


TCF20 dysfunction leads to cortical neurogenesis defects and autistic-like behaviors in mice

Chao Feng^{1,2,3}, Jinyue Zhao^{1,2}, Fen Ji^{1,2}, Libo Su^{1,2}, Yihui Chen^{4,*} & Jianwei Jiao^{1,2,5,**} 

Abstract

Recently, *de novo* mutations of transcription factor 20 (TCF20) were found in patients with autism by large-scale exome sequencing. However, how TCF20 modulates brain development and whether its dysfunction causes ASD remain unclear. Here, we show that TCF20 deficits impair neurogenesis in mouse. TCF20 deletion significantly reduces the number of neurons, which leads to abnormal brain functions. Furthermore, transcriptome analysis and ChIP-qPCR reveal that the DNA demethylation factor TDG is a downstream target gene of TCF20. As a nonspecific DNA demethylation factor, TDG potentially affects many genes. Combined TDG ChIP-seq and GO analysis of TCF20 RNA-Seq identifies T-cell factor 4 (TCF-4) as a common target. TDG controls the DNA methylation level in the promoter area of TCF-4, affecting TCF-4 expression and modulating neural differentiation. Overexpression of TDG or TCF-4 rescues the deficient neurogenesis of TCF20 knockdown brains. Together, our data reveal that TCF20 is essential for neurogenesis and we suggest that defects in neurogenesis caused by TCF20 loss are associated with ASD.

Keywords autism spectrum disorder; DNA demethylation; neurogenesis; TCF20; TDG

Subject Categories Development; Molecular Biology of Disease; Neuroscience

DOI 10.15252/embr.201949239 | Received 8 September 2019 | Revised 11 May 2020 | Accepted 14 May 2020 | Published online 8 June 2020

EMBO Reports (2020) 21: e49239

Introduction

During cerebral cortex development, neuron generation from neural stem cells and progenitor cells is precisely controlled. Neural stem cells located in the ventricle zone and subventricle zone use symmetrical and asymmetrical division to self-renew and generate neuroblast cells. After terminal mitosis, newborn neurons migrate to the appropriate regions following the inside-out rule, at which point they form functional neural circuits [1–3]. Precise control of

neurogenesis is indispensable for normal brain development and function. Abnormal regulation of neurogenesis usually leads to neurodevelopmental disorders comprising ASD, ADHD, and intellectual disability [4].

Autism spectrum disorders describe a series of neurodevelopmental disorders with deficits in social communication and interaction, as well as stereotypical repetitive behavior patterns [5]. Because they are neurodevelopmental disorders, ASDs are lifelong diseases that have a huge impact on not only patients' lives but also their families and society. Since being first reported by Kanner in 1943, the worldwide prevalence of autism has significantly increased from 1/2,500 to 1/100, and it is even higher in some countries, such as 1/86 in the UK and 1/59 in the United States [6–8]. Due to the current global state of ASD, many efforts have been made to investigate its pathogenetic mechanism. Currently, genetic factors are regarded as the most important risk factors for the development of ASDs. For parents of a child with autism, the next child has a 20–50 times higher risk of suffering ASD than the next child of parents who have a child without autism [7,9–11]. However, the causes of autism are still unclear. Studies of autism candidate genes are urgently needed.

Recently, many gene mutations were identified as high-risk factors in the pathogenesis of neurodevelopmental disorders through large-scale exome sequencing technology [12,13]. Among these novel gene mutations, *de novo* mutations of TCF20 (transcription factor 20) were identified in the sequencing data of autism patients, which might imply that TCF20 dysfunction is related to the development of ASD [14–16]. TCF20, also called stromelysin-1 platelet-derived growth factor (PDGF)-responsive element binding protein, is a transcription factor involved in the regulation of stromelysin-1 transcription [17]. TCF20 was also found to play a role as transcriptional coactivator in modulating the transcriptional activity of Sp1, c-Jun, Est1, and RNF4 [18–20]. TCF20 is located in the chromosome 22q13.2 region, which is close to the Phelan-McDermid syndrome (PMS) candidate gene shank3 [21]. PMS, known as 22q13.3 deletion syndrome, involves a range of phenotypes, including global developmental delay, intellectual disability, neonatal hypotonia, autism, and autistic-like behaviors, which has caused PMS to be considered a syndromic form of ASD [22–24]. TCF20

1 State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

2 University of Chinese Academy of Sciences, Beijing, China

3 Sino-Danish College at University of Chinese Academy of Sciences, Beijing, China

4 Department of Ophthalmology, Yangpu Hospital, Tongji University School of Medicine, Shanghai, China

5 Innovation Academy for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing, China

*Corresponding author. Tel: +86 021 65690520; E-mail: cyh80h@163.com

**Corresponding author. Tel: +86 010 64806335; E-mail: jwjiao@ioz.ac.cn

mutations or microdeletions were also found in PMS patients [21,23,25]. All these studies suggest that TCF20 dysfunction could be related to ASDs.

However, because of a lack of direct evidence, it is still unclear whether TCF20 dysfunction causes ASDs. In addition, the function of TCF20 in the brain development process also remains unclear. To understand the biological function of TCF20 in neurogenesis and the possible relationship between TCF20 mutations and ASD, we generated TCF20 knockout mice and found that the deletion of TCF20 results in neurogenesis defects as well as autistic-like behavioral patterns. Specific depletion of TCF20 in NSCs leads to a decrease in neuronal differentiation and an increase in NPCs proliferation, which might be responsible for abnormal behaviors. In terms of molecular mechanisms, TCF20 regulates the expression of TDG, a DNA demethylation enzyme, by binding to the TDG promoter region and promoting its transcription. Furthermore, TDG reduces the methylation level in the promoter CpG island region of the downstream neural differentiation gene T-cell factor 4 (TCF-4), which enhances its expression. Overexpression of TDG or TCF-4 could rescue the phenotype caused by TCF20 dysfunction. Together, our work reveals that TCF20 is essential for neurogenesis in the developing brain, and we provide insight into the relevance of DNA methylation of neural differentiation genes to the etiology of ASDs.

Results

TCF20 is expressed in the cerebral cortex of mouse embryos

To investigate the function of TCF20 in neurogenesis, we first detected the expression of TCF20 in the brain of E13 and E16 mouse embryos. The immunofluorescence results showed that TCF20 was ubiquitously expressed in cortex. Co-immunostaining with the neural progenitor cell (NPC) marker PAX6, NESTIN, and SOX2 in isolated neural progenitor cells from E13 embryos also revealed the expression of TCF20 in NPCs (Fig 1A and B). Then, we used absolute quantitative RT-PCR combined with Western blotting to detect the expression level during brain development. The results showed that expression increased throughout development until it reached an expression peak at E16, and then, expression slightly decreased (Fig 1C–E).

Next, we generated 2 shRNA plasmids targeting the TCF20 CDS region. First, their knockdown efficiencies were confirmed through Western blotting, RT-qPCR, and immunostaining in isolated NSCs (Fig EV1A–E). Results indicated that TCF20-sh2 had a better efficiency. We next generated an overexpression vector of TCF20 CDS region. Western blotting showed that this overexpression vector had ability to overexpress TCF20 (Fig EV1F and G). Then, TCF20 knockdown plasmids were transferred into NPCs located in the VZ through *in utero* electroporation (IUE) at E13.5. After 3 days, these embryos were harvested and sectioned. We found that two TCF20 knockdown plasmids both resulted in the reduction of GFP⁺ cells proportion in the CP and the increase of GFP⁺ cells proportion in the IZ and VZ/SVZ (Fig 1F and G). The abnormal cell distribution might result from three possibilities: the imbalance of NSCs proliferation and differentiation, migration defects, and cell death. The increase proportion of GFP⁺ cells in VZ/SVZ might be ascribed to enhanced cell proliferation, while the decrease percentage of cells in

CP might come from the cell differentiation defects or migration problems or both. These three possibilities would be separately analyzed in the following experiments.

Then, a rescue experiment was performed through co-IUE TCF20 knockdown vectors with TCF20 overexpression vectors. This result indicated that TCF20 shRNA did not have off-target effect (Fig EV2D and E). The abnormal cell distribution phenotype indeed resulted from TCF20 dysfunction. Because TCF20-sh2 had better knockdown efficiency and a more dramatic phenotype, further explorations of cell proliferation, differentiation, and migration were mainly carried out using the TCF20-sh2 group.

TCF20 knockdown promotes neural progenitor cell proliferation and inhibits differentiation

After we observed the abnormal GFP distribution of TCF20 knockdown mice, 2 h of BrdU (bromodeoxyuridine) labeling experiments was conducted to label proliferating cells. BrdU is an analog of thymidine that can label cells in the S-phase. Immunostaining of BrdU showed an increase in BrdU and GFP double-positive cells (Fig 2A–C). Furthermore, Western blotting of isolated NSCs with TCF20 knockdown also revealed that loss of TCF20 leads to an increase of SOX2 protein level (Fig 2D). Then, immunostaining of PAX6 and TBR2 was performed to investigate whether the cell proportions of radial glia cells (RGs) and intermediate progenitor cells (IPs) were disturbed after the loss of TCF20. We found the proportions of GFP⁺ PAX6⁺ cells and GFP⁺ TBR2⁺ cells both had different degrees of increase when TCF20 was depleted. Meanwhile, the percentage of GFP and proliferation marker SOX2 double-positive cells was also increased in TCF20 knockdown group (Fig 2E–I). The results suggested that loss of TCF20 enhances NPC proliferation. Since we already found that TCF20 knockdown affects cell proliferation, we next wondered what the effect of loss of TCF20 was on the cell cycle? Therefore, we performed a cell cycle exit experiment by co-immunostaining with Ki67 and BrdU. BrdU was injected into pregnant mice at E15, 24 h before sacrificing mother mice. Costaining of BrdU and Ki67 helped us to detect the percent of cells exiting cell cycles in the following 24 h after BrdU administration. The knockdown of TCF20 reduced the proportion of GFP⁺ BrdU⁺ Ki67⁻ cells among the GFP⁺ BrdU⁺ cells population (Fig EV2A–C). This result indicated that the loss of TCF20 results in more cells remaining in the cell cycle rather than undergoing differentiation.

Given that the knockdown of TCF20 leads to a reduction in GFP⁺ cells in the CP, a neuronal birthdate experiment was performed to investigate whether the neural differentiation was changed after TCF20 knockdown. IUE was performed with TCF20 knockdown plasmids in E13 embryonic brains. After 24 h, BrdU was injected into mother mice. All electroporated brains were collected at P0 for analysis of terminal mitosis. Staining with a BrdU antibody, we found that the percentage of GFP and BrdU double-positive cells in the CP was dramatically decreased. The data demonstrated that many NPCs could not proceed terminal mitosis and differentiate into neurons on time, which is consistent with the cell cycle exit. Furthermore, co-immunostaining was performed for SATB2, CTIP2, and TBR1. A significant reduction in GFP-positive cells that colocalizing with SATB2, CTIP2, or TBR1 was observed, which means that neural differentiation was disturbed after TCF20

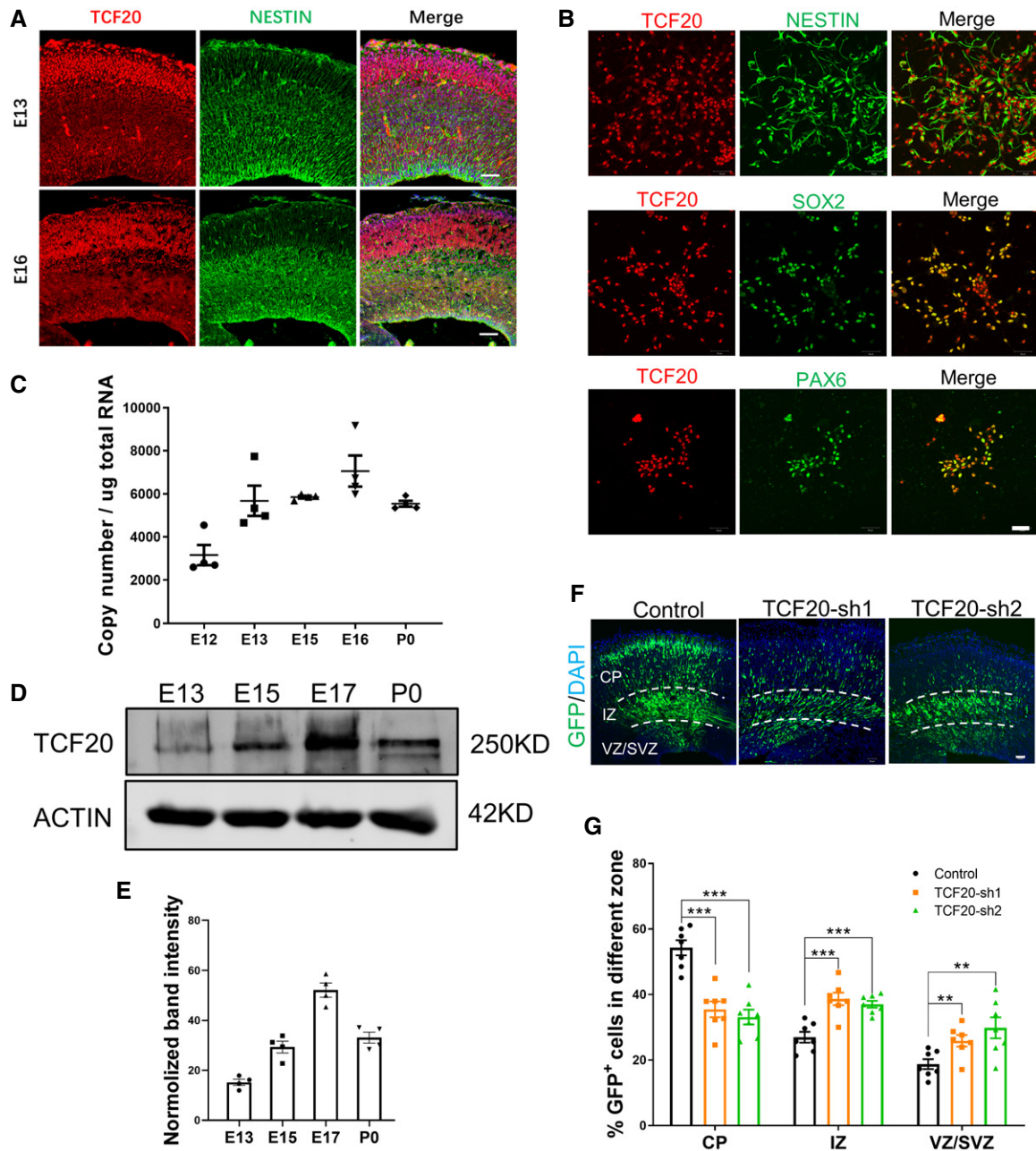


Figure 1. TCF20 is expressed in the developing cerebral cortex and neural stem cells.

A Immunostaining of TCF20, NESTIN, and merged with DAPI in E13 and E16 embryonic brain sections. TCF20 is widely expressed in the brain.
 B Costaining with the NSC markers NESTIN, SOX2, and PAX6 reveals that TCF20 is expressed in NSCs.
 C Absolute quantitative PCR of TCF20 in different developmental stages from E12 to P0. $n = 4$ independently repeated tests.
 D Western blotting results show the changes in TCF20 protein expression level during brain development.
 E Statistical analysis of normalized band intensity in the Western blot results. $n = 4$ samples for each group.
 F In utero electroporation of TCF20 shRNAs in the E13 cerebral cortex led to abnormal cell distribution. These electroporated embryos were collected at E16.
 G Statistical analysis of the distributions of GFP-positive cells in different brain regions. The percentage of GFP-positive cells in the CP (cortical plate), IZ (intermediate zone), and VZ/SVZ (ventricle zone and subventricle zone) was analyzed. $n = 7$ brains.

Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired t-tests were used to analyze the data, n.s. (no significant difference), $P < 0.01$ (**), $P < 0.001$ (***). Scale bar in (A) e13, 50 μ m, e16 100 μ m; (B) 50 μ m; (F) 50 μ m. Source data are available online for this figure.

knockdown (Fig 3D–H). Co-immunostaining with CUX1, a marker of upper layer neurons, also revealed that NPCs did not differentiate into neurons as a result of the loss of TCF20 (Fig EV2F and G).

Taken together, these data showed that TCF20 knockdown in the cerebral cortex promotes cell proliferation but inhibits cell differentiation. Here, we thought the proliferation and differentiation

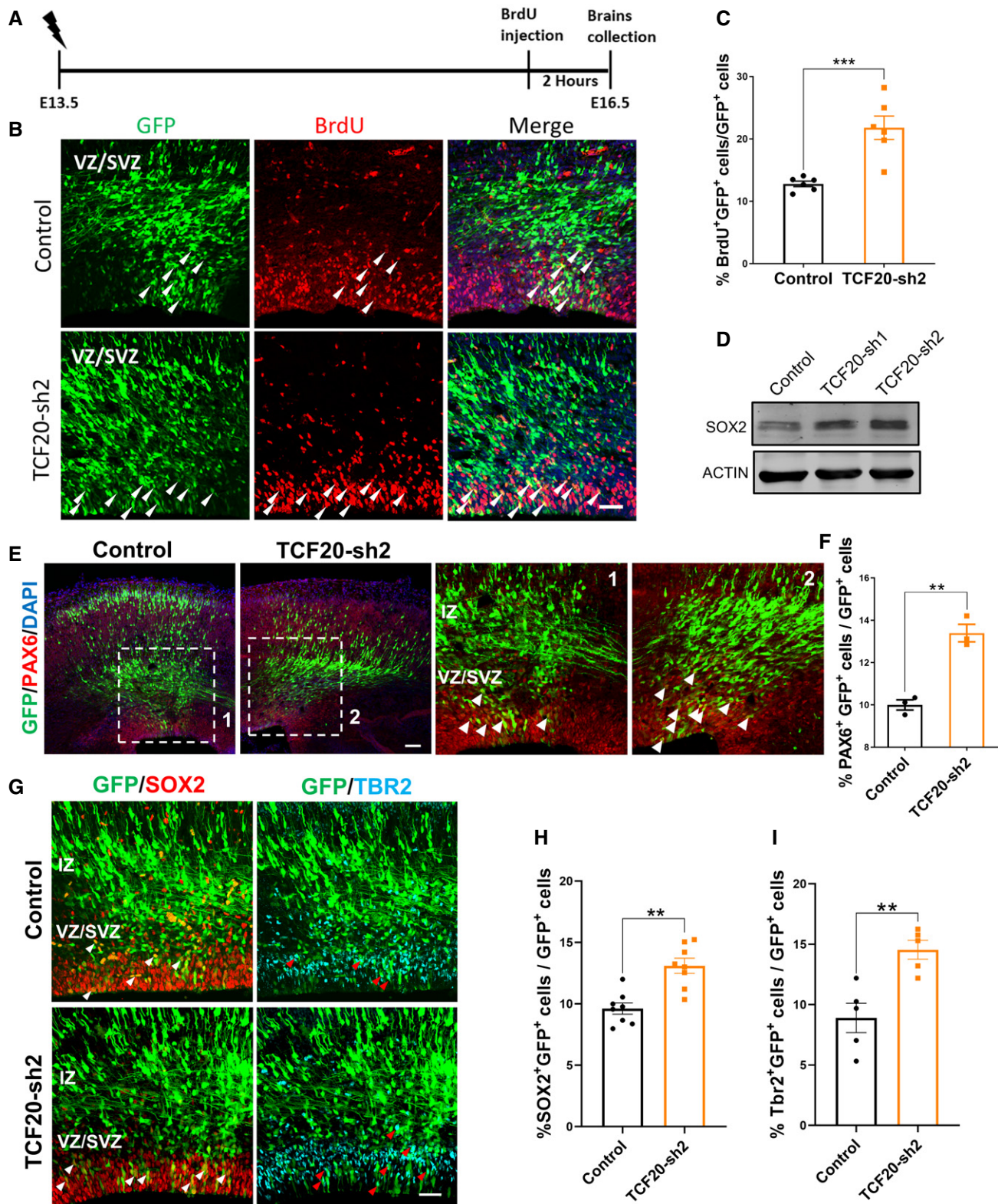


Figure 2.

Figure 2. TCF20 deficits result in abnormal cell proliferation.

- A A diagram of the BrdU labeling time point in electroporation experiments.
 B Immunostaining of BrdU. White arrows represent BrdU GFP double-positive cells. Scale bar: 50 μ m.
 C Statistical analysis of 2-h BrdU labeling experiments. $n = 6$ brains.
 D Western blotting analysis indicated an increase in the proliferation marker SOX2 in TCF20 depleted NSCs.
 E Immunostaining with PAX6 in E16 electroporated embryo brains. White arrows represent PAX6, GFP double-positive cells. 1,2 are 2x enlarged as the dashed boxes showed. Scale bar: 50 μ m.
 F Statistical analysis of the percentage of PAX6⁺GFP⁺ cells among the total GFP⁺ cells. $n = 3$ brains for each group.
 G Immunostaining with SOX2 and PAX6 in E16 electroporated embryo brains. White arrows represent GFP, SOX2 double-positive cells; red arrows represent GFP, TBR2 double-positive cells. Scale bar: 50 μ m.
 H Statistical analysis of the percentage of SOX2⁺GFP⁺ cells among the total GFP⁺ cells. $n = 8$ brains for each group.
 I Statistical analysis of the percentage of TBR2⁺GFP⁺ cells among the total GFP⁺ cells. $n = 5$ brains for each group.
- Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired t-tests were used to analyze the data, n.s. (no significant difference), $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***)
 Source data are available online for this figure.

dysregulation might be the main factor resulting in GFP⁺ cell abnormal distribution. The other two possibilities were analyzed in further experiments.

TCF20 knockout impairs neurogenesis in the embryo cerebral cortex

TCF20 knockout mice were generated to further investigate the biological function of TCF20 during brain development. We designed 2 guide RNAs targeting the 2 ends of exon 2 of TCF20, and this strategy would result in the deletion of the majority of the TCF20 CDS area. Approximately 5kbp of genomic DNA was removed through CRISPR technology (Fig 4A). Western blotting and RT-qPCR results both confirmed that TCF20 was successfully deleted in TCF20 KO mice. In addition, immunostaining of brain slices confirmed this result (Fig 4B–F).

We first detected several proliferation and differentiation makers in the TCF20 knockout mice. The differentiation markers SATB2 and NeuN were decreased with dosage dependence of the loss of TCF20, while the proliferation markers PAX6, SOX2, and PCNA had opposite results along with the loss of TCF20 (Fig 4G).

Second, different layer markers CUX1, SATB2, CTIP2, and TBR1 were used to detect the differentiation of E16 TCF20 KO mice embryos. The layer thickness or cell numbers in different layers were decreased in TCF20 KO and HET mice (Figs 4H–K and EV3A and B). Third, *in utero* electroporation of E13–E16 TCF20 KO embryos with GFP plasmids, and the mice showed that TCF20 dysfunction resulted in abnormal cell distribution, which was similar to what was observed in the TCF20 knockdown experiments. After TCF20 deletion, the proportion of GFP⁺ cell located in the CP was significantly decreased, while GFP⁺ cells in the VZ/SVZ and IZ were increased, which indicated that TCF20 deletion promoted NPCs proliferation and inhibited differentiation (Fig 4L and M). And IUE experiment of GFP plasmids in TCF20 KO mice from E15 to P0 showed that most cells with GFP labeling could normally migrate into upper layer, which indicated that cell migration ability might not be affected in the late period after TCF20 deletion. However, because the differentiation and migration were strongly linked, it was hard to clearly distinguish them in our research system. Thus, we thought migration deficits might also be involved in the whole brain development process (Fig EV3C). TUNEL staining of E16 embryo cerebral cortex slices showed no obvious difference

between WT and TCF20 KO mice. This result implied that TCF20 deletion did not cause apoptosis (Fig EV3D and E). These results suggested that TCF20 is essential for the balance between neural differentiation and proliferation. The cell cycle exiting result in TCF20 KO mice was similar to that of TCF20 knockdown experiments. Many NPCs still stayed in the cell cycle after TCF20 deletion (Fig EV3H–J). Then, we detected changes of microglia marker IBA1, and no significant differences were found after TCF20 knock out (Fig EV3K and L).

The dysfunction of TCF20 leads to autistic-like behaviors in adult and pup mice

Since the deletion of TCF20 disturbed neurogenesis in the mouse embryo cortex, we wondered whether TCF20 dysfunction would have phenotypes in mice similar to those in humans. Unfortunately, our homozygous TCF20 knockout (KO) mice could not survive to adulthood. During development, the homozygous TCF20 KO mice showed severe global development delay, decreased body weight and brain size from embryo to P14. Over 80% of homozygous TCF20 KO mice died before P14, so we used heterozygous TCF20 KO mice instead to conduct the adult behavioral test (Fig EV4A–I). Besides, heterozygous TCF20 KO mice are closer to the ASD patients in genotype. First, we performed an ultrasonic vocalization test on wild-type (WT), heterozygous TCF20 KO (HET), and homozygous TCF20 KO (KO) pups. Compared with the WT pups, the call numbers and total call durations were significantly reduced in HET and KO pups. However, for the single call duration, only the KO pup group showed a significant decrease. Although many HET pups also had shorter call durations, there was no significant difference. It seemed that KO pups had a worse ability to communication than HET and WT pups (Fig 5A–D). These results suggested that not only KO pups but also HET pups show autism-like behaviors.

Because most of the TCF20 KO mice could not survive to adulthood, we chose HET mice for the adult behavioral test (Fig EV4D). Open field tests showed that TCF20 HET mice did not have movement ability defects. We did not observe a significant change in either total moving distance and time or moving distance in the center area, which means that HET mice exhibited normal movement activity and exploration ability (Fig 5E–H). Then, Y-Maze tests were conducted to measure working memory ability. The results showed that HET mice had a decreased tendency to enter the new

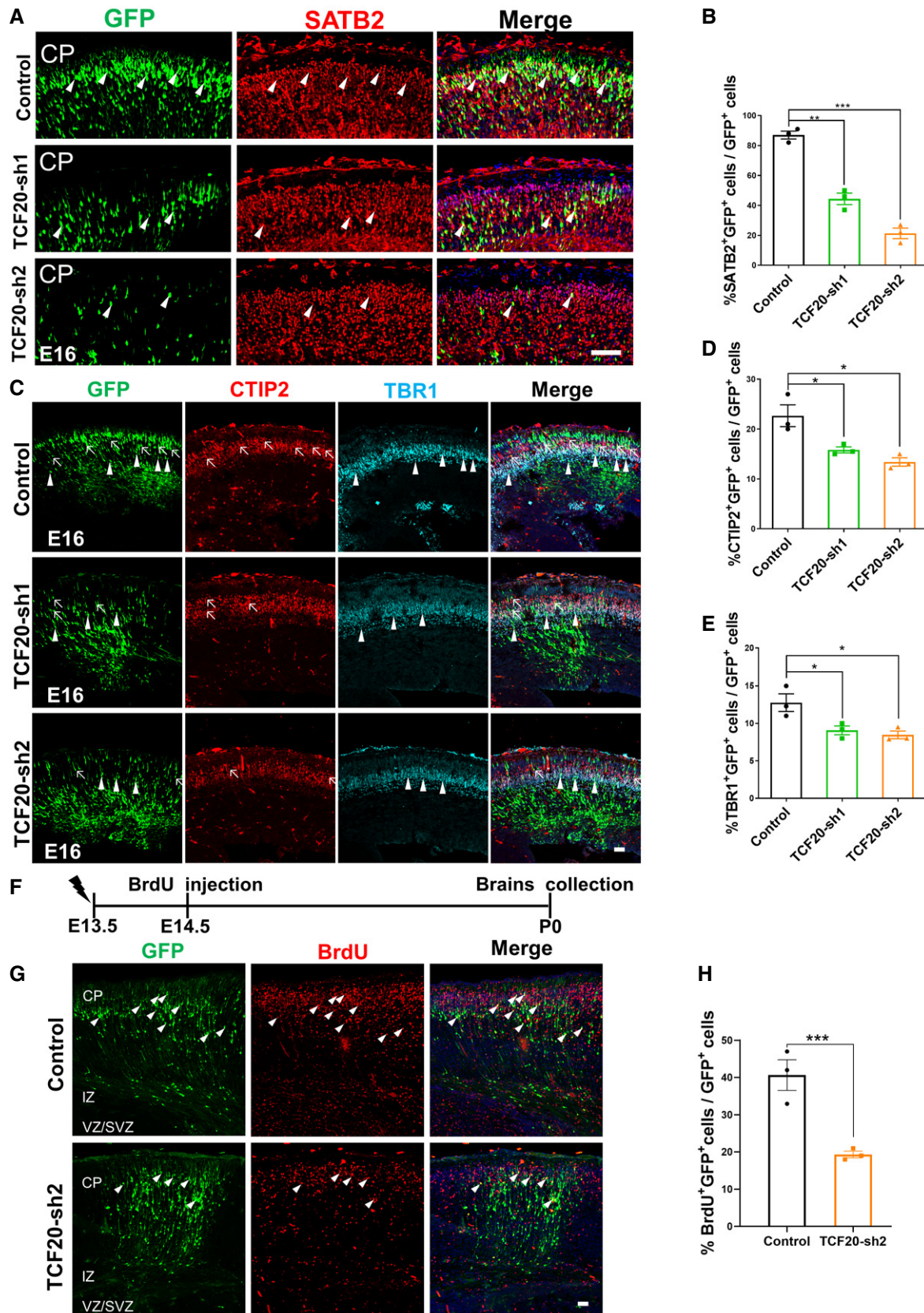


Figure 3.

Figure 3. TCF20 depletion impairs neuronal differentiation.

- A SATB2 staining revealed a reduction in SATB2 and GFP colocalization when TCF20 was depleted. White arrows represent SATB2, GFP double-positive cells.
 B Statistical analysis of the SATB2 GFP double-positive cell ratio in GFP-positive cells. $n = 3$ brains.
 C Staining with CTIP2 and TBR1 antibodies shows TCF20 knockdown decreases the colocalization of CTIP2 and TBR1 with GFP in E13.5–E16.5 brains. White arrows represent CTIP2, GFP double-positive cells. White arrow heads represent TBR1, GFP double-positive cells.
 D, E Statistical analysis of CTIP2, GFP, and TBR1, GFP double-positive cell ratio in GFP-positive cells. $n = 3$ brains.
 F A diagram of the BrdU labeling time point in electroporation experiments.
 G Immunostaining of BrdU. White arrows represent BrdU GFP double-positive cells.
 H Percentage of BrdU and GFP double-positive cells in GFP-positive cells $n = 3$ brains for each group.

Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired *t*-tests were used to analyze the data, n.s. (no significant difference), $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***). Scale bar: (C, G) 25 μm ; (A) 50 μm .

arms, and the total time of HET mice remaining in each arm was similar. This indicated that working memory was disturbed after the deletion of TCF20 (Figs 5I and EV4J). Anxiety is a commonly shared phenotype in different ASDs. Next, elevated-plus maze tests were performed to measure the anxiety of WT and HET mice. Compared with WT mice, the moving time of HET mice in the open arms was significantly reduced, which suggested that TCF20 loss resulted in anxiety (Figs 5J and K and EV4K). To confirm the reduction in social interaction ability in adult mice, three-chamber tests were also conducted with WT and HET mice. As expected, the HET mice did not prefer to interact with new novel wild-type mice, which revealed that TCF20 dysfunction could result in social interaction defects or social interest reduction (Fig 5M–Q). In addition to social interaction deficits, rigid repetitive behavior is another typical behavioral symptom in patients with ASD. The marble burying test results showed that HET mice buried more marbles than WT mice, which indicated that HET mice exhibited repetitive behaviors (Figs 5L and EV4L). Short videos of WT and HET activities in home cages also showed HET mice would run in circles (Movies EV1 and EV2). This result clearly showed that HET mice have a rigid repetitive behavioral pattern. Together, TCF20 dysfunction could cause mice to exhibit autistic-like behaviors.

TCF20 regulates neurogenesis through TDG

To further understand of the mechanism behind TCF20 modulating neurogenesis, we performed RNA sequencing to investigate the transcriptome differences in the cerebral cortex between E13.5 TCF20 KO mice and WT mice. Approximately 500 genes were changed

after TCF20 was deleted. GO analysis revealed that the downregulated genes were enriched in biological processes of neuron fate specification, neuron differentiation, negative regulation of cell proliferation, and embryonic morphogenesis. The upregulated genes were mainly enriched in relation to hypersensitivity and inflammatory response (Fig 6A–C). Among the downregulated genes, thymine-DNA glycosylase (TDG) was significantly reduced, which suggests that TDG could be a downstream target of TCF20 (Fig 6D and E). In addition, some studies have shown that TDG is involved in various tissue differentiation and neuron fate specifications [26–28]. Therefore, we first detected the expression pattern of TDG in the developing cerebral cortex. TDG exhibited an expression pattern that was similar to that of TCF20, and it was also coexpressed with NSC markers and the immature neuron marker DCX (Fig EV5A and C).

Then, we generated 2 shRNAs targeting TDG and an overexpression vector (Fig EV5B). As expected, knockdown of TDG in embryo brains caused a cell distribution phenotype that was similar to TCF20 knockdown, which suggested TDG could be a key modulator involved in neurogenesis modulated by TCF20. TDG knockdown increased the cell numbers in the IZ and VZ/SVZ but decreased the cell number in the CP (Fig 6F and G). We found that TCF20 could bind with the TDG promoter within 1,500 bp of the TSS (Fig 6H).

T-cell factor-4 is responsible for TDG controlling neural differentiation

As a DNA demethylation enzyme, TDG could affect expression of many genes. Which gene is responsible for neural differentiation

Figure 4. Deletion of TCF20 impairs neurogenesis.

- A A diagram of the construction of TCF20 knockout mice. Two guide RNAs targeted the ends of TCF20 exon 2. With the help of cas9, TCF20 exon 2 was deleted from the genome.
 B Western blotting analysis of TCF20 in TCF20 WT, HET, KO mice reveals that TCF20 was deleted.
 C Statistical analysis of the normalized band intensity of TCF20. $n = 3$ samples for each genotype.
 D RT–qPCR analysis of TCF20 in TCF20 WT, HET, and KO mice reveals that TCF20 was deleted. $n = 4$ samples.
 E Immunostaining of E16 WT and KO mouse brain slices revealed that TCF20 was successfully knocked out. Scale bar: 50 μm .
 F Statistical analysis of the normalized fluorescence intensity of TCF20. $n = 6$ samples for each genotype.
 G Western blotting analysis of neural proliferation and differentiation markers in WT HET and KO mice showed that TCF20 deficits promote cell proliferation and inhibit cell differentiation.
 H Immunostaining of different cortical layer markers TBR1, CTIP2, and SATB2, White dash line represents the positive area. Scale bar: 50 μm .
 I–K Statistical analysis of different cells layers in TCF20 WT, HET, and KO mice. $n = 7$ brains of each group.
 L TCF20 deletion impairs neurogenesis and results in abnormal cell distribution. Scale bar: 50 μm .
 M Statistical analysis of GFP⁺ distribution in different zones. $n = 3$ brains for each genotype.

Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired *t*-tests were used to analyze the data, n.s. (no significant difference), $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***).

Source data are available online for this figure.

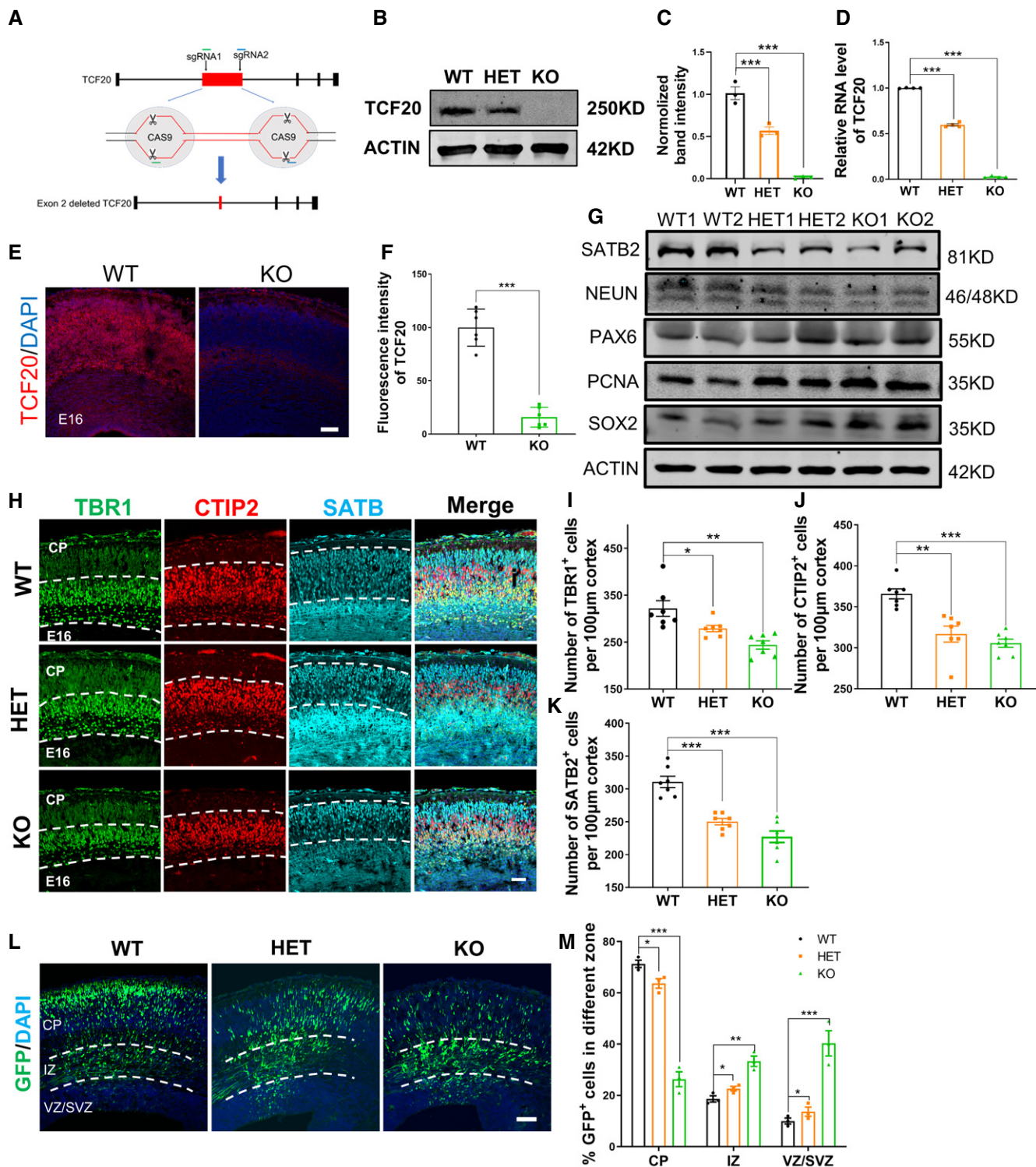


Figure 4.

controlled by TDG? Comparing previously reported TDG ChIP-seq results with our RNA-seq data, we found T-cell factor 4 (TCF-4), which was found in a GO term named neural fate specification, has TDG binding sites in its the promoter region [29–32] (Fig 6I). Besides, Western blotting analysis revealed that TCF-4 expression

level was disturbed in TDG depleted NSCs (Fig EV5E and F). ChIP-qPCR experiments were performed to confirm whether TDG could bind to the TCF-4 promoter area. The results showed that TDG could bind to the region of the TCF-4 promoter and also that there is a CpG island in this promoter (Fig 7A and B). Since CpG islands are

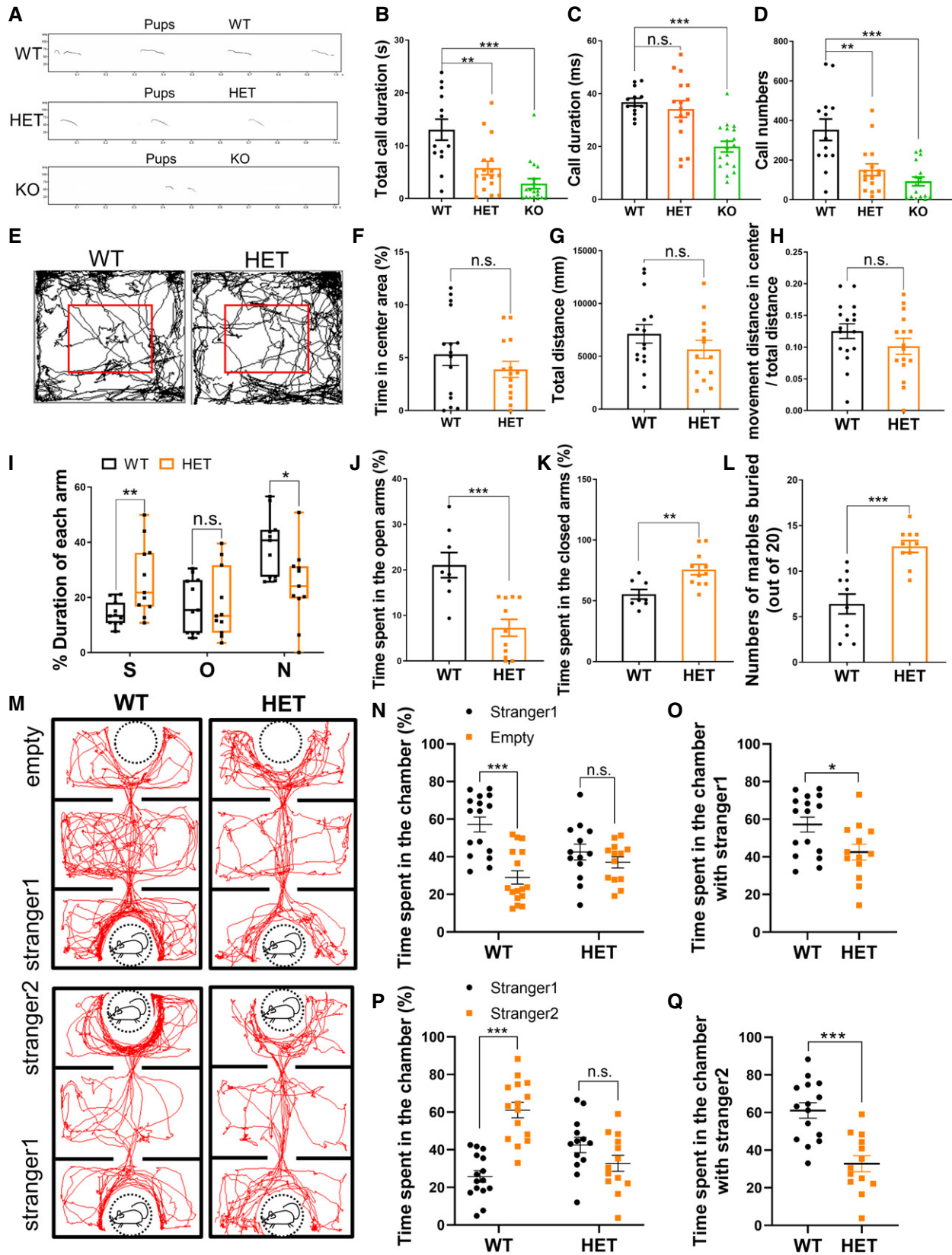


Figure 5.

Figure 5. The dysfunction of TCF20 leads to autistic-like behaviors in adult and pup mice.

- A Diagram of pup ultrasonic vocalization at the frequency level after Fourier transforms
- B–D (B) Total call durations, (C) call durations, and (D) call numbers of WT, HET, and KO mouse pups. nWT = 13 mice, nHET = 16 mice, and nKO = 18 mice.
- E A sketch of the mouse trace in the open field test.
- F–H (F) Statistical analysis of time spent in the center area, (G) statistical analysis of total mouse movement distance in 5 min, (H) the ratio of mice in the center area movement distance to the total distance. *n* for WT = 15 adult mice, *n* for HET = 12 adult mice.
- I Statistical analysis of the Y-Maze test. S, O, and N represent the start arm, old arm, and new arm. The percent of time that mice stayed in each arm was recorded and analyzed; *n* for WT = 11 adult mice, *n* for HET = 11 adult mice. Whiskers represent the minimum to maximum value; central lines represent median values; boxes represent the values between upper quartile and lower quartile.
- J Results of the elevated-plus maze. Time that mice spent in open arms; *n* for WT = 8 adult mice, and *n* for HET = 11 adult mice.
- K Results of elevated-plus maze. Time that mice spent in closed arms; *n* for WT = 8 adult mice, and *n* for HET = 11 adult mice.
- L Numbers of marbles buried in the marble burying test; *n* for WT = 10 adult mice, and *n* for HET = 10 adult mice.
- M Diagram of the three-chamber social interaction test and mouse movement trace. HET mice did not show preference or stay longer in the stranger1 or stranger2 chambers than the WT mice did in the two test phases.
- N Statistical analysis of the percentage of time spent in each chamber with one empty chamber and another chamber containing stranger1; *n* for WT = 16 adult mice, *n* for HET = 13 adult mice.
- O Comparison between the time that WT and HET mice spent in stranger1 chamber. *n* for WT = 16 adult mice, *n* for HET = 13 adult mice.
- P Time spent in each chamber when one contains familiar mouse stranger1, as well as new mouse stranger2. *n* for WT = 16 adult mice, *n* for HET = 13 adult mice.
- Q Comparison between the time of WT and HET mice staying in the stranger2 chamber. *n* for WT = 16 adult mice, *n* for HET = 13 adult mice.
- Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired *t*-tests were used to analyze the data, n.s. (no significant difference), $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***)

the main active regions in relation to DNA methylation, it is reasonable that TDG mainly binds this region. Next, we detected the 4 DNA methylation products in this CpG island through MeDIP-qPCR. We used four antibodies to separately enrich the DNA fragments containing 5mC, 5hmC, 5fC, and 5caC, and then, RT-qPCR was conducted to quantify the amount of each methylation product. We found that after TCF20 deletion, 5mC, 5fC, and 5caC accumulated but 5hmC was decreased in this CpG island (Fig 7C–F). Immunostaining of 5fC and 5caC in TCF20 knockdown NSCs also revealed an increase in 5fC and 5caC (Fig 7G–J). As reported previously, 5mC, 5fC, and 5caC had a repression function on transcription, so the accumulation of 5mC, 5fC, and 5caC in the TCF-4 promoter resulted in a reduction of its expression and further impaired neural differentiation. Next, overexpression of TDG or TCF-4 rescued the phenotype caused by the knockdown of TCF20, respectively (Fig 7K and L). Taken together, TCF20 could modulate TDG expression, which could further affect the expression of the neural differentiation gene TCF-4.

Discussion

Neurogenesis is under the precise temporal and spatial control of many transcription factors and epigenetic factors [2,3]. Abnormal neurogenesis often results in neurodevelopmental disorders. As the most common neurodevelopmental disorders, ASDs have a large population prevalence of 1%, which means that 70 million patients suffer from an ASD [6]. This places a great burden on families and society. However, the pathogenesis of ASD is poorly understood.

Recently, exome sequencing studies reported that *de novo* mutations of TCF20 exist in patients with ASD and other neurodevelopmental disorders [12–16,25,33]. Since TCF20 mutations were recently found in these patients, there is no existing report to explain the biological function of TCF20 in neurogenesis and the possible mechanism of TCF20 involvement in ASDs.

Here, we first report that TCF20 deletion impairs neurogenesis in the embryonic cerebral cortex. TCF20 dysfunction leads to autistic-

like behaviors in mice. Loss of TCF20 promotes NPCs proliferation. In TCF20 knockdown experiments, TCF20 deficits cause abnormal cell distribution. The decreased proportion of GFP⁺ cells in CP comes from cell differentiation defects or migration defects or both. Western blotting and RNA sequencing results indicate that TCF20 deletion disturbs neural differentiation. The E15-P0 IUE experiment in KO mice shows that majority of cells could migrate into CP layer. This result implies migration ability may not be affected in the late period of neurogenesis. However, there may still exist a migration delay that also contributes this phenotype. The GO analysis of RNA sequencing data shows that there are several neuron migration-related genes slightly affected after TCF20 dysfunction. Due to the closely connection between cell fate specification and migration, we also cannot clearly distinguish these two possibilities [34]. Here, we can only conclude that the differentiation defect is one reason that leads to our phenotype. There still need to do more works to understand the phenotype caused by TCF20 dysfunction.

In TCF20 KO mice, markers of different layers, CUX1, SATB2, CTIP2, and TBR1, were significantly decreased. In addition, these knockout mice have global developmental delay with small brain size, which might result from the abnormal cell cycle. The reduction of cell cycle exiting leads to a reduction in neuronal output and these reasons cause small brain size, which is similar to a previous report of *insm1* knockout mice [34]. Besides, there are several reports showed that some patients with TCF20 mutations have microcephaly, which is constant to our mouse model [21,35]. We further performed RNA-seq to explore the possible target of TCF20. TDG is significantly reduced after TCF20 deletion in E13.5 embryo brains. TDG, also called thymine-DNA glycosylase, is responsible for converting 5fC and 5caC to normal cytosine, which is an important part of DNA demethylation and BER repair [28,29,36]. DNA demethylation is very important in neurogenesis. Many DNA demethylation related genes, such as *Mecp2* and *TET1-3*, are involved in neurogenesis and the pathogenesis of neurodevelopmental disorders [37–40]. As a key modulator of the DNA demethylation pathway modulated by TCF20, TDG likely participates in the development of ASD. Furthermore, TDG is also found to be

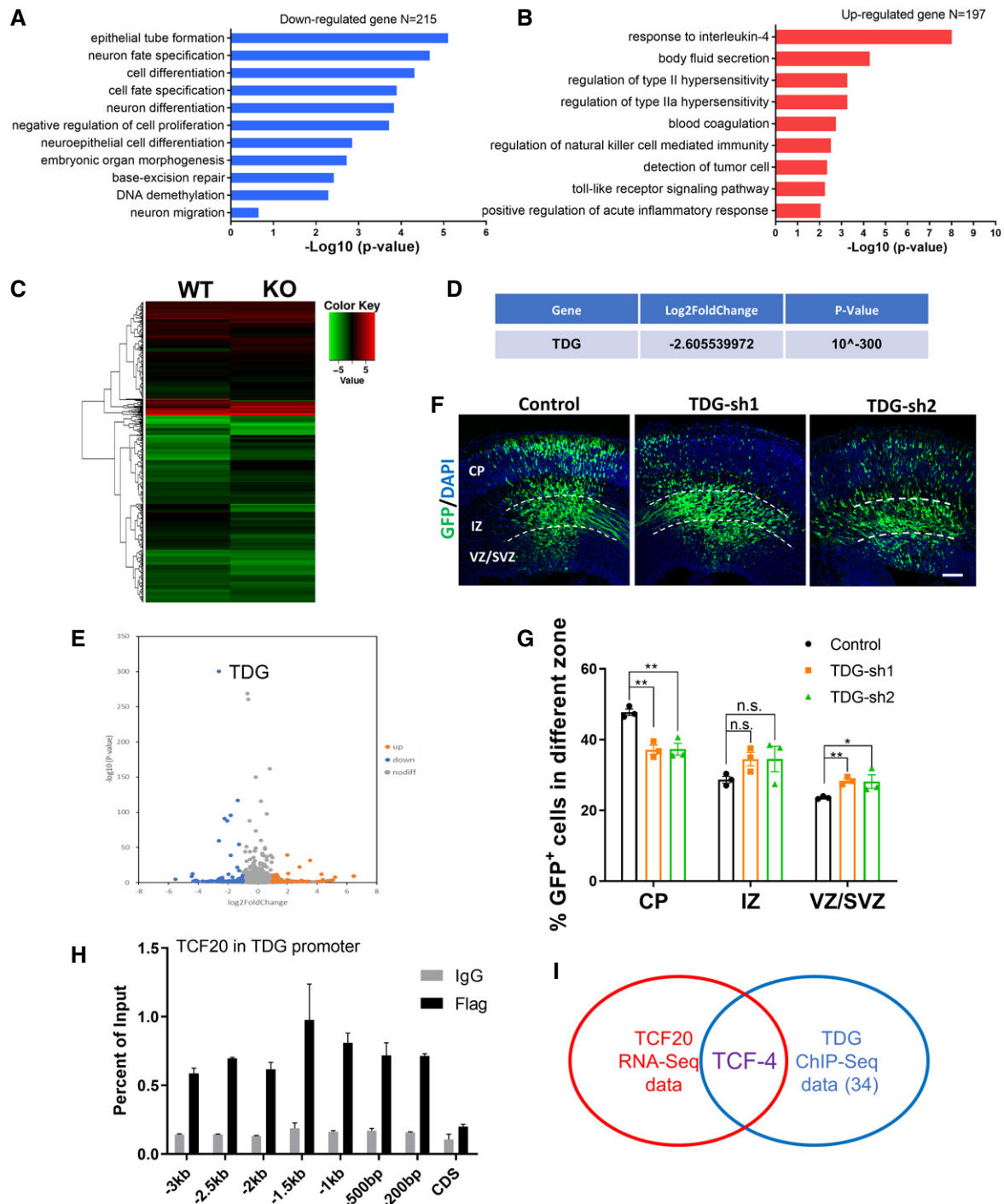


Figure 6. TCF20 regulates neurogenesis through TDG.

A GO analysis of 215 significantly downregulated genes (Log2 fold change < -1, -P value < 0.05).

B GO analysis of 197 significantly upregulated genes (Log2 fold change > 1, -P value < 0.05).

C Heat map of significantly different genes.

D TDG is significantly decreased after TCF20 deletion.

E Volcano map showing that TDG is significantly decreased.

F, G Electroporation of GFP at E13.5 to E16.5 in TCF20 KO mice impairs cell distribution. *n* = 3 brains. Scale bar 25 μ m.

H ChIP-qPCR of TCF20 binding to the TDG promoter. *n* = 4 biological repeats.

I Combining RNA-seq and TDG ChIP-seq data, TCF-4 is a downstream target gene.

Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired *t*-tests were used to analyze the data, n.s. (no significant difference), *P* < 0.05 (*), *P* < 0.01 (**).

important for neurogenesis and neural differentiation [26,27,41]. However, as a DNA demethylase, it could affect the expression of thousands of genes. Therefore, we combined TCF20 RNA-seq and TDG ChIP-seq data from a previous study [29], and we find that TCF-4 could be the gene affecting of TCF20 modulation of neurogenesis. TCF-4 clusters with the GO terms neural fate specification and neural differentiation. In addition, TDG could bind to CpG islands in the TCF-4 promoter area, as confirmed by ChIP-qPCR. T-cell factor-4 (TCF-4), also known as TCF7 l2, is an ASD-related gene. Some

works have reported that TCF-4 is essential for neural differentiation. TCF-4 knockdown or deletion inhibits neural differentiation [30–32,42], which is a phenotype that is similar to loss of TCF20 and its downstream effector TDG. Overexpression of TCF-4 could rescue the TCF20 knockdown phenotype, which suggests that TCF-4 functions downstream of TCF20 in modulating neural differentiation.

At the behavioral level, as we expected, mice with TCF20 deleted indeed had autistic-like behaviors. Most Phelan-McDermid

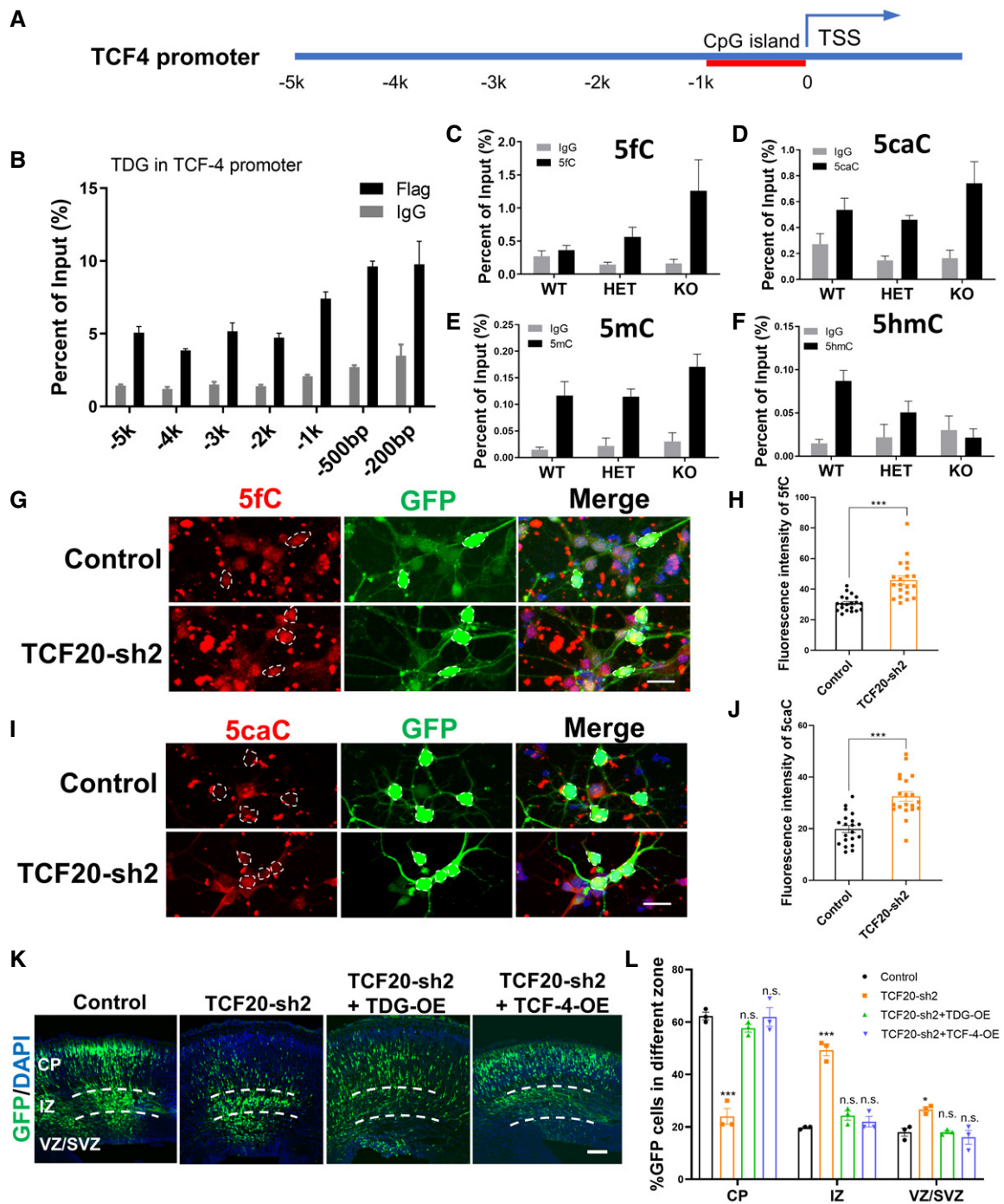


Figure 7.

Figure 7. T-cell factor-4 is responsible for TDG controlling neural differentiation.

- A Diagram of the TCF-4 promoter region.
- B TDG ChIP-qPCR revealed that TDG mainly binds to TCF-4 CpG islands. Seven regions (from -5 kbp to -200 bp) of the promoter were analyzed. -500 bp and -200 bp regions were located in the CpG island of the TCF-4 promoter ($n = 4$ biological repeats).
- C-F MeDIP-qPCR showed that 5fC, 5caC, and 5mC are increased in the TCF-4 promoter, while 5hmC is decreased when TCF20 is insufficient. The MeDIP assay measures the region -500 bp from the TCF-4 promoter ($n = 4$ biological repeats).
- G NSCs were isolated from C57B6/J E13.5 embryonic brains and then were infected with a lentivirus containing a TCF20 shRNA or with a control plasmids. Immunostaining of 5fC shows that TCF20 depletion increased 5fC levels. White dashed lines show the cells infected with lentivirus. Scale bar 25 μ m.
- H Statistical analysis of 5fC fluorescence level. $n = 20$ cells, and $P < 0.0001$.
- I Immunostaining of 5caC shows that TCF20 depletion increased 5caC level. White dashed lines show the cells infected with lentivirus. Scale bar 25 μ m.
- J Statistical analysis of the 5caC fluorescence level. $n = 20$ cells, and $P < 0.0001$.
- K, L Overexpression of TDG and TCF-4 rescued the TCF20 knockdown phenotype in brains; $n = 3$ brains. Scale bar 50 μ m.
- Data information: Bars and error bars represent the means \pm SEM. Two-tailed unpaired t -tests were used to analyze the data, n.s. (no significant difference), $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***)

syndrome patients have autism and autistic-like behaviors as well as some other phenotypes, such as global development delay, intellectual disability, and neonatal hypotonia [22,24]. In mice with TCF20 deletion, there are many phenotypes that are similar to those observed in patients with PMS. TCF20 knockout mice have global developmental delays. Their body size and organ size are significantly decreased compared with those of wild-type mice. In addition, behavioral tests also reveal that TCF20 knockout mice have social communication and interaction defects, as well as repetitive rigid behavioral patterns which are typical ASD phenotypes. In the meanwhile, a clinical report showed PMS patients have a microdeletion in TCF20 [21]. These results suggest that TCF20 is a strong candidate gene or having a role in PMS or ASD.

Taken together, we reveal that TCF20 is essential for neurogenesis during embryonic brain development. TCF20 affected the DNA methylation level in the neural differentiation-related gene TCF-4 promoter region by regulating TDG expression and further affect TCF-4 expression level. TCF20 dysfunction leads to deficits in neurogenesis, which further results in the development of ASD. Of note, this work investigates the role of TCF20 in neurogenesis for the first time and provides insight into the mechanisms of TCF20 dysfunction and ASD, which is useful for developing treatments for patients with ASD or PMS and studying the pathogenesis of ASD.

Materials and Methods

Animals

All the ICR mice and C57BL/6 mice were obtained from Beijing Vital River Laboratory Animal Technology Co., Ltd. The mice were raised under 12-h/12-h day-night cycle (7 a.m. to 7 p.m.). Besides, animal experiments conformed to the Guide for the Care and Use of Laboratory Animals of Chinese academy of sciences. All the animal experiments were also permitted by animal committee of Chinese academy of sciences.

Construction of TCF20 knockout mice

CRISPR-Cas9 technology was used to generate TCF20 knockout mice. Two guide RNAs, which separately target both ends of exon2 of TCF20 gene, were cloned to pUC57-kan-T7-gRNA vector. The

sequences of these 2 gRNAs are upstream-gRNA: GATGAGCTGTGCACCTCCTGTGG, downstream-gRNA: AAGACTGTGGTGGAGGTCCTCGG. And the gRNA vectors and Cas9 vector were injected into fertilized eggs. Then, these treated zygotes were transferred into surrogate mother mice.

Plasmids constructs

Short hairpin RNA sequences are TCF20-sh1: GCTGAAGAGAAAGAGAACGAT, TCF20-sh2: CCACGAAATGTCAGTGGTTAT. TDG-sh1: CCAAGACTCTTCTG-ACATTT, TDG-sh2: GCCACGAATAGCGGTGTTTAA. These shRNA templates were cloned into pSicoR-GFP vector. And TCF20 and TDG over expression were generated by PCR and cloned into pCDH-CMV-GFP vector.

Knockdown and overexpression efficiency verification

Knockdown efficiency experiment was performed in Neuro 2a cell line. 2 μ g of shRNA and control vectors was transfected to Neuro 2a cell lines. After 3 days, these cells were collected and analyzed through Western blotting with TCF20 antibody. Overexpression efficiency was performed in HEK293 cell line. 2 μ g of overexpression vector and control vectors was transfected to HEK293 cell lines. After 3 days, these cells were collected and analyzed through Western blotting with anti-FLAG antibody.

Genotyping of TCF20 mice

Genotyping primers were TCF20-55391F: GCTTGATCTTTGTGTCAAGCAGA, TCF20-61253R: ACCAAGCTGGTAAGGGTTTC, which could produce a 445 bp positive band. Because the wild-type band was as big as 5 kb, another pair of primers was used to get negative PCR products which located in the 5 kb area.

Cell culture

Neuro 2a or HEK 293FT cell line was cultured with DMEM that contained 10% FBS, 100 U/ml penicillin, and 0.1 mg/ml streptomycin. Neural stem cells were isolated from cerebral cortex of E13 mouse embryos. These cells were cultured with proliferation medium (DMEM/F12, Neurobasal, Glutamax, B27 supplement, FGF, EGF, penicillin/streptomycin) or differentiation medium (low-glucose DMEM, B27 supplement, 1% FBS, penicillin/streptomycin).

In Utero Electroporation

After anesthetized by pentobarbital sodium solution, the pregnant mice were attached on a surgery plate and opened the abdomen. Then, the uterine horns were gently taken out. Recombinant plasmids mixed with tiny fast green solution (5 mg/ml) were injected to fetal mouse brain ventricle through glass capillaries. Then, these plasmids were electroporated into neural stem cells, which located in mouse brain ventricle zone through electroporator (Manual BTX ECM830). The electroporation condition was total 5 pulses. Each pulse was 36V with 50 ms length and a 950-ms interval between every two pulses. After 3 days or other proper time, the pregnant mice were sacrificed for further experiments. The embryo brains were fixed in 4% PFA solution at 4°C for 24 h and dehydrated in 30% sucrose for 24 h.

BrdU labeling

Basically, BrdU (100 mg/kg) was injected to pregnant mice 2 or 24 h before sacrifice the mother mice. For cell proliferation analysis, 2 h of BrdU injection was used at E16. And for cell cycle exiting analysis, 24 h of BrdU injection was performed at E15.

Immunostaining

These electroporated brains were cut into 15- μ m slices by using Lecia CM 1950 freezing microtome. Next, the brain slices or cultured cells were fixed in 4% PFA for 30 min. After washed 3 times by 1% PBST (1% Triton X-100 in PBS), these samples were blocked by 5% BSA solution for 1 h. Then, they were incubated with primary antibody at 4°C overnight. And fluorescence labeling secondary antibody was used to visualize target molecules. After incubation with secondary antibody solutions at room temperature for 1 h, these brain sections were washed three times with 1% PBST. And DAPI solution (2 μ g/ μ l; Sigma; D9542) was used for nuclear staining and 50% glycerin was used to seal the slide. Finally, these immunostained slices were scanned on Zeiss LSM 780 confocal microscope and these images were analyzed through ImageJ software. For those fluorescence-intensity-comparing experiments, the slices were scanned at same scanning parameters. After drawing field of interested, the fluorescence intensities of these images were measured through ImageJ mean grayscale measurement function. And for the cell number statistics (e.g., SATB2⁺), only signals with clear cell boundary and co-localized with DAPI were counted. All the groups were analyzed under same standards. All the information of antibodies used for immunostaining was in Table EV2.

Western blotting

Tissue or cells were lysed in RIPA buffer with 10 mM PMSF and proteinase inhibitor cocktail. After centrifuge, the protein extraction mixed with loading buffer was analyzed by SDS-PAGE. Then protein bands were transferred onto NC or PVDF membranes. After 1 h blocking with 5% not fat milk in PBST (0.05% Tween-20 in PBS) at room temperature, the membranes were incubated with primary antibody at 4°C overnight. 800CW Donkey Anti-Rabbit/Mouse IgG and 680LT Donkey Anti-Rabbit/Mouse IgG (LI-COR

Biosciences) were used to visualize the target bands through Odyssey Infrared Imaging System. The band intensities were also detected through this imaging system. The target band intensity was divided by ACTIN band intensity of the same sample to obtain normalized intensity value. Then, we set the control group intensity into 1 and calculate the relative intensity dividing by control intensity. All the information of antibodies used for Western blotting was in Table EV2.

Apoptosis detection assay

Cells under apoptosis were detected through TUNEL assay. *In Situ* Cell Death Detection Kit-TMR Red (Roche, 12156792910) was used in this experiment. All the procedures were performed according to the protocol.

Real-time PCR

The total RNA of tissue or cells was extracted through TRIzol method (Invitrogen). Then, the cDNA was generated by using the Fast Quant RT kit (TIANGEN). RT-qPCR was performed using the Super Real SYBR Green PreMix Plus kit (TIANGEN) on ABI 7500 real-time PCR system. For relative quantification, all the data were analyzed through $2^{-\Delta\Delta C_t}$ methods. Actin was used as reference gene for normalization. For absolute quantification, the TCF20 overexpression plasmids were used as standard substance to generate standard curve. The copy number of each sample was calculated through standard curve with corresponding Ct value. All the information of primers used for RT-qPCR and PCR was in Table EV1.

RNA sequencing

Total RNA was extracted from E13.5 TCF20 knockout and wild-type embryo cerebral cortex through TRIzol method. After quality quantification, the total RNA was converted to cDNA library and analyzed by Illumina HiSeq 2500 platform. The RNA-seq data are available in GEO with accession number GSE135483.

Chromosome immunoprecipitation

ChIP protocols were described previously [43]. After transfected with Flag tagged TCF20 or TDG overexpression vectors and cultured in DMEM medium for 48 h, these N2a cells were crosslinked by incubating with 1% formaldehyde solution (50 mM HEPES-KOH, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 1% formaldehyde) for 15 min at room temperature. And 2.5 M glycine is used to end the reaction. Then, these cells were washed by ice-cold PBS for three times and scraped into 1 ml lysis buffer (10 mM EDTA, 1% SDS, 50 mM pH 8.1 Tris-HCl, and 1% PMSF). The crosslinked chromosome complexes were ultrasonicated into around 300 bp length. The ultrasonic products were incubated with anti-Flag beads and IgG beads (Dynabeads, Invitrogen) at 4°C overnight. The beads were washed six times with buffer W (50 mM HEPES-KOH, 500 mM LiCl, 10 mM EDTA, 1% NP-40, and 0.7% sodium deoxycholate). After incubated in 65°C water bath overnight to dissociate covalent bonds of antibodies, the DNA was purified by DNA purified kit (TIANGEN).

Methylated DNA immunoprecipitation

The total DNA was extracted from E13 embryo mouse cerebral cortex and purified by DNA purified kit (TIANGEN). After 10-min ultrasonic treatment, the DNA was incubated at 4°C overnight with protein G beads which already bind with 5mC, 5hmC, 5fC, and 5caC antibodies. Then, the beads were washed three times by IP buffer and 200 µl digestion buffer with 5 µl proteinase K was added to the beads incubating at 50°C 4 h. Finally, the DNA was purified and further analyzed through RT-qPCR [44].

Behavioral test

All the behavioral tests were performed in a quiet environment at day period from 9:00 am to 5:00 pm. And mice adapted to the test environment 2 h before behavioral tests. All the behavioral experiments were performed in a double-blind manner.

Open field test

Adult male mice were gently transferred into the center of a 40 cm long × 40 cm wide × 40 cm high open filed box and allowed freely exploring the area for 5 min. All the movement of mice was recorded and analyzed by Topscan behavioral analysis software (Clever Sys Inc., Reston, VA, USA). The total distance and moving time or distance in center were analyzed [43].

Elevated-plus maze

Test was conducted following the steps described before [43,45]. Elevated-plus maze contains 2 open arms (40 cm long × 9.5 cm wide) and 2 close arms (40 cm long × 9.5 cm wide with 9.5 cm high wall) located 70 cm above the ground. Mice were placed in the center of Elevated-plus maze. Then, 5-min free exploration was recorded and analyzed by Topscan software.

Three-chamber social interaction test

Three-chamber social interaction test is the common method to assess the social interaction ability [45,46]. The three-chamber box has 3 chambers (each chamber is 40 cm long × 20 cm wide × 20 cm high) with 2 round wire mesh cages. The inner walls are transparent which allow the mice to see 2 terminal chambers when placed in the middle chamber. The experiment could be divided into 3 phases. First, the mice had a 5-min free exploration of the empty three-chamber box as a habituation period. Then, a novel mouse was put into a wire mesh cage and placed in one terminal chamber and another chamber had an empty wire mesh cage. The test mouse was gently put into the middle chamber and allowed to freely explore the three chambers for 10 min. After that, another novel mouse was put into the empty wire mesh cage and the test mouse was allowed to freely explore the three-chamber box for 10 min. All the traces and activities were recorded and analyzed by Topscan software.

Y-Maze

All the experimental procedures were described previously [43]. Y-Maze has 3 arms with a 120° angle between each neighbor arm.

At the beginning, we define an arm as a start arm and randomly choose one arm from the rest two arms as the old arm, another as the new arm with clues on each arm walls. Then, the mice were allowed freely explore the maze with the new arm closed for 5 min. After 15 min, the mice were placed into the start arm again with the new arm open and a 5-min free activity of the mice was recorded and analyzed by Topscan software.

Marble burying test

All procedures were performed as described before [46]. Standard cages and same size marbles were used in this test. Before each trial, 20 marbles were washed with 70% ethanol and dried. Then, these glass balls were placed in 4 rows in a cage with 5 cm high bedding. Place the mice on the corner of the test cage and keep the mice far from the marbles at the beginning and allow the mice to freely act for 30 min, then count the numbers of marble that were buried. We consider a marble was buried when its 2/3 surface was covered by bedding.

Ultrasonic vocalization test of isolation mice pups

The experimental procedures were described previously [46,47]. The p6-8 pups were separately isolated from their mother and littermates for 5 min. Then each pup was gently transferred to a soundproof chamber with a small box. The ultrasonic vocalizations were recorded by Ultra Sound Gate Condenser Microphone CM16 (Avisoft Bioacoustics, Berlin, Germany) and analyzed by Avisoft SASLab Pro (Avisoft Bioacoustics, Berlin, Germany). For the analysis part, a fast Fourier transform was performed to the original sound data with parameters containing 512 FFT length, 100% frame, Hamming window, and 75% time-window overlap. Besides, the spectrums were generated in a 488-Hz frequency resolution and 0.512-ms time resolution. Call above -40 dB and 10 ms was considered to be a vocalization. An amplitude filter of -60 dB with a high-pass filter of 30 kHz was used to reduce the background noise. The total recording time was 5 min.

Statistical analysis

All bars and error bars represent mean value ± SEM. Significance comparisons were performed using unpaired two-tailed *t*-test and one-way ANOVA. Differences were regarded as significant by *P* value, n.s. (no significant difference), **P* < 0.05, ***P* < 0.01, ****P* < 0.001. All the statistical analysis and diagram were performed on GraphPad Prism 8.0.

Data availability

The RNA sequencing data have been uploaded to NCBI's GEO. All the data could be accessed through GEO accession number GSE135483 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135483>).

Expanded View for this article is available online.

Acknowledgements

We would like to thanks all the members in our laboratory for their valuable comments on our project. This work was supported by grants obtained from

the National Key R&D Program of China (2019YFA0110300), the National Science Fund for Distinguished Young Scholars (81825006), the CAS Strategic Priority Research Program (XDA16010301), the National Science Foundation of China (31730033 and 31621004), and K.C. Wong Education Foundation.

Author contributions

CF initiated and performed the experiments, analyzed the data, and wrote the manuscript. JZ provided help in the behavioral tests and knockout mice experiments. FJ and LS helped the knockdown experiments. YC provided suggestions and co-supervised the project. JJ supervised this project and provided the funding support.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Jabaudon D (2017) Fate and freedom in developing neocortical circuits. *Nat Commun* 8: 16042
- Molyneaux BJ, Arlotta P, Menezes JR, Macklis JD (2007) Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* 8: 427–437
- Noctor SC, Martínez-Cerdeño V, Ivic L, Kriegstein AR (2004) Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci* 7: 136–144
- Ernst C (2016) Proliferation and differentiation deficits are a major convergence point for neurodevelopmental disorders. *Trends Neurosci* 39: 290–299
- Frith U, Happe F (2005) Autism spectrum disorder. *Curr Biol* 15: R786–R790
- Lai M-C, Lombardo MV, Baron-Cohen S (2014) Autism. *Lancet* 383: 896–910
- Levy SE, Mandell DS, Schultz RT (2009) Autism. *Lancet* 374: 1627–1638
- Bhat S, Acharya UR, Adeli H, Bairy GM, Adeli A (2014) Autism: cause factors, early diagnosis and therapies. *Rev Neurosci* 25: 841–850
- Brainstorm C, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, Escott-Price V, Falcone GJ, Gormley P et al (2018) Analysis of shared heritability in common disorders of the brain. *Science* 360: eaap8757
- Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J et al (2007) Strong association of *de novo* copy number mutations with autism. *Science* 316: 445–449
- Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, Jian M, Liu G, Greer D, Bhandari A et al (2012) Whole-genome sequencing in autism identifies hot spots for *de novo* germline mutation. *Cell* 151: 1431–1442
- Deciphering Developmental Disorders S (2017) Prevalence and architecture of *de novo* mutations in developmental disorders. *Nature* 542: 433–438
- Lelieveld SH, Reijnders MR, Pfundt R, Yntema HG, Kamsteeg EJ, de Vries P, de Vries BB, Willemsen MH, Kleefstra T, Lohner K et al (2016) Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci* 19: 1194–1196
- Babbs C, Lloyd D, Pagnamenta AT, Twigg SRF, Green J, McGowan SJ, Mirza G, Naples R, Sharma VP, Volpi EV et al (2014) *De novo* and rare inherited mutations implicate the transcriptional coregulator TCF20/SPBP in autism spectrum disorder. *J Med Genet* 51: 737–747
- Schafgen J, Cremer K, Becker J, Wieland T, Zink AM, Kim S, Windheuser IC, Kreiss M, Aretz S, Strom TM et al (2016) *De novo* nonsense and frameshift variants of TCF20 in individuals with intellectual disability and postnatal overgrowth. *Eur J Hum Genet* 24: 1739–1745
- Smeland OB, Frei O, Kauppi K, Hill WD, Li W, Wang Y, Krull F, Bettella F, Eriksen JA, Witoelar A et al (2017) Identification of genetic loci jointly influencing schizophrenia risk and the cognitive traits of verbal-numerical reasoning, reaction time, and general cognitive function. *JAMA Psychiatry* 74: 1065
- Sanz L, Moscat J, Diaz-Meco MT (1995) Molecular characterization of a novel transcription factor that controls stromelysin expression. *Mol Cell Biol* 15: 3164–3170
- Sjottem E, Rekdal C, Svineng G, Johnsen SS, Klenow H, Uglehus RD, Johansen T (2007) The ePHD protein SPBP interacts with TopBP1 and together they co-operate to stimulate Ets1-mediated transcription. *Nucleic Acids Res* 35: 6648–6662
- Lyngso C, Bouteiller G, Damgaard CK, Ryom D, Sanchez-Munoz S, Norby PL, Bonven BJ, Jorgensen P (2000) Interaction between the transcription factor SPBP and the positive cofactor RNF4. An interplay between protein binding zinc fingers. *J Biol Chem* 275: 26144–26149
- Rekdal C, Sjottem E, Johansen T (2000) The nuclear factor SPBP contains different functional domains and stimulates the activity of various transcriptional activators. *J Biol Chem* 275: 40288–40300
- Upadia J, Gonzales PR, Atkinson TP, Schroeder HW, Robin NH, Rudy NL, Mikhail FM (2018) A previously unrecognized 22q13.2 microdeletion syndrome that encompasses TCF20 and TNFRSF13C. *Am J Med Genet A* 176: 2791–2797
- Phelan K, McDermid HE (2011) The 22q13.3 deletion syndrome (Phelan-McDermid syndrome). *Mol Syndromol* 2: 186–201
- Kurtas N, Arrigoni F, Errichiello E, Zucca C, Maghini C, D'Angelo MG, Beri S, Giorda R, Bertuzzo S, Delledonne M et al (2018) Chromothripsis and ring chromosome 22: a paradigm of genomic complexity in the Phelan-McDermid syndrome (22q13 deletion syndrome). *J Med Genet* 55: 269–277
- Phelan MC (2008) Deletion 22q13.3 syndrome. *Orphanet J Rare Dis* 3: 14
- Torti E, Keren B, Palmer EE, Zhu Z, Afenjar A, Anderson IJ, Andrews MV, Atkinson C, Au M, Berry SA et al (2019) Variants in TCF20 in neurodevelopmental disability: description of 27 new patients and review of literature. *Genet Med* 21: 2036–2042
- Wheldon LM, Abakir A, Ferjentsik Z, Dudnakova T, Strohbuecker S, Christie D, Dai N, Guan S, Foster JM, Correa IR Jr et al (2014) Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. *Cell Rep* 7: 1353–1361
- von Meyenn F, Iurlaro M, Habibi E, Liu NQ, Salehzadeh-Yazdi A, Santos F, Petrini E, Milagre I, Yu M, Xie Z et al (2016) Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol Cell* 62: 848–861
- Niederreither K, Harbers M, Chambon P, Dollé P (1998) Expression of T: G mismatch-specific thymidine-DNA glycosylase and DNA methyl transferase genes during development and tumorigenesis. *Oncogene* 17: 1577–1585
- Kolendowski B, Hassan H, Krstic M, Isovich M, Thillainadesan G, Chambers AF, Tuck AB, Torchia J (2018) Genome-wide analysis reveals a role for TDG in estrogen receptor-mediated enhancer RNA transcription and 3-dimensional reorganization. *Epigenetics Chromatin* 11: 5

30. Chodelkova O, Masek J, Korinek V, Kozmik Z, Machon O (2018) Tcf7L2 is essential for neurogenesis in the developing mouse neocortex. *Neural Dev* 13: 8
31. Lee M, Yoon J, Song H, Lee B, Lam DT, Yoon J, Baek K, Clevers H, Jeong Y (2017) Tcf7 l2 plays crucial roles in forebrain development through regulation of thalamic and habenular neuron identity and connectivity. *Dev Biol* 424: 62–76
32. Wang H, Matisse MP (2016) Tcf7 l2/Tcf4 transcriptional repressor function requires HDAC activity in the developing vertebrate CNS. *PLoS ONE* 11: e0163267
33. Doan RN, Lim ET, De Rubeis S, Betancur C, Cutler DJ, Chiacchetti AG, Overman LM, Soucy A, Goetze S, Freitag CM et al (2019) Recessive gene disruptions in autism spectrum disorder. *Nat Genet* 51: 1092–1098
34. Monaghan CE, Nechiporuk T, Jeng S, McWeeney SK, Wang J, Rosenfeld MG, Mandel G (2017) REST corepressors RCOR1 and RCOR2 and the repressor INSM1 regulate the proliferation-differentiation balance in the developing brain. *Proc Natl Acad Sci USA* 114: E406–E415
35. Vetrini F, McKee S, Rosenfeld JA, Suri M, Lewis AM, Nugent KM, Roeder E, Littlejohn RO, Holder S, Zhu W et al (2019) De novo and inherited TCF20 pathogenic variants are associated with intellectual disability, dysmorphic features, hypotonia, and neurological impairments with similarities to Smith-Magenis syndrome. *Genome Med* 11: 12
36. Kohli RM, Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502: 472
37. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, van Iterson M, van Dijk F, van Galen M, Bot J et al (2016) Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 49: 131
38. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD et al (2013) Global epigenomic reconfiguration during mammalian brain development. *Science* 341: 1237905
39. Ciernia AV, LaSalle J (2016) The landscape of DNA methylation amid a perfect storm of autism aetiologies. *Nat Rev Neurosci* 17: 411
40. Tatton-Brown K, Seal S, Ruark E, Harmer J, Ramsay E, del Vecchio Duarte S, Zachariou A, Hanks S, O'Brien E, Aksglaede L et al (2014) Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat Genet* 46: 385
41. Kitsera N, Allgayer J, Parsa E, Geier N, Rossa M, Carell T, Khobta A (2017) Functional impacts of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine at a single hemi-modified CpG dinucleotide in a gene promoter. *Nucleic Acids Res* 45: 11033–11042
42. Forrest MP, Waite AJ, Martin-Rendon E, Blake DJ (2013) Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS ONE* 8: e73169
43. Shen T, Ji F, Wang Y, Lei X, Zhang D, Jiao J (2018) Brain-specific deletion of histone variant H2A.z results in cortical neurogenesis defects and neurodevelopmental disorder. *Nucleic Acids Res* 46: 2290–2307
44. Weng Y-I, Huang THM, Yan PS (2009) Methylated DNA immunoprecipitation and microarray-based analysis: detection of DNA methylation in breast cancer cell lines. In *Molecular endocrinology: methods and protocols*, Park-Sarge O-K, Curry TE (eds), pp 165–176. Totowa, NJ: Humana Press
45. Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, Suyama M, Takumi T, Miyakawa T, Nakayama KI (2016) CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature* 537: 675–679
46. Wang S, Tan N, Zhu X, Yao M, Wang Y, Zhang X, Xu Z (2018) Sh3rf2 haploinsufficiency leads to unilateral neuronal development deficits and autistic-like behaviors in mice. *Cell Rep* 25: 2963–2971 e6
47. Wöhr M, Roullet FI, Hung AY, Sheng M, Crawley JN (2011) Communication impairments in mice lacking Shank1: reduced levels of ultrasonic vocalizations and scent marking behavior. *PLoS ONE* 6: e20631