Supporting Information for

BRCA1 and ELK-1 regulate Neural Progenitor Cell Fate in the Optic Tectum in response to Visual Experience in *Xenopus laevis* tadpoles

Lin-Chien Huang*, Caroline R. McKeown*, Haiyan He, Aaron C. Ta, Hollis T. Cline#

* Equal contributions

Corresponding author:
Hollis T. Cline
Email: cline@scripps.edu

This PDF file includes:

Supporting text

Figures S1-S6

Tables S1-S3

SI References

Other supporting materials for this manuscript include the following:

Data Tables S1-S4

Supporting Methods and Results

Animals

Albino *Xenopus laevis* tadpoles of both sexes were obtained from an in-house colony at the Department of Animal Resources at the Scripps Research Institute, La Jolla or purchased from Xenopus Express (Brooksville, FL, USA, RRID:XEP_Xla200). Animals were reared in 0.1X Steinberg's solution at 22^{o}C under a 12 hour light/12 hour dark cycle, and anesthetized before all procedures in 0.02% tricaine methanesulfonate (MS-222). When experiments were completed, animals were euthanized with 0.1% MS-222. Animals were staged according to Nieuwkoop and Faber (1). All animal protocols were approved by the Institutional Animal Care and Use Committee of Scripps Research (protocol # 08-0083-3).

Isolation of enriched populations of neural progenitor cells and immature neurons from the optic tectum

Cell samples enriched in neural progenitor cells and immature neurons were collected from tadpole midbrain as follows. Stage 46 tadpoles (1) were taken directly from their 12h dark cycle, anesthetized in 0.02% MS-222, and their brains were electroporated with plasmids expressing Gal4-UAS-turboGFP (tGFP) driven by the oct4/sox2 enhancer from the minimal FGF promoter, called pSOX2-bd::turboGFP. pSOX2-bd::tGFP (2µg/µl) was injected into the ventricle and electrodes were positioned next to the midbrain to electroporate plasmid into ventricular layer cells. After 24h, tGFP was detected in neural progenitor cells in the ventricular layer and immature neurons in the neuronal layers (2, 3). The HMG-Box transcription factor, SOX2, is expressed in neural progenitor cells in vertebrate brain, including Xenopus (3, 4) and is necessary to maintain progenitor cell identity (5). Expression of tGFP from this plasmid requires pre-existing SOX2 protein to be present in the cell, and tGFP expression levels are amplified by the Gal4-UAS elements. This process birthdates SOX2+ neural progenitor cells lining the ventricle with tGFP.

We tested the birthdating ability of pSOX2bd::tGFP by co-electroporating the plasmid along with translation-blocking morpholinos targeted against the *sox2* transcript, which do not affect levels of pre-existing SOX2 protein. Co-electroporating *sox2* morpholinos with pSOX2bd::tGFP did not interfere with tGFP expression indicating that SOX2 protein present prior to introduction of morpholinos was sufficient to drive

expression of pSOX2bd::tGFP in progenitors. Furthermore, tGPP expression persisted in neuronal progeny (Figure S1). Therefore, pSOX2bd::tGFP is a powerful tool for birthdating progenitors and their neuronal progeny with tGFP expression in the Xenopus brain, even under conditions that might subsequently alter the generation of additional neural progenitors.

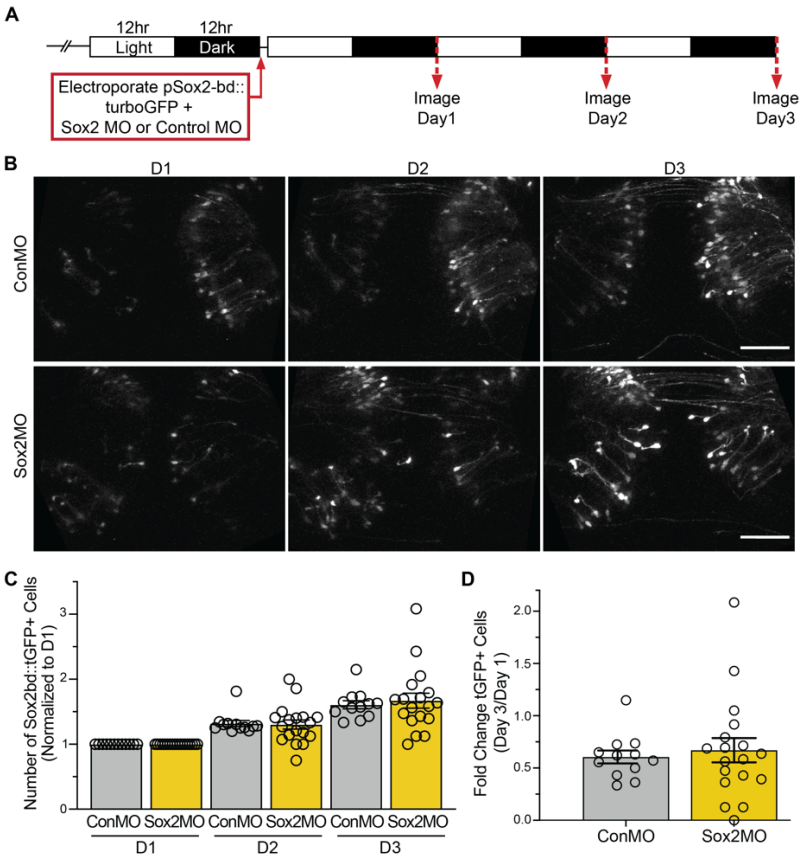

Figure S1. pSOX2bd::tGFP labels cells even in the absence of new SOX2 synthesis.

A. Schematic diagram of *in vivo* imaging protocol. After co-electroporation with pSOX2bd::tGFP and either control or *sox2* morpholinos (MO), animals were reared in a 12h light/12h dark cycle and imaged *in vivo* for 3 consecutive days by confocal microscopy. **B.** Representative images of tGFP+ cells in the optic tectum of control (top, ConMO) and *sox2* (bottom, M *sox2*MO) morpholino-treated animals, imaged daily for 3 days. **C.** Number of Sox2bd::tGFP-labeled cells in animals treated with control (gray) and *sox2* (yellow) morpholinos over 3 days. Cell numbers are normalized within animals to day 1. Bar graphs indicate average ±SEM, with circles indicating individual animal numbers. At each day, the number of cells is the same between groups (Mann-Whitney test). Note that labeled cells continue to divide over the 3 day experiment in both treatment groups. **D.** Difference in fold-change of Sox2bd::tGFP positive cells over the 3 day experiment is not significant between control (gray) and *sox2* (yellow) morpholino-treated animals (Mann-Whitney test). Scale bar, 100 µm.

Immediately following electroporation, animals were exposed to either a simulated motion stimulus, referred to as enhanced visual stimulation, to enrich for immature neurons, or dark for 24 hours to enrich for neural progenitor cells, as described (2, 3, 6). Midbrains were collected from 100 animals reared in each condition and dissociated into single cells with amphibian PBS (NaCl 113mM, Na2HPO4 8mM, KH2PO4 1.5mM, EDTA 0.1%, EGTA 2mM). Approximately 40,000 tGFP+ cells were collected from ~100 animals from each condition, using Fluorescence Activated Cell Sorting (FACS; FACSAria II, BD Biosciences, USA; RRID:SCR_018934). For FACS, forward scatter was used to set the threshold for cell size and side scatter was used to set the threshold for cellular granularity. The background fluorescence in the FITC channel was set according to fluorescence from cells dissociated from midbrains without electroporation and cells with green fluorescence higher than the background were collected. Forward and side scatter plots indicated no difference in the size or granularity of the tGFP+ cells compared to the non-electroporated cells, which are a mixture of neural progenitor cells and immature neurons. There was no difference in size and granularity between tGFP+ neural progenitor cells and immature neurons. In addition, there was no overlap between tGFP+ cells and cells labeled with SytoxRed (S34859, Life Technologies), a nuclear dye that labels unfixed dead cells, suggesting that the tGFP+ cells are healthy. Total RNA was extracted using the mirVana kit (Life Technologies, USA), followed by DNase treatment to remove genomic DNA and followed by clean-up using RNeasy mini kit (Qiagen, USA). The total RNA isolated from these sorted cells was high quality (Figure S2). Samples with RIN values of 9.4 for neural progenitor cells and 9.2 for immature neurons, measured with a Bioanalyzer, were used for subsequent analysis. 2ng of total RNA was amplified to 2–3 μg of double-stranded cDNA, using the Ovation RNA-Seq System V2 (NuGEN, USA). The amplified cDNA was purified, using Agencourt AMPure XP beads (Beckman Coulter, Inc.), quantified by NanoDrop using Agilent Bioanalyzer. Three biological replicates were analyzed for each condition.

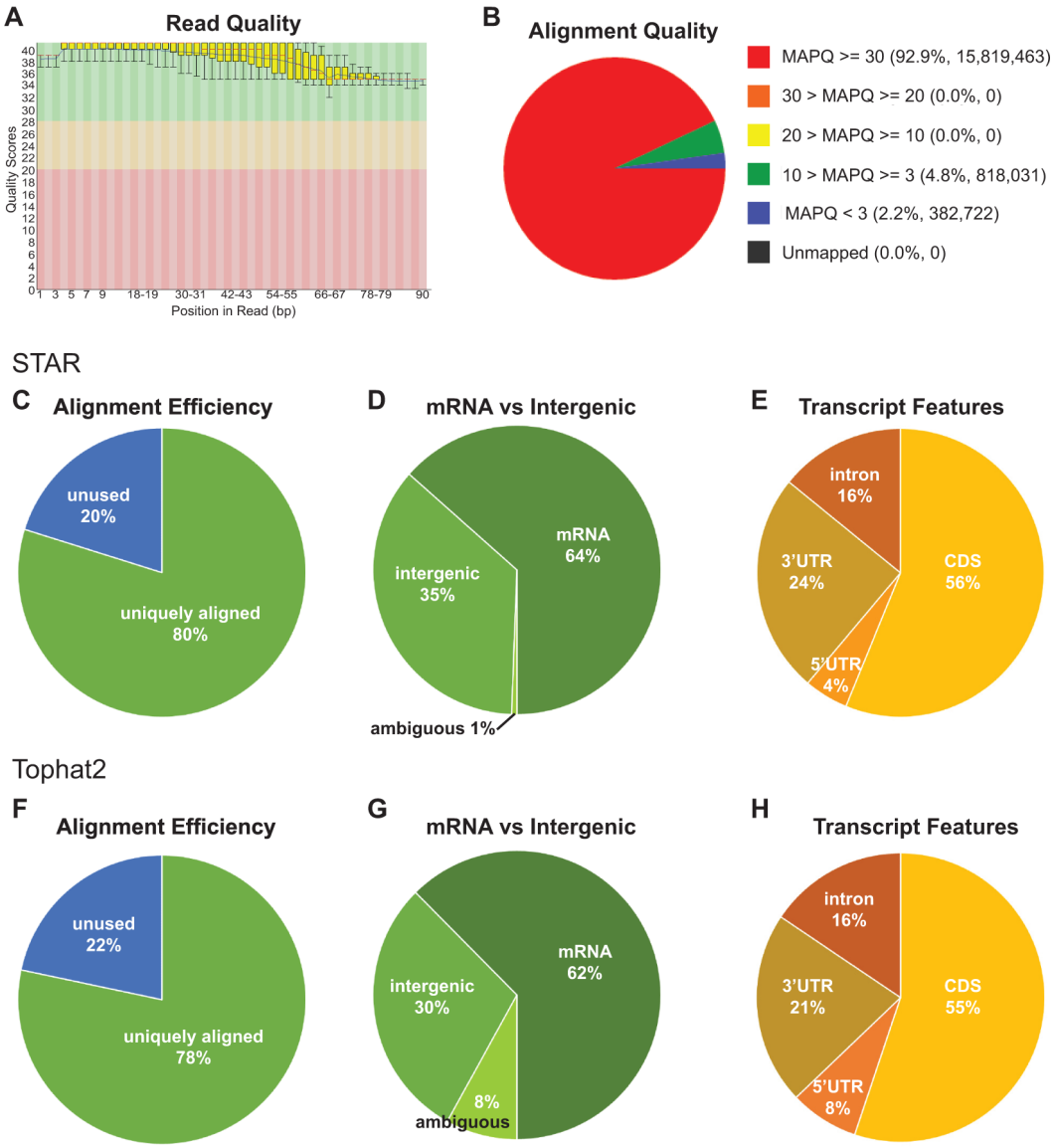

Figure S2. Read quality and alignment of differential expression dataset.

A. Read quality of entire dataset including 3 biological replicates of both NPCs and immature neurons. **B.** MAPQ alignment quality using SAMStat. **C-E.** Details of alignment of reads against *Xenopus laevis* genome, using STAR. **C.** 80% of RNA-seq reads in average are uniquely aligned to the genome scaffolds, indicating the specificity of the reads to *Xenopus laevis* genome. **D.** Against genome scaffold (J-strain v9.1) and gff3 file, 64% of the aligned reads in average belong to mRNA; while 35% of the aligned reads in average belong to the regions between mRNA, ie. intergenic region. **E.** Percentage of reads aligned to the features in the transcripts. **F-H.** Details of alignment of reads against *Xenopus laevis* genome, using TopHat2. **F.** 78% of RNA-seq reads in average are uniquely aligned to the genome scaffolds, indicating the specificity of the reads to *Xenopus laevis* genome. **G.** Against genome scaffold (J-strain v9.1) and gff3 file, 62% of the aligned reads in average belong to mRNA; while 30% of the aligned reads in average belong to the regions between mRNA, ie. intergenic region. **H.** Percentage of reads aligned to the features in the transcripts.

RNA-Seq of neural progenitor cells and immature neurons

1µg of cDNA was sheared in microTube (Covaris) and then used for library preparation (KAPA Taq PCR kits). The size selection for the final PCR product between 200 – 500bp was done by gel purification. The next generation sequencing was done using HiSeq2000 platform (Illumina; RRID:SCR_020132) for single-end reads at size 100bp. Samples were multiplexed in one lane at the Next Generation Sequencing Core at the Scripps Research Institute (La Jolla). Each sample has between 17 and 20 million reads (Table S1). We used the differential expression analysis package, DESeq2 (RRID:SCR_015687) (7), and graphics are performed under R (v3.1.2; cran.r-project.org; RRID:SCR_001905) through Bioconductor (RRID:SCR_006442)(8).

Analysis of Read Quality: Methods: The quality of raw reads was reviewed, using FASTQC (v0.11.4) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/; RRID:SCR_014583). Two aligners, STAR (v2.4.0j; RRID:SCR_004463) (9) and TopHat2 (v2.0.13; RRID:SCR_013035) (10), were used to align the reads against the J-strain, v9.1 *Xenopus laevis* genome assembly (Xenbase; RRID:SCR_003280) with the annotation using gene model (JGI v1.8; Xenbase)(11, 12). The v9.1 of the *Xenopus laevis* genome release incorporated >90% of the genome sequence into 18 pseudomolecules representing the 18 chromosomes of *X. laevis*. These 18 pseudomolecules can be categorized as 9 pairs, each pair containing one L and one S pseudomolecule. 45,099 primary transcripts are annotated in v9.1, with primary transcripts being defined as the longest splice variant of a particular gene. Not all the annotated transcripts have a published gene symbol. 17,409 of the transcripts have a published gene symbol; 20,685 have a gene symbol starting with xelaev, which means that this transcript is specific to *X. laevis*; 4,293 start with LOC, which have no published symbols and no orthologs; 2,712 start with xetrop which are homologous to *Xenopus tropicalis* but do not have any published gene symbol. Approximately 38.6% of the primary transcripts have a known gene symbol. These transcripts were used for functional classification with existing databases. For alignment with STAR, no trimming on the raw reads was performed before alignment since STAR itself has a soft clipping function. For TopHat2 alignment, dynamic trimming was performed, using Trimmomatic (v0.32; RRID:SCR_011848) (13). Alignment quality was examined, using SAMStat (v1.09; RRID:SCR_005432) (14). Only the reads with MAPQ score higher than 20 were included in the differential expression analysis. The reads were counted by HTSeq (HTSeq-count; v0.6.1p1; RRID:SCR_005514;

RRID:SCR_011867) (15). Differential expression analysis package, DESeq2 (RRID:SCR_015687) (7), and graphics are performed under R (v3.1.2; cran.r-project.org; RRID:SCR_001905) through Bioconductor (RRID:SCR_006442)(8).

Differential expression analysis with DESeq2: Out of 45,099 transcripts, 27,027 and 26,137 transcripts were detected, using aligner STAR and TopHat2, respectively, based on the criteria that the total number of normalized counts is larger than 6 across 6 samples, i.e. one count per sample on average. The difference in the number of transcripts detected between STAR and TopHat2 likely results from a difference in the number of reads that are classified as ambiguous. We used two methods to detect changes in transcript expression: one calculates differential expression based on the reads aligned to the whole transcript/mRNA, while the other calculates the reads aligned to the coding sequence (CDS) only. The statistical analysis based on CDS only shows which transcripts are differentially expressed between neural progenitor cells and immature neurons and have the potential to be translated into proteins, while the analysis based on mRNAs incorporates the information from 5'UTR, 3'UTR and introns in addition to CDS. Our alignment statistics and quality analysis are further detailed in Figures S2, S3. We identified 487 or 464 differentially expressed transcripts based on the mRNAs, and 738 or 677 differentially expressed transcripts based on the CDS using STAR or TopHat2, respectively, as the aligner (Figure S3, Table S1). The statistical significance was determined using a false discovery rate less than 0.1 and based on a requirement for a minimum log2 fold change of 2 for the transcript expression to be called as differentially expressed between cell populations. To examine how closely correlated the output is from two different aligners, we compared the fold change calculated based on the output from STAR vs TopHat2. The fold change of the transcript expression between STAR and TopHat2 is well correlated in a linear relationship. In addition, 67.4% (383/568) of differentially expressed transcripts, calculated based on the mRNAs (Figure S2A) and 71.7% (591/824) of differentially expressed transcripts based on CDS (Figure S2 B) were overlapped between two aligners. The high percentage of transcripts found differentially expressed in the analyses based on both aligners further supports a strong correlation between the respective outputs from STAR and TopHat2. To further assess whether the analyses performed based on mRNA and CDS, using the same aligner, are well-correlated, a comparison between fold change of transcript expression was performed. A linear relationship was observed in the fold change of the transcript expression between CDS and mRNA, using STAR (Figure S2C). In addition, an MA plot of the mean

expression and fold change of the transcript expression between neural progenitor cells and immature neurons based on the CDS, using STAR, shows that the mean expression (the number of normalized reads) of differentially expressed transcripts ranged from low to medium expression level (Figure S3D), indicating the analysis is unbiased based on expression level. From the alignment data (Data S2), we identified a total of 1,130 transcripts that were differentially expressed between neural progenitor cells and immature neurons.

Table S1. Read alignment against *X. laevis* genome using TopHat2 and STAR

tophat 2	neural progenitor cells			immature neurons		
	sample 1	sample 2	sample 3	sample 1	sample 2	sample 3
raw reads	17,590,066	19,990,928	16,804,899	17,523,604	19,423,625	19,859,802
trimmed reads	14,464,497	16,044,500	14,953,754	15,741,735	17,349,543	17,800,922
uniquely mapped reads	11,444,705	13,710,391	11,517,049	12,503,563	13,537,406	13,777,726
reads with MAPQ >= 30	10,477,365	12,427,637	10,572,169	11,509,693	12,276,215	12,545,706
reads counted for mRNA	6,947,287	8,048,157	7,111,316	7,779,759	8,132,974	8,324,947
reads counted for intergenic region	3,637,589	4,467,038	3,520,628	3,829,538	4,244,592	4,307,845
reads counted for CDS	3,118,975	4,485,633	4,392,286	3,628,992	4,439,246	5,125,655
reads counted for 3UTR	1,257,093	1,807,527	1,660,929	1,446,179	1,774,019	1,865,073
reads counted for 5UTR	589,252	843,930	590,526	676,117	779,413	668,988
reads counted for intron	1,911,967	911,067	467,575	2,028,471	1,140,296	665,231

STAR	neural progenitor cells			immature neurons		
	sample 1	sample 2	sample 3	sample 1	sample 2	sample 3
raw reads	17,590,066	19,990,928	16,804,899	17,523,604	19,423,625	19,859,802
uniquely mapped reads	13,273,346	15,819,463	13,406,592	14,520,781	15,568,590	15,937,605
reads counted for mRNA	8,422,751	9,655,249	8,480,161	9,388,103	9,747,548	9,914,191
reads counted for intergenic region	4,767,796	6,071,063	4,853,417	5,045,312	5,728,889	5,935,165
reads counted for CDS	3,796,644	5,316,712	5,172,041	4,304,968	5,259,977	6,033,870
reads counted for 3UTR	1,561,072	2,245,761	2,060,704	1,792,487	2,201,242	2,312,324
reads counted for 5UTR	757,445	1,084,443	764,621	868,924	1,001,476	862,520
reads counted for intron	2,307,590	1,008,330	482,795	2,421,724	1,284,853	705,477

Accessing RNA-seq read alignment against *Xenopus laevis* genome scaffold v9.1

We profiled the transcript expression using v9.1 of the *Xenopus laevis* genome assembly and annotation (JGI gene model v1.8) for alignment. Approximately 39% of the transcripts have a known gene symbol, and only these transcripts were used for functional classification with the existing databases. In each of our 3 biological replicates, there are

approximately 17-20 million single-end reads at 100bp (Table S1), and the reads were determined to be high-quality by FASTQC (Figure S2A). To include all the potential changes in transcript expression, two aligners or mappers were used, Spliced Transcripts Alignment to a Reference (STAR) and TopHat2. Each aligner, due to the nature of different algorithms they incorporate, has different true positive rates and false positive rates. The high quality of alignment was summarized by SAMStat with 92.9% of the reads having a MAPQ score ≥30 (Figure S2B). Including the output from both aligners provides a more inclusive picture of the transcriptome analysis. The efficiency of alignment against the genome assembly is 80% and 78% on average among all the samples, using STAR and TopHat2, respectively (Figure S2C-H). The reads with an alignment score < 20 were not included in the differential expression analysis, to ensure the specificity of the read count for each transcript. Over 14 million reads on average were uniquely aligned to the genome, which provided enough depth in the sequencing for a reliable differential expression analysis. We further characterized where these reads were aligned using STAR and found that 64% of the reads on average were aligned to transcript regions annotated on the genome scaffolds, while 35% of the reads on average were aligned to intergenic regions (Figure S2C-E). The alignment of 1% of the reads could not be determined since they were aligned to multiple transcripts. In a similar analysis using TopHat2 as the aligner, 62% of the aligned reads were assigned to a transcript; 30% aligned to intergenic regions and 8% were classified as ambiguous (Figure S2F-H). There was no apparent difference in the percent of aligned reads against intergenic region between neural progenitor cells and immature neurons. To examine where the reads are aligned in the transcripts, we categorized the transcript alignment into different regions: 56% of the reads align to coding domains (CDS), 16% to introns, 4% to 5'UTR, and 24% to 3'UTR using STAR (Figure S2E), and 55% CDS, 16% introns, 8% 5'UTR, and 21% 3'UTR using TopHat2 (Figure S2H). Taken together, these data indicate the high quality of our reads and read alignments to the *Xenopus* genome and demonstrate the distribution of read alignments in the coding and non-coding regions. High quality of reads and read alignments are essential for a reliable differential expression analysis and subsequent analysis.

Differential expression analysis of transcripts expressed by neural progenitor cells and immature neurons.

We profiled the differences in neural progenitor cell and immature neuron transcriptomes and identified the transcripts that are enriched in one cell population or the other, by conducting differential expression analysis on transcript expression with DESeq2 on the alignment output from STAR and TopHat2. Out of 45,099 transcripts, 27,027 and 26,137 transcripts were detected, using aligner STAR and TopHat2, respectively, based on the criteria that the total number of normalized counts is larger than 6 across 6 samples, i.e. one count per sample on average. The difference in the number of transcripts detected between STAR and TopHat2 likely results from a difference in the number of reads that are classified as ambiguous. We used two ways to detect changes in transcript expression: one calculates differential expression based on the reads aligned to the whole transcript/mRNA, while the other way calculates the reads aligned to the CDS only. The statistical analysis based on CDS only shows which transcripts are differentially expressed between neural progenitor cells and immature neurons and have the potential to be translated into proteins, while the analysis based on mRNAs incorporates the information from 5'UTR, 3'UTR and introns in addition to CDS. We identified 487 or 464 differentially expressed transcripts based on the mRNAs and 738 or 677 differentially expressed transcripts based on the CDS using STAR or TopHat2, respectively, as the aligner. The statistical significance was determined using a false discovery rate less than 0.1 and based on a requirement for a minimal fold change of 4 (log2(2.0)) for the transcript expression to be called as differentially expressed between cell populations. To examine how closely correlated the output is from two different aligners, we compared the fold change calculated based on the output from STAR vs TopHat2. The fold change of the transcript expression between STAR and TopHat2 is well correlated in a linear relationship. In addition, 67.4% (383/568) of differentially expressed transcripts, calculated based on the mRNAs (Figure S3A) and 71.7% (591/824) of differentially expressed transcripts based on CDS (Figure S3B) were overlapped between two aligners. The high percentage of transcripts found differentially expressed in the analyses based on both aligners further supports a strong correlation between the respective outputs from STAR and TopHat2. To further assess whether the analyses performed based on mRNA and CDS, using the same aligner, are well-correlated, a comparison between fold change of transcript expression was performed. A linear relationship was observed in the fold change of the transcript expression between CDS and mRNA, using STAR (Figure S3C). In addition, an MA plot of the mean expression and fold change of the transcript expression between neural progenitor cells and immature neurons based on the CDS, using STAR,

shows that the mean expression (the number of normalized reads) of differentially expressed transcripts ranged from low to medium expression level (Figure S3D), indicating the analysis is unbiased based on expression level. From these data, we identified a total of 1,130 transcripts that were differentially expressed between neural progenitor cells and immature neurons from both aligners based on mRNA and CDS (Data S1).

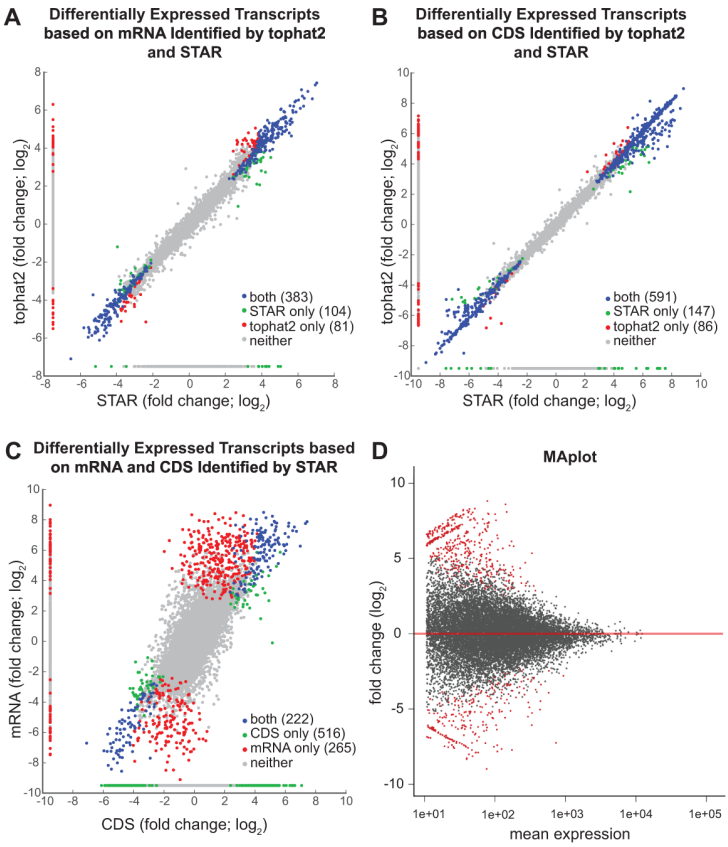

Figure S3. Comparison of the differentially expressed transcripts between neural progenitor cells and immature neurons, identified using STAR and TopHat2.

A. A scatter plot showing the correlation of the fold change of transcript expression in immature neurons in comparison to neural progenitor cells based on mRNA between STAR and TopHat2. **B.** A scatter plot showing the correlation of the fold change of transcript expression in immature neurons in comparison to neural progenitor cells based on coding region (CDS) between STAR and TopHat2. **C.** A scatter plot showing correlation of the fold change of transcript expression in immature neurons in comparison to neural progenitor cells based on mRNA and CDS, using STAR. **D.** MA plot shows the mean expression of transcripts vs its fold change (log2) between neural progenitor cells and immature neurons. The differentially expressed genes are indicated in red with adjusted p-value < 0.1.

We validated the enrichment of NPCs and immature neurons by testing for differential expression of genes known to be involved in neuronal differentiation or proliferation. The immature neuron population has higher expression of transcripts involved in neural patterning and differentiation, such as *neurod1, wnt1, fgf2, vegfa, nfkb1,* and *smad9*. Conversely, NPCs expressed transcripts known to be involved in proliferation, including e*lk-1, e4f1, sstr4, bmp4, jak2* and *nr2f5* (Table S2). Taken together, the DE analysis of the transcripts recovered from cell populations enriched in neural progenitor cells and immature neurons demonstrated a well-correlated fold change in transcript expression. *sox2* did not emerge as a differentially expressed transcript in our dataset, likely because GFP+ cells isolated following VE included a minority of SOX2+ NPCs (3, 6), and our experimental design to identify early master regulators of cell fate decisions.

Table S2. Canonical progenitor and neuronal genes are enriched in isolated neural progenitor and neuronal samples.

Enriched in Neural Progenitor Cells	**Fold Change (log_2)**	**Enriched in Immature Neurons**	**Fold Change (log_2)**
elk1	6.4	neurod1	8.0
e4f1	6.3	wnt1	7.6
sstr4	5.9	fgf2	6.8
bmp4	4.8	vegfa	6.2
jak2	3.5	nfkb1	4.5
nr2f5	3.0	smad9	3.5

Bioinformatic analysis

STRING (v10; RRID:SCR_005223) (16) was used for protein-protein interaction analysis; Cytoscape (v3.2.1; RRID:SCR_015784) (http://www.cytoscape.org/) for network analysis and visualization; ClueGO (v2.1.7; RRID:SCR_005748) (17) for functional network analysis; PANTHER (RRID: SCR_004869) (18) for Gene Ontology (GO) analysis; ENCODE (RRID:SCR_015482) (19) for transcription factor analysis. Fragment Per Kilobase of transcript per Million mapped reads (FPKM) was calculated using Cufflinks suite (v2.2.1; RRID:SCR_014597) (10).

PANTHER protein class analysis identified functional categories of differentially expressed genes

To investigate the involvement of the differentially expressed transcripts in cell proliferation and neuronal differentiation, we functionally categorized these transcripts using the PANTHER (Protein Analysis Through Evolutionary Relationships) classification system. Of the 1,130 transcripts that were differentially expressed between neural progenitor cells and immature neurons, 635 were annotated with a published gene symbol in the genome assembly (v9.1), which then could be used for functional analysis (Figure S4, Data S2). 630 out of the 635 differentially expressed transcripts were recognized by the PANTHER database, and 367 out of 630 genes were clustered based on PANTHER protein classification. The PANTHER protein classification categorized more genes than any other classification scheme in the PANTHER system, such as GO slim biological process, cellular component and molecular function, and PATHWAY analysis. The 367 PANTHER-classified transcripts were clustered into 4 major protein categories: catalytic activity (220 genes), DNA binding (98), receptor-mediated signaling (146) and structural proteins (89) (Figure S4, Data S2). The GO protein classes with the most DE genes were nucleic acid binding (52), enzyme modulator (47), and transcription factor (46). These results indicate that catalytic activities and transcription are heavily involved in either maintaining self-renewal capacity or neuronal differentiation. Kinases and phosphatases, proteases, receptors and signaling molecules, and transcription factors are enriched in our differential expression dataset. Graphing these data according to their fold differential expression indicates that there is no obvious bias in the direction of fold changes between progenitors and neurons (Figure S5, Data S2).

PANTHER's functional categorization of the DE transcripts identified molecular components and signaling pathways induced in response to activity and extracellular signaling events that affect cell proliferation or cell cycle exit and neuronal differentiation. Prominent categories of the DE transcripts include proteases and transcription factors which regulate progenitor cell fate in other systems, validating our approach. The transcription factors enriched in NPCs, such as *e4f1*, *znf217* and *limx1b*, can promote proliferation or cell survival (20, 21). LMX1B, cooperatively with LMX1A, regulates proliferation and specification of midbrain dopaminergic neurons (22). In contrast, the transcription factors we identified as enriched in immature neurons, such as *neuroD1*,

foxg1 and *evx1*, can induce the differentiation of NPCs into neurons. NeuroD1 induces terminal neuronal differentiation (23), while FOXG1 and EVX1 function in the specification of neuronal cell types (24, 25).

Other DE transcripts which are less well known with respect to NPC fate regulation may reveal additional mechanisms involved in this context. For instance, 3 of the 15 differentially expressed proteases encode members of the ADAMTS (a dis-integrin and metalloproteinase with thrombospondin motifs) family of extracellular proteases: *adamts-5*, *adamts-14* and *adamts-17*. *Adamts-17* is involved in cell survival and proliferation (26), suggesting it may have a role in NPC proliferation. On the other hand, ADAMTS-5 enhanced neurite extension in immature neurons (27). *Adamts-14*, a newer and less studied member of the ADAMTS family, may influence genetic predisposition for multiple sclerosis (28), but its potential role in neurogenesis is unknown.

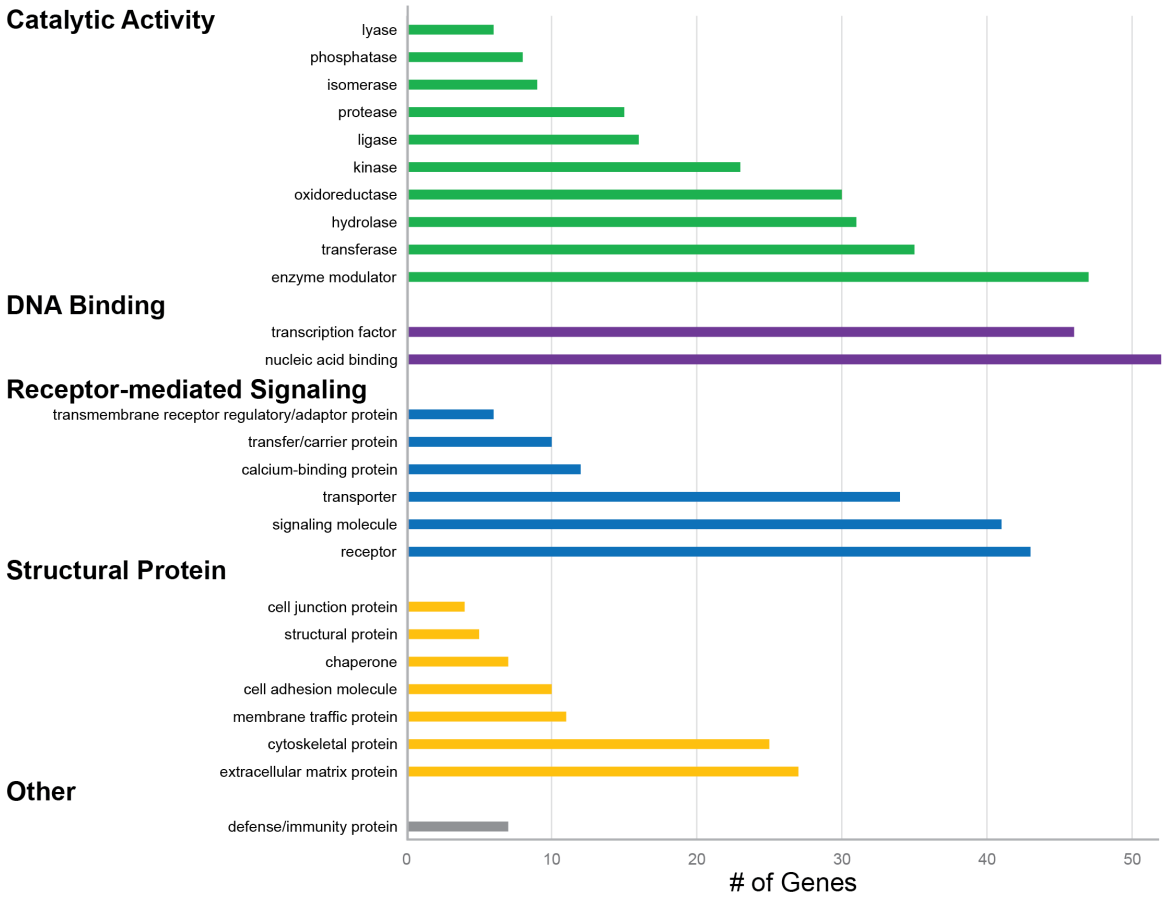

Figure S4. Gene Ontology (GO) analysis of differentially expressed transcripts. Differentially expressed transcripts are categorized based on GO protein classes. Gene number per GO class is indicated on the X-axis.

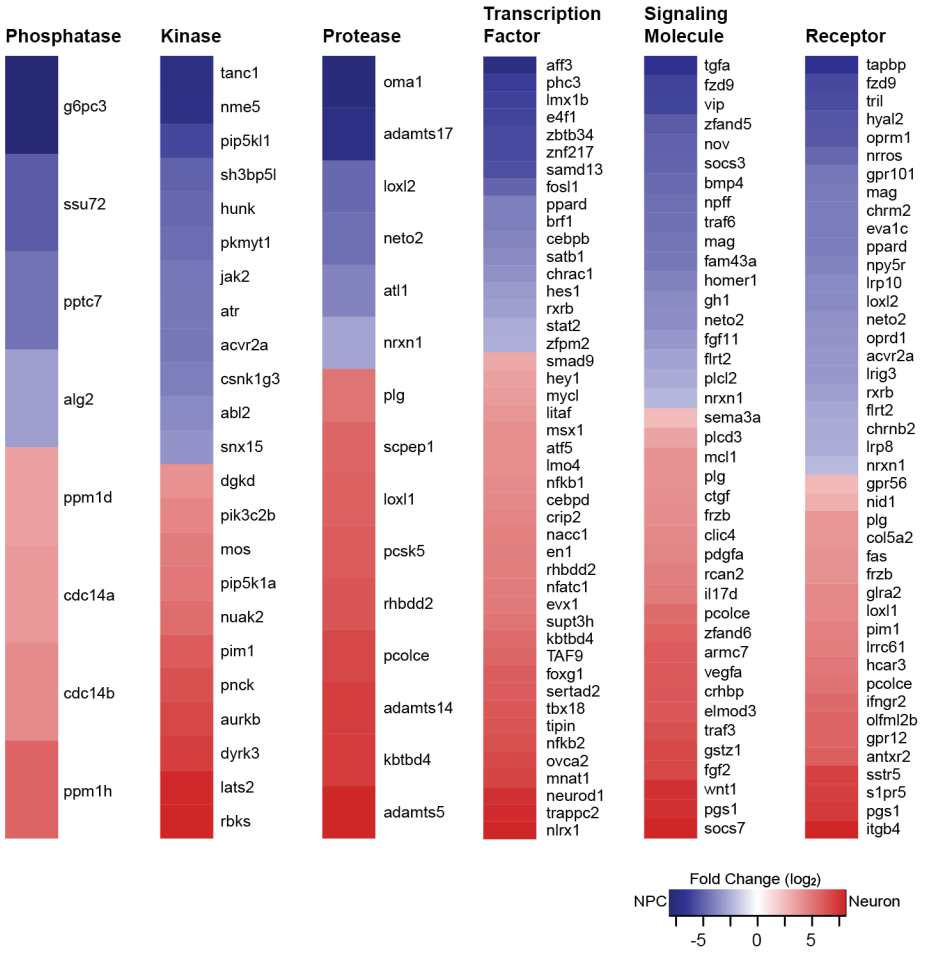

Figure S5. Bioinformatic analysis of differentially expressed transcripts identifies functional categories and interaction networks. Fold change of transcript expression between neural progenitor cells and immature neurons in the most prominent protein categories. Transcription factors, signaling molecules and receptors were the most enriched sub-categories in our differential expression dataset. Graphing the data according to their fold differential expression indicates that there is no obvious bias in the direction of fold changes between progenitors and neurons. Protein expression is graphed relative to expression NCPs (blue) and neurons (red).

Protein-Protein interaction networks identify key players in neurodevelopment among the differentially-expressed transcripts

We identified a network of DE transcripts that are predicted to be important for NPC fate and neuronal differentiation, by applying a strategy to identify the encoded proteins which have the most protein-protein interactions, based on the likelihood that these predicted protein interaction networks would play a regulatory role in the fates of NPCs and neurons. This analysis identified nine DE transcripts whose protein products had more

than 20 interactions: ACTA2, BMP4, JAK2 and BRCA1, which were enriched in NPCs, and ITGA2, VEGFA, FGF2, AURKB and NFKB1, which were enriched in immature neurons. Focusing on the neuronal candidates, ITGA2 (Integrina2) has 40 protein interaction partners among the network of 458 proteins generated from our DE transcripts. Integrina2 facilitates migration of differentiating embryonic stem cells and iPSCs (induced pluripotent stem cells) by remodeling extracellular matrix and initiating intracellular signaling cascades (29). In addition to ITGA2, FGF2 (fibroblast growth factor 2) and VEGFa (vascular endothelial growth factor A), with 23 and 24 interactions in the network, respectively, are growth factors that regulate neuronal differentiation (30-32). VEGFa may also be involved in migration of newly differentiated neurons (33). We demonstrated that FGF2 KD increased NPC proliferation in *Xenopus* optic tectum, (2), consistent with FGF2 maintaining NPC capacity for self-renewal (34). NF-kB1 (nuclear factor of kappa light polypeptide gene enhancer in B-cells 1) regulates neuronal differentiation in the adult mouse hippocampus and maintains cell survival (35). It is interesting to note that NF-kB1 signaling activity is regulated by Aurkb (Aurora kinase B) (36), another highly interconnected protein whose transcript is enriched in the immature neurons.

Bioinformatic identification of Transcription factor networks regulating differential gene expression in neural progenitor cells and neurons.

To identify master regulators that could control differential transcript expression in neural progenitor cells and immature neurons, we evaluated candidates based on two characteristics: the number of the target genes and the number of interactions they have with other transcription factors, assuming that a transcription factor with more protein-protein interactions could indirectly regulate transcription of more genes. As a test case, we mined the ENCODE database and identified 126 transcription factors that could regulate the expression of our 635 differentially expressed transcripts (Figures S4, S5, Data S5). Each transcription factor can regulate between 3 and 488 of the differentially expressed target genes, indicated as the size of the circle in Figure S6. Three transcription factors, TAF1, SIN3A and MAX, regulate the most differentially expressed transcripts, with 488, 484, and 484 gene targets, respectively.

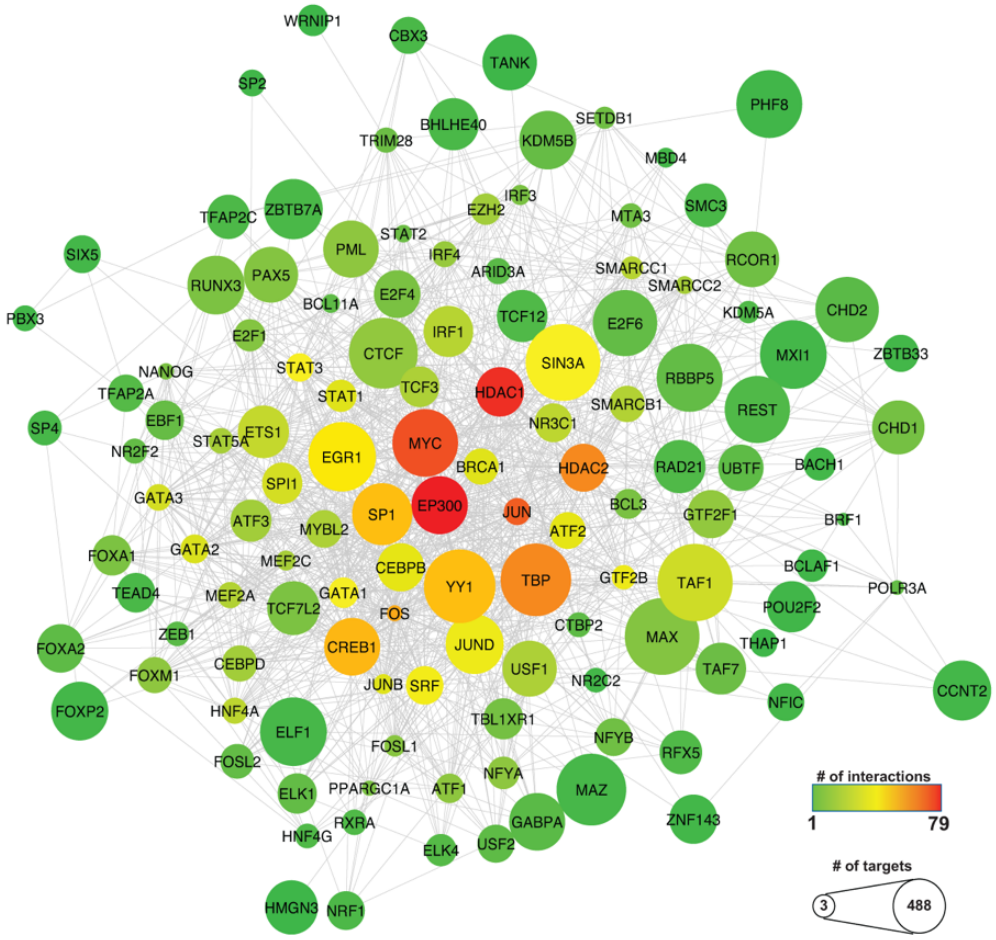

Figure S6. Mining ENCODE identified master regulatory network controlling differential gene expression in neural progenitor cells and neurons.

Network analysis of the 126 transcription factors identified in the ENCODE database that could regulate differential expression of Xenopus genes in our dataset. The size of the node reflects the number of differentially expressed transcripts that each transcription factor can potentially regulate. The color reflects the number of transcription factor binding partners.

We then conducted a protein-protein interaction network analysis, using STRING and Cytoscape. STRING is a database of known and predicted protein interactions based on bioinformatic analysis of genomic and proteomic data and on published work. Starting again from the 1,130 differentially expressed transcripts, of which 635 have a published gene symbol, 629 of the 635 were recognized by the STRING protein database. Of the 629 proteins predicted from the annotated transcripts, STRING identified 458 proteins as having one or more interaction partners (Data S3). We graphed the number of protein-protein interactions versus the degree of differential expression of the transcripts, considering degree centrality and closeness centrality as two key metrics of the

importance of nodes within networks (Figure S6) (37). Degree centrality in the network reflects the number of interactions (i.e., degree) each protein has within a node. Closeness centrality is a measure of the shortest path between proteins within a node. In this case, a high closeness score indicates that the interactions of a protein and its neighboring partner protein are less likely to be by-passed by other proteins in the node. This means that the protein plays an irreplaceable role in the network.

STRING analysis revealed that only a subset of 4 of the 126 transcription factors is well connected in the network, indicated with hot colors in Figure S6. EP300, with the highest network connectivity, has 79 interactions, followed by HDAC1 (73), MYC (67) and JUN (65). Indeed, these most highly connected transcription factors are recognized as master regulators, validating our strategy of using protein-protein interactions combined with the number of target genes to identify master regulators for cell proliferation and neuronal differentiation in our dataset of differentially expressed transcription factors. Each of these transcriptional regulators has been shown to play roles in cell proliferation, differentiation, and cell fate determination. EP300 (E1A binding protein p300) acts as both a transcriptional co-activator and a histone acetyltransferase, positively regulating histone acetylation to initiate transcription. EP300 regulates transcription of both pluripotency genes (*c-myc, c-myb, creb, c-jun, and c-fos*) and neural genes (*pax6*, *sox1*, *zic2*, and *znf521*) (38-40), indicating that EP300 can affect both cell proliferation and neural differentiation. HDAC1 (histone deacetylase 1) is a transcriptional regulator that epigenetically represses gene transcription. Inhibition of HDAC activity can both maintain pluripotency of human embryonic stem cells and inhibit neural differentiation, depending on its targets (38). HDAC1 indirectly increases in c-Myc protein levels (41), which then increases self-renewability of neural progenitor cells (42). c-Myc is thought to regulate expression of 15% of all genes (43). Both c-Myc and Jun regulate cell proliferation by regulating expression of cyclins, and over-expression of c-Jun represses p53 expression, enhancing cell proliferation (44). These highly-connected transcription factors are recognized as master regulators of various cellular processes during development, validating our approach using protein-protein interactions and the number of target genes to identify master regulators in cell proliferation and neuronal differentiation.

We applied this strategy to our DE dataset and identified 6 candidates that regulated other DE transcripts: BRCA1, ELK-1, CEBPB, CEBPD, FOSL1 and BRF1. A Venn diagram of the genes targeted by CEBPB (268), ELK-1 (209), CEBPD (178), BRCA1 (177)

and FOSL1 (66) (Figure 3B, Table S2) identified potential co-regulated transcripts, including six genes that could be regulated by all 5 transcriptional regulators. Four of these targets are enriched in immature neurons: APBA3 (fold change (FC) = 6.9), THUMPD (FC = 6.25), ELMOD3 (FC = 6.2), SLC39A3 (FC = 4.5), and 2 are enriched in NPCs: C12orf57 (FC = -3.9) and MTMR4 (FC = -2.1).

In vivo time-lapse imaging

For *in vivo* live-cell time-lapse imaging, brain of late stage 46 tadpoles were electroporated with 2 µg/µl pSOX2-bd::tGFP. For analysis of BRCA1 or ELK1 function, we co-electroporated pSOX2-bd::tGFP with 0.4mM antisense translation-blocking morpholino oligonucleotides tagged with lissamine fluorophores (45), targeted against *brca1* transcripts (GGTTCCATTTGTGTCAGCTCTCAGC) and against *elk-1* transcripts (GGTCATTTTACTTTGTCCTGTCCCT), or a control non-specific sequence (GeneTools, Philomath, OR). We have previously shown that pSOX2-bd::tGFP can be used to birthdate progenitors that express SOX2 protein, and that tGFP persists for several days in their neuronal progeny (3). Moreover, we demonstrated that even when SOX2 levels are knocked down with *sox2* morpholinos (AGCTCGGTCTCCATCATGCTGTA), the pSOX2-bd::tGFP reporter expresses tGFP due to pre-existing SOX2 protein (Figure S1). Therefore, we reasoned that pSOX2-bd::tGFP could be used in conjunction with BRCA1 KD to birthdate cells prior to any possible downstream effects of BRCA1 on SOX2. Animals were randomly divided into two groups for the 3-day duration of the time-lapse imaging; one group housed in the control 12h light/12h dark condition and the other housed in the dark in a light impermeable black chamber. Animals were maintained in their respective housing conditions throughout the experiment except when imaging. Tadpoles were screened for consistent morpholino fluorescent labeling in the tectum. For imaging, tadpoles were anesthetized in 0.01% MS-222, placed in a custom-built chamber and imaged with 20X (Olympus XLUMPlanFL 0.95 NA) water immersion lens on a custom-built two-photon microscope modified from an Olympus FV300 system (46, 47). A stack of images for each tectal lobe was acquired at 1µm intervals ranging from 120µm to 160µm depending on the distribution of tGFP+ tectal cells over three days. All samples were imaged in parallel using identical image acquisition parameters. Data files were coded prior to analysis and all analysis was conducted blind to condition. Analysis was conducted using Cell Counter plugin in FIJI, an image processing package of ImageJ

(RRID:SCR_002285) (48, 49). tGFP-labelled cells were identified and categorized into three groups, neural progenitor cells, neurons or unidentifiable based on their morphology, using criteria as described (2, 3). While the majority of tGFP+ cells could be classified as either neural progenitor cells or neurons, a small population of cells was unclassified or unidentifiable. In control groups, the fraction of unidentifiable cells was 0.5-4.4% across all experiments over the course of the 3 days. For BRCA1 KD, the unidentifiable cell population ranged from 0.4-5.9%, and for ELK-1 the fraction of unidentifiable cells was 3.1-10.8% across all experiments and timepoints. The proliferation and survival rates were calculated based on the changes in number of tGFP-labelled cells at day 2 and day 3 normalized against the number at day 1. The changes in the fate of neural progenitor cells as the result of the treatment is presented as the percentage of each cell type compared to the total cell population each day. For experiments testing for interactions between the effects of visual experience and BRCA1 or ELK-1 expression on neural progenitor cell fate, we used a factorial experimental design (50) which allows us to compare results between multiple experimental conditions.

Immunohistochemistry

Animals were anesthetized and then fixed with freshly made 4% paraformaldehyde (Electron Microscopy Sciences, Fort Washington, PA) in 1x phosphate-buffered saline (PBS; pH 7.4) with a brief microwave pulse (150mV on-off-on, 1min each; Pelco BioWave Pro microwave, Model 36500, Ted Pella, Redding, CA) and were post-fixed at 4^{o}C overnight. Whole brains were dissected and incubated in blocking solution (5% normal donkey serum and 1% Bovine Serum Albumin (BSA; Sigma) in PBS with 0.1% Triton-X100 (PBS-T)) for 1 hour at room temperature before transferred to the anti-pH3 antibody solution (1:200 in blocking solution; #9706, Cell Signaling; RRID:AB_331748) at 4^{o}C for 3 days. After washes with PBS-T, brain tissues were incubated in secondary antibody solution (Alexa488 donkey anti-mouse secondary antibody, 1:1,000; A21202, Life Technologies; RRID:AB_141607) at 4^{o}C overnight. After PBS-T washes, for cell death analysis in fixed tissue, nuclear labeling using Sytox Orange (SytoxO, 1:5000; S11368, Life Technologies) was applied to the brain tissues for 20 mins at room temperature (51). After several washes with PBS-T, brain tissues were mounted in 6M urea in 50% glycerol for imaging. 36µm Z-series we collected at 1µm intervals using Nikon C2 (20x Plan Apo lens with 0.75 NA), and ImageJ Cell Count plugin was used for analysis. Apoptotic cells

were identified based on the morphology of small granular structure and the high labeling intensity (52), and categorized into two groups, neural progenitor cells (NPC) and neurons based on their location in the tectum.

Western blots

To test the effect of different visual experience conditions, animals were reared under enhanced visual stimulation or dark for 30 hours as described previously (6) and the midbrains were dissected for homogenization. To test the effect of morpholino knockdown, brains of late stage 46 tadpoles were electroporated with 0.4mM morpholinos, and midbrains were dissected 2 days later. Because BRCA1 requires different extraction and detection conditions, tissues were homogenized in different lysis buffers for different antibodies. Experimental and paired control samples were prepared and processed side by side. For ELK-1 (Abcam #ab188316; RRID:AB_2890919) and SOX2 (Cell Signaling Technology #3579S; RRID:AB_2195767) antibodies, tissues were homogenized in RIPA buffer with brief sonication, and concentration was measured by BCA assay. 10µg of lysate was run on a Mini-Protean TGX precast gels (BioRad). For BRCA1 antibody (SCBT #SC-646; RRID:AB_630945), tissue was homogenized in lysis buffer documented in (53); HEPES 100mM (pH 7.5), NaCl 200mM, EDTA 40mM, EGTA 4mM, NaF 100mM, β-glycerophosphate 20mM, sodium orthovanadate 2mM, Nonidet P-40 1%, Complete Protease Inhibitor mixture 1:50), and concentration measured by DC Protein Assay (BioRad). 40mg of lysate was loaded onto an in-house made 7% gel. Proteins were transferred to a nitrocellulose membrane and blotted with standard protocols. Antibodies were detected by goat anti-mouse/rabbit HRP-conjugated secondaries (BioRad) followed by ECL (Pierce, Thermo Fisher Scientific, 32209). Quantification was performed using densitometry (ImageJ), different exposures were used to avoid saturation, and bands were normalized to total protein using Ponceau S (54). BRCA1 and ELK-1 detected in western blots of the midbrain does not reflect their NPC or neuronal localization.

Table S3 (for Figure 3B)
Gene list of transcripts regulated by the identified network of 5 transcription factors.

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
abca3	X	X			
abl2		X			
acaa2		X			
acta2	X	X			
acvr2a	X			X	
adamts5		X			
adsl			X	X	
aff3		X			
akap2		X			
alg2			X		
alkbh4				X	
amn			X		
anapc5			X		
ankrd16			X	X	
ano6		X			
antxr2		X			
ap4e1	X	X		X	
ap4s1		X	X		
ap5s1	X	X	X		
apba3	X	X	X	X	X
arap3		X			X
arf3	X			X	
arhgef1			X		
arih2		X	X	X	
arl14ep		X		X	
arl6ip5		X			
armc7			X	X	
art5					X
asmtl	X				
atf5		X	X	X	X
atg13	X	X		X	
atic		X		X	

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
atl1	X	X		X	
atr	X	X		X	X
atrip	X				
atxn1l	X	X	X		
aurkb	X	X			
b3galt4				X	X
b3galt6			X	X	
b9d1	X			X	
bag5				X	
bcar3	X	X		X	
bin3		X		X	
bnip3			X		
brca1	X	X			
c12orf57	X	X	X	X	X
c16orf13	X	X		X	
c21orf2	X		X	X	
c2orf82		X	X		
c3orf17		X		X	
cacul1	X	X		X	
cant1			X	X	X
carhsp1		X	X		
ccbl2					X
ccdc125				X	
ccdc127			X		
ccdc134			X	X	
ccdc43			X		X
ccdc50		X	X	X	X
ccdc90b	X	X	X	X	
ccm2	X	X		X	
cdc14b	X	X			
cdc37l1	X	X		X	
cdc42se1		X	X	X	
cdc45	X	X		X	X
cdca7l	X	X		X	
cebpb	X	X	X		X
cebpd	X	X	X		

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
cenpa				X	
cenph		X		X	
cep104		X		X	
cep76	X	X			
cfl2		X	X		
chrac1	X			X	
clic4		X	X	X	
cnep1r1	X	X	X		
cnnm2	X			X	
cnppd1		X	X		X
cnpy2	X	X		X	
col4a3bp		X	X		
commd6	X	X		X	
cops4		X			
cox18	X	X	X		
creld1	X	X	X		
crtap			X		
csgalnact2		X	X	X	
csnk1g3		X	X	X	
ctdsp2			X		
ctgf		X	X		
ctr9	X	X		X	
cxadr		X	X	X	
cxcl14		X			
cxxc1	X	X		X	
cycs	X	X	X	X	
cyth2	X			X	
dbf4b	X	X			
dbndd1			X		
dbr1	X			X	
dcaf12		X			
dcakd	X	X	X	X	
dclre1c		X		X	
dctd		X			
ddx56			X		
desi2	X	X	X	X	

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
dgkd		X			
dhrs4		X	X		
dlg2		X			
dnah11		X			
dnaja4				X	
dnajb1	X	X	X	X	
dnajc3		X			
dpy30	X	X	X	X	
dusp7	X				
dydc2		X			
dyrk3	X				
e4f1	X	X	X	X	
eaf2	X	X	X		
eci2		X	X		
ecsit	X	X	X		
edem2		X			
edem3	X			X	
egln2	X	X	X		
elk1		X		X	X
elmod3	X	X	X	X	X
emd	X				
eogt				X	
eri1	X	X	X	X	
eva1c	X				
eya3				X	
ezh1			X		
faah2		X			X
fahd2a	X	X	X		
fam174b	X	X		X	
fam177a1		X			
fam213a			X		
fam222b	X				
fam43a		X			
fam64a			X		
fancd2	X	X		X	
fas	X	X			

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
fdxr	X	X	X	X	
fem1c		X	X		
fgd6	X				
fgf2		X			
fgfrl1		X			
fitm2	X	X			
fkbp10			X		X
fnip1			X	X	X
fosl1	X	X		X	X
foxn4		X			
frs3			X		
frzb		X			
fscn2		X			
g2e3		X	X	X	
g6pc3	X	X	X	X	
gale		X	X	X	
gatc	X	X	X	X	
gdpgp1		X			X
gga1	X	X	X	X	
glce	X				
gnai1		X			
gnpat				X	
gnpda1				X	X
golph3l	X	X			
gpatch3	X	X	X		
gpr56	X	X			
grn			X		X
gstz1	X		X	X	
gyg1					X
has2		X			
haus6		X	X	X	
hcar3		X			
hebp2				X	
hemk1			X	X	
hes1	X	X	X	X	
hiat1			X		

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
homer1	X	X	X	X	
hsd17b6		X			
hsd17b7		X			
iars		X	X		
icmt	X				
idh3g	X			X	
ier2	X	X		X	
ifltd1		X			
ifngr2		X			
ift172	X	X			
igfbp4		X			
ipo9		X			
iqch	X	X		X	X
irs2		X		X	
isg20l2	X	X		X	
itga2		X			
jak2					X
kbtbd4	X	X	X	X	
kiaa0141				X	
krt17	X	X		X	
lace1				X	
lats2	X	X			
leprel4			X		X
letm2	X	X			
lmo4		X			
loxl1		X			X
lpcat3			X		
lrp10		X	X	X	X
lrrc41	X	X	X	X	
lrrc61	X	X		X	
lrrc8d			X		
mad2l1	X	X		X	
malsu1	X	X			
march8	X				X
mcl1		X	X	X	
mdm1		X		X	

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
mecr	X			X	
mien1				X	X
mmadhc		X	X		
mnat1		X	X	X	
mrpl10		X	X	X	
mrpl15		X	X		
mrpl3	X	X	X	X	
mrpl51		X	X	X	
mrps16	X	X		X	X
mrps2	X	X		X	
mrps34			X	X	
mrrf	X	X		X	
msmo1	X	X	X		X
msx1			X		
mtmr4	X	X	X	X	X
naa16		X	X	X	X
nags			X		X
narf		X	X	X	
narfl	X	X		X	
ndrg1		X			
ndufa12		X	X		
ndufa5	X			X	X
neurod1	X				
nfkb1		X			
nfkb2	X			X	
nhlrc2		X		X	
nid1			X		
nle1					X
nob1	X	X	X		
nol9	X	X	X	X	
npff		X			
nuak2		X			
nufip1	X	X			
nvl	X	X		X	
oat			X		
oaz2			X		

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
olfml2b		X			
oma1		X		X	
orc6	X	X	X		
ormdl1			X		X
ovca2			X	X	
pank4		X	X		
pawr		X			X
pcsk5		X			X
pde5a		X			
pdf		X		X	
pet112				X	
pfdn4	X		X	X	
pgap1		X			
pgs1		X	X	X	X
phc3		X	X		
pign	X			X	
pigp	X	X		X	X
pim1	X		X		X
pip5k1a		X	X		
pip5kl1		X			
plcd3	X				
plekhj1		X		X	X
plk2	X		X		
plk4	X	X	X	X	
pms1			X		X
ppard	X	X	X	X	
ppm1d	X		X	X	
ppm1h		X	X		
ppp1r11		X		X	X
ppp1r13l	X	X		X	X
ppp1r2	X			X	
prc1	X	X			
prkacb				X	
pros1		X			
prosc				X	
psme1			X		

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
psmg3		X	X	X	
ptcd2	X	X		X	
ptdss1	X		X	X	
ptgr1		X			
pttg1ip			X	X	
pxmp2	X		X	X	X
rap2b		X	X		
rasl10b			X		X
rbks	X	X	X		
rbm48				X	
rcc1	X	X	X	X	
recql	X			X	
rexo4	X	X		X	
rhbdd2		X			X
rhob		X	X		
rhog	X	X	X		X
ric3				X	
rnf103			X		
rpain			X		
rpp38	X	X		X	
rpusd3				X	
rraga	X	X		X	
rrp9		X	X	X	
rrs1		X		X	
rtkn		X	X		
rwdd4		X		X	
rxrb	X		X	X	
scrn2				X	
scyl2	X			X	
sdccag3	X	X	X	X	
sec11a		X	X	X	
sept2		X	X	X	
sertad2		X	X		X
sh3bp5l		X	X	X	X
slc13a5			X		
slc25a20				X	

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
slc25a44	X			X	X
slc2a9		X			
slc30a7	X	X		X	
slc35e3			X	X	X
slc39a13		X			
slc39a3	X	X	X	X	X
slc8a2					X
snx15	X				
snx2		X			
snx24		X			
snx33		X	X		
snx5	X		X	X	
snx9		X	X		
socs3		X	X		
socs7	X	X	X	X	
spata5l1		X	X	X	
spata7	X	X			
srr		X	X		
sssca1	X	X		X	
ssu72	X	X		X	
stam2				X	
stat2	X			X	
stim1	X				X
sts		X			
stx12	X	X		X	
sumf1					X
supt3h			X	X	X
tapbp		X	X		
tatdn3	X	X	X	X	
tbc1d9b				X	
tbccd1	X	X	X	X	
tbx18		X			
tfb2m	X	X	X	X	
tgfa			X		
thumpd1	X	X	X	X	X
timm21		X	X		

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
tipin	X	X			
tm2d3	X			X	
tmbim4		X	X	X	
tmem11	X			X	
tmem138				X	
tmem150a		X	X		
tmem170b			X		
tnnc2	X	X	X		
tob2	X	X	X	X	
tomm20		X		X	
tor1b	X			X	
tp53inp2	X	X	X		
tpmt		X		X	
traf6			X		
trappc2	X	X		X	
trim44	X			X	
trmt13		X		X	
trmt5	X	X		X	
trnt1	X	X		X	
ube2c	X				
ubfd1	X	X	X	X	
ubqln1	X	X		X	
ufm1		X		X	
ugcg			X	X	
usp12			X		X
usp2	X	X			
usp38					X
usp45		X			
utp23	X	X		X	
vegfa	X		X		
vprbp	X		X	X	
vps26b			X	X	
vta1	X	X	X	X	
vti1b	X			X	
vwa3b		X			
wbp4		X	X	X	

	BRCA1	CEBPB	CEBPD	ELK1	FOSL1
wdr16		X		X	
wdr48	X	X	X	X	
wnt1				X	
xkr9		X			
xrcc3	X	X		X	
yars2		X	X	X	
ydjc	X	X		X	
yrdc	X			X	
zbtb1	X	X			
zbtb34				X	X
zfand5	X	X	X	X	
zfand6	X				
znf217		X			
znf318					X
znf804b		X			
znhit3		X		X	
zwilch	X	X	X		

Additional Supporting Files:

Data S1. List of DE transcripts.

Data S2. Corresponding to Figures S4, S5 (list of genes in PANTHER GO categories)

Data S3. Corresponding to Figure 2C (STRING analysis of DE genes)

Data S4. Corresponding to Figure S6 (ENCODE data on transcription factors)

References

1. P. D. Nieuwkoop, J. Faber, *Normal table of Xenopus laevis (Daudin); a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis* (North-Holland Pub. Co., Amsterdam,, 1956), pp. 243 p.

2. J. E. Bestman, L. C. Huang, J. Lee-Osbourne, P. Cheung, H. T. Cline, An in vivo screen to identify candidate neurogenic genes in the developing Xenopus visual system. *Dev Biol* **408**, 269-291 (2015).

3. J. E. Bestman, J. Lee-Osbourne, H. T. Cline, In vivo time-lapse imaging of cell proliferation and differentiation in the optic tectum of Xenopus laevis tadpoles. *J Comp Neurol* **520**, 401-433 (2012).

4. M. Gaete *et al.*, Spinal cord regeneration in Xenopus tadpoles proceeds through activation of Sox2-positive cells. *Neural Dev* **7**, 13 (2012).

5. M. Gotz, M. Nakafuku, D. Petrik, Neurogenesis in the Developing and Adult Brain-Similarities and Key Differences. *Cold Spring Harb Perspect Biol* **8** (2016).

6. P. Sharma, H. T. Cline, Visual activity regulates neural progenitor cells in developing xenopus CNS through musashi1. *Neuron* **68**, 442-455 (2010).

7. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

8. R. C. Gentleman *et al.*, Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).

9. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

10. C. Trapnell *et al.*, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578 (2012).

11. J. B. Bowes *et al.*, Xenbase: gene expression and improved integration. *Nucleic Acids Res* **38**, D607-612 (2010).

12. M. Fisher *et al.*, Xenbase: key features and resources of the Xenopus model organism knowledgebase. *Genetics* **224** (2023).

13. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

14. T. Lassmann, Y. Hayashizaki, C. O. Daub, SAMStat: monitoring biases in next generation sequencing data. *Bioinformatics* **27**, 130-131 (2011).

15. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

16. D. Szklarczyk *et al.*, STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447-452 (2015).

17. G. Bindea *et al.*, ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091-1093 (2009).

18. H. Mi, A. Muruganujan, P. D. Thomas, PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* **41**, D377-386 (2013).

19. E. P. Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

20. C. Paul *et al.*, The LIM-only protein FHL2 is a negative regulator of E4F1. *Oncogene* **25**, 5475-5484 (2006).

21. J. J. Cowger, Q. Zhao, M. Isovic, J. Torchia, Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. *Oncogene* **26**, 3378-3386 (2007).

22. H. E. Marei *et al.*, Gene expression profiling of embryonic human neural stem cells and dopaminergic neurons from adult human substantia nigra. *PLoS One* **6**, e28420 (2011).

23. A. Pataskar *et al.*, NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J* 10.15252/embj.201591206 (2015).

24. C. Thaeron *et al.*, Zebrafish evx1 is dynamically expressed during embryogenesis in subsets of interneurones, posterior gut and urogenital system. *Mech Dev* **99**, 167-172 (2000).

25. Y. Yang *et al.*, Impaired Interneuron Development after Foxg1 Disruption. *Cereb Cortex* 10.1093/cercor/bhv297 (2015).

26. Z. Jia, S. Gao, N. M'Rabet, C. De Geyter, H. Zhang, Sp1 is necessary for gene activation of Adamts17 by estrogen. *J Cell Biochem* **115**, 1829-1839 (2014).

27. M. G. Hamel *et al.*, Multimodal signaling by the ADAMTSs (a disintegrin and metalloproteinase with thrombospondin motifs) promotes neurite extension. *Exp Neurol* **210**, 428-440 (2008).

28. R. Goertsches, M. Comabella, A. Navarro, H. Perkal, X. Montalban, Genetic association between polymorphisms in the ADAMTS14 gene and multiple sclerosis. *J Neuroimmunol* **164**, 140-147 (2005).

29. H. Y. Li *et al.*, Collagen IV significantly enhances migration and transplantation of embryonic stem cells: involvement of alpha2beta1 integrin-mediated actin remodeling. *Cell Transplant* **20**, 893-907 (2011).

30. J. M. Rosenstein, N. Mani, A. Khaibullina, J. M. Krum, Neurotrophic effects of vascular endothelial growth factor on organotypic cortical explants and primary cortical neurons. *J Neurosci* **23**, 11036-11044 (2003).

31. J. F. Cavanagh, M. C. Mione, I. S. Pappas, J. G. Parnavelas, Basic fibroblast growth factor prolongs the proliferation of rat cortical progenitor cells in vitro without altering their cell cycle parameters. *Cereb Cortex* **7**, 293-302 (1997).

32. A. Louvi, S. Artavanis-Tsakonas, Notch signalling in vertebrate neural development. *Nat Rev Neurosci* **7**, 93-102 (2006).

33. H. Wang *et al.*, VEGF Enhances the Migration of MSCs in Neural Differentiation by Regulating Focal Adhesion Turnover. *J Cell Physiol* **230**, 2728-2742 (2015).

34. R. Sanalkumar, S. Vidyanand, C. Lalitha Indulekha, J. James, Neuronal vs. glial fate of embryonic stem cell-derived neural progenitors (ES-NPs) is determined by FGF2/EGF during proliferation. *J Mol Neurosci* **42**, 17-27 (2010).

35. S. Denis-Donini *et al.*, Impaired adult neurogenesis associated with short-term memory defects in NF-kappaB p50-deficient mice. *J Neurosci* **28**, 3911-3919 (2008).

36. J. Y. He *et al.*, Knockdown of Aurora-B alters osteosarcoma cell malignant phenotype via decreasing phosphorylation of VCP and NF-kappaB signaling. *Tumour Biol* **36**, 3895-3902 (2015).

37. S. Wang, J. Zhao, Multi-attribute intergrated measurement of node importance in complex networks. *Chaos* **25** (2015).

38. Y. Qiao, R. Wang, X. Yang, K. Tang, N. Jing, Dual roles of histone H3 lysine 9 acetylation in human embryonic stem cell pluripotency and neural differentiation. *J Biol Chem* **290**, 2508-2520 (2015).

39. Y. L. Chen, N. Monteith, P. Y. Law, H. H. Loh, Dynamic association of p300 with the promoter of the G protein-coupled rat delta opioid receptor gene during NGF-induced neuronal differentiation. *Biochem Biophys Res Commun* **396**, 294-298 (2010).

40. Y. Chen *et al.*, HDAC-mediated deacetylation of NF-kappaB is critical for Schwann cell myelination. *Nat Neurosci* **14**, 437-441 (2011).

41. J. Luo, F. Su, D. Chen, A. Shiloh, W. Gu, Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature* **408**, 377-381 (2000).

42. H. Zheng *et al.*, p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. *Nature* **455**, 1129-1133 (2008).

43. J. Gearhart, E. E. Pashos, M. K. Prasad, Pluripotency redux--advances in stem-cell research. *N Engl J Med* **357**, 1469-1472 (2007).

44. M. Schreiber *et al.*, Control of cell cycle progression by c-Jun is p53 dependent. *Genes Dev* **13**, 607-619 (1999).

45. J. E. Bestman, H. T. Cline, Morpholino Studies in Xenopus Brain Development. *Methods Mol Biol* **2047**, 377-395 (2020).

46. J. E. Bestman, H. T. Cline, The RNA binding protein CPEB regulates dendrite morphogenesis and neuronal circuit assembly in vivo. *Proc Natl Acad Sci U S A* **105**, 20494-20499 (2008).

47. E. S. Ruthazer, J. Li, H. T. Cline, Stabilization of axon branch dynamics by synaptic maturation. *J Neurosci* **26**, 3594-3603 (2006).

48. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

49. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675 (2012).

50. L. M. Collins, *Optimization of behavioral, biobehavioral, and biomedical interventions: The multiphase optimization strategy (MOST).* (Springer, New York, 2018).

51. R. L. Faulkner, T. J. Wishard, C. K. Thompson, H. H. Liu, H. T. Cline, FMRP regulates neurogenesis in vivo in Xenopus laevis tadpoles. *eNeuro* **2**, e0055 (2015).

52. C. K. Thompson, H. T. Cline, Thyroid Hormone Acts Locally to Increase Neurogenesis, Neuronal Differentiation, and Dendritic Arbor Elaboration in the Tadpole Visual System. *J Neurosci* **36**, 10356-10375 (2016).

53. V. Joukov, J. Chen, E. A. Fox, J. B. Green, D. M. Livingston, Functional communication between endogenous BRCA1 and its partner, BARD1, during Xenopus laevis development. *Proc Natl Acad Sci U S A* **98**, 12078-12083 (2001).

54. I. Romero-Calvo *et al.*, Reversible Ponceau staining as a loading control alternative to actin in Western blots. *Anal Biochem* **401**, 318-320 (2010).