




Transcriptome Analysis Identifies *SenZfp536*, a Sense lncRNA that Suppresses Self-renewal of Cortical Neural Progenitors

Kuan Tian^{1,2} · Andi Wang^{1,2} · Junbao Wang^{1,2} · Wei Li^{1,2} · Wenchen Shen^{1,2} · Yamu Li^{1,2} · Zhiyuan Luo^{1,2} · Ying Liu^{1,2} · Yan Zhou^{1,2,3} 

Received: 28 November 2019 / Accepted: 12 August 2020 / Published online: 16 November 2020
© Shanghai Institutes for Biological Sciences, CAS 2020

Abstract Long non-coding RNAs (lncRNAs) regulate transcription to control development and homeostasis in a variety of tissues and organs. However, their roles in the development of the cerebral cortex have not been well elucidated. Here, a bioinformatics pipeline was applied to delineate the dynamic expression and potential *cis*-regulating effects of mouse lncRNAs using transcriptome data from 8 embryonic time points and sub-regions of the developing cerebral cortex. We further characterized a sense lncRNA, *SenZfp536*, which is transcribed downstream of and partially overlaps with the protein-coding gene *Zfp536*. Both *SenZfp536* and *Zfp536* were predominantly expressed in the proliferative zone of the developing cortex. *Zfp536* was *cis*-regulated by *SenZfp536*, which facilitates looping between the promoter of *Zfp536* and the genomic region that transcribes *SenZfp536*. Surprisingly, knocking down or activating the expression of *SenZfp536*

increased or compromised the proliferation of cortical neural progenitor cells (NPCs), respectively. Finally, overexpressing *Zfp536* in cortical NPCs reversed the enhanced proliferation of cortical NPCs caused by *SenZfp536* knockdown. The study deepens our understanding of how lncRNAs regulate the propagation of cortical NPCs through *cis*-regulatory mechanisms.

Keywords *Zfp536* · Sense lncRNA · Self-renewal · Cortical development · Neural progenitor

Introduction

The mammalian cerebral cortex, also known as the neocortex, is a six-layer structure comprising a large cohort of neurons and glial cells, which is the basis for sophisticated cognitive functions. Cortical projection neurons (PNs) are functional units that make long-distance connections with other parts of the central nervous system. Cortical PNs in mouse and human are generated directly and indirectly from radial glial progenitor cells (RGs), including apical and basal RGs, lying in the dorsal part of the embryonic forebrain [1, 2]. Some cortical PNs are direct progeny of RGs, but most are the progeny of symmetrical cell divisions of basal neural progenitors, which are also derivatives of RGs and reside in the subventricular zone (SVZ) [3, 4]. With the exception of the first-born layer I neurons, deep layer PNs are born early and superficial layer PNs are born later. Newborn PNs migrate outward (radially), passing through deep layer PNs to settle in their final destination (an inside-out pattern, Fig. S1A) [5–7].

Extensive work has been focused on the molecular events underlying cell-fate specification of cortical neural progenitor cells (NPCs), including both RGs and basal

Kuan Tian and Andi Wang have contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12264-020-00607-2>) contains supplementary material, which is available to authorized users.

✉ Ying Liu
y.liu@whu.edu.cn

✉ Yan Zhou
yan.zhou@whu.edu.cn

¹ College of Life Sciences, Renmin Hospital of Wuhan University, Wuhan University, Wuhan 430072, China

² Frontier Science Center for Immunology and Metabolism, Medical Research Institute, School of Medicine, Wuhan University, Wuhan 430071, China

³ Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

neural progenitors, as well as the postmitotic maturation and migration of layer-specific PNs. Both extrinsic and intrinsic mechanisms balance the self-renewal and differentiation of NPCs during cortical development. Moreover, changes in the competency of NPCs over time during cortical development are likely due to temporal shifts of sets of transcription factors expressed in NPCs, but it is largely unknown how these changes occur [8–12]. Molecular dissection of these key properties will not only help to better understand how neural development proceeds, but also have implications for manipulating NPCs *in vivo* and *in vitro* to tackle neurological disorders such as neurodegenerative diseases and malignant brain tumors.

The transient expression, flexible structure, and dynamic localization of RNA molecules enable the fine-tuning of genome arrangement, scaffolding, and transcription functions, thus precisely regulating gene expression in a time- and site-specific manner. Recent work indeed points to critical roles of long noncoding RNAs (lncRNAs), transcripts longer than 200 nucleotides (nt) but without protein-coding potential, in controlling transcription by associating with epigenetic regulatory machinery and chromosomal architectural organization [13, 14]. Therefore, lncRNAs participate in numerous physiological and pathological processes including the maintenance of pluripotency, lineage specification, organogenesis, tumorigenesis, and metabolism [15–19]. Moreover, spatiotemporal co-expression of tissue-specific lncRNAs with nearby protein-coding genes is prevalent in mammals. lncRNAs can modulate neighboring transcription (*cis*-regulation) and participate in biological processes similar to those controlled by neighboring protein-coding genes [20]. Recent high-throughput studies have characterized spatiotemporally-regulated lncRNAs expressed in the developing brain of human and mouse [21, 22]. However, very few lncRNAs, particularly those with *cis*-regulatory roles, have been found to regulate the fate choices of cortical NPCs [23].

Here, we analyzed transcriptome data from developing mouse cerebral cortex to identify lncRNAs with spatiotemporal expression patterns and potential *cis*-regulatory functions. An example is *SenZfp536*, a sense lncRNA that partially overlaps with the 3' end of *Zfp536* and suppresses the self-renewal of cortical NPCs by *cis*-activating *SenZfp536*. The study advances the understanding of molecular mechanisms, particularly those mediated by lncRNAs, in controlling the self-renewal of cortical NPCs.

Materials and Methods

Mouse

All CD-1 mice were obtained from the Hunan SJA Laboratory Animal Co., Ltd (Changsha, China). The

husbandry of CD-1 mice was carried out under specific pathogen-free (SPF) housing conditions and mouse experiments were approved by the Animal Care and Ethics Committee at Wuhan University. The day when a vaginal plug was detected was counted as embryonic (E) day 0.5.

Dissection and RNA-Sequencing of Dorsal Forebrain at E10.5 and E12.5

The dorsal forebrain (neocortex) was microdissected from E10.5 and E12.5 mouse embryos under a dissecting microscope and flash-frozen in liquid nitrogen. Total RNAs were extracted twice using TRIzol (Thermo Fisher Scientific, Waltham, United States) and treated with DNase I (NEB Biolabs, Ipswich, United States). The integrity of RNA was analyzed with an Agilent Bioanalyzer 2100 (Agilent, Santa Clara, United States). Removal of ribosomal RNAs and construction of libraries for strand-specific RNA-seq were performed in BGI Tech (Shenzhen, China). Totally ~410 million clean reads of putative polyadenylated and non-polyadenylated transcripts were generated from the cortices of the two developmental stages.

Processing of RNA-Seq Data

We used the Galaxy platform [24] to integrate our in-house RNA-seq data with the ENCODE RNA-seq data of mouse cortex at E10.5, E11.5, E12.5, E13.5, E14.5, E15.5, E16.5, and P0. Reads were mapped to the mouse genome (mm10) using HISAT2 [25] and then subjected to transcripts assembly using StringTie [26] with default parameters, followed by generating gene counts using the featureCount [27] software. Differentially-expressed genes of Tis21+ versus Tis21– apical radial glial cells were detected by edgeR [28]. Heatmaps and cluster analysis were performed with the R package ‘pheatmap’ [29]. See Table S1 for RNA-seq data used in this study.

lncRNA Catalog Assembly

The GTF file generated from StringTie was used as the input for GffCompare. The complete GENCODE gene annotation GTF file (release M22 GRCm38.p6) and NONCODE v5 mouse gene annotation GTF file were used as references. Completely matched protein-coding and non-coding genes were excluded. For transcripts not matched with the reference lncRNA database, another filtering was performed by Blastx in the *kegg* [30], *nr* [31], *cog* [32], and *swissprot* [33] protein databases. Transcripts mostly matched with known protein-coding genes (identity > 0.9 and coverage > 0.8) were removed from the final lncRNA list. For the remaining transcripts, intrinsic sequence feature-based software, CPC2 (Coding Potential

Calculator) [34] and CPAT (Coding-Potential Assessment Tool) [35] were used to predict novel lncRNAs. The final lncRNA catalog contained known and novel lncRNAs (see Table S2).

Co-Expression Analysis

For co-expression analysis and module construction, WGCNA (Weighted Gene Co-Expression Network Analysis) [36], a topology distance-based R package, was employed as previously described [37]. The gene expression matrix was processed by a variance stabilizing transformation of the RNA-seq count data followed by the workflow of WGCNA to generate a gene co-expression network with the parameter ‘power = 5’. *K*-means clustering of RNA-seq data from E14.5 ventricular zone (VZ), SVZ/intermediate zone (IZ), and cortical plate (CP) was carried out by iDEP (integrated web application for differential expression and pathway analysis of RNA-seq data), a web-based RNA-seq data analysis tool [38]. See Tables S3 and S4 for WGCNA co-expression and *K*-means clustering results.

Gene Ontology Enrichment Analysis and Cis-Regulatory Position Analysis

Gene Ontology analysis was performed by DAVID [39, 40]. Further enrichment analysis was carried out on ShinyGO [41], which includes the Gene Signature Database (GeneSigDB) [42] and the L2L database [43]. GeneSigDB and L2L consist of lists of differentially-expressed genes, with their signatures extracted and manually curated from the literature and mammalian microarray studies respectively. Genomic Regions Enrichment of Annotations Tools (GREAT) [44] was employed for the *cis*-regulatory analysis.

Cell Cultures

The murine Neuro-2a neuroblastoma cell line was from the Cell Bank of the Chinese Academy of Sciences, and the human embryonic kidney (HEK) 293 cell line was a gift from Dr. Hongbing Shu at Wuhan University. Cells were cultured in the indicated culture media (Dulbecco’s modified Eagle’s medium or minimum essential medium) containing 10% fetal bovine serum and penicillin/streptomycin (Life Technologies, Carlsbad, United States). For neurosphere culture assays, cortical NPCs dissociated from E12.5 mouse cortex were cultured on ultra-low-attachment plates (Corning, New York, United States) and maintained in F12 medium (Life Technologies) with N2 and B27 supplements (1×, Life Technologies), 1 mmol/L Na-pyruvate, 1 mmol/L N-acetyl-L-cysteine (NAC), human

recombinant FGF2, and EGF (20 ng/mL each; Life Technologies) [45].

Plasmid Construction

Full-length mouse *SenZfp536* and *Zfp536* were amplified from cDNAs of E13.5 mouse cortex and then cloned into pCAGGS. The short hairpin RNA (shRNA) vectors were cloned by inserting annealed oligonucleotides targeting *SenZfp536* or *Zfp536* into pLKO.1-ZsGreen (LKO shRNA) or pCAG-mir30 (mir30 shRNA) vectors. For luciferase reporter assays, 5×UAS-TK-Luc, pcDNA3-Gal4-λN, and pcDNA3-BoxB-LacZ were gifts from Dr. Xiang Lv (Chinese Academy of Medical Sciences & Peking Union Medical College), with the *LacZ* cassette in pcDNA3-BoxB-LacZ replaced by the full-length *SenZfp536* sequence. For *SenZfp536* tethering assays, the sequences for *SenZfp536*, antisense *SenZfp536*, or *Gfp* were tagged with an sgRNA sequence to target the promoter of *Zfp536*. The fused cassettes were cloned into pcDNA3.1. For CRISPRa assays, sgRNA sequences targeting the *SenZfp536* promoter were cloned into sgRNA (MS2) cloning vector (Addgene, #61424). See Table S5 for the oligonucleotide sequences used.

Lentivirus Packaging

HEK 293T cells were seeded at 3.8×10^6 cells per 10-cm plate and incubated at 37°C under 5% CO₂ for ~20 h. Twelve micrograms of pLKO.1-ZsGreen, 6 μg psPAX2 and 6 μg pMD2.G vectors were transfected into 293T cells using calcium phosphate. The supernatant containing lentivirus particles was harvested at 48 h post-transfection and filtered through a Millex-GP filter unit (0.45 μm pore size, Merck, Darmstadt, Germany). The viral supernatant was aliquoted, snap-frozen in liquid nitrogen, and stored at –80°C.

Cell Proliferation Assay (CCK-8 Assay)

Vectors expressing shRNAs against *SenZfp536* were transfected into 70%–80% confluent Neuro-2a cells. After two days, the cells were re-suspended and seeded into 96-well culture plates (1000 cells/well). For detection, 10 μL CCK-8 solution (C0039, Beyotime, Shanghai, China) was added to each well, followed by 37°C and 5% CO₂ incubation for 1 h. The absorbance at 450 nm was measured with a microplate reader (ELx800, BioTek, Winooski, United States).

Immunofluorescence and Immunoblotting

Mouse brains were fixed overnight in 4% formaldehyde (Merck), then equilibrated in 20% sucrose-PBS and frozen-

embedded in O.C.T. (Sakura, Torrance, United States). Cryosections (14 μm) were blocked in blocking buffer (3% heat-inactivated normal goat serum, 0.1% bovine serum albumin, and 0.1% Triton-X 100 in 10 mmol/L Tris-HCl, pH 7.4; 100 mmol/L NaCl) for 1 h at room temperature. The sections were incubated with diluted primary antibodies overnight at room temperature. After washing three times (5 min–10 min each) with 1 \times PBS, and incubation with secondary antibodies for 1 h at room temperature, the sections were mounted with anti-fade reagent with DAPI and stored at 4°C. BrdU immunoblotting assays were carried out according to a previous procedure [46]. Briefly, slides were treated with 20 $\mu\text{g}/\text{mL}$ proteinase K (Merck) for 10 min followed by 2 mol/L HCl for 30 min prior to blocking. See Table S5 for antibodies used.

5' and 3' Rapid Amplification of cDNA Ends (5' and 3' RACE)

We designed nested primers to target regions of *SenZfp536* according to previous studies [47] and our RNA-seq data. For 5' RACE, two different approaches, a 5'-Full RACE Kit with TAP (Takara, Kusatsu, Japan) and a SMARTtm RACE Kit (Clontech, Mountain View, United States) were applied according to their manuals. 3' RACE was carried out according to the manufacturer's guide (Takara). PCR products were cloned into pGEM-T easy vectors (Promega, Madison, United States) and analyzed by Sanger sequencing to identify the 5' and 3' ends of *SenZfp536*.

Northern Blotting

Total RNAs were extracted using TRIzol (Thermo Fisher Scientific) from Neuro-2a cells. Twenty micrograms of total RNAs were subjected to formaldehyde denaturing agarose electrophoresis followed by transfer to a positively-charged nylon membrane (Beyotime) with 20 \times SSC buffer (3.0 mol/L NaCl and 0.3 mol/L sodium citrate, pH 7.0) through an ascending capillary transfer system. Wet membranes were cross-linked by UV irradiation (254 nm for 1 min 45s, 1.5 J/cm²) and treated with pre-warmed DIG Easy Hyb Hybridization solution (Roche, Carlsbad, United States) at 65°C for 1 h. DIG-labeled RNA probes (*SenZfp536* and *Gapdh*) generated by *in vitro* transcription were denatured at 85°C for 5 min and chilled on ice. Hybridization was carried out at 65°C overnight for at least 10 h. Membranes were washed three times in wash buffer 1 (0.1 \times SSC and 0.1% SDS) for 15 min at 65°C, then rinsed in wash buffer 2 [0.1 mol/L maleic acid, 0.15 mol/L NaCl, 0.3% Tween 20 (pH 7.5)]. After incubation in blocking reagent (Roche) for 1 h at room temperature, the membrane was incubated with a 50,000-fold dilution of anti-DIG-AP Fab fragment (Roche) in blocking reagent for 30

min at room temperature, and washed three times in wash buffer 2 for 10 min at room temperature. After immersion in detection buffer [0.1 mol/L Tris-HCl, 0.1 mol/L NaCl (pH 9.5)] for 5 min, bands on the membrane were detected using the CDP-star chemiluminescent substrate for alkaline phosphatase (Roche) and X-ray film exposure. Primers used in generating probes for Northern blots are listed in Table S5.

Luciferase Reporter Assay

For BoxB-Gal4- λN luciferase reporter assays, Neuro-2a cells were seeded in 24-well plates for 24 h and transfected with mixed vectors of 50 ng 5 \times UAS-TK-Luc, 50 ng pcDNA3-Gal4- λN ; and 50 ng pcDNA3-BoxB-*LacZ/SenZfp536* along with 5 ng pTK-Ren vectors. Cells were harvested and delivered to the reporter activity testing assay by a Dual-Glo luciferase Assay System on the GloMax Luminometer System (Promega) at 24 h post-transfection. Data were gathered by calculating the Luc/Ren ratio of triplicate wells.

CRISPR/dCas9-Mediated Transcription Activation (CRISPRa) Assay

CRISPRa assays were performed according to published procedures [48, 49]. In short, sgRNA sequences targeting the promoter of *SenZfp536* were cloned into the sgRNA (MS2) cloning vector (Addgene, #61424). Neuro-2a cells seeded for 24 h were transfected with mixed vectors of sgRNA, dCAS9-VP64-GFP, and MS2-P65-HSF1-GFP. Cells were harvested to extract RNAs for assessing *SenZfp536* and *Zfp536* RNA levels by RT-qPCR at 48 h post-transfection and for assessing ZFP536 protein levels by immunoblotting at 72 h post-transfection.

RNA Tethering

SgRNAs that activate *Zfp536* transcription were identified by CRISPRa assays. The most efficient sgRNA was cloned into the pcDNA3.1 vector and tagged with *Gfp*, antisense *SenZfp536* or *SenZfp536* sequences. Neuro-2a cells were transfected with mixed vectors that expressed dCas9 along with sgRNA tagged with *Gfp*, antisense *SenZfp536*, or *SenZfp536*. Transfected cells were harvested to extract RNAs for assessing *Zfp536* RNA levels by RT-qPCR at 48 h post-transfection.

Chromosome Conformation Capture (3C)

The 3C assay was performed according to the published protocol [50]. In short, Neuro-2a cells were harvested (1 \times 10⁷ cells) and crosslinked with 2% formaldehyde for

10 min at room temperature followed by quenching with 0.125 mol/L glycine for 5 min. Then extracted nuclei were digested by BamH I (NEB Biolabs) overnight at 37°C, followed by treatment with T4 ligase (Takara) for 30 min at room temperature.

DNAs were purified by phenol-chloroform extraction and analyzed by qPCR. The specificity and efficiency of all 3C primers were agarose-gel-verified by digestion and ligation of the BAC DNA (RP23-454K9) that contained the region of interest. For each primer pair (test primer to constant primer), we performed a triplicate quantification (Ct1, Ct2, and Ct3), and then calculated the mean Ct. The crosslinking value was calculated using parameters of the standard curve of BAC DNA: $\text{value} = 10^{(Ct-b/a)}$ (b : intercept; a : slope). These values were finally normalized to *Gapdh* to generate relative crosslinking frequency.

In Situ Hybridization

cDNA templates for antisense RNA probes were PCR amplified from mouse cDNA with primers specific to *SenZfp536* and *Zfp536*. PCR products were gel-extracted and cloned into the pGEM-T Easy vector (Promega) that contains T7 and SP6 promoters. The vectors were linearized by restriction enzyme and then transcribed using a DIG-RNA labeling kit (Roche). The DIG-labeled probes were stored at -80°C. CD-1 mouse brains at E12.5, E14.5, and E16.5 were fixed overnight in RNase-free 4% formaldehyde (Sigma), then equilibrated in 20% sucrose-PBS and frozen-embedded in O.C.T. (Sakura). Cryosections (14 µm) were incubated with DIG-labeled RNA antisense probe (0.1 ng/µL–0.2 ng/µL in hybridization buffer) overnight at 65°C in humidified boxes after treatment with proteinase K (2 µg/mL, 10 min) and acetylation in 0.1 mol/L TEA (triethanolamine) for 10 min. Slides were washed four times (20 minutes each) in 0.1×SSC at 65°C followed by blocking with 10% heat-inactivated sheep serum in buffer B1 (0.1 mol/L Tris-HCl, pH 7.4, 150 mmol/L NaCl) at room temperature for 2 h. Sections were incubated with the anti-DIG antibody (Roche; 1:5,000) overnight at 4°C and then washed three times with buffer B1. Slides were equilibrated with buffer B3 (0.1 mol/L Tris-HCl, 0.1 mol/L NaCl, 50 mmol/L MgCl₂, 0.1% Tween-20, pH 9.5) twice for 10 min, followed by staining with NBT/BCIP (Roche; 4.5 µL/mL each in buffer B3) overnight in the dark. Slides were dehydrated using gradient ethanol and xylene sequentially, and mounted with neutral balsam.

In Utero Electroporation (IUE)

In utero electroporation was carried out as published [51]. In brief, pregnant CD-1 mice with E13.5 embryos were

anesthetized by injection of pentobarbital sodium (70 mg/kg), and the uterine horns were exposed for microinjection of mixed plasmids with 0.05% Fast Green (Sigma) into the lateral ventricles of brains. Five pulses (36 V; 50 ms on/950 ms off) were released across the heads of embryos by 5-mm platinum plate-electrodes connected to a CUY21-VIVO-SQ electroporator (BEX, Tokyo, Japan). The uteri were replaced and the incisions were sutured. Mice were treated with analgesia (ibuprofen in drinking water) until sacrifice at E15.5 or E16.5.

Statistics

Statistical plots were generated and analyzed using Graphpad Prism 8.0 software. Results are presented as the mean ± SEM or ± SD. The two-tailed *t*-test was used to compare two groups. One-way ANOVA followed by *Dunnnett's* multiple comparison test was used for comparison of three or more groups with control groups. One-way ANOVA followed by Tukey's multiple comparison test was used for comparison of three or more groups among each other. All animals and biochemical experiments were conducted with at least three replicates per group.

Data Access

Strand-specific RNA-seq of E10.5 and E12.5 CD-1 mouse forebrains have been deposited in the NCBI Gene Expression Omnibus under accession number GSE55600.

Results

Cataloging lncRNAs in Developing Mouse Cerebral Cortex

To systematically identify lncRNAs associated with cortical development, we used an *ab initio* approach (Fig. 1A). First, public RNA-seq data for E10.5, E11.5, E12.5, E13.5, E14.5, E15.5, E16.5, and P0 mouse forebrains (developing cerebral cortices, Fig. S1A) were obtained from ENCODE (Encyclopedia of DNA Elements) [52, 53]; and RNA-seq data for the VZ, SVZ/IZ, and CP at E14.5 [54] and the cerebral cortex at E17 [55] were downloaded. To supplement lncRNA without a 3' polyadenylation (polyA) tail, strand-specific RNA-seq on ribosomal-depleted total RNA isolated from the dorsal forebrain at E10.5 and E12.5 was implemented at a sequencing depth of >200 million reads per sample. Moreover, RNA-seq data acquired from ESCs and ESC-derived NPCs [56] were incorporated into our transcriptome database for developing murine cerebral cortex (see also Table S1). Subsequently, we aligned >1,800 million RNA-seq clean reads to the reference

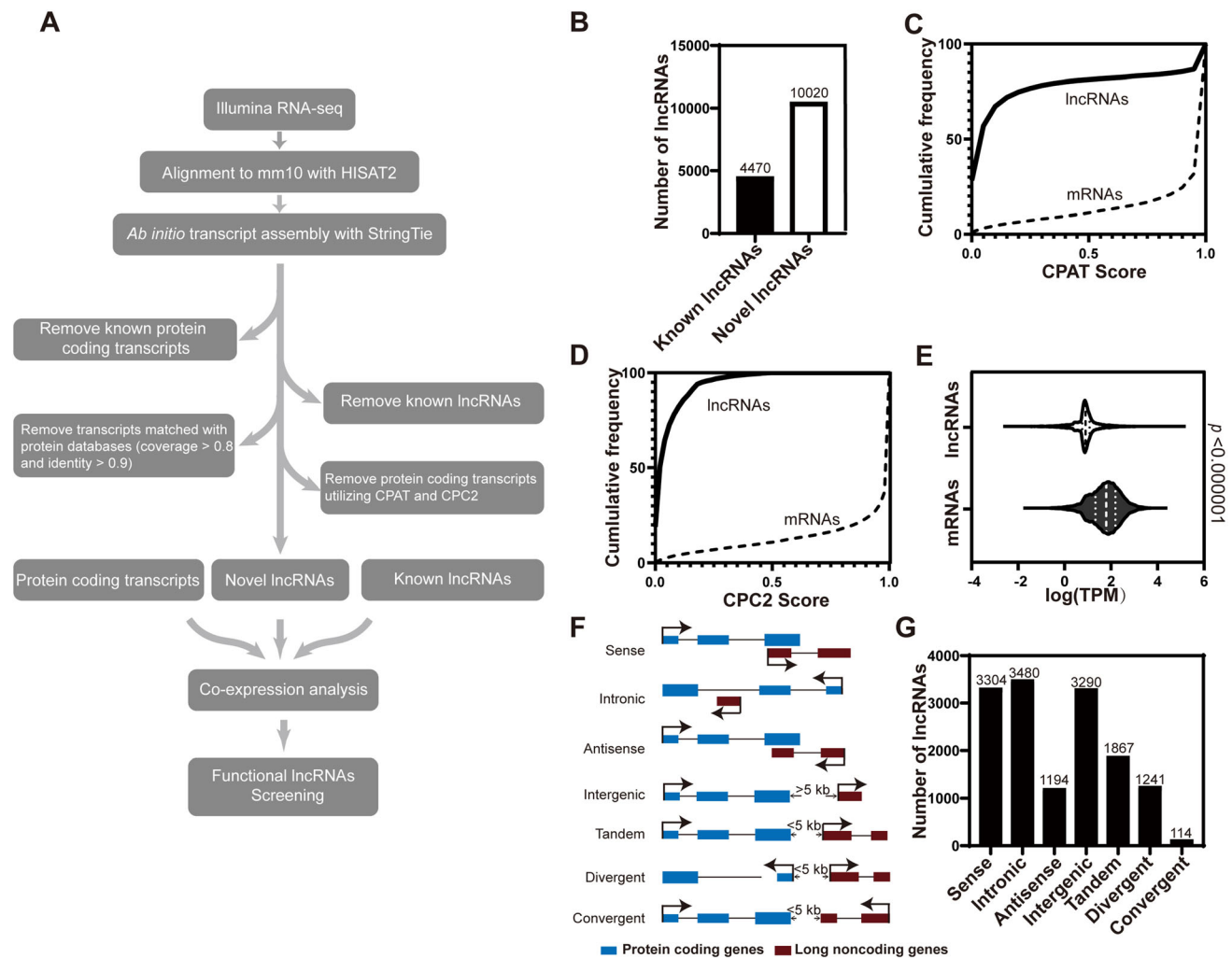


Fig. 1 Comprehensive cataloging of lncRNAs in developing mouse cerebral cortex. **A** The bioinformatics pipeline for identification and annotation of lncRNAs expressed in developing mouse cerebral cortex. See also Table S1. **B** Numbers of known and novel lncRNAs identified in this study. Known lncRNAs have been annotated in GENCODE v25 and NONCODE v5. **C** Cumulative frequency of coding potential score for lncRNAs and mRNAs calculated by CPAT.

genome (GRCm38/mm10) by HISAT2 [25], followed by *ab initio* transcript assembly using StringTie [26, 57]. In conformity with this pipeline (Fig. 1A), a total of 31,067 gene loci and 54,076 transcripts were eventually identified.

Using GffCompare [26] and Blastx [58], protein-coding genes and known lncRNAs were successfully characterized, which were also annotated in GENCODE M22 [59] and NONCODE v5 [60] respectively. After filtering with CPAT [35] and CPC2 [34], two protein-coding potential calculators, 4,470 known and 10,020 novel lncRNAs were authenticated (Fig. 1B). The re-use of CPC2 and CPAT validated the classification of lncRNA and protein-coding (mRNA) transcripts (Fig. 1C, D). Consistent with previous studies [56, 61, 62], lncRNAs associated with developing

neocortex were expressed at significantly lower levels than mRNAs (Fig. 1E). By cross-referencing with transcriptome data from apical radial glia (aRGs), basal intermediate progenitors (bIPs), basal radial glia (bRGs), and projection neurons [63] at E14.5, we identified numerous cell-type-specific mRNA and lncRNA transcripts. Interestingly, aRGs expressed the most cell-type-specific transcripts (Fig. S1B). Moreover, many mRNAs and lncRNAs were also differentially expressed between proliferative (Tis21⁻) and differentiative (Tis21⁺) mouse aRGs (Fig. S1C) [63]. Thus, the expression of lncRNAs is spatiotemporally regulated during cortical development and among cell types.

neocortex were expressed at significantly lower levels than mRNAs (Fig. 1E).

By cross-referencing with transcriptome data from apical radial glia (aRGs), basal intermediate progenitors (bIPs), basal radial glia (bRGs), and projection neurons [63] at E14.5, we identified numerous cell-type-specific mRNA and lncRNA transcripts. Interestingly, aRGs expressed the most cell-type-specific transcripts (Fig. S1B). Moreover, many mRNAs and lncRNAs were also differentially expressed between proliferative (Tis21⁻) and differentiative (Tis21⁺) mouse aRGs (Fig. S1C) [63]. Thus, the expression of lncRNAs is spatiotemporally regulated during cortical development and among cell types.

Cis-Regulatory Prediction of lncRNAs

Recent studies have shown that lncRNAs often *cis*-regulate their adjacent protein-coding genes [20], which would facilitate functional annotation of lncRNAs. To categorize lncRNAs and explore their potential biological roles, we described all lncRNAs according to their genomic distribution patterns relative to protein-coding loci. Among them, 22.8% (3,304) lncRNAs partially overlapped with adjacent protein-coding genes (PCGs) in the sense orientation, while the transcription of 8.2% (1,194) of lncRNAs were antisense and partially overlapped with adjacent PCGs. Twenty-four percent (3,480) of lncRNAs were transcribed from intronic regions of PCGs, and 22.7% (3,290) had no PCGs within their 5-kb flanking regions (intergenic). The rest, 22.2% (3,222), were tandem, divergent, and convergent lncRNAs that did not overlap with, but were transcribed within the 5-kb flanking regions of their PCG neighbors (Fig. 1F, G).

To predict the *cis*-regulatory roles of lncRNAs, the GREAT program [44] was applied to analyze the enrichment of molecular functions and pathways for their neighboring PCGs, excluding sense and intronic lncRNAs. Most enriched terms were associated with the cell cycle and regulation of mitosis, suggesting that lncRNAs have vigorous effects on cell proliferation during cortical development (Fig. S1D, E). In addition, the distances between lncRNA genes and transcription start sites (TSSs) of nearby PCGs ranged from 5 kb to 500 kb (Fig. S1F). Numerous HiC-Seq studies have shown that non-coding regions of the genome regulate the transcription of nearby PCGs *in cis* [64, 65]. Thus, Pearson correlation coefficients for expression levels between lncRNAs and nearby PCGs were calculated. As expected, the expression of > 20% of lncRNAs were strongly correlated (Pearson > 0.6) with their PCG neighbors during cortical development (Fig. S1G).

Co-Expression Analysis of Transcripts Expressed in Developing Cerebral Cortex

To associate expression patterns with developmental stages, WGCNA (Weighted Gene Co-Expression Network Analysis) [36, 66] was used to categorize groups of transcripts, or ‘Modules’. Transcripts in each module showed either positively or inversely correlated expression patterns, thus were probably involved in common biological processes or had regulatory relationships (Fig. 2A, B).

Through analyzing eight developmental time points of the forebrain (two samples for each time point from E10.5 to P0) from ENCODE, we successfully established 46 co-expression modules, each of which exhibited unique stage-specific characteristics. For instance, Modules 1 and 2 were enriched with transcripts whose expression levels gradually

increased or decreased during cortical development. Genes essential for the stemness properties of neural progenitors such as *Sox2*, *Notch1*, and *Hes1*, and the proneural gene *Neurog1* were assigned to Module1. Moreover, transcripts in Module 5 had the highest or lowest expression level at E10.5. Modules 2, 3, and 5 were enriched with genes involved in transit-amplifying and neuronal differentiation of the neocortex (Fig. 2C–F, S2A–D). To discover the functional significance of genes in Modules 1 and 2, enrichment analysis of PCGs in both modules was performed in L2L [43] and GeneSigDB [42], which consist of lists of differentially-expressed genes with their signatures extracted and manually curated from mammalian microarray studies and the literature, respectively. The results showed that genes in Modules 1 and 2 highly correlated with terms of stem cells and organ development (Fig. S2E–H).

Clustering of Transcripts Based on Their Spatial Distribution to Identify Genes Associated with Self-Renewal of Cortical NPCs

We next analyzed spatial expression data from the dorsal forebrain at E14.5 [54]. RNA-seq data from the VZ, SVZ/IZ, and CP at E14.5 were analyzed and an unsupervised learning algorithm, *K*-means clustering, was performed to classify long noncoding and protein-coding transcripts based on their spatial distribution [67–69]. A total of 9,677 genes were categorized into 10 main clusters (Clusters 0–9), each exhibiting a spatially-specific pattern (Fig. 3A). Clusters 1, 2, and 4 contained the most NPC-related genes and shared similar expression trends, namely, highest expression levels in the VZ but reduced in the SVZ/IZ and CP at E14.5. Moreover, phenotype enrichment analysis based on the Gene List Analysis and Visualization (Vlad) [70] and MGI phenotypes [71] revealed that PCGs in Clusters 1, 2, and 4 were enriched for phenotypes related to abnormal cell cycle, mitosis, and cell lethality (Fig. 3B–D).

Aiming to identify lncRNAs with regulatory roles in controlling the self-renewal of cortical NPCs, we integrated temporal and spatial co-expression data. Specifically, 1,930 mRNAs and 722 lncRNAs were at intersections of Modules 1 and 2 with Clusters 1, 2, and 4, including PCGs implicated in stemness and self-renewal of cortical NPCs - *Sox2*, *Pax6*, *Numb*, and *Fabp7* (Fig. 3E). Gene Ontology enrichment analysis indicated that PCGs at these intersections were associated with functions for cell cycle, DNA repair, and modulation of transcription (Fig. 3F). Collectively, we sorted out lncRNA candidates with similar spatiotemporal patterns and potential roles in regulating the self-renewal of NPCs during cortical development.

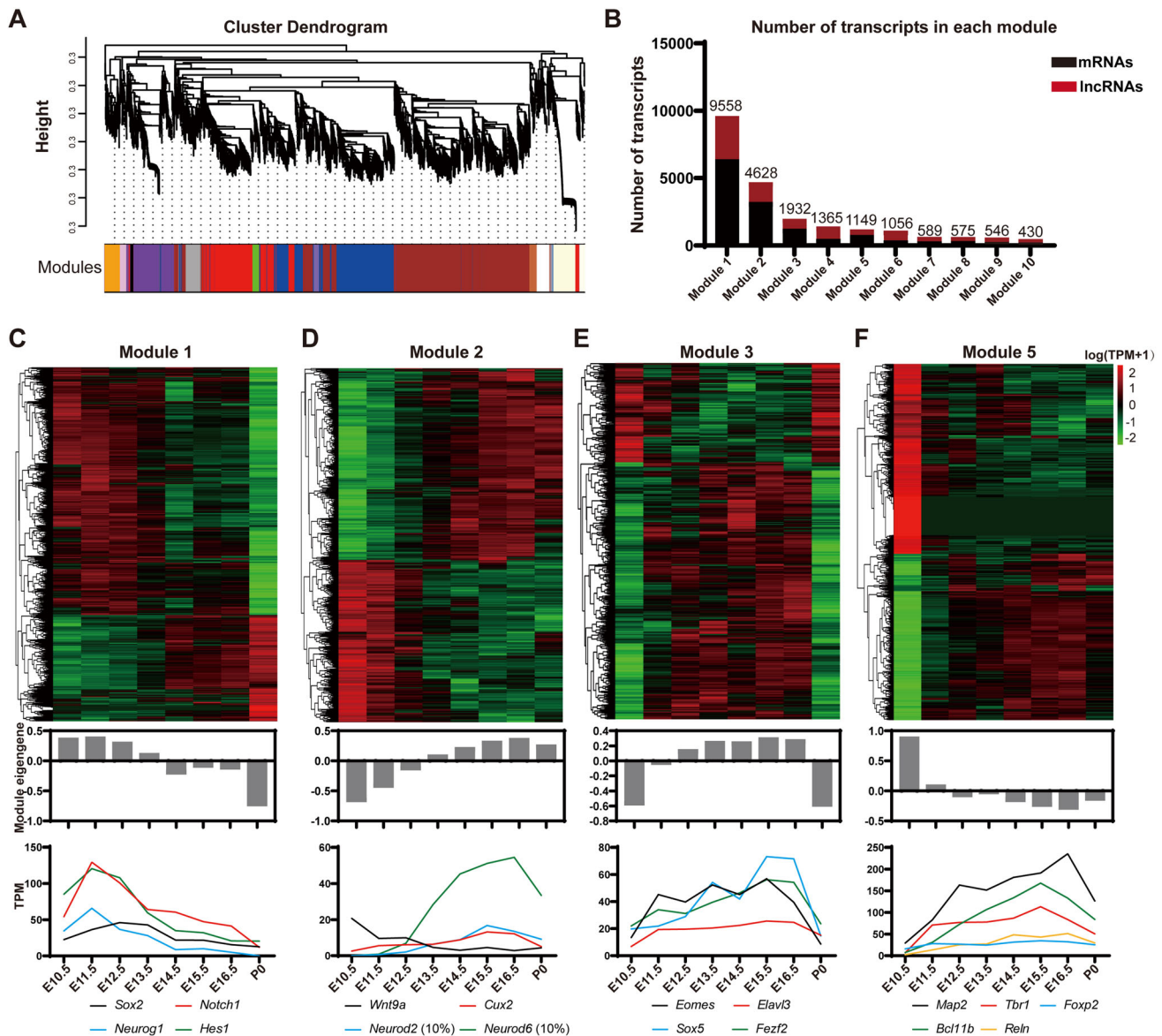


Fig. 2 Co-expression modules categorize transcripts implicated in cortical development. **A** Hierarchical cluster dendrogram of co-expression modules for lncRNA genes and mRNA genes. A total of 46 modules are labeled with different color bands beneath corresponding tree branches. **B** Numbers of mRNAs and lncRNAs in Module 1–10. **C–F** Heatmaps and Eigengene bar plots of four modules. Upper panels: expression levels of lncRNAs and PCGs

normalized to $\log_{10}(\text{TPM}+1)$ (red, higher expression; black, neutral expression; green, lower expression); middle panels: module Eigengene barplots representing expression levels of all transcripts, calculated by *moduleEigengenes* in the WGCNA package and in line with the first principal component obtained by singular value decomposition of each module; lower panels: broken line graphs exhibit expression patterns of representative protein-coding genes.

Characterization of *SenZfp536*

Among the 722 lncRNAs at intersections of Modules 1 and 2 with Clusters 1, 2, and 4 with potential roles in regulating NPC self-renewal, the 20 most VZ-enriched transcripts were screened using neurosphere culture assays. We identified a lncRNA that inhibits the proliferation of cortical NPCs. This lncRNA was previously annotated as AK163177 (NCBI) and subsequently termed *SenZfp536* due to its partial overlap with the last exon of *Zfp536* in the

sense direction. The genomic region of *SenZfp536* is evolutionarily conserved and is associated with RNA polymerase II (Pol II), H3K4me3, and H3K36me3 in mouse brain at E14.5 (Fig. 4A), suggesting its active transcription in the developing brain. Northern blot, 5' and 3' rapid amplification of cDNA end (RACE) revealed that *SenZfp536* is a 2,542-base-long transcript, with its TSS and the first exon located inside the last exon of *Zfp536* (Fig. 4A–C). Moreover, no protein-coding potential was detected in *SenZfp536* using CPC2 and CPAT (Fig. 4D).

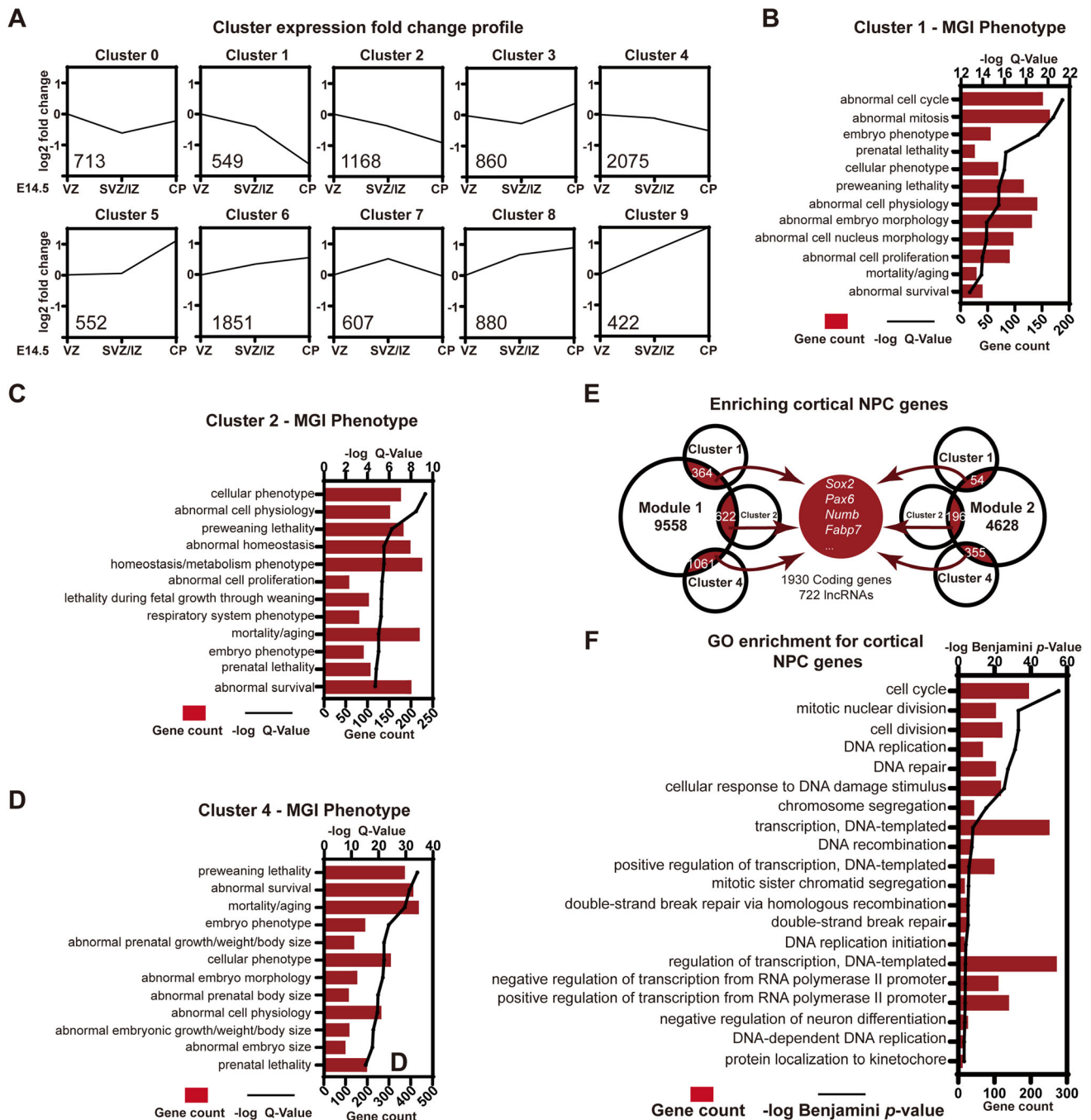


Fig. 3 Spatial clustering of lncRNAs and characteristics of genes associated with cortical NPCs. **A** *K*-means clustering based on the spatial distribution of transcripts. Expression fold-changes are plotted to exhibit trends from E14.5 VZ, SVZ/IZ to the CP. Number of genes in each cluster is marked in the lower-left corner. **B–D** Top 12 MGI phenotype terms enriched in Clusters 1, 2, and 4. Bars display the gene count in each term and lines reflect $-\log Q$ values. **E** Venn

diagrams showing intersections between Modules 1 and 2 and Clusters 1, 2, and 4. Intersected transcripts are identified as NPC-associated genes, with a few key genes listed. **F** Top 20 Gene Ontology (GO) Biological Process terms enriched in NPC-related PCGs, as determined by DAVID GO analysis. Bars display the gene count in each GO term and the line reflects $-\log$ Benjamin *P* values.

Quantitative analysis of RNA-seq data confirmed the correlated expression trends between *Zfp536* and *SenZfp536* in the developing cerebral cortex, with *SenZfp536* expressed at 1/2 to 2/3 the levels of *Zfp536* (Fig. 4E).

Consistent with their relatively high expressions in the VZ and SVZ/IZ (Fig. 4F), *in situ* hybridization demonstrated that both *SenZfp536* and *Zfp536* were dominantly expressed in the VZ/SVZ of the developing cerebral

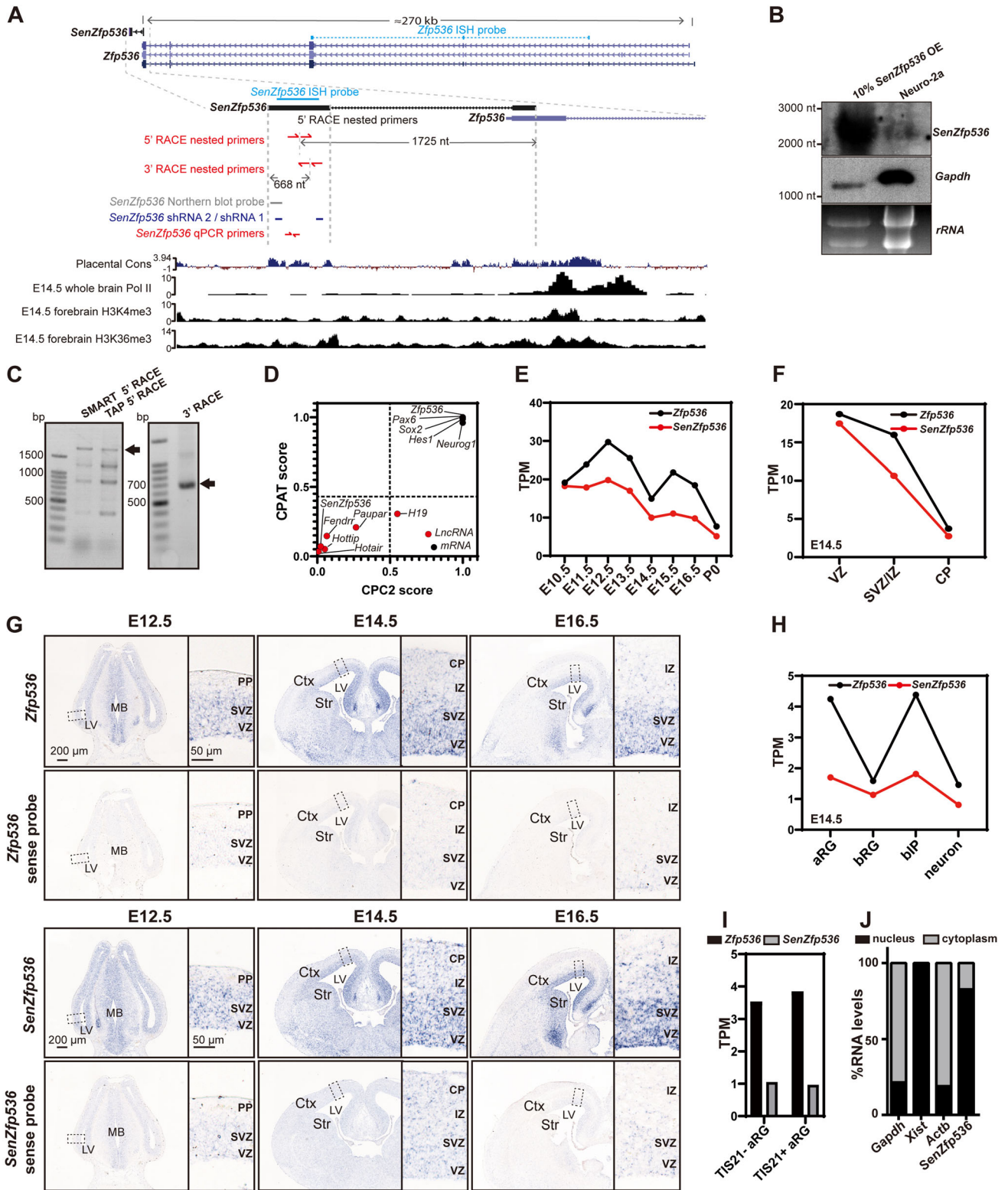


Fig. 4 The lncRNA *SenZfp536* is a sense lncRNA overlapping with *Zfp536*. **A** Schematic of the murine *SenZfp536* locus (UCSC mm10). Upper track: the sequence conservation (Cons) among placental animals; lower tracks: ChIP-seq signals for Pol II, H3K4me3, and H3K36me3 in whole brain or forebrain at E14.5. Locations of the ISH probes, Northern probe, shRNA targets, qPCR primers and nested primers for 5' and 3' RACE are also shown. **B** Northern blot analysis of RNA extracted from Neuro-2a cells expressing exogenous *SenZfp536* (left lane, 10% input) and wild-type Neuro-2a cells (right lane). **C** Agarose gels showing products of 5' and 3' RACE nested PCR for *SenZfp536*. Two 5' RACE approaches, SMART and TAP, were applied. All visible bands were collected for DNA sequencing. Arrows indicate specific *SenZfp536* products. **D** Coding potential scores for *SenZfp536* and other known long non-coding and protein-coding transcripts. Dotted lines indicate cutoffs for coding and non-coding values. **E** Line plots for the temporal expression of *SenZfp536* and *Zfp536* in developing cortex. **F** Line plots for the spatial expression of *SenZfp536* and *Zfp536*, using RNA-seq data from the VZ, SVZ/IZ, and CP at E14.5. **G** *In situ* hybridization of *SenZfp536* and *Zfp536* in coronal sections of embryonic mouse brains using anti-sense (upper panels) and sense probes (lower panels). Magnified images are displayed in the right panels. Ctx, cortex; Str, striatum; LV, lateral ventricle; PP, preplate; CP, cortical plate; IZ, intermediate zone; SVZ, subventricular zone; VZ, ventricular zone. **H** Line plots for expression of *SenZfp536* and *Zfp536* in apical radial glia (aRG), basal radial glia (bRG), basal intermediate progenitors (bIP), and neurons at E14.5. **I** Expression of *SenZfp536* and *Zfp536* in proliferative (Tis21-GFP-) and differentiative (Tis21-GFP+) aRG of cortex at E14.5. **J** Subcellular distribution of *SenZfp536* in Neuro-2a cells. Most *SenZfp536* RNAs reside in the nucleus. - *Gapdh*, *Xist*, and *Actb* are reference RNAs.

cortex, whereas sense probes did not elicit specific signals (Fig. 4G). Quantitative analysis of high-throughput data from aRGs, bIPs, bRGs, and neurons at E14.5 [63] revealed that both *Zfp536* and *SenZfp536* were expressed at higher levels in aRGs and bIPs than bRGs and neurons (Fig. 4H). Moreover, *Zfp536* and *SenZfp536* were expressed at comparable levels in proliferative (Tis21-) and differentiative (Tis21+) mouse aRGs (Fig. 4I). Subcellular fractionation followed by quantitative PCR showed that ~80% of *SenZfp536*'s transcripts were located in the nucleus, indicating related roles (Fig. 4J).

SenZfp536* Positively Regulates the Expression of *Zfp536

Since *SenZfp536* is transcribed from the last exon of *Zfp536*, we then explored whether *SenZfp536* regulates the transcription of *Zfp536*. Knocking down *SenZfp536* by shRNAs significantly downregulated the expression of *Zfp536* in Neuro-2a cells (Fig. 5A). Accordingly, knocking down *SenZfp536* using two types of shRNA-expressing vectors, pCAG-mir30 and pLKO.1, attenuated ZFP536 at the protein level (Fig. 5B).

Next, we evaluated whether activation of *SenZfp536*'s transcription up-regulates the expression of *Zfp536* using the CRISPR/dCas9-mediated transcription activation

system (CRISPRa) [48, 49]. Activation of *SenZfp536* also elevated the levels of transcripts and proteins for *Zfp536* in Neuro-2a cells (Fig. 5C–D). Notably, the extent of transcriptional repression and activation of *Zfp536* correlated with that of *SenZfp536*, ruling out off-target effects by shRNAs or CRISPRa. However, episomal expression of *SenZfp536* had no effect on *Zfp536* expression (Fig. 5E), implying that *SenZfp536* is a positive *cis*-regulator of *Zfp536* in neural cells.

SenZfp536* is a *Cis* Transcriptional Activator of *Zfp536

To determine whether *SenZfp536* has an intrinsic ability to promote gene expression, we first used the Gal4- λ N/BoxB system to tether *SenZfp536* to the promoter of a heterologous reporter (Fig. 5F). Luciferase activity was measured in Neuro-2a cells 24 h after transfection. *SenZfp536* enhanced the luciferase activity by ~25% in comparison with the *LacZ* control (Fig. 5G). Next, transcripts for *SenZfp536*, antisense *SenZfp536*, and *Gfp* were tagged with *Zfp536*-TSS-targeted sgRNA for the RNA tethering assay. The data showed that tethering *SenZfp536*, but not antisense *SenZfp536*, to the TSS of *Zfp536* markedly induced the expression of *Zfp536* (Fig. 5H, I).

Since *SenZfp536* activates transcription of *Zfp536* *in cis* but not *in trans*, we went on to investigate if the locus that transcribes *SenZfp536* can physically associate with the promoter of *Zfp536*. We thus performed chromosome conformation capture (3C) assays using one constant primer on the promoter of *Zfp536* along with six testing primers around *SenZfp536* (Fig. 5J, K). qPCR showed that T2 and T3, the loci nearest *SenZfp536*, were significantly crosslinked with the promoter of *Zfp536* (Fig. 5L). Furthermore, depletion or activation of *SenZfp536* significantly attenuated or enhanced, respectively, the association between T2/T3 and the *Zfp536* promoter, suggesting that transcribed *SenZfp536* maintains a chromatin loop to bring close the *cis* elements around *SenZfp536* with the promoter of *Zfp536* (Fig. 5M–R).

Together, these results underscored that *SenZfp536* *cis*-activates *Zfp536*, probably by facilitating the chromatin loop between the *SenZfp536* locus and the *Zfp536* promoter.

***SenZfp536* Negatively Regulates the Self-renewal of Cortical NPCs**

SenZfp536 is highly expressed in the VZ/SVZ, the proliferative zone, of the developing neocortex. To determine if *SenZfp536* controls the self-renewal of cortical NPCs, suspension-cultured cortical NPCs at E12.5 were transduced with lentiviruses expressing shRNAs against

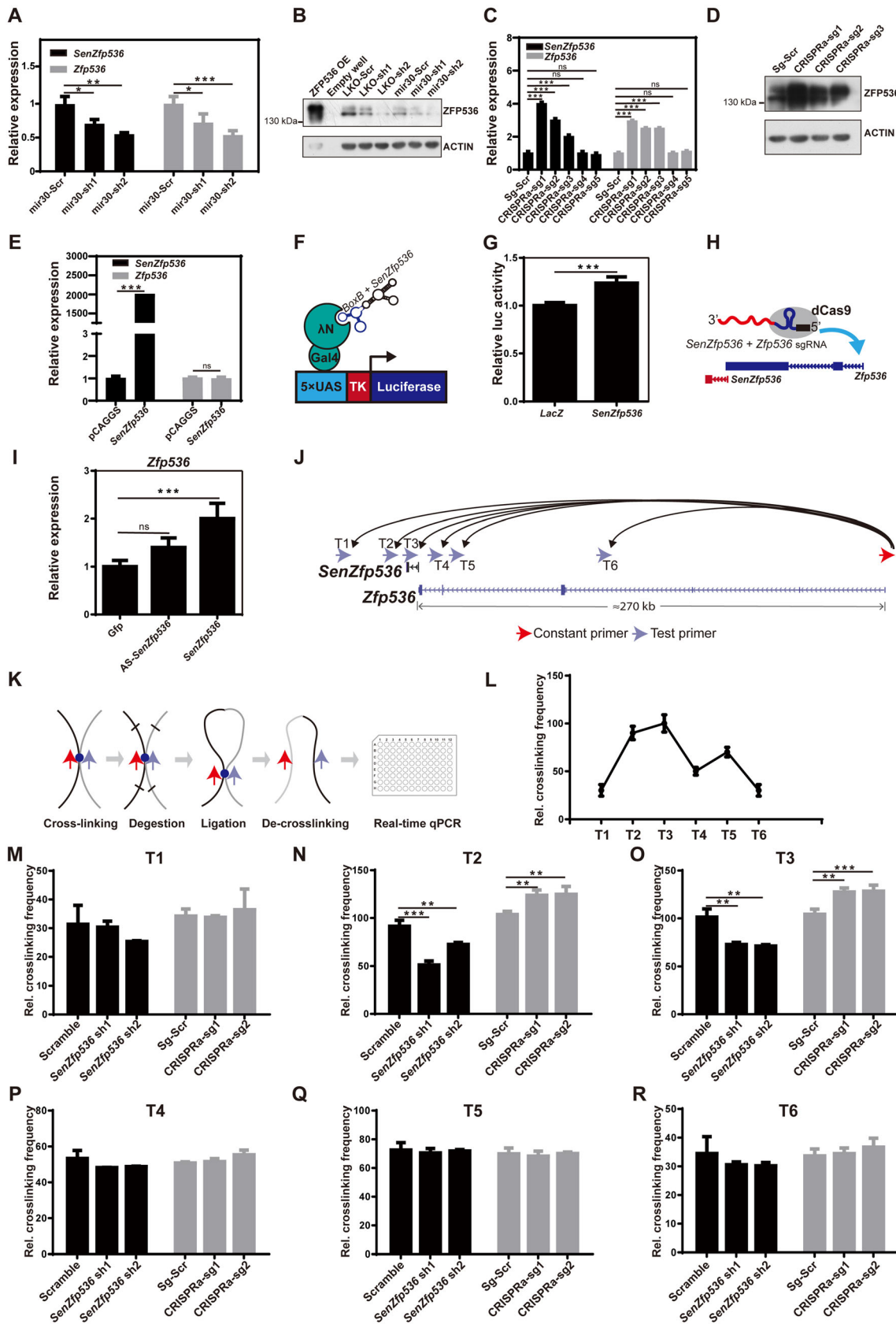


Fig. 5 *SenZfp536* maintains the expression of *Zfp536* *in cis*. **A** RNA levels of *SenZfp536* and *Zfp536* in Neuro-2a cells transfected with vectors expressing shRNAs against *SenZfp536* or scrambled control for two days. For *SenZfp536* expression, $F = 15.56$, $P = 0.0078$; for *Zfp536* expression, $F = 24.86$, $P = 0.0056$ (one-way ANOVA). **B** Immunoblots of ZFP536 and ACTIN of Neuro-2a cells transfected with vectors expressing shRNA against *SenZfp536* for three days. **C** RNA levels of *SenZfp536* and *Zfp536* in Neuro-2a cells transfected with indicated CRISPRa vectors for two days. For *SenZfp536* expression, $F = 102.4$, $P < 0.0001$; for *Zfp536* expression, $F = 61.54$, $P < 0.0001$ (one-way ANOVA). **D** Immunoblots of ZFP536 and ACTIN of Neuro-2a cells transfected with vectors that *cis*-activate *SenZfp536* expression for three days. **E** RNA levels of *SenZfp536* and *Zfp536* in Neuro-2a cells transfected with indicated vectors for two days. For *SenZfp536* expression, $P < 0.0001$; for *Zfp536* expression, $P = 0.83$ (two-tailed unpaired Student's *t*-test). **F** Schematic of the Gal4- λ N/BoxB RNA tethering system. **G** Relative luciferase activity of Neuro-2a cells transfected with plasmids expressing BoxB-tagged *LacZ* or *SenZfp536* along with Gal4- λ N and 5 \times UAS-TK-Luciferase for 24 h. Two-tailed unpaired student's *t*-test was used. **H** Schematic of dCas9-mediated RNA tethering assay. **I** Relative expression of *Zfp536* in Neuro-2a cells transfected with plasmids expressing *Zfp536*-TSS-targeted sgRNA tagged with *Gfp*, *SenZfp536*, or antisense-*SenZfp536* along with MS2-P65-HSF1-GFP and dCAS9-VP64-GFP for 48 h. $F = 16.64$, $P = 0.0076$ (one-way ANOVA). **J** Design of the chromosome conformation capture (3C) assay. Red arrow, constant primer at the *Zfp536* promoter; blue arrows; test primers; curved arrows, potential interactions between test loci and the *Zfp536* promoter. **K** Schematic of 3C-qPCR. **L** Relative crosslinking values of loci T1-T6 with the *Zfp536* promoter in Neuro-2a cells. **M-R** Relative crosslinking values of loci T1-T6 in Neuro-2a cells with depletion or enhanced expression (CRISPRa) of *SenZfp536* for two days. was performed. No significant differences were detected in T1, T4, T5, and T6. For short hairpin RNAs in (N), $F = 155.1$, $P < 0.0001$; for CRISPRa in (N), $F = 35.05$, $P = 0.0005$; for short hairpin RNAs in (O), $F = 45.35$, $P = 0.0003$; for CRISPRa in (O), $F = 120.2$, $P < 0.0001$. Replicates in each experiment, $n = 3$. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, not significant (one-way ANOVA). Results are presented as mean \pm SD.

SenZfp536 or scrambled control. The results showed that depletion of *SenZfp536* greatly increased the sphere sizes and cell numbers of cortical NPCs (Fig. 6A–C). Interestingly, knockdown of *SenZfp536* in Neuro-2a cells did not enhance their proliferation, suggesting a cell type-specific function of *SenZfp536* (Fig. S3A).

We next address whether *SenZfp536* suppresses the self-renewal of cortical NPCs *in vivo*. E13.5 cortices were electroporated with plasmids expressing shRNAs against *SenZfp536* or scrambled control for phenotypic analysis at E15.5, with transduced cells labeled with co-expressed EGFP (Fig. 6D). BrdU was administered at E14.5, 24 h prior to sacrifice. Significantly more *SenZfp536*-shRNA-transduced cortical cells expressed PAX6 and SOX2, markers for RGCs (Fig. 6E, F, H, and I). Moreover, significantly more *SenZfp536*-shRNA-transduced cortical cells expressed Ki67, with more cells double positive for Ki67 and 24-h BrdU, suggesting that *SenZfp536*-depleted

cells tend to remain as proliferative neural progenitors (Fig. 6G, J, and K).

We next extended the phenotypic analyses to E16.5. More *SenZfp536*-depleted cortical cells resided in the VZ/ SVZ and IZ and fewer in the CP than in scrambled-shRNA-transduced cells. Moreover, *SenZfp536*-depleted cells were more likely to co-localize with PAX6, SOX2, and BrdU (30 min), further supporting the idea that loss of *SenZfp536* promotes the self-renewal of cortical NPCs (Fig. S3B–I). We next determined whether loss of *SenZfp536* affects cortical layer specification. Although more *SenZfp536*-depleted cells resided in the IZ, the percentages of *SenZfp536*-depleted cells that expressed SATB2 or CTIP2, markers for superficial and deep cortical layers, respectively, were the same as for scrambled controls, indicating unaltered cortical layer fate upon loss of *SenZfp536* (Fig. S3J–M).

We further used CRISPRa to enhance the transcription of *SenZfp536* in the developing cortex [48, 49], and this led to decreased co-localization of *SenZfp536*-activated cells with PAX6 and SOX2 (Fig. S4A–D). In contrast, episomal expression of *SenZfp536* did not have an effect on the self-renewal of cortical NPCs (Fig. S4E–J). These data support the idea that *SenZfp536* controls cortical NPC self-renewal *in cis* but not *in trans*.

Zfp536 Expression Rescues *SenZfp536* Knockdown-Associated Phenotypes

Similar to the phenotypes found in *SenZfp536* knockdown, depletion of *Zfp536* in cortical NPCs also led to more transduced cells expressing PAX6 and SOX2, indicating enhanced self-renewal (Fig. S4K–N). In contrast, overexpression of *Zfp536* caused compromised self-renewal of transduced cortical NPCs, as significantly fewer *Zfp536*-expressing cells were labeled with BrdU (30 min), PAX6, and SOX2 without causing more cell death (Figs 7A–H and S4O). Notably, *Zfp536* overexpression also hampered the radial migration of cortical projection neurons, evidenced by fewer transduced cells located in the CP, indicating that proper expression levels of *Zfp536* are essential for neuronal differentiation. Last, overexpressing *Zfp536* reversed the enhanced self-renewal of cortical NPCs caused by *SenZfp536* depletion, supporting the idea that *SenZfp536* inhibits the propagation of cortical NPCs by *cis*-activating *Zfp536* (Fig. 7A–H).

To sum up, we characterized lncRNAs with spatiotemporal expression patterns in developing cerebral cortex. Among them, *SenZfp536* restricts the proliferation of cortical NPCs by *cis*-activating the expression of *Zfp536*. Mechanistically, *SenZfp536* promotes a local chromatin conformation to bring close the *SenZfp536* locus and the *Zfp536* promoter (Fig. 7I).

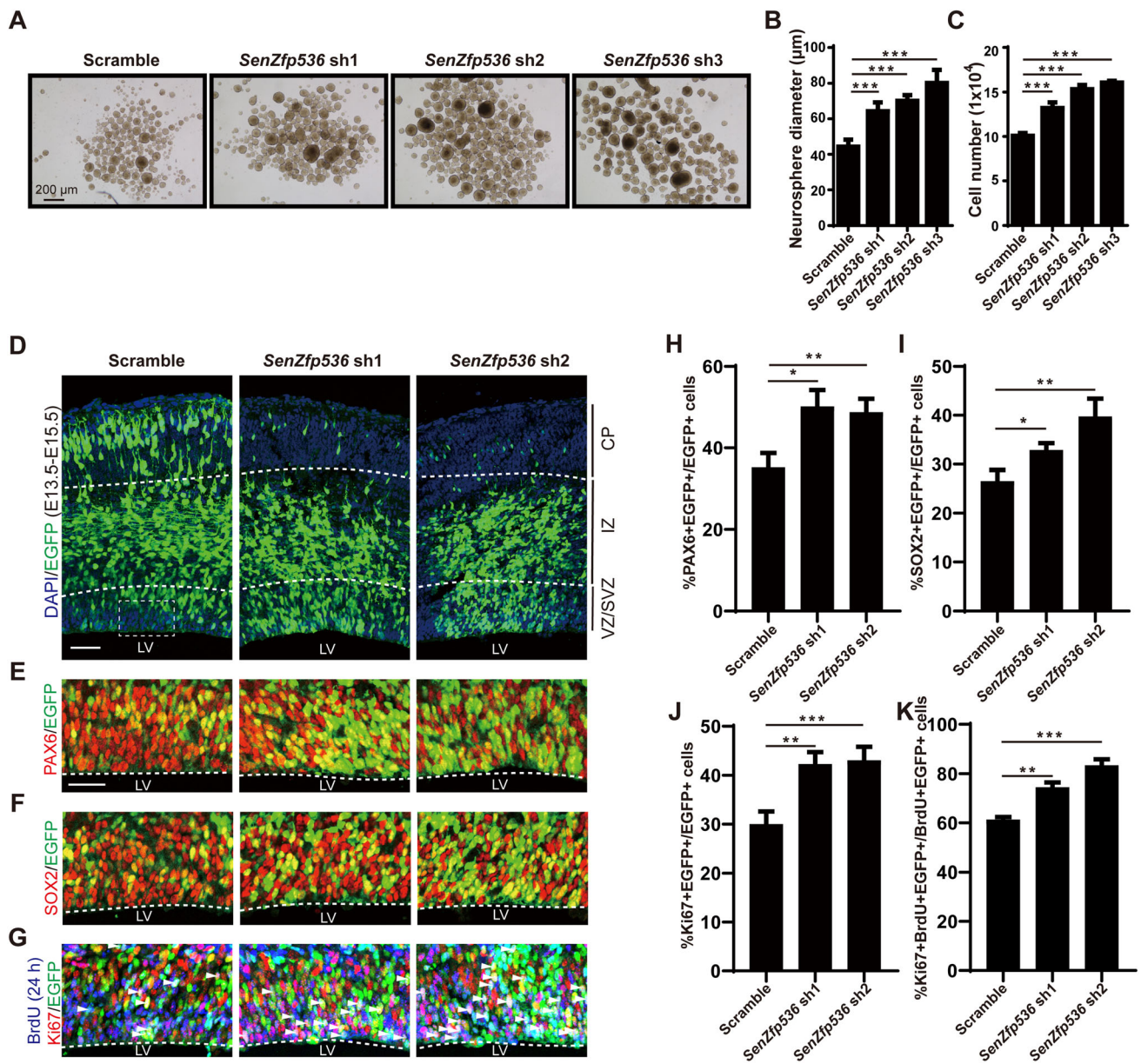


Fig. 6 *SenZfp536* regulates the proliferation of cortical NPCs. **A** Representative images showing E12.5 cortices cultured under sphere conditions and transduced with the indicated shRNAs for 7 days. **B** Bar plots of the average diameter of neurospheres as in **A**. $F = 43.86$, $P < 0.0001$ (one-way ANOVA followed by *Dunnnett's* multiple comparison test). **C** Bar plots of the cell numbers of neurospheres as in **A**. $F = 112.4$, $P < 0.0001$ (one-way ANOVA followed by *Dunnnett's* multiple comparison test). **D–K** E13.5 mouse cortices were electroporated with a mix of shRNA-expressing and EGFP-expressing vectors; embryos were sacrificed at E15.5 for immunofluorescent analyses. Spatial distributions of transduced cells (**D**). VZ/SVZ immunofluorescent images and quantification of

PAX6+ (**E**, **H**), SOX2+ (**F**, **I**), and Ki67+ (**G**, **J**) transduced cells (EGFP+) are displayed. Co-immunostaining and quantification for 24-h BrdU (E14.5–E15.5), Ki67, and EGFP reveal that *SenZfp536* knockdown leads to more cells staying in the cell cycle as measured by the percentages of BrdU+/Ki67+/EGFP+ among BrdU+/EGFP+ cells (**G**, **K**). For **H**, $F = 30.30$, $P < 0.001$; for **I**, $F = 34.05$, $P < 0.001$; for **J**, $F = 103.9$, $P < 0.0001$; for **K**, $F = 123.6$, $P < 0.0001$ (one-way ANOVA followed by *Dunnnett's* multiple comparison test). Embryos in each experiment: scrambled, $n = 4$; *SenZfp536* sh1, $n = 4$; *SenZfp536* sh2, $n = 3$. Scale bars, 50 μm . * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, not significant. Results are presented as mean \pm SEM.

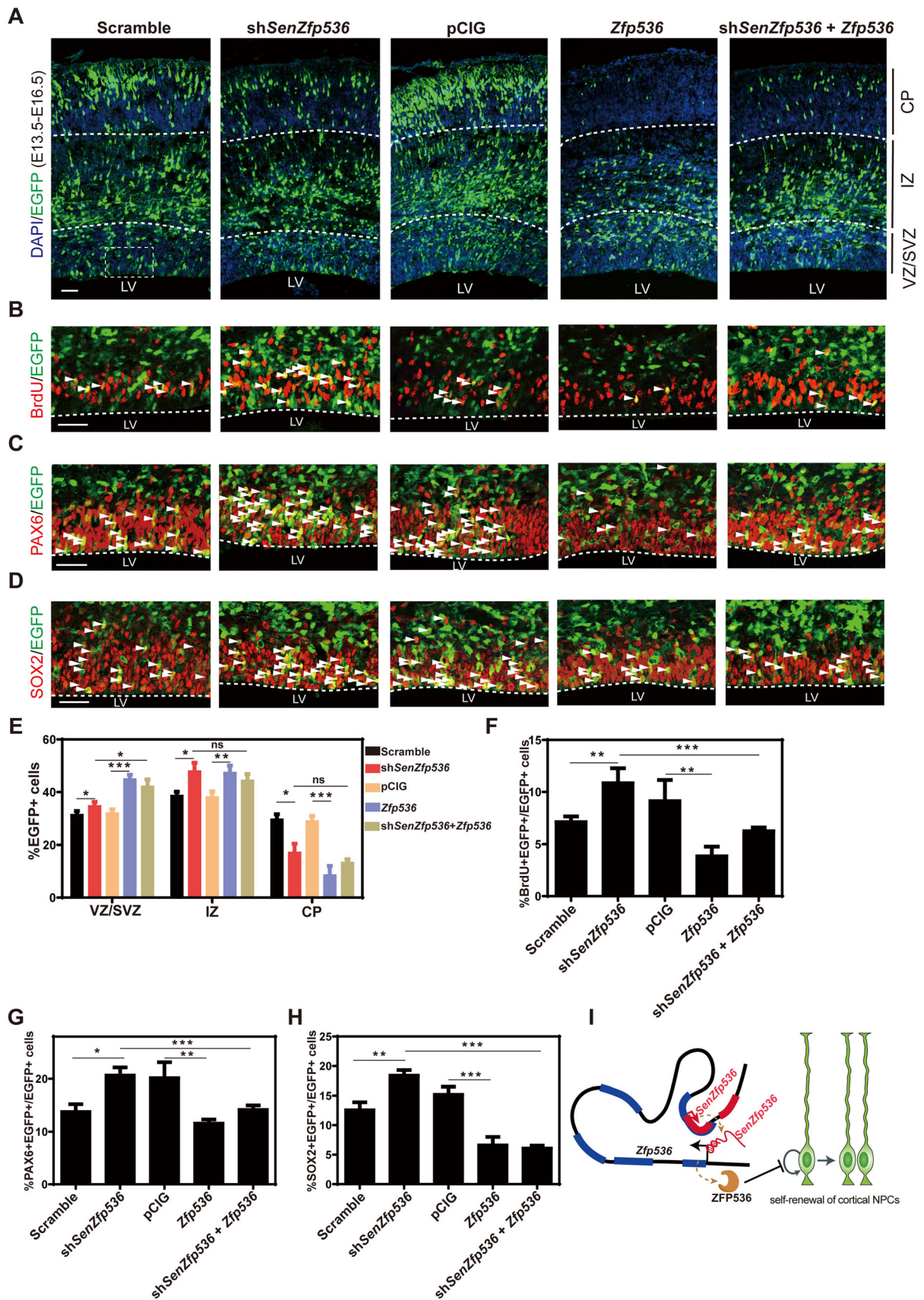


Fig. 7 *Zfp536* expression rescues the enhanced proliferation of cortical NPCs caused by *SenZfp536* knockdown. **A–H** E13.5 mouse cortices were electroporated with indicated vectors and transduced cells were labeled with EGFP; embryos were sacrificed at E16.5 for immunofluorescent analyses. **A, E** Spatial distribution of transduced cells. VZ/SVZ immunofluorescent images and quantification for BrdU (**B, F**), PAX6 (**C, G**), and SOX2 (**D, H**) are displayed. White arrows, co-labeled cells. For VZ/SVZ in **E**, $F = 36.22$, $P < 0.0001$; for IZ in **E**, $F = 30.23$, $P < 0.0001$; for CP in **E**, $F = 46.32$, $P < 0.0001$. For **F**, $F = 33.35$, $P < 0.0001$; for **G**, $F = 38.23$, $P < 0.0001$. for **H**, $F = 44.24$, $P < 0.0001$ (one-way ANOVA followed by Tukey's multiple comparison test). Samples per experiment: Scrambled, $n = 6$; sh*SenZfp536* $n = 4$; pCIG, $n = 7$; *Zfp536*, $n = 3$; sh*SenZfp536* + *Zfp536*, $n = 5$. Scale bars, 50 μm . * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, not significant. Results are presented as mean \pm SEM. **I** Model showing that *SenZfp536* negatively regulates the self-renewal of cortical NPCs by *cis*-activating *Zfp536*.

Discussion

Although many lncRNAs are expressed during neural development, very few have been reported to play regulatory roles. In general, the expression levels of lncRNAs are much lower than mRNAs, which argues that lncRNAs could be a result of transcriptional noise. However, lncRNAs with relatively high expression and/or displaying spatiotemporal distributions are more likely to have a functional capacity. Moreover, current knowledge of epigenetic mechanisms, including lncRNA-mediated transcriptional regulation, in regulating cell-fate choices during cortical development is limited. In the present study, we integrated in-house and public transcriptome data of developing mouse neocortex to characterize lncRNAs with potential functional roles in regulating the self-renewal of cortical NPCs. In fact, aRGs, the predominant cortical NPCs in developing mouse neocortex, specifically express the most lncRNAs. We further found that *SenZfp536*, a lncRNA overlapping the 3' of *Zfp536* in the sense direction, positively regulates the expression of *Zfp536* and inhibits the self-renewal of cortical NPCs both *in vitro* and *in vivo*. *SenZfp536* facilitates DNA looping and the association between the *Zfp536* promoter and the genomic region that transcribes *SenZfp536*.

Recent progress in single-cell transcriptome studies identified lncRNAs involved in brain development and homeostasis [72, 73]. However, due to the current technical bottleneck, low-expression transcripts cannot be efficiently resolved. Thus, by integrating temporal and spatial transcriptome data from the developing mouse cortex, our approach managed to characterize both coding and non-coding transcripts implicated in the self-renewal of cortical NPCs. Moreover, ribosomal-depleted and strand-specific sequencing of the dorsal forebrain at E10.5 and E12.5

greatly supplemented the lncRNA annotation, as many non-polyadenylated transcripts were also identified.

Although many lncRNAs overlap with mRNAs in the sense direction, very few have been functionally dissected. Sense lncRNAs are commonly regarded as non-coding transcripts transcribed from alternative TSSs of PCGs, and are functionally related to their host PCGs, but the molecular mechanisms underlying the actions of lncRNAs are largely unknown. *SenZfp536* is a sense lncRNA that is transcribed from an independent TSS; this is supported by Northern blotting, 5' RACE experiments, and enriched Pol II and H3K4me3 around the *SenZfp536* TSS. *SenZfp536* is polyadenylated and largely resides in the nucleus.

Ablation of *SenZfp536* downregulated the expression of *Zfp536*, whereas transcriptional activation of *SenZfp536* elevated the *Zfp536* level, and the expression levels of *SenZfp536* and *Zfp563* were correlated in the developing cortex and in manipulated Neuro-2a cells. However, episomal expression of *SenZfp536* had no effect on *Zfp536* expression. In addition, both reporter assay and RNA tethering experiments indicated that the transcript of *SenZfp536* has transactivating activity. All these data support the conclusion that *SenZfp536* maintains *Zfp536* expression *in cis* but not *in trans* and rule out possible off-target effects by shRNAs or CRISPRa. Using 3C assays, we further revealed that the genomic region that transcribes *SenZfp536* interacts with the promoter of *Zfp536* to form an enhancer-promoter loop structure, which is facilitated by *SenZfp536*. CTCF and cohesion might work with *SenZfp536* to support the loop formation [74]. Moreover, the locus that transcribes *SenZfp536* could be an enhancer for *Zfp536*, a scenario similar to the role of the lncRNA *CCAT1-L* in controlling *MYC* expression [75].

Both *SenZfp536* and *Zfp536* are highly expressed in VZ/ SVZ, the germinal zone of the developing cerebral cortex, with almost identical patterns. Surprisingly, they suppressed but did not promote the self-renewal of cortical NPCs both *in vitro* and *in vivo*, probably affecting NPC cell-cycle length. A previous study demonstrated that ZFP536 negatively modulates the neural differentiation of P19 cells by participating in retinoic acid receptor-mediated gene regulation [76]. In the current study, however, we largely used primary cortical NPCs and developing cortex rather than the P19 carcinoma cells in Qin's study [76]; they more faithfully reflect the physiological function of ZFP536, namely, limiting cortical NPC self-renewal without affecting cortical layer specification.

Together, we found large amounts of lncRNAs are expressed during cortical development. Furthermore, we illustrated the *cis*-regulatory role of the sense lncRNA *SenZfp536* to control self-renewal of cortical NPCs. Given that numerous sense lncRNAs are expressed in radial glial

NPCs, this could be a general theme for balancing the self-renewal and differentiation of cortical NPCs.

Acknowledgements We thank members of the Zhou lab for critical reading of the manuscript. This work was supported by grants from the National Key R&D Program of China (2018YFA0800700), the National Natural Science Foundation of China (31970770, 31970676, and 31671418), the Natural Science Foundation of Hubei Province, China (2018CFA016), Fundamental Research Funds for the Central Universities, the Medical Science Advancement Program (Basic Medical Sciences) of Wuhan University (TFJC2018005) and State Key Laboratory Special Fund 2060204.

Conflict of interest All authors claim that there are no conflicts of interest.

References

- Rakic P. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* 2009, 10: 724–735.
- Greig LC, Woodworth MB, Galazo MJ, Padmanabhan H, Macklis JD. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat Rev Neurosci* 2013, 14: 755.
- Noctor SC, Martinez-Cerdeno V, Ivic L, Kriegstein AR. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci* 2004, 7: 136–144.
- Haubensak W, Attardo A, Denk W, Huttner WB. Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis. *Proc Natl Acad Sci U S A* 2004, 101: 3196–3201.
- Caviness VS, Jr., Sidman RL. Time of origin or corresponding cell classes in the cerebral cortex of normal and reeler mutant mice: an autoradiographic analysis. *J Comp Neurol* 1973, 148: 141–151.
- Rakic P. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. *Science* 1974, 183: 425–427.
- Liu X, Zheng J, Qi S, Shen Q. NONO regulates cortical neuronal migration and postnatal neuronal maturation. *Neurosci Bull* 2019, 35: 1097–1101.
- Kageyama R, Ohtsuka T, Hatakeyama J, Ohsawa R. Roles of bHLH genes in neural stem cell differentiation. *Exp Cell Res* 2005, 306: 343–348.
- Guillemot F. Spatial and temporal specification of neural fates by transcription factor codes. *Development* 2007, 134: 3771–3780.
- Imayoshi I, Kageyama R. bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. *Neuron* 2014, 82: 9–23.
- Cadwell CR, Bhaduri A, Mostajo-Radji MA, Keefe MG, Nowakowski TJ. Development and arealization of the cerebral cortex. *Neuron* 2019, 103: 980–1004.
- Tzortzopoulos A, Thomaidou D, Gaitanou M, Matsas R, Skoulakis E. Expression of mammalian BM88/CEND1 in *Drosophila* affects nervous system development by interfering with precursor cell formation. *Neurosci Bull* 2019, 35: 979–995.
- Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* 2016, 17: 47–62.
- Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*, 2012, 81: 145–166.
- Yang YW, Flynn RA, Chen Y, Qu K, Wan BB, Wang KC, *et al.* Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife* 2014, 3.
- Ramos AD, Diaz A, Nellore A, Delgado RN, Park KY, Gonzales-Roybal G, *et al.* Integration of genome-wide approaches identifies lncRNAs of adult neural stem cells and their progeny in vivo. *Cell Stem Cell* 2013, 12: 616–628.
- Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 2011, 472: 120.
- Li L, Chang HY. Physiological roles of long noncoding RNAs: insight from knockout mice. *Trends Cell Biol* 2014, 24: 594–602.
- Wu T, Chen C, Yang L, Zhang M, Zhang X, Jia J, *et al.* Distinct lncRNA expression profiles in the prefrontal cortex of SD rats after exposure to methylphenidate. *Biomed Pharmacother* 2015, 70: 239–247.
- Yan P, Luo S, Lu JY, Shen X. Cis- and trans-acting lncRNAs in pluripotency and reprogramming. *Curr Opin Genet Dev* 2017, 46: 170–178.
- Lipovich L, Tarca AL, Cai J, Jia H, Chugani HT, Sterner KN, *et al.* Developmental changes in the transcriptome of human cerebral cortex tissue: long noncoding RNA transcripts. *Cerebral Cortex* 2013, 24: 1451–1459.
- Wang A, Wang J, Liu Y, Zhou Y. Mechanisms of long non-coding RNAs in the assembly and plasticity of neural circuitry. *Front Neural Circuits* 2017, 11: 76.
- Li W, Shen W, Zhang B, Tian K, Li Y, Mu L, *et al.* Long non-coding RNA LncKdm2b regulates cortical neuronal differentiation by cis-activating Kdm2b. *Protein Cell* 2020, 11: 161–186.
- Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, Čech M, *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018, 46: W537–W544.
- Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 2015, 12: 357.
- Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015, 33: 290.
- Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2013, 30: 923–930.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009, 26: 139–140.
- Kolde R. pheatmap: Pretty Heatmaps. 2015.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000, 28: 27–30.
- Benson DA, Boguski MS, Lipman DJ, Ostell J, Ouellette BF. GenBank. *Nucleic Acids Res* 1998, 26: 1–7.
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, *et al.* The COG database: an updated version includes eukaryotes. *BMC Bioinf* 2003, 4: 41.
- Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, *et al.* The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 2003, 31: 365–370.
- Kang YJ, Yang DC, Kong L, Hou M, Meng YQ, Wei L, *et al.* CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res* 2017, 45: W12–W16.
- Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Res* 2013, 41: e74–e74.

36. Langfelder P, Horvath S. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf* 9: 559. 2009.
37. Heng X, Guo Q, Leung AW, Li JYH 2017 Analogous mechanism regulating formation of neocortical basal radial glia and cerebellar *Bergmann glia*. *eLife* 6: e23253.
38. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinf* 2018, 19: 534.
39. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, 4: 44.
40. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2008, 37: 1–13.
41. Ge S, Jung D. ShinyGO: a graphical enrichment tool for animals and plants. *bioRxiv* 2018: 315150.
42. Culhane AC, Schröder MS, Sultana R, Picard SC, Martinelli EN, Kelly C, *et al.* GeneSigDB: a manually curated database and resource for analysis of gene expression signatures. *Nucleic Acids Res* 2011, 40: D1060–D1066.
43. Newman JC, Weiner AM. L2L: a simple tool for discovering the hidden significance in microarray expression data. *Genome Biol* 2005, 6: R81–R81.
44. McLean CY, Bristol D, Hiller M, Clarke SL, Schaar BT, Lowe CB, *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* 2010, 28: 495.
45. Luo Z, Mu L, Zheng Y, Shen W, Li J, Xu L, *et al.* NUMB enhances Notch signaling by repressing ubiquitination of NOTCH1 intracellular domain. *J Mol Cell Biol* 2020, 12: 345–358.
46. Gallagher S, Chakavarti D. Immunoblot Analysis. *JoVE* 2008: e759.
47. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, *et al.* Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 2002, 420: 563–573.
48. Luo S, Lu JY, Liu L, Yin Y, Chen C, Han X, *et al.* Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. *Cell Stem Cell* 2016, 18: 637–652.
49. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 2015, 517: 583–588.
50. Hagège H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, *et al.* Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* 2007, 2: 1722.
51. Xie Z, Moy LY, Sanada K, Zhou Y, Buchman JJ, Tsai LH. Cep120 and TACCs control interkinetic nuclear migration and the neural progenitor pool. *Neuron* 2007, 56: 79–93.
52. Consortium EP, Dunham K, Aldred SF C. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012, 489: 57–74.
53. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucl Acids Res* 2017, 46: D794–D801.
54. Ayoub AE, Oh S, Xie Y, Leng J, Cotney J, Dominguez MH, *et al.* Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc Natl Acad Sci* 2011, 108: 14950–14955.
55. Dillman AA, Hauser DN, Gibbs JR, Nalls MA, McCoy MK, Rudenko IN, *et al.* mRNA expression, splicing and editing in the embryonic and adult mouse cerebral cortex. *Nat Neurosci* 2013, 16: 499–506.
56. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, *et al.* Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotech* 2010, 28: 503–510.
57. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc* 2016, 11: 1650–1667.
58. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990, 215: 403–410.
59. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 2018, 47: D766–D773.
60. Fang S, Zhang L, Guo J, Niu Y, Wu Y, Li H, *et al.* NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res* 2017, 46: D308–D314.
61. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, *et al.* Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011, 25: 1915–1927.
62. Hudson WH, Prokhnjevskaya N, Gensheimer J, Akondy R, McGuire DJ, Ahmed R, *et al.* Expression of novel long noncoding RNAs defines virus-specific effector and memory CD8+ T cells. *Nat Commun* 2019, 10: 196.
63. Marta F, Mareike A, Elena T, Takashi N, Holger B, Eric L, *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* 2015, 347: 1465–1470.
64. de la Torre-Ubieta L, Stein JL, Won H, Opland CK, Liang D, Lu D, *et al.* The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* 2018, 172: 289–304.e218.
65. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, *et al.* The accessible chromatin landscape of the human genome. *Nature* 2012, 489: 75–82.
66. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005, 4: Article17.
67. Forgy EW. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 1965, 21: 768–769.
68. Le Cam L, Neyman J, Scott EL. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability: Held at the Statistical Laboratory, University of California, June 21–July 18, 1970. Univ of California Press, 1972.
69. Lloyd S. Least squares quantization in PCM. *IEEE Trans Inf Theory* 1982, 28: 129–137.
70. Richardson JE, Bult CJ. Visual annotation display (VLAD): a tool for finding functional themes in lists of genes. *Mamm Genome* 2015, 26: 567–573.
71. Bult CJ, Eppig JT, Kadin JA, Richardson JE, Blake JA. The Mouse Genome Database (MGD): mouse biology and model systems. *Nucleic Acids Res* 2008, 36: D724–D728.
72. Liu SJ, Nowakowski TJ, Pollen AA, Lui JH, Horlbeck MA, Attenello FJ, *et al.* Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol* 2016, 17: 67.
73. Luo YP, Coskun V, Liang AB, Yu JH, Cheng LM, Ge WH, *et al.* Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell* 2015, 161: 1175–1186.
74. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet* 2014, 15: 234.
75. Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res* 2014, 24: 513–531.
76. Qin Z, Ren F, Xu X, Ren Y, Li H, Wang Y, *et al.* ZNF536, a novel zinc finger protein specifically expressed in the brain, negatively regulates neuron differentiation by repressing retinoic acid-induced gene transcription. *Mol Cell Biol* 2009, 29: 3633–3643.