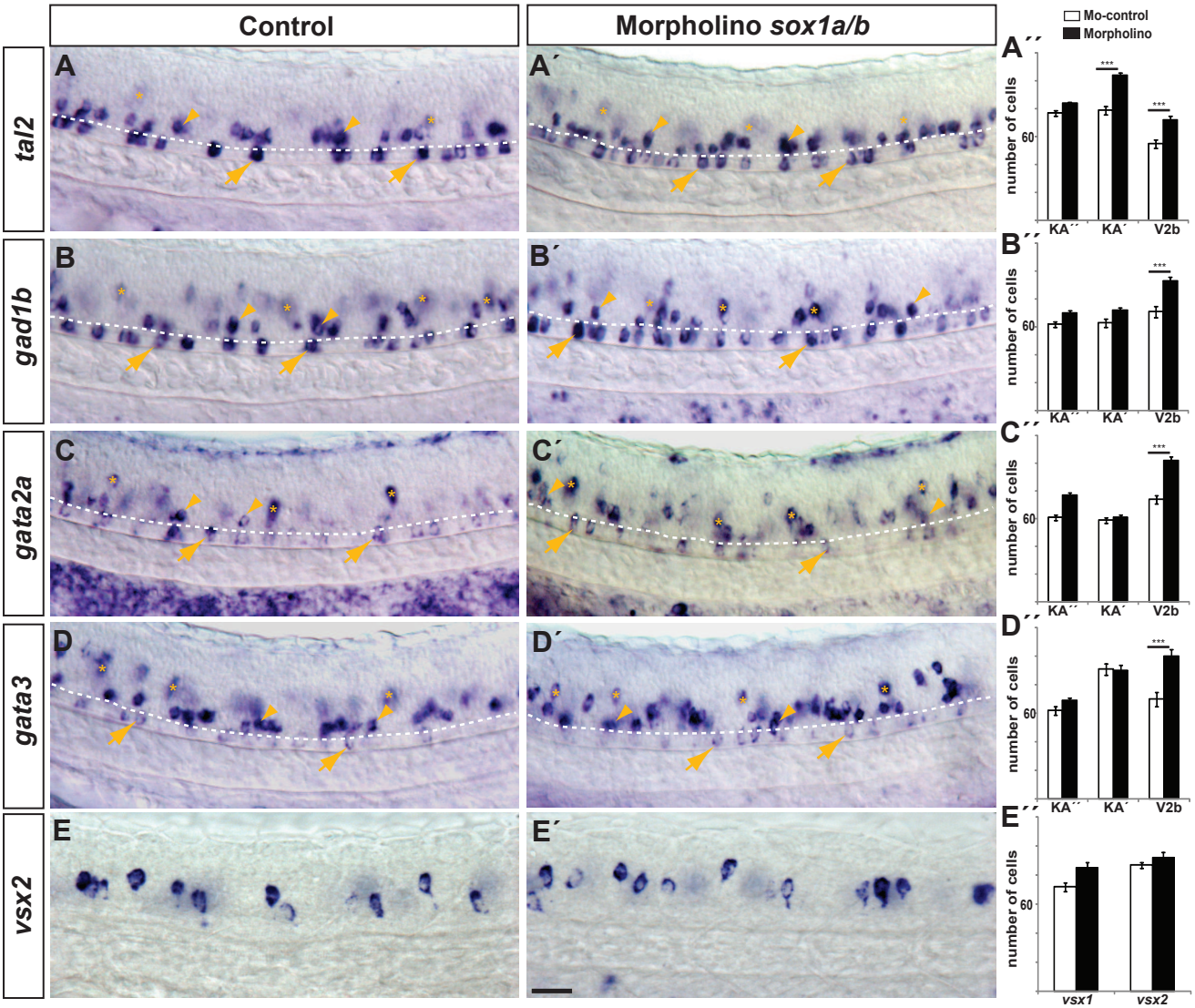
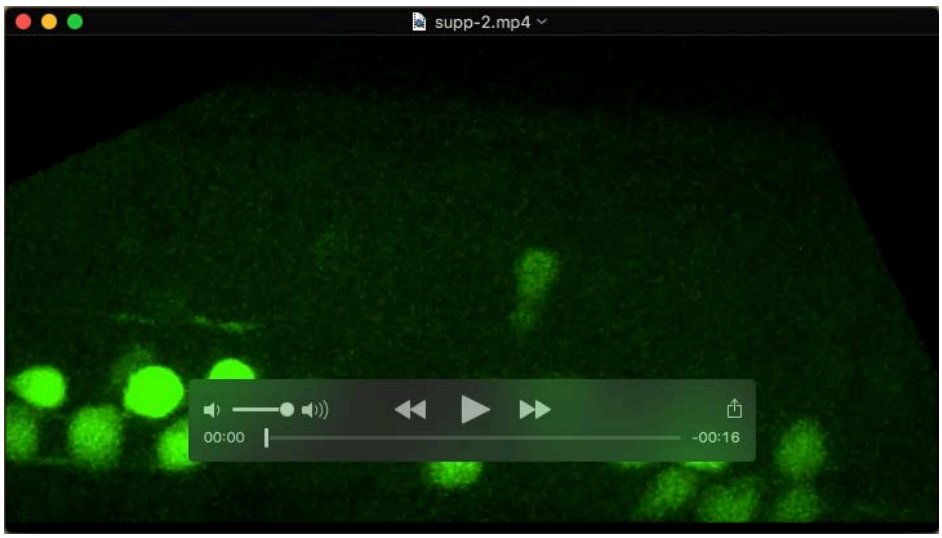


Figure S8

Knockdown of Sox1a and Sox1b results in an increase of V2b cells.

A-E´´: Control injected with mismatch morpholinos (A, B, C, D, E), *sox1a/b*-morpholino injected embryos (A´, B´, C´, D´, E´) and quantification (A´´, B´´, C´´, D´´, E´´) of KA´´ (arrows), KA´ (arrowheads), V2b (asterisks) and V2a cells. Data are presented as mean±s.e.m.. **A-A´´:** Loss of function of Sox1a/b results in a ~30% increase of *tal2*+ KA´ and V2b cells (control: 77±2 KA´´, 79±3 KA´ and 55±3 V2b cells, 23 embryos; knock-down: 84±1 KA´´, 104±2 KA´ and 72±3 V2b cells, 44 embryos). **B-B´´:** Knock-down of Sox1a/b resulted in a 30% increase of *gad1b*+ V2b cells (control: 62±2 KA´´, 63±3 KA´ and 71±4 V2b cells, 28 embryos; knock-down: 70±2 KA´´, 72±2 KA´ and 93±3 V2b cells, 33 embryos). **C-C´´:** An increase of 38% *gata2a*+ V2b cells was observed in morpholino injected embryos (control: 61±2 KA´´, 59±2 KA´ and 74±3 V2b cells, 26 embryos; knock-down: 77±2 KA´´, 61±2 KA´ and 102±3 V2b cells, 29 embryos). **D-D´´:** Knock-down of Sox1a/b increased *gata3*+ V2b cells by 43% (control: 62±3 KA´´, 91±4 KA´ and 70±5 V2b cells, 32 embryos; knock-down: 69±2 KA´´, 90±4 KA´ and 100±5 V2b cells, 23 embryos). **E-E´´:** Sox1a/b loss-of-function did not affect *vsx1* (ISH not shown) or *vsx2* expressing V2a cells (control: 72±3 *vsx1*+ and 87±2 *vsx2*+ V2a cells, 25 and 38 embryos, respectively; knock-down: 85±4 *vsx1*+ and 92±4 *vsx2*+ V2a cells, 12 and 16 embryos, respectively). All embryos are 24 hpf and cells were counted in the spinal cord along the entire trunk. Data are from at least two independent experiments. Statistical significance was assessed using the unpaired two-tailed Student´s *t*-test ***$p \leq 0.001$. Dorsal up, anterior left. Scale bar: 25 µm.



Movie 1

Time lapse movie of *sox1a:eGFP* transgenic embryo from 20 hpf to about 36 hpf in lateral view of the spinal cord. *Sox1a:eGFP*+ V2 neurons initiate GFP expression in the intermediate region of the spinal cord and send their axons downwards to the tail. Ventrally located cells are KA neurons. Note that the *sox1a:eGFP* expressing cells never divided. We thus believe that GFP accumulation relative to mRNA accumulation is delayed in the transgene precluding the observation of very early events during and immediately after birth of the neuron.

Table S1: Markers of different V2 interneuron subtypes in the zebrafish spinal cord at 24 hpf.

24 hpf	V2a/b,s progenitor	V2a	V2b,s precursor	V2b	V2s
gata2a	+		+	+	
gata3			+	+	
tal2			+	+	
tal1			+	+	
gad1b			+	+	+
sox1a			+		+
sox1b			+		+
nkx1.2lb					+
slc6a5					+
vsx1		+			
vsx2		+			

Table S2: Gene names, previous names and references for *in situ* hybridization probes used in this study.

Gene name	Previous name	Reference for probe
gata2a	*gata2*	(Detrich et al., 1995)
gata3		(Yang et al., 2010)
tal2		(Pinheiro et al., 2004)
tal1		(Armant et al., 2013)
sox1a		(Armant et al., 2013)
sox1b		(Armant et al., 2013)
nkx1.2lb		(Armant et al., 2013)
slc6a5	*glyt2* (mix of *glyt2a* and *b* probes)	(Higashijima et al., 2004)
gad1b	*gad67* (mix of *gad67a* and *b* probes)	(Higashijima et al., 2004)
vsx1		(Passini et al., 1997)
vsx2		(Passini et al., 1997)
eGFP		this study

Supplementary references

Armant, O., Marz, M., Schmidt, R., Ferg, M., Diotel, N., Ertzer, R., Bryne, J. C., Yang, L., Baader, I., Reischl, M., et al. (2013). Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Dev Biol* **380**, 351-362.

Briscoe, J., Pierani, A., Jessell, T. M. and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* **101**, 435-445.

Detrich, H. W., 3rd, Kieran, M. W., Chan, F. Y., Barone, L. M., Yee, K., Rundstadler, J. A., Pratt, S., Ransom, D. and Zon, L. I. (1995). Intraembryonic hematopoietic cell migration during vertebrate development. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10713-10717.

Higashijima, S., Masino, M. A., Mandel, G. and Fetcho, J. R. (2004). Engrailed-1 expression marks a primitive class of inhibitory spinal interneuron. *J Neurosci* **24**, 5827-5839.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Passini, M. A., Levine, E. M., Canger, A. K., Raymond, P. A. and Schechter, N. (1997). Vsx-1 and Vsx-2: differential expression of two paired-like homeobox genes during zebrafish and goldfish retinogenesis. *The Journal of comparative neurology* **388**, 495-505.

Pinheiro, P., Gering, M. and Patient, R. (2004). The basic helix-loop-helix transcription factor, Tal2, marks the lateral floor plate of the spinal cord in zebrafish. *Gene expression patterns : GEP* **4**, 85-92.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26.

Yamada, T., Pfaff, S. L., Edlund, T. and Jessell, T. M. (1993). Control of cell pattern in the neural tube: motor neuron induction by diffusible factors from notochord and floor plate. *Cell* **73**, 673-686.

Yang, L. X., Rastegar, S. and Strahle, U. (2010). Regulatory interactions specifying Kolmer-Agduhr interneurons. *Development* **137**, 2713-2722.