Neurosci. Bull. December, 2019, 35(6):1045–1057
https://doi.org/10.1007/s12264-019-00400-w

www.neurosci.cn
www.springer.com/12264

ORIGINAL ARTICLE

De Novo and Inherited *SETD1A* Variants in Early-onset Epilepsy

Xiuya Yu[1] · Lin Yang[2,3] · Jin Li[4] · Wanxing Li[1] · Dongzhi Li[5] · Ran Wang[6] · Kai Wu[6] · Wenhao Chen[6] · Yi Zhang[6,7] · Zilong Qiu[8] · Wenhao Zhou[1,3,9]

Received: 21 November 2018 / Accepted: 29 January 2019 / Published online: 13 June 2019

Abstract Early-onset epilepsy is a neurological abnormality in childhood, and it is especially common in the first 2 years after birth. Seizures in early life mostly result from structural or metabolic disorders in the brain, and the genetic causes of idiopathic seizures have been extensively investigated. In this study, we identified four missense mutations in the *SETD1A* gene (SET domain-containing 1A, histone lysine methyltransferase): three *de novo* mutations in three individuals and one inherited mutation in a four-generation family. Whole-exome sequencing indicated that all four of these mutations were responsible for the seizures. Mutations of *SETD1A* have been implicated in schizophrenia and developmental disorders, so we examined the role of the four mutations (R913C, Q269R, G1369R, and R1392H) in neural development. We found that their expression in mouse primary cortical neurons affected excitatory synapse development. Moreover, expression of the R913C mutation also affected the migration of cortical neurons in the mouse brain. We further identified two common genes (Neurl4 and Usp39) affected by mutations of *SETD1A*. These results suggested that the mutations of *SETD1A* play a fundamental role in abnormal synaptic function and the development of neurons, so they may be pathogenic factors for neurodevelopmental disorders.

Keywords Early-onset epilepsy · Whole-exome sequencing · *SETD1A* · Neural development

Introduction

Early-onset epilepsy is a common and distinctive manifestation of neurological abnormalities in childhood, especially in the first 2 years, with an incidence of 70.1 per 100

Xiuya Yu and Lin Yang have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12264-019-00400-w) contains supplementary material, which is available to authorized users.

✉ Yi Zhang
zy@eulertechnology.com

✉ Zilong Qiu
zqiu@ion.ac.cn

✉ Wenhao Zhou
zwhchfu@126.com

[1] Division of Neonatology, Children's Hospital of Fudan University, Shanghai 201102, China

[2] Clinical Genetic Center, Children's Hospital of Fudan University, Shanghai 201102, China

[3] Key Laboratory of Birth Defects, Children's Hospital of Fudan University, Shanghai 201102, China

[4] Institute of Biomedicine Sciences, Fudan University, Shanghai 200032, China

[5] Department of Prenatal Diagnosis, The Women and Children's Medical Center, Guangzhou 510623, China

[6] Euler Genomics, Beijing 102206, China

[7] School of Life Sciences, Peking University, Beijing 100191, China

[8] Institute of Neuroscience, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

[9] Laboratory of Neonatal Diseases, Ministry of Health, Children's Hospital of Fudan University, Shanghai 201102, China

000 [1]. Early-onset epilepsy is clinically defined according to the onset age, etiology, seizure type, frequency, duration, and prognosis. The etiologies that cause early-onset epilepsy mostly include hypoxic ischemic encephalopathy, congenital brain malformations, and metabolic disorders. A few decades ago, many infantile epileptic syndromes were characterized as "idiopathic", until genetic testing became an effective means of investigating the contribution of genes [2]. Although most "idiopathic" seizures are benign and have a good prognosis, such as benign familial neonatal seizures, benign familial neonatal-infantile seizures, benign familial infantile seizures, and some febrile seizures, some malignant epileptic syndromes can have long-lasting effects such as neurodevelopmental deficits, epileptic encephalopathies, or even death, as reported in Ohtahara syndrome, Dravet syndrome, and West syndrome [3]. Due to the technological advances of next-generation sequencing, molecular genetics research on epilepsy has made rapid progress in recent years [4]. Because different types of early-onset epilepsy have a large variety of genetic etiologies, finding the genetic basis of an early-onset epilepsy is of great value in both diagnosis and treatment [5–7].

Many reported genes that cause seizures encode ion channels with relatively clear mechanisms, such as *KCNQ2*, *KCNT1*, and *SCN2A* (sodium voltage-gated channel alpha subunit 2). These genes commonly encode a subunit of the voltage-gated channel and affect important currents in neurons, and mutations in these genes may interfere with neuronal excitability in the immature brain, leading to seizures [8, 9]. Other types of genes that may be responsible for seizures include those encoding proteins associated with synaptic function (dynamin 1 and NECAP endocytosis associated 1) and signal transduction (dedicator of cytokinesis 7 and WW domain containing oxidoreductase), neurotransmitter membrane receptors (gamma-aminobutyric acid type A receptor alpha1 subunit and glutamate ionotropic receptor NMDA type subunit 2B), intracellular transporters (solute carrier family 13 member 5), and enzymes (SZT2, KICSTOR complex subunit) [8].

In this study, we idenidentified four SETD1A variants in four 1-day- to 2-year-old probands with early-onset epilepsy in order to determine the role of variants in interfering brain development.

Materials and Methods

Ethics Approval

This study was approved by the Ethics Committee of the Children's Hospital of Fudan University and the Guangzhou Women and Children Medical Center. Informed consent was given by the parents, who agreed to joining this study and the use of the data for scientific research and publication. The methods used in this study were performed in accordance with the approved guidelines.

Case Presentation

Patient 1: The proband with variant of SETD1A p.R913C was a girl born at full term without complications, but she experienced a paroxysmal seizures characterized by blinking, staring, foaming and twisting at the mouth, and facial cyanosis averaging 2–3 times per day from the second day after birth. Physical examination showed no obvious symptoms except mild jaundice. The video electroencephalogram revealed slight abnormalities including multifocal sharp wave and sharp slow wave discharges predominantly in the right central area and left parietal region while asleep. Serum tandem mass spectroscopy showed that the amino-acid and carnitine levels were normal, while α-ketoglutarate acid was increased in the urine. Assessment of general movements indicated a poor repertoire—monotonous movements without complexity. An MRI of the brain was normal (Fig. S2E–G). The patient was treated with phenobarbital and the seizure frequency was reduced. When she was 15 months old, a follow-up examination revealed that she had normal intellectual level. Her family history was remarkable. The patient's mother was healthy, but her father had a history of seizures after birth. Her grandfather and great-grandmother also had seizures during childhood, but the details were not clear.

Patient 2: A girl with variant of SETD1A p.Q269R was born following an uneventful pregnancy, and first developed symptoms at the age of 2 years. She suffered from poor spirits, a pale complexion, and vomiting, followed by a persistent twitch which moved from the left finger, to the arm, to the left lower extremity and mouth. Seizures were reduced after she was hospitalized, but she appeared to lose consciousness with salivation. No abnormality was observed in the physical examination, and the patient had normal growth and development. No seizure history was reported in her family. An electroencephalogram (EEG) showed paroxysmal sharp waves and sharp slow waves on both sides, especially at the back (Fig. S2A, B). A brain MRI revealed no abnormality (Fig. S2H–J), except for a high signal on T2WI in the right maxillary sinus and mastoid process. A developmental screen test showed that the developmental quotient was 110, and the mental index was 107.

Patient 3: This was a boy with variant of SETD1A p.G1369R; his first seizure happened at the age of 9 months. He experienced shrug and wry-neck without apparent cause lasting 1 s, several times per day. The symptoms occurred frequently upon waking, and less

frequently when in a sedated condition. A physical examination showed development delay. He was born at full term following an uneventful pregnancy. The patient had a previous history of hypoglycemia, hypothyroidism, and growth retardation. No family history of seizures or neurological disease was found. An EEG revealed frequent multifocal sharp waves, spike waves, and spike slow waves, which extended predominantly on the left side during sleep (Fig. S2C, D). An MRI showed cerebral dysplasia, decreases in myelinated white matter, and mild ventriculomegaly (Fig. S2K–M).

Patient 4: This was a male child with the variant SETD1A p.R1392H. His mother was gravida 2 para 2, his gestational age was 36 weeks and birth weight was 2.99 kg. On the first day of life, he began to suffer from severe tonic-clonic seizures. An EEG showed bilateral sharp waves and spike-and-waves. A brain MRI showed a high T1W1 signal in bilateral cerebral occipital areas (Fig. S2N–P). His older brother had neonatal seizures and died at $\sim$ 1 month of age without detailed clinical information. This patient was lost to follow-up, so there was no information for progress or prognosis.

Human Sample Preparation

Blood samples were collected from the participants, and genomic DNA was extracted with a QIAamp DNA Blood Mini kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. This study was approved by the Ethics Committee of the Children's Hospital, Fudan University.

Whole-exome Sequencing and Confirmation

Genomic DNA was sequenced on a HiSeq 2000 sequencer according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Clean reads were aligned to the reference human genome (GRCh37) using the Burrows–Wheeler Aligner (v.0.5.9-r16). The average sequencing depth ranged from 35.73 $\times$ to 38.45 $\times$. The mapping rate was from 97.03% to 97.27%, and the coverage of the genome was from 99.83% to 99.85%. The aligned data was then processed by Freebayes to call variants. A filter of SAR > 1 (number of alternate observations on the reverse strand), SAF > 1 (number of alternate observations on the forward strand), QUAL > 20 was applied to the variants for initial filtration. A second filtration was applied with stringent population frequency of AC_ExAC_EAS < 2 (variant frequency in East Asian), Hom_ExAC = 0 (homozygosis sample in ExAC database), 1000 Genome AF < 0.001(allele frequency in 1000 Genome database) and an in-house database with phenotype-related frequency. The Mendelian inheritance vector was calculated with the logical operator from SnpSift software (omictools.com/snpsift-tool). The final variant calling format (VCF) was aggregated to a gene-mutation-burden table. A natural language process (NLP)-based search program "langya" was applied to match each gene name in the gene-mutation-burden table with the clinical phenotype, resulting in a likelihood score for prioritizing the mutation list.

Sanger Sequencing Validation

The detected variants were confirmed using PCR, and the PCR-amplified DNA products were subjected to direct automated sequencing (3500XL Genetic Analyzer, Applied Biosystems, Carlsbad, CA, USA) according to the manufacturer's specifications.

Plasmids

The human gene *SETD1A* was purchased from GenScript (Clone ID: C97170) and cloned into pcDNA3.1 + -DYK using CloneEZ plasmid. The mutants were cloned using the following primers: *SETD1A*-R913C, 5′-TGTCCCTCCAC TCCTGCTGAGGAAG-3′ (forward) and 5′-AGGCCTCT TTTCCTCACTGGCCTCG-3′ (reverse); *SETD1A*-Q269R, 5′-GAGGAACCCCCTACACGTCTCG-3′ (forward) and 5′-GGGAGGACTGAGGTGTGAACTGG-3′ (reverse); *SETD1A*-G1369R, 5′-AGGGAGGAAGAGGAGGAGGA GTC-3′ (forward) and 5′-CTCTTCCTCCCCCTCCTCT TCGC-3′ (reverse); and *SETD1A*-R1392H, 5′-ACAGCCT CCGCTCCCACGC-3′ (forward) and 5′-GCCTCCGGA GGGCGCCCTC-3′ (reverse). The *SETD1A* wild-type (WT) and mutant products were digested at Nhe I/EcoR I sites and inserted into the FUGW-EGFP plasmid at the Xba I/EcoR I sites by T4 ligase. The Setd1a-short-hairpin RNA(shRNA) vector was FUGW-H1-GFP. sh1# forward: 5′-ctagaCCCCACCCAAGAAACGCCGGAATTCAAGAG ATTCCGGCGTTTCTTGGGTGTTTTTg-3, sh1# Reverse: 5′-gatccAAAAACACCCAAGAAACGCCGGAATCTCT TGAATTCCGGCGTTTCTTGGGTGGGGt-3′; sh2# Forward: 5′-ctagaCCCCCACCCAAGAAACGCCGGATT CAAGAGATCCGGCGTTTCTTGGGTGGTTTTTg-3′, sh2# reverse: 5′-gatccAAAAACCACCCAAGAAACGCC GGATCTCTTGAA TCCGGCGTTTCTTGGGTGGGGGt-3′; sh3# forward: 5′-ctagaCCCCTTGGAACACAACTATG CCTTCAAGAGAGGCATAGTTGTGTTCCAAGTTTTT g-3′, sh3# reverse: 5′-gatccAAAAACTTGGAACACAAC-TATGCCTCTCTTGAAGGCATAGTTGTGTTCCAAGG GGt-3′. All constructs were verified by sequencing.

Cell Culture

Primary cortical neurons were prepared from the brains of C57BL/6J mice on embryonic day 15 (E15). Sterile chamber slides (154941, Thermo, Rochester, USA) were used for long-term culture, and calcium phosphate transfection was used for primary neuron culture. Dissociated cortical cells were plated at 1×10^5 cells/well.

Calcium Phosphate Transfection

Primary cortical neurons were transfected at 5 days *in vitro* (DIV). For each well, 1.9 μL of 2 mol/L $CaCl_2$, 2 μg DNA, and sterile water added to 15 μL were mixed in a 1.5 ml Eppendorf tube. After adding 15 μL pH 7.05 2 × HBS (HEPES buffered saline) to the tube, the mixture was incubated for 30 min at room temperature. Before transfection, 50% of the medium from each well was removed to mix with new medium at 37 °C. Then, 30 μL of the transfection mixture was added slowly to each well, and the cells were incubated for 90 min. Finally, the neurons were washed 3 times with washing solution and added to the prepared medium. None of the culture media contained serum.

In Utero Electroporation

Pregnant C57BL/6J mice were purchased from the Shanghai Laboratory Animal Center (Chinese Academy of Sciences) for *in utero* electroporation. Plasmids were injected into the ventricles of E14.5 embryonic mouse brains with a green fluorescent protein (GFP)-expressing plasmid (FUGW) at a 4:1 ratio. Plasmids were prepared in a final volume of 15 μL/mouse. After anesthetization with isoflurane and intraperitoneal phenobarbital, each embryo was subjected to electroporation (50 ms square pulses at 1000 ms intervals) using an Electro Square Porator (ECM 830, San Diego, CA, USA) at 30 mV. The pups were sacrificed at birth, and whole brains were dissected out. Longitudinal brain sections (40 μm) were cut on a freezing microtome (Leica CM1950, Germany) for later immunofluorescence experiments.

Immunohistochemistry

Mouse hippocampal neurons 4 DIV (for determining the length of axons) or 15 DIV (for dendrites and spines) were fixed with 4% (wt/vol) paraformaldehyde in 0.1 mol/L phosphate-buffered saline (PBS). After blocking in 3% bovine serum albumin and 0.1% Triton, the cells were incubated in anti-GFP (1:1000, Invitrogen, Waltham, MA, USA) at 4 °C overnight. After 3 washes in 1 × PBS the secondary antibody was used at 1:1000. Cells were also stained with DAPI (1:500) to visualize nuclei.

Neurite Length Analysis

GFP-positive neurons were randomly selected from each experiment and the total lengths of dendrites or axons were analyzed using Fiji software (imagej.net/fiji). All quantifications were tested with an unpaired *t*-test and expressed as the mean ± SD. The results were considered to be significantly different if $P < 0.05$. At least three independent experiments were performed for each condition.

Spine Density Analysis

An LSM510 (Heidelberg, Germany) laser scanning confocal microscope and an oil-immersion 60 × objective lens were used for acquiring images. GFP-positive neurons were selected randomly in each view, and the length of secondary branches of dendrites and the number of spines were analyzed with Fiji software. The average spine density was expressed as the number of spines/10 μm dendritic length. At least 20 neurons were analyzed in each group.

Neuronal Migration Analysis

Confocal z-stack images were acquired on a Nikon A1 (Tokyo, Japan) confocal laser microscope system. GFP-positive cells were counted using Fiji software. The ratios of GFP-positive cells in the outside zone (Cortical plate), middle zone and the inside zone (Ventricular zone) were analyzed. All statistical analyses were conducted using unpaired *t*-tests. At least three independent experiments were performed for each condition.

Real-Time Quantitative PCR

Samples of neurons were prepared after 4 days of transfection, and total RNAs were isolated using Trizol reagent according to the manufacturer's instructions. First-strand cDNAs were obtained using PrimeScript II kit (RR036A, TaKaRa, Japan). The mRNA levels of Neurl4 and Setd1a were measured by SYBR green-based real-time quantitative PCR (RT-qPCR) on a Roche 480 sequence detection system (Basel, Switzerland). The Gapdh gene was used as control. The calculation of the difference cycle threshold (ΔCT) was used to compare the target gene and the Gapdh. The forward and reverse primers used were as follows:

Setd1a: forward: 5′-CTTTTCCGGGTGGATCACGA-3′
reverse: 5′-TGGGCTGCTTCGAGTAGATG-3′
Neurl4: forward: 5′-CAGTATGCCGGGGCAGAGATGC-3′
reverse: 5′-CTCAAGTTCCGATGGGCTGAAGTG-3′
Gapdh: forward : 5′-ACGGCCGCATCTTCTTGTGCAGTG-3′

reverse: 5′-GGCCTTGACTGTGCCGTTGAATTT-3′

Statistical analysis of RT-qPCR data: All data are expressed as the mean ± SEM, using GraphPad Prism 8 software (www.graphpad.com). One-way ANOVA with Dunnett's multiple comparisons test was used for comparison of more than two groups. The unpaired *t* test was used for comparison of two groups.

Results

Identification of Mutations in SETD1A in Four Families

We first analyzed 4 generations of a Chinese family with 4 members affected by neonatal seizures. Whole-exome sequencing (WES) was performed in 3 affected members (IV-1, III-1, II-1) and 1 unaffected family member (II-2) to search for candidate genetic variants. The WES data for the probands consisted of 113,734,522 effective base reads for each person, with an average sequencing depth of × 113.33 of the target. We found that 99.6% of the target region was covered, of which 97.6% was covered at × 10 and 95.5% w at × 20. In the data of the affected individuals, we found a total of 102,245 genetic variants, 19,204 if which were in coding regions, distributed as follows: 8,693 missense, 227 frameshift, 91 stop-gain or stop-loss, 12 start-loss, 259 non-frameshift insertions or deletions (indels), and 9,922 synonymous mutations. We also identified 64 splice variants and 204 start-gain variants. After mutation effect annotation, a Mendelian inheritance screen and population frequency filtration, we determined that the p.R913C variant was absent from current public databases (dbSNP, 1000 Genome, ExAC, and gnomAD). Finally, we focused on 46 variants (Table S1) that overlapped between the 3 affected family numbers (III-1, IV-1, II-1) and were absent in the unaffected grandmother (II-2) (Fig. 1A, B). We screened these candidate genes for their relevance to neurological disorders and found that *SETD1A* has been strongly implicated in neurological and psychiatric disorders [10, 11].

We went on to investigate more sporadic, isolated early-onset epilepsy cases and identified three variants in *SETD1A* (defined by their absence from dbSNP, 1000 Genome, ExAC, and gnomAD and an allele count (AC) < 2), which were *de novo* variants in the early-onset epilepsy trio in our studies. Although the p.R1392H variant had a very low frequency in ExAC (AF: 0.0001034) and gnomAD (AF: 0.00003303), and the p.G1369R variant also showed an extremely low frequency in gnomAD (AF: 0.000007896), they were both novel in East Asians. Therefore, we considered these variants to be rare mutations (Fig. 1C–E, Table 1).

Functional Studies of SETD1A Mutations

Epilepsy reflects an abnormal and excessively active state of a set of neurons in the brain. First, we collected mouse cortical tissue from embryonic day 12.5 to 4 months old and measured the expression of *SETD1A* using real-time quantitative PCR. The results showed a relative consistent expression pattern of *SETD1A* during mouse brain development (Fig. S1A). To investigate the role of the *SETD1A* mutations in neural development, we assessed the effects of four *SETD1A* mutations on the neuronal morphology and spine development of primary mouse cortical neurons. We cloned the human (WT) *SETD1A* gene and performed point mutagenesis for the R913C, Q269R, G1369R, and R1392H mutants. We next transfected a vector control, *SETD1A* WT, and the four mutations into mouse cortical primary neurons and measured the axon length *via* immunostaining for anti-SIM312, an axonal marker (Fig. 2A–F). We found that the expression of *SETD1A* mutations did not significantly decrease the axon length of cortical neurons compared to the *SETD1A* WT, suggesting that *SETD1A* mutations do not affect normal axonal development (Fig. 2G). These results indicate that *SETD1A* mutations contribute little to axon growth and neuronal morphology.

Role of SETD1A Mutations in Synaptic Development

We then further investigated whether *SETD1A* mutations affect the development of synapses. We transfected GFP-expressing vector, along with vector control, *SETD1A* WT, R913C, Q269R, G1369R or R1392H constructs, into cultured mouse cortical primary neurons for 14 days. The dendritic spines, which represent excitatory synapses, became mature after ∼ 2 weeks in culture. Mature spines had mushroom-like shapes. We then calculated the average density of mushroom-like spines on the secondary branches of dendrites and found that their density was markedly lower in the groups expressing *SETD1A* mutations than in the WT (Fig. 3A–M). Together, these results indicated that *SETD1A* mutations play a role in compromising neural development, particularly in the normal development of neurons and synapses, thus suggesting that *SETD1A* mutations may contribute to the pathology of seizures by interfering with normal brain development.

Role of SETD1A R913C in Mouse Cortical Development

Given the solid family history of *SETD1A* R913C in the 4-generation pedigree, we then using an *in utero* electroporation assay to address whether expression of this inherited variant could interfere with cortical development.

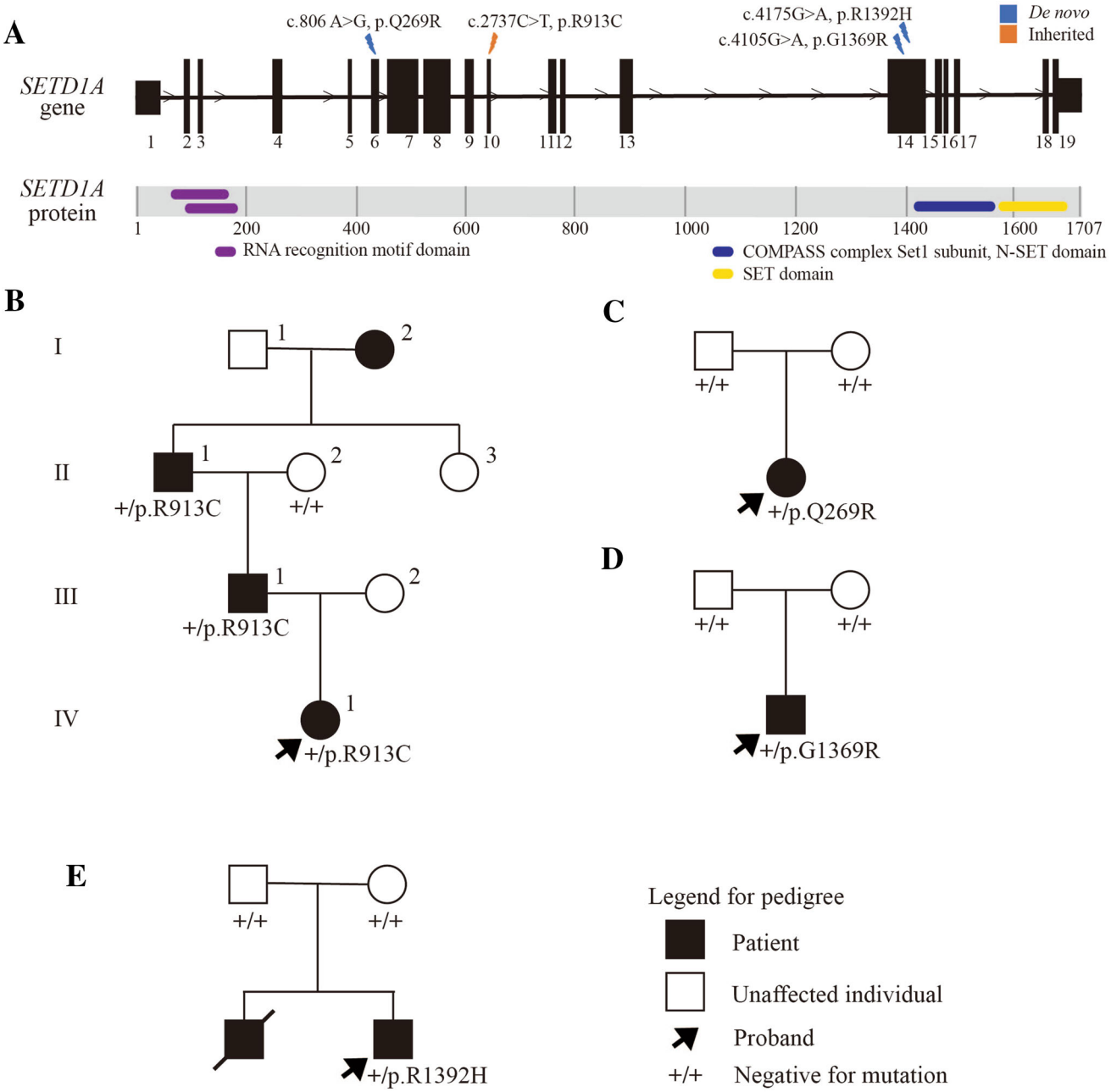

Fig. 1 Familial cohorts with early-onset epilepsy. **A** Identification of four missense mutations in the *SETD1A* gene. **B** Family pedigree of neonatal seizures and the p.R913C mutation in *SETD1A*. **C**–**E** Family pedigrees of the p.Q269R (**C**), p.G1369R (**D**), and p.R1392H (**E**) mutations in *SETD1A*.

We electroporated vector control *versus SETD1A* WT and control *vs SETD1A* R913C-expressing vector along with GFP, into E14.5 fetuses. Then, we collected the brains at birth. After immunostaining against GFP, we analyzed the migration of neurons expressing vector, WT, or *SETD1A* R913C. We found that neurons expressing *SETD1A* WT appeared to migrate more slowly than vector-expressing neurons but counts no statistical difference (Fig. 4A–C), while neurons expressing *SETD1A* R913C migrated faster than vector-expressing neurons (Fig. 4D–F), suggesting that *SETD1A* R913C indeed disturbs the normal process of cortical development and may lead to seizures.

A Common Gene Set Differentially Expressed Between the Four Mutations and the WT

Finally, to further define the transcriptomic differences between each mutation and the *SETD1A* WT group, we

Table 1 Summary of cases with mutations in *SETD1A* and their mutation frequency in different databases.

Gender	Gene variant	Mutation frequency					Inheritance	Onset	Seizures type	EEG	MRI (Brain)	Additional features	Age at follow-up	Development follow-up
		dbSNP	1000 Genome	ExAC	gnomAD									
					Total	East Asian								
Female	*SETD1A* Exon 10 c.2737C > T p.R913C	0	0	0	0	0	Inherited	2d	Clonic and tonic	Some sharp waves and sharp-slow-waves	Normal	No	1y3m	Normal
Female	*SETD1A* Exon 6 c.806A > G p.Q269R	0	0	0	0	0	*De novo*	2y	Clonic and tonic	Bilateral spike waves and Sharp-and-slow - waves	Normal	No	2y6m	Normal DQ = 110 MI = 107
Male	*SETD1A* Exon 14 c.4105G > A p.G1369R	0	0	0	0.000007896	0	*De novo*	9 m	Clonic and tonic	Multiple spike-and-slow-waves, slow-spike-and-waves, especially in sleeping	Cerebral dysplasia	No	1y9m	Developmental delay
Male	*SETD1A* Exon 14 c.4175G > A p.R1392H	0	0	0.0001034	0.00003303	0	*De novo*	1d	Clonic and tonic	Bilateral sharp waves and spike-and-waves	subdural hemorrhage	Lung dysplasia, sepsis, patent foramen ovale, bilateral hydrocele of testis	NA	NA

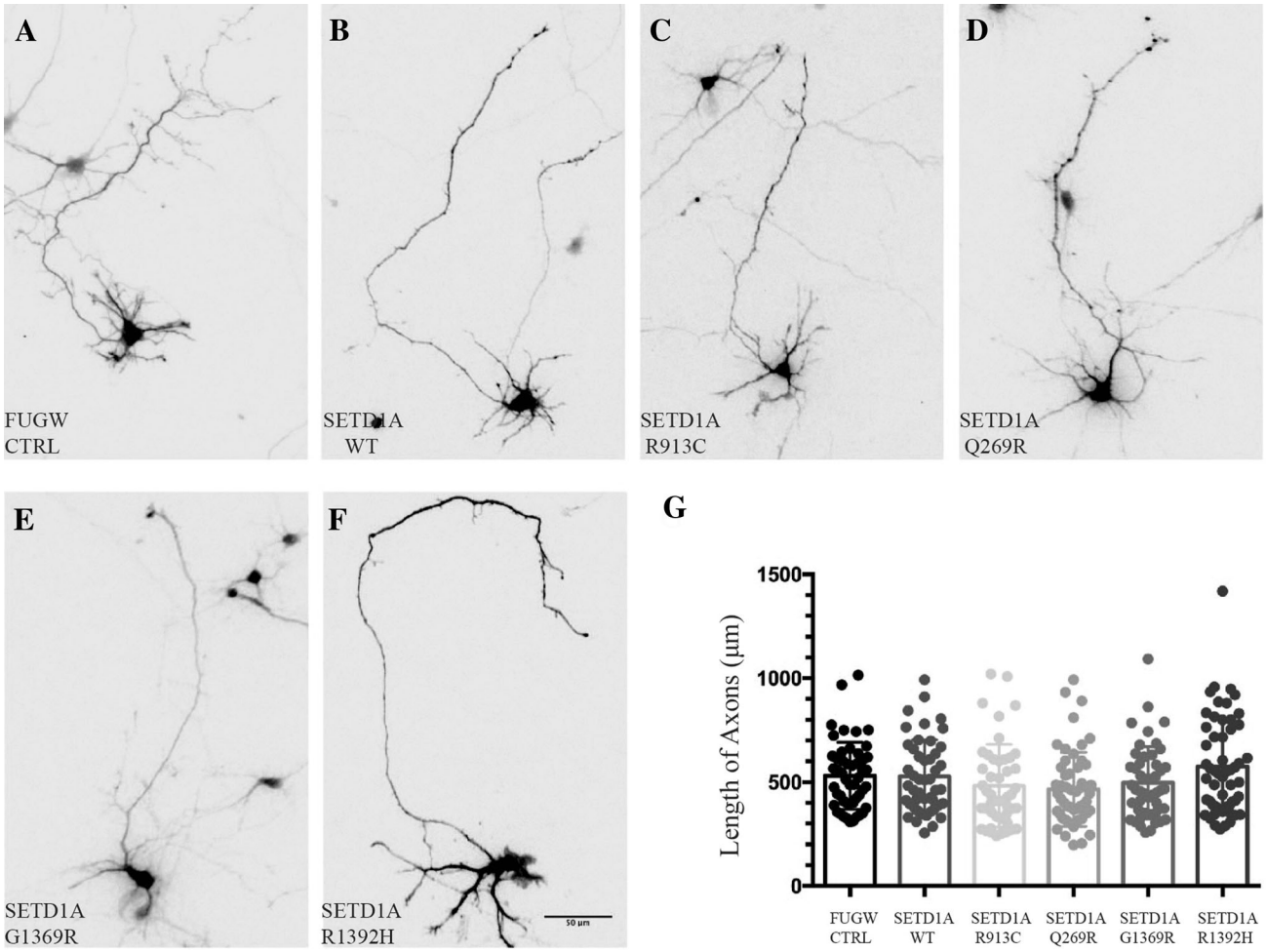

Fig. 2 Mutations of *SETD1A* do not regulate axon growth of mouse primary cortical neurons. **A**–**F** Mouse primary cortical neurons after transfection with GFP, *SETD1A* wild-type, and *SETD1A* R913C, Q269R, G1369R, and R1392H mutants. Neurons were fixed for immunohistochemistry at 4 DIV with GFP antibody for axon length determination. Scale bar, 50 μm. **G** Total axon length in each group (one-way ANOVA with Dunnett's multiple comparisons test).

collected cortical neuron samples from fetal mice transfected with control, *SETD1A* WT, R913C, Q269R, G1369R, or R1392H vectors and sent them for RNA sequencing. We compared each mutant group with the WT group to identify common differential gene expression between the mutation and WT groups (Fig. 5A, Table S2). Then we performed Venn Analysis and interestingly found that two common genes (Neurl4 and Usp39) were significantly downregulated after overexpression of the four mutations (Fig. 5B, C). We tested the samples by RT-qPCR and verified the decreased expression of Neurl4 in the G1369R and R1392H variants (Fig. 5D). The predicted target genes between each mutant group and the WT group were further analyzed using Gene Ontology analysis. Among all, the R1392H group had the largest number of genes that differed from the WT group (Fig. 5E).

Then we further investigated the relationship between Setd1a and its downstream genes (Neurl4 and Usp39). We designed three shRNAs against the mouse Setd1a gene. We transfected these three shRNAs along with control vector into mouse primary cortical neurons and total RNAs were collected after culture for 5 days. The expression of Setd1a, Neurl4, and Usp39 was determined by RT-qPCR and normalized to Gapdh. The results showed that knockdown of Setd1a strongly increased the expression levels of Neurl4 and Usp39 (Fig. 6A–C).

Discussion

In the current study, we used next-generation sequencing-based whole-exome sequencing to investigate the genetic etiology of early-onset epilepsy. By screening genetic mutants with early-onset epilepsy from 1 day to 2 years old in four families with early-onset epilepsy, we first identified one inherited and three *de novo* mutations of

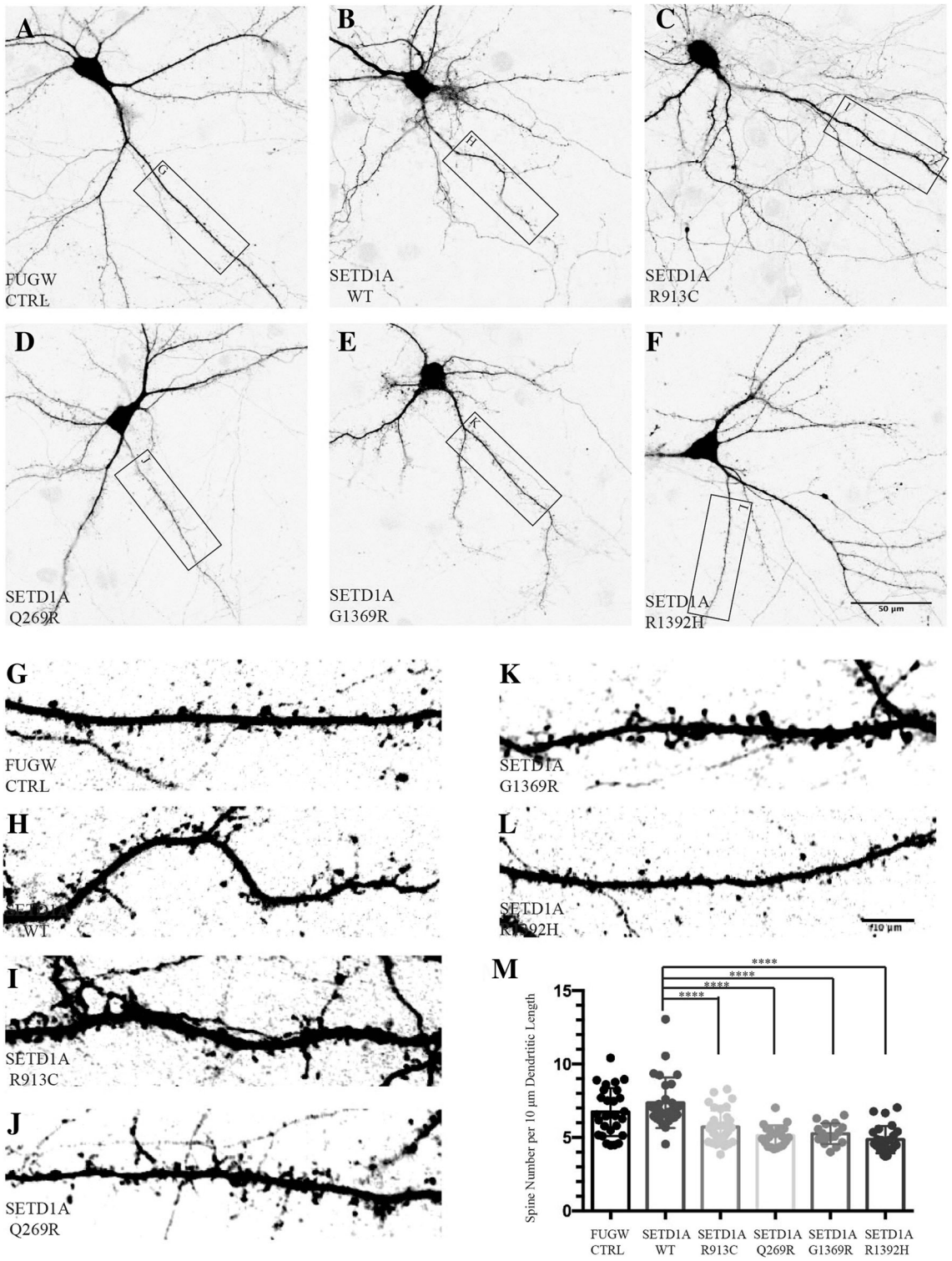

Fig. 3 Mutations of *SETD1A* affect spine development. **A–L** Mouse primary cortical neurons transfected at 5 DIV with GFP alone, *SETD1A* WT, or the *SETD1A* R913C, Q269R, G1369R, or R1392H mutations. Neurons were fixed for immunohistochemistry at 14 DIV with GFP antibody for average spine density determination. The secondary branches of dendrites were chosen for spine number counting. Scale bars, 50 µm (**A–F**) and 10 µm (**G–L**). **M** Average spine density measured as the number of spines 10 µm dendritic length (one-way ANOVA with Dunnett's multiple comparisons test, $****P < 0.0001$).

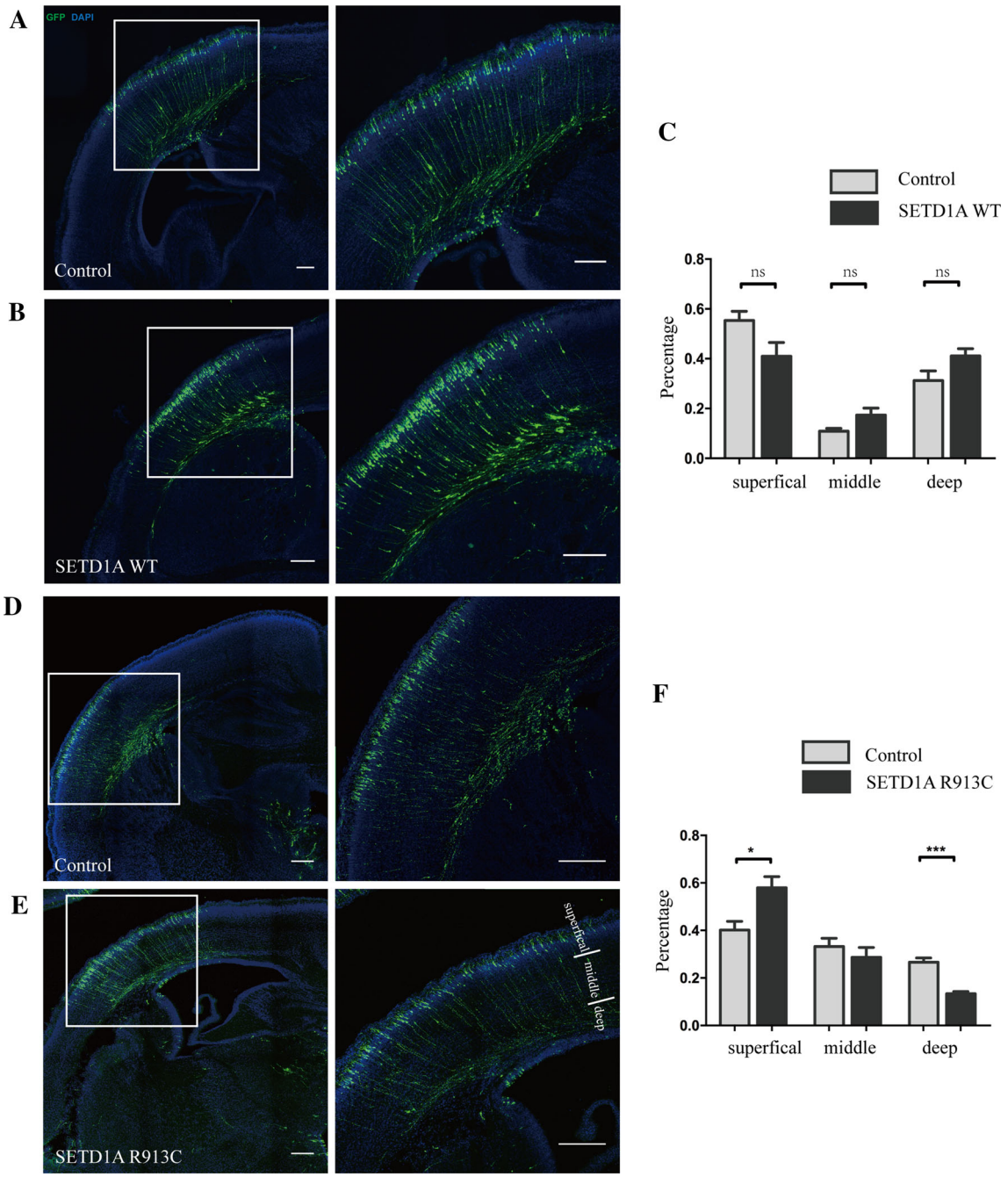

Fig. 4 R913C of *SETD1A* affects cortical development. **A–B** Examples of cortical migration in control and *SETD1A* WT overexpression groups. **C** Quantification of data as in panels A and B. Percentages of GFP-positive cells in the superficial, middle, and deep layers of cortex. **D–E** Examples of cortical migration in control and *SETD1A* R913C overexpression groups. F. Quantification of data as in panels **D** and **E**. Percentages of GFP-positive cells in the superficial, middle, and deep layers. $*P = 0.0392$, $***P = 0.0003$ (unpaired t test). Scale bars, 200 μm.

SETD1A, which were considered to be pathogenic variants. In addition, the R913C mutation of *SETD1A* was found in 4 members affected by neonatal seizures in a four-generation family.

We then further found that the four mutations of *SETD1A* significantly reduced the average density of spines, while the R913C mutation also changed the migration of cortical neurons in fetal mice, suggesting that all four are functional in neuronal development. The results reflected a subtle change in the neuronal development, which may affect the synaptic function of neural circuits and have long-term consequences for development.

Furthermore, we compared the total RNA sequencing results between each mutant group and the WT group, and

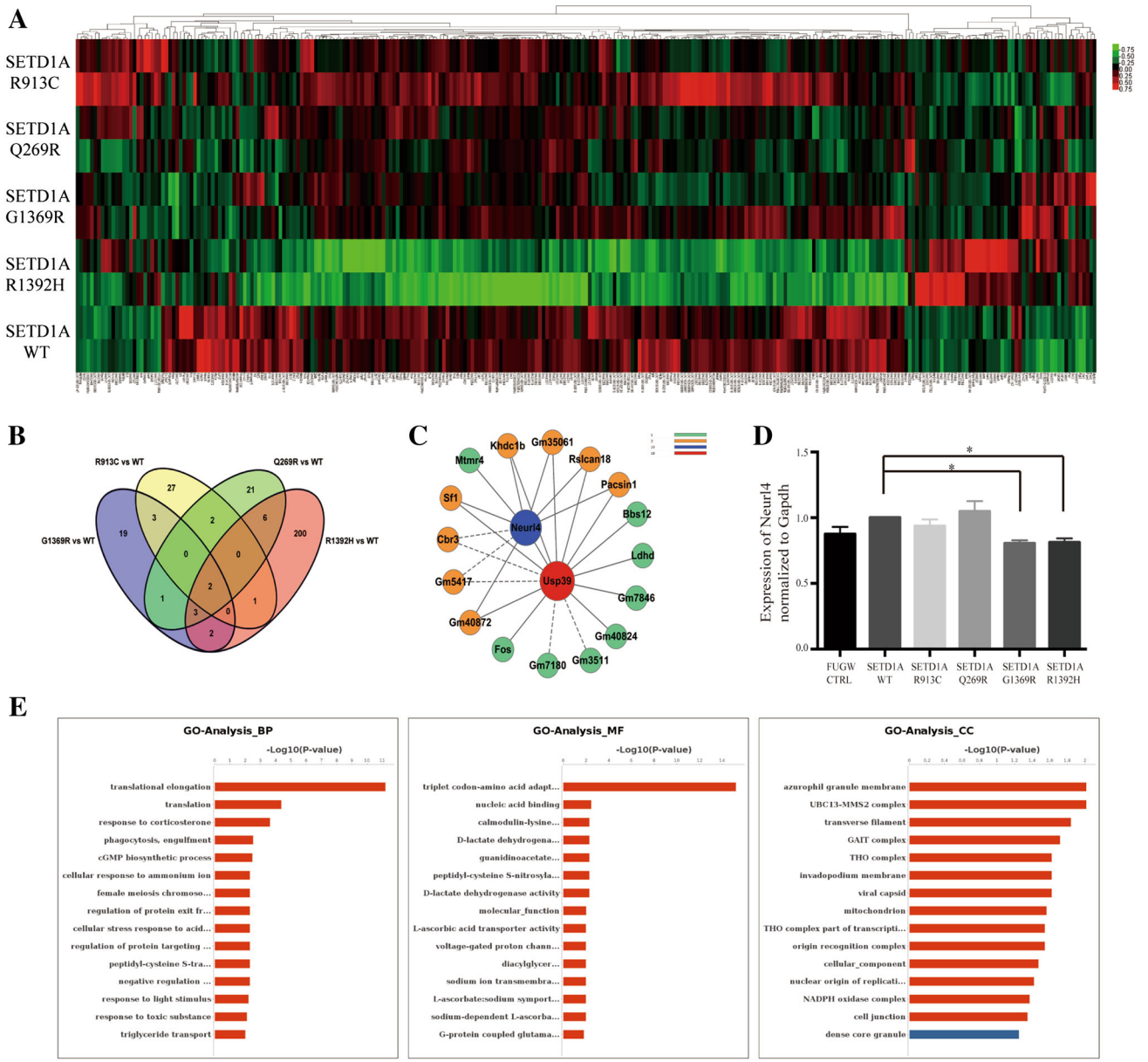

Fig. 5 Genes were differentially expressed between each mutant and the wildtype. Primary mouse cortical neurons were transfected with control, *SETD1A* WT, R913C, Q269R, G1369R or R1392H vectors. **A** Heatmap showing the common differentially-expressed genes in the total transcriptome of each mutant group *versus* the WT group. Each column represents one gene. **B**–**C** Venn Analysis of the common differentially-expressed genes (**B**) and the core network (**C**). **D** RT-qPCR results of Neurl4 normalized to Gapdh (one-way ANOVA with Dunnett's multiple comparisons test, $*P < 0.5$). **E** GO analysis of the target genes (*SETD1A* WT *vs* R1392H).

discovered two common differentially-expressed genes, Neurl4 and Usp39. We then followed this up and found that knockdown of Setd1a in neurons significantly increases the expression of Neurl4 and Usp39. These are both ubiquitin-related proteins. Past research has shown a relationship between ubiquitylation and H3K4 methylation [12]. And recent studies have shown that knockdown of Neurl4 is important for presynaptic bouton formation in the rodent cerebellum [13].

The SET Domain-Containing Protein 1A (*SETD1A*, OMIM: 611052) gene, also named Lysine-Specific Methyltransferase 2F, located on chromosome 16p11.2, encodes a catalytic component of SET/COMPASS (for a complex of proteins associated with Set1), which is capable of mono-, di-, and tri-methylation of H3K4. In humans, SET/

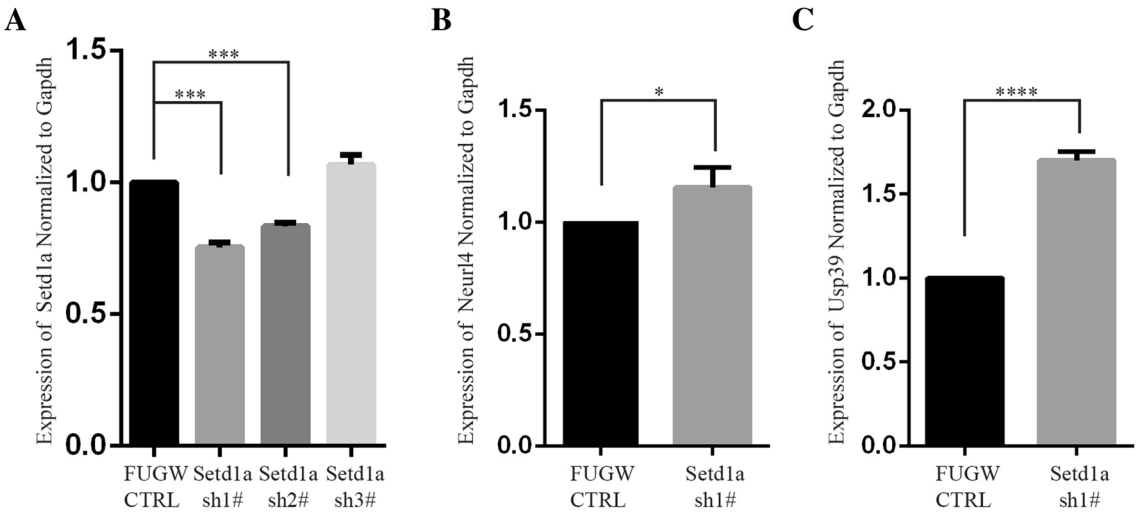

Fig. 6 Knockdown of Setd1a significantly increased the mRNA expression of Neurl4 and Usp39. **A** Setd1a mRNA levels by RT-qPCR showing that sh1# was the most effective (one-way ANOVA, $^{***}P < 0.001$). **B** Neurl4 mRNA levels by qPCR (unpaired *t* test, $^{*}P = 0.0415$). **C** Usp39 mRNA levels by qPCR (unpaired *t* test, $^{****}P < 0.0001$).

COMPASS includes four MLL homologues (MLL1 to MLL4), another Set-related protein SET1B, and several other types of protein [14, 15]. Although SET1/MLL family members and the functional components were discovered years ago, the crystal structure of the COMPASS catalytic core of yeast was not clearly shown until 2018 [16, 17]. Histone methylation at lysine residues, especially methylation of lysine 4 on histone 3 (H3K4), is considered to be a marker of transcriptional activity and is essential for neuronal differentiation [18]. Loss-of-function (LOF) variants of methyltransferases can lead to severe developmental diseases such as autism, Wiedemann–Steiner syndrome, Kabuki syndrome, Kleefstra syndrome, and some intellectual disability [11, 19].

In 2014, Takata *et al*. used exome sequencing to identify two *de novo* LOF variants in the *SETD1A* gene in 231 schizophrenia cases, indicating that the *SETD1A* gene is a risk factor in neurodevelopmental disorders [10]. By analyzing the whole-exome sequence data for 1077 published schizophrenia families, Singh and colleagues found a significant association between the *SETD1A* gene and schizophrenia (Fisher's combined $P = 3.3 \times 10^{-9}$), suggesting that LOF mutations are a risk factor for schizophrenia. In addition, the authors also found seven *SETD1A* LOF variants with learning difficulties and four *SETD1A* LOF carriers in children with developmental disorders [11] (Table S3). These recent findings both indicate that LOF mutation of *SETD1A* can lead to neuropsychiatric disorders like schizophrenia, but the mechanism by which the *SETD1A* gene regulates neuronal development is still poorly understood.

Although gene-testing methods like WES and WGS provide insights into neurobiological diseases, as gene mutations are frequently discovered, determining the relationship between genotype and phenotype remains problematic. The genetic landscape is complex, and phenotypic heterogeneity adds further complexity [20]. For example, 80% of cases of benign familial neonatal seizures are caused by mutation of the *KCNQ2* gene [21–23]. However, this gene is also associated with severe early-onset epileptic encephalopathies, such as Ohtahara syndrome, which may lead to severe developmental deficits [24]. A similar phenomenon has also been reported for the *SCN1A* (sodium voltage-gated channel alpha subunit 1) and *KCNT1* genes [20]. Similarly, in our study, mutations of the *SETD1A* gene were found in different types of early-onset epileptic syndromes and led to distinct outcomes. We then examined the locations of mutations in the *SETD1A* gene (Fig. 1A), and interestingly found that the R913C (on exon 6) and the Q269R (on exon 10) variants were far from the SET domain, so this may result in a relatively benign prognosis. On the other hand, the G1369R and R1392H variants were on exon 14, close to the SET domain, leading to more severe clinical phenotypes and outcomes. Recently, next-generation sequencing was performed in a large cohort of children with epilepsy in China, and the results indicated that genetic data are important for precision medicine approaches [6].

In 1995, Lee *et al*. developed a model of early-onset epilepsy by injecting tetanus toxin into rats during early life [25]. Later, Jiang *et al*. found a considerable reduction in the density of dendritic spine and anatomical alterations of hippocampal cells in this model of epilepsy [26].

In conclusion, we have identified four novel mutations in the *SETD1A* gene in patients with early-onset epilepsy using WES. Although further study is needed to determine

the precise role of the *SETD1A* gene in the occurrence of seizures, our work reveals *SETD1A* as a new candidate gene for early-onset epilepsy in early childhood.

Acknowledgements This work was supported by the National Natural Science Foundation of China (81741087, 91432111, 31625013, and 81471484), the Science and Technology Commission of Shanghai Municipality, China (14411950402), a Shanghai Municipal Science and Technology Major Project (#018SHZDZX05), and the Postdoctoral Science Foundation of China (2017M621361).

Conflict of interest The authors declare that they have no competing interests.

References

1. Eltze CM, Chong WK, Cox T, Whitney A, Cortina-Borja M, Chin RF, *et al.* A population-based study of newly diagnosed epilepsy in infants. Epilepsia 2013, 54: 437–445.
2. Hauser WA, Kurland LT. The epidemiology of epilepsy in Rochester, Minnesota, 1935 through 1967. Epilepsia 1975, 16: 1–66.
3. Deprez L, Jansen A, De Jonghe P. Genetics of epilepsy syndromes starting in the first year of life. Neurology 2009, 72: 273–281.
4. Picco G, Garnett MJ. A road map for precision cancer medicine using personalized models. Cancer Discov 2017, 7: 456–458.
5. Lee EH. Epilepsy syndromes during the first year of life and the usefulness of an epilepsy gene panel. Korean J Pediatr 2018, 61: 101–107.
6. Yang L, Kong Y, Dong X, Hu L, Lin Y, Chen X, *et al.* Clinical and genetic spectrum of a large cohort of children with epilepsy in China. Genet Med 2018, 21:564–571
7. Lemke JR, Riesch E, Scheurenbrand T, Schubach M, Wilhelm C, Steiner I, *et al.* Targeted next generation sequencing as a diagnostic tool in epileptic disorders. Epilepsia 2012, 53: 1387–1398.
8. Mastrangelo M. Novel genes of early-onset epileptic encephalopathies: from genotype to phenotypes. Pediatr Neurol 2015, 53: 119–129.
9. Maljevic S, Lerche H. Potassium channel genes and benign familial neonatal epilepsy. Prog Brain Res 2014, 213: 17–53.
10. Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, Karayiorgou M. Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. Neuron 2014, 82: 773–780.
11. Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, *et al.* Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci 2016, 19: 571–577.
12. Kim J, Kim JA, McGinty RK, Nguyen UT, Muir TW, Allis CD, *et al.* The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. Mol Cell 2013, 49: 1121–1133.
13. Valnegri P, Huang J, Yamada T, Yang Y, Mejia LA, Cho HY, *et al.* RNF8/UBC13 ubiquitin signaling suppresses synapse formation in the mammalian brain. Nat Commun 2017, 8: 1271.
14. Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, *et al.* COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A 2001, 98: 12902–12907.
15. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu Rev Biochem 2012, 81: 65–95.
16. Qu Q, Takahashi YH, Yang Y, Hu H, Zhang Y, Brunzelle JS, *et al.* Structure and conformational dynamics of a compass histone H3K4 methyltransferase complex. Cell 2018, 174: 1117–1126.
17. Hsu PL, Li H, Lau HT, Leonen C, Dhall A, Ong SE, *et al.* Crystal structure of the compass H3K4 methyltransferase catalytic module. Cell 2018, 174: 1106–1116.
18. Wynder C, Hakimi MA, Epstein JA, Shilatifard A, Shiekhattar R. Recruitment of MLL by HMG-domain protein iBRAF promotes neural differentiation. Nat Cell Biol 2005, 7: 1113–1117.
19. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, *et al.* Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 2015, 87: 1215–1233.
20. McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol 2016, 15: 304–316.
21. Singh NA, Charlier C, Stauffer D, DuPont BR, Leach RJ, Melis R, *et al.* A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nat Genet 1998, 18: 25–29.
22. Biervert C, Schroeder BC, Kubisch C, Berkovic SF, Propping P, Jentsch TJ, *et al.* A potassium channel mutation in neonatal human epilepsy. Science 1998, 279: 403–406.
23. Charlier C, Singh NA, Ryan SG, Lewis TB, Reus BE, Leach RJ, *et al.* A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. Nat Genet 1998, 18: 53–55.
24. Weckhuysen S, Ivanovic V, Hendrickx R, Van Coster R, Hjalgrim H, Moller RS, *et al.* Extending the KCNQ2 encephalopathy spectrum: clinical and neuroimaging findings in 17 patients. Neurology 2013, 81: 1697–1703.
25. Lee CL, Hrachovy RA, Smith KL, Frost JD, Jr., Swann JW. Tetanus toxin-induced seizures in infant rats and their effects on hippocampal excitability in adulthood. Brain Res 1995, 677: 97–109.
26. Jiang M, Lee CL, Smith KL, Swann JW. Spine loss and other persistent alterations of hippocampal pyramidal cell dendrites in a model of early-onset epilepsy. J Neurosci 1998, 18: 8356–8368.