ZFHX3 Associated with Partial Epilepsy/Spasms and Correlation between Outcome & Gene Expression Stage

Ming-Feng He,[1,†] Li-Hong Liu,[1,†] Sheng Luo,[1,†] Juan Wang,[1] Jia-Jun Guo,[1] Peng-Yu Wang,[1] Qiong-Xiang Zhai,[2] Su-Li He,[3] Dong-Fang Zou,[4] Xiao-Rong Liu,[1] Bing-Mei Li,[1] Hai-Yan Ma,[5] Jing-Da Qiao,[1] Peng Zhou,[1] Na He,[1,*] Yong-Hong Yi,[1,*] and Wei-Ping Liao[1]

For the China Epilepsy Gene 1.0 Project

†These authors contributed equally to this work and share co-first authorship.

Abstract

The *ZFHX3* gene is highly expressed in the developing brain and plays vital roles in embryonic development, cell proliferation, neuronal differentiation, and neuronal death. The association between the *ZFHX3* gene and human disease was undefined. This study aims to explore the relationship between *ZFHX3* variants and epilepsy.

Whole-exome sequencing was performed in patients with partial epilepsy without acquired causes. *Drosophila* model of *Zfh2* (ortholog of *ZFHX3*) knockdown was used to validate the association between *ZFHX3* and epilepsy. The expression level of *ZFHX3* in different developmental stages was analyzed by using the data in humans from the Brainspan database and in flies and mice determined by RT-qPCR.

Eight pairs of compound heterozygous variants in *ZFHX3* were identified in eight unrelated cases of childhood epilepsies. All patients presented partial epilepsy, including two with early spasms and one with frequent nonconvulsive status epilepticus. The three cases with severe epilepsies also had neurodevelopmental abnormities. However, all patients became seizure-free, including the patients with spasms. These variants had no or low allele frequencies in controls and presented

statistically higher frequency in the case cohort than the controls. The number of recessive *ZFHX3* variants identified in this cohort was significantly more than the expected number of the East Asian population and that of the control of 1942 asymptomatic parents. Variants were predicted to be damaging by protein modeling and/or *in silico* prediction tools. Genotype-phenotype correlation analysis revealed that the patients with missense variants in C-terminus or C-terminus-truncated variant exhibited mild phenotype (only partial epilepsy). Knockdown of *Zfh2* in flies led to increased susceptibility to seizures and abnormal firing of excitability neurons. In *Drosophila*, the expression of *Zfh2* is high in larvae, decreased in pupae and early adults, and increased in later adults. In mice, *Zfhx3* is predominantly expressed in fetuses and decreased dramatically after birth. In humans, the data from the Brainspan database showed that *ZFHX3* is highly expressed in the embryonic period and decreased after birth with a nadir at approximately 10 years old. The dramatical decreasing of *ZFHX3* in early life correlated with the natural course of the illness.

ZFHX3 variants were potentially associated with partial epilepsy of childhood and infant spasms. The correlation between the outcome and gene expression stage provided insight into the underlying mechanism of the natural course of illness, potentially being helpful in management of the patients.

Author affiliations:

1. Department of Neurology, Institute of Neuroscience, Key Laboratory of Neurogenetics and Channelopathies of Guangdong Province and the Ministry of Education of China, the Second Affiliated Hospital, Guangzhou Medical University, Guangzhou 510260, China.

2. Department of Pediatrics, Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510120, China.

3. Department of Pediatrics, Shantou Chaonan Minsheng Hospital, Shantou 515000,

China.

4. Epilepsy Center and Department of Neurology, Shenzhen Children's Hospital, Shenzhen 518029, China.

5. Department of Functional Neurosurgery-epilepsy center, Nanjing Brain Hospital affiliated to Nanjing Medical University, Nanjing 210029, China.

*Correspondence to:

Dr. Yong-Hong Yi, E-mail: yyh168@sina.com

Institute of Neuroscience and Department of Neurology of the Second Affiliated Hospital of Guangzhou Medical University, Chang-Gang-Dong Road 250, Guangzhou, 510260, China

Dr. Na He, E-mail: henachilli@163.com

Institute of Neuroscience and Department of Neurology of the Second Affiliated Hospital of Guangzhou Medical University, Chang-Gang-Dong Road 250, Guangzhou, 510260, China

Running title: *ZFHX3* and Partial Epilepsy/Spasms

Keywords: *ZFHX3* gene, partial epilepsy, trios-based WES, neurodevelopmental disorder, gene expression stage

Introduction

The zinc finger homeobox 3 gene (*ZFHX3*, OMIM *104155), also known as AT motif-binding transcription factor 1 gene (*ATBF1*), encodes a transcription factor with 4 homeodomains and 23 zinc finger motifs. It is ubiquitously expressed, especially in the developing brain. *ZFHX3* plays vital roles in multiple biological processes, including embryonic development, cell proliferation, neuronal differentiation, and neuronal death.[1-4] In mice, homozygous knockout of *Zfhx3* caused prenatal lethality with complete penetrance, while heterozygous knockout mice displayed growth restriction and partial postnatal lethality.[5] Knockdown of *Zfhx3* in zebrafish led to increased susceptibility to seizures.[6] However, the association between the *ZFHX3* gene and human disease was undefined.

In this study, we conducted trios-based whole-exome sequencing (WES) in a cohort of 378 unrelated children with partial (focal) epilepsy. Compound heterozygous variants of *ZFHX3* were identified in eight unrelated patients with partial epilepsy with favorable outcomes. *Drosophila* model with *Zfh2* knockdown was established to investigate the association between *ZFHX3* and epilepsy. The expression levels of *ZFHX3* in different developmental stages were analyzed using the data in the human from the Brainspan database and determined by RT-qPCR in flies and mice. This study suggests that the *ZFHX3* gene is potentially a novel causative gene of partial epilepsy of childhood and infant spasms. The develop-dependent expression pattern of *ZFHX3* is correlated with the natural course of the illness.

Materials and methods

Patients

Patients were identified in a cohort of 378 unrelated children with partial epilepsy without acquired causes, which were recruited from five hospitals from 2019 to 2022, including the Second Affiliated Hospital of Guangzhou Medical University,

Guangdong General Hospital, Shenzhen Children's Hospital, Shantou Chaonan Minsheng Hospital, and the Affiliated Brain Hospital of Nanjing Medical University. The clinical information of affected individuals was collected, including gender, seizure onset age, seizure types and frequency, response to antiepileptic drugs (AEDs), family history, general, and neurological examinations. Long-term video-electroencephalography (EEG) was performed to monitor epileptic discharges. Brain magnetic resonance imaging (MRI) scans were performed to detect structural abnormalities. Epileptic seizures and epilepsy syndromes were diagnosed according to the criteria of the Commission on Classification and Terminology of the ILAE (1989, 2001, 2010, 2017, and 2022).

This study was approved by the Ethics Committee of The Second Affiliated Hospital of Guangzhou Medical University (2020-h5-49), and written informed consent was obtained from all patients or their parents.

Trios-based whole-exome sequencing and genetic analysis

Genomic DNA was extracted from blood samples of the probands and their parents (trios), and other available family members, by using the Flexi Gene DNA kit (Qiagen, Hilden, Germany). Trios-based WES was performed with NextSeq500 sequencing instruments (Illumina, San Diego, CA, USA). Detailed sequencing methods were described in our previous studies.[7-9] A case-by-case analytical approach was adopted to identify candidate causative variants in each trio.[10, 11] Primarily, the rare variants were prioritized with a minor allele frequency (MAF) < 0.005 in the gnomAD database (gnomad.broadinstitute.org). Then, potentially pathogenic variants were retained, including frameshift, nonsense, canonical splice site, initiation codon, in-frame variants, missense, and synonymous variants predicted to impact splicing. Lastly and importantly, possibly disease-causing variants in each individual were screened under five models: 1) epilepsy-associated gene; 2) *de novo* variant dominant; 3) autosomal recessive inheritance, including homozygous and compound

heterozygous variants; 4) X-linked; and 5) co-segregation analysis if available. To identify novel potential epilepsy genes, genes with recurrently identified *de novo*, bi-allelic, hemizygous, and co-segregated variants were selected for further studies. *ZFHX3* appeared as a candidate gene with recurrent compound heterozygous variants in this cohort. All variants in *ZFHX3* in this study were validated by Sanger sequencing and annotated based on the transcript NM_006885.

Repetitiveness analysis

Three statistical analysis methods were used to analyze the association between *ZFHX3* and epilepsy, including recessive burden analysis,[12] aggregate frequency of variants,[13] and frequency of compound heterozygous variants.

For the burden of recessive variants, the *P*-value was calculated as [1-cumulative binomial probability [(n-1, N, R)] according to the previous report,[12] where n is the observed biallelic variant number for *ZFHX3*, N is the number of trios (378 in this cohort), and R is the rate of *ZFHX3* variants by chance in populations. Considering all patients were Han Chinese, the cutoff is set according to MAF in the population of ExAC-East Asian.

For aggregate frequency analysis, frequencies of identified variants between the case cohort and the controls were compared, including general and the East Asian populations in the gnomAD database,[14] general populations in the ExAC database,[15] and the 33,444 persons without known neuropsychiatric conditions in the Epi25 WES Browser (https://epi25.broadinstitute.org/).[16]

For the control of compound heterozygous variants, we established a cohort of 1942 asymptomatic parents from trios, in whom the compound heterozygous variants were identified by detecting one of the paired variants in the child, based on that one of the paired variants in a parent would transmit to the child. The number of identified compound heterozygous *ZFHX3* variants between the case cohort and the 1942 asymptomatic parents control were compared.

Molecular structural analysis

Protein modeling was performed to predict the effects of missense variants on the molecular structure, by using the I-TASSER tool (https://zhanggroup.org/I-TASSER/). Protein structures were visualized and analyzed by using PyMOL Molecular Graphics System (Version 2.5; Schrödinger, LLC; New York, USA). Protein stability changes of missense variants were predicted by using I-Mutant Suite (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi).

Drosophila experiments

Flies were reared at 25□ and 60-70% humidity with a standard cornmeal diet under a 12:12 h light and dark cycle. The *UAS-Zfh2-RNAi* line (CG1449, TH01656.N) was purchased from the TsingHua Fly Center. The *tub-Gal4* driver line and *Canton-s* line were donated by Prof. Liu Ji-Yong (Guangzhou Medical University, Guangzhou, China). *Zfh2* knockdown flies were generated by crossing the *UAS-Zfh2-RNAi* line with the *tub-Gal4* driver line. *Canton-s* were the wild type line. The knockdown efficiency was detected by using reverse transcription quantitative PCR (RT-qPCR).

To evaluate the role of Z*fh2* deficiency in development, the body length of fly larvae was assessed, as in our previous study.[17] The larval stages of flies can be divided into three stages, i.e., the first (24 h), second (48 h), and third (72 h) instar larvae. The length of the larvae was measured using the ImageJ NIH software (https://imagej.nih.gov/ij/).

To evaluate seizure-like behavior, the bang-sensitivity assay was conducted on flies 3-5 days after eclosion.[18, 19] Briefly, flies were anesthetized using CO_2 and transferred to clean vials for 0.5-1 h before testing. Three to seven flies were placed in each vial and stimulated mechanically using a vortex mixer (Lab Dancer, IKA, Germany) at maximum speed for 10 seconds.[18, 20] The duration and percentage of seizure-like behavior were recorded.

To determine the impact of *Zfh2* knockdown on neuronal excitability, the

electrophysiology activity of projection neurons was recorded. The fly brains were prepared for electrophysiological recording as described previously.[21] The standard external & internal solution and patch pipette was prepared as described previously.[22] Spontaneous activities were acquired using the 700B amplifier, 1440B Digital-Analog converter, and pClamp 10.5 software (Molecular Devices, San Jose, CA, USA). The cell with access resistance of <30 MΩ sample was used for analysis. The spontaneous EPSPs (sEPSPs) data > 1 mV were analyzed by the Mini Analysis software.

Assessment of *ZFHX3* orthologs expression profile

The mRNA expression level of *ZFHX3* orthologs in different developmental stages was determined by RT-qPCR. For flies, the whole mRNA was extracted in five developmental stages, including third instar larvae, pupae, early adult (day 1), middle adult (day 5), and later adult (day 10). For mice, the mRNA of the frontal cortex was extracted in eight developmental stages, including fetus, neonate (1 day), infant (1 week), toddle-period (2 weeks), pre-school (4 weeks), juvenile-adult (10 weeks), middle-age (15 weeks), and old-age (32 weeks). The primer sequences used in this study were listed in **Table S1**. Total RNA was extracted by using the HiPure Universal RNA Mini Kit (Magen Biotechnology, Guangzhou, China). Reverse transcription was performed with the HiScript III RT SuperMix for qPCR (+gDNA wiper) (Vazyme, Nanjing, China) kit, then qPCR was performed by the Taq Pro Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) with the LightCycler 480 System (Roche) tool.

Human RNA-seq data at different developmental stages (from 8 post-conceptional weeks to 40 years) for multiple brain areas were obtained from the Brainspan database (http://www.brainspan.org/). RNA expression was normalized to reads per kilobase million (RPKM). The temporal expression curve is fitted by the third-order polynomial least squares for interpreting the expression pattern of *ZFHX3* by GraphPad Prism 9.

Statistical analysis

R statistical software (v4.0.2) and GraphPad Prism 9 were used for statistical analysis. All quantitative data are presented as mean ± standard error of the mean (SEM). The burden of recessive variants was analyzed according to the previous report.[12] The number and aggregated frequencies of *ZFHX3* variants of this cohort and controls were compared by a two-tailed Fisher's exact test. The student's t-test was used to compare two independent samples and the Mann-Whitney test was used to assess non-parametric data. The P value < 0.05 was considered to be statistically significant.

Data availability

Raw data were generated at Institute of Neuroscience, The Second Affiliated Hospital of Guangzhou Medical University. Derived data supporting the findings of this study are available from the corresponding author on request.

Results

Identification of *ZFHX3* variants

Compound heterozygous variants in *ZFHX3* were identified in eight unrelated cases with partial epilepsies, including c.314C>T/p.Pro105Leu & c.2282G>C/p.Gly761Ala, c.2419G>A/p.Glu807Lys & c.10853C>A/p.Pro3618Gln, c.2671T>C/p.Phe891Leu & c.9583_9584insT/p.Pro3195LeufsTer44, c.2686G>A/p.Ala896Thr & c.10439C>T/p.Ala3480Val, c.4125_4127del/p.Ala1376del & c.10853C>A/p.Pro3618Gln, c.5152A>C/p.Met1718Leu & c.10510G>C/p.Val3504Leu, c.6161C>T/p.Ala2054Val & c.10445G>T/p.Ser3482Ile, and c.10445G>T/p.Ser3482Ile & c.10853C>A/p.Pro3618Gln (**Table 1 and Figure 1**). The eight pairs of compound heterozygous variants consisted of eleven missense variants, one frameshift truncation, and one in-frame deletion. The missense variant

p.Ser3482Ile was recurrently identified in two cases (Cases 7 & 8), and the variant p.Pro3618Gln was recurrently identified in three cases (Cases 2, 5, & 8). All of the *ZFHX3* compound heterozygous variants were inherited from their asymptomatic parents, consistent with the Mendelian autosomal recessive inheritance pattern.

The identified variants presented no or low frequencies in the gnomAD database (MAF < 0.005) (**Table 2**). Six variants did not present in the normal control of Epi25 WES Browser, while other seven variants presented extremely low frequencies (MAF < 0.0005). None of the variants presented as homozygous states in gnomAD.

When the burden of recessive variants was analyzed,[12] the number of recessive *ZFHX3* variants identified in this cohort was significantly more than the expected number by chance in the East Asian population (MAF < 0.005, $P = 1.60\times10^{-4}$). The aggregate frequency of variants identified in the case cohort was significantly higher than that in the controls, including the controls of gnomAD-East Asian population ($P = 4.36\times10^{-2}$), the controls of gnomAD-all population ($P = 2.30\times10^{-13}$), the general population of ExAC database ($P = 2.34\times10^{-13}$), and the normal controls of Epi25 WES Browser ($P = 3.56\times10^{-18}$) (**Table 2**). The frequency of compound heterozygous *ZFHX3* variants of the case cohort was significantly higher than that in the control of 1942 asymptomatic parents. (8/378 vs 5/1942, $P = 2.80\times10^{-4}$).

None of the eight patients had pathogenic or likely pathogenic mutations in the genes known to be associated with epileptic phenotypes.[23]

Clinical features of the cases with *ZFHX3* variants

The detailed clinical characteristics of the eight unrelated cases with *ZFHX3* variants were summarized in **Table 1**. The onset age of seizures ranged from 5 months to 13 years old, with a median onset age of 5 years. The patients were all diagnosed with partial epilepsy, including three cases with neurodevelopmental abnormities. Two of the three cases presented frequent spasms (Cases 4 & 6) at the early stage of the illness

and one had nonconvulsive status epilepticus (Case 7). The other five cases exhibited only infrequently (yearly) focal seizures or focal-origin generalized tonic-clonic seizures (Cases 1-3, 5, 8). Focal, multifocal, and/or diffuse epileptic discharges were recorded in their EEG (**Figure 1E**). The brain MRI was normal in all cases. Seizure-free was achieved in all eight patients.

Molecular effect and functional prediction of *ZFHX3* variants

ZFHX3 protein contains 23 zinc finger motifs and 4 homeodomains. Two variants p.Glu807Lys and p.Ala1376del were located in the zinc finger motifs, and other variants were scattered between zinc fingers and/or homeodomains (**Figure 2A**). The variants p.Pro3618Gln, located in the last part of the C-terminus, were repetitively identified in three of five mild cases (Case 2, 5, & 8). Similarly, the frameshift variant (p.Pro3195LeufsTer44), which truncated the C-terminus, was also identified in a mild case (Case 3). The remaining mild case (Case 1) had two variants located in the N-terminus.

The damaging effect of missense variants was analyzed by protein modeling. Two missense variants p.Glu807Lys and p.Ala896Thr were predicted to alter hydrogen bonding with surrounding residues. Six variants were predicted to decrease the protein stability with a □□G value less than -0.5 kcal/mol (**Figure 2B**). All the variants were predicted to be damaged by at least one *in silico* tool (**Table S2**).

Knockdown of *Zfh2* in *Drosophila* led to increased susceptibility to seizures

To validate the association between *ZFHX3* and epilepsy, the *Drosophila* model of *Zfh2* knockdown was established, with a knockdown efficiency of approximately 66% (**Figure 3A**). The larval development of *Zfh2* knockdown flies was initially

investigated. The larval body length of *Zfh2* knockdown flies was similar to that of the *UAS-Zfh2-RNAi* control flies (**Figure 3B**).

Bang-sensitivity assay was examined in *Zfh2* knockdown flies and the *UAS-Zfh2-RNAi* control flies to assess the susceptibility to seizures. The percentage of seizure-like behaviors in the *Zfh2* knockdown flies was significantly higher than that of *UAS-Zfh2-RNAi* control ($P < 0.0001$), *Canton-s* control ($P = 0.0017$), and *tub-Gal4* control ($P < 0.0001$) (**Figure 3C**). The duration of seizure-like behavior in *Zfh2* knockdown flies was also longer than that of the controls (**Figure 3D**).

The effect of *Zfh2* deficiency on the electrophysiologic activity of projection neurons was examined, which are important excitatory neurons in the central nervous system of *Drosophila* (**Figure 3E**).[24] The frequency of sEPSPs in *Zfh2* knockdown flies was significantly higher than wide type flies ($P < 0.0001$) (**Figure 3F)**. There was no significant difference in the amplitude of sEPSPs between the *Zfh2* knockdown flies and WT flies ($P = 0.0888$) (**Figure 3G**).

The temporal expression stage of *ZFHX3* orthologs

Infantile spasms is generally a severe form of epilepsy with poor prognosis in the majority.[25] In this study, the two patients with early spasms presented favorable outcomes. Recent studies indicated that the gene dependent (expression) stage (GDS) is associated with the natural course and outcomes of illness.[26-28] We thus analyzed the temporal expression pattern of *ZFHX3* orthologs. In *Drosophila*, the expression of *Zfh2* was high in larvae, decreased in pupae and early adults, and increased in later adults (**Figure 4A**). In mice, the expression level of *Zfhx3* was highly expressed in the fetus, decreased dramatically after birth, and slightly increased in later adults (32 weeks) (**Figure 4B**). We further analyzed the temporal expression pattern of *ZFHX3* in the human brain by using the data from BrainSpan. *ZFHX3* was also highly expressed in the embryonic stage, decreased dramatically in childhood with a nadir at approximately 10 years old, and slightly increased around 30 years old (**Figure 4C**).

Discussion

ZFHX3 (*ATBF1*) is located at chromosomal locus 16q22.2-22.3 and composed of 10 exons. It encodes a transcription factor with 3703 amino acids, which plays an important role in neuronal differentiation and neuronal death in response to DNA damage or oxidative stress.[1-4] In this study, we identified compound heterozygous *ZFHX3* variants in eight unrelated patients with partial epilepsy, including two with earlier spasms. All patients achieved seizure-free. The aggregated frequency of *ZFHX3* variants identified in the case cohort was significantly higher than that in the controls. The number of recessive variants identified in the case cohort was significantly more than the expected number by chance in the East Asian population and the control of asymptomatic parents. Knockdown of *Zfh2* in flies caused increased susceptibility to seizures and abnormal firing of neurons. The analysis of temporal expression profile indicated that *ZFHX3* orthologs were highly expressed in the embryonic stage and decreased dramatically in childhood, which is correlated with the favorable outcomes of the patients. This study suggested that *ZFHX3* is potentially a novel causative gene of partial epilepsy of childhood and infant spasms.

The *ZFHX3* gene is highly conserved with homologs in flies, zebrafish, mice, and humans. In mice, homozygous knockout *Zfhx3* led to prenatal lethality with complete penetrance, while heterozygous knockout resulted in growth restriction and postnatal lethality with incomplete penetrance.[5] In zebrafish, knockdown *Zfxh3* caused significantly increased pentylenetetrazol-induced seizures.[6] In flies, previous studies suggested that the majority of knockout/knockdown *Zfh2* lines exhibited preadult lethality (http://flybase.org/reports/FBgn0004607#phenotypes). This study identified biallelic *ZFHX3* variants in eight unrelated patients with partial epilepsy with/without neurodevelopmental abnormities; and the knockdown of *Zfh2* in *Drosophila* led to increased susceptibility to seizures and abnormal firing of neurons (**Figure 3**). The recurrent epileptic and/or neurodevelopmental phenotypes across species supported the

association between *ZFHX3* and epilepsy and neurodevelopmental abnormities.

Data from variants in the human genome reveal that *ZFHX3* is intolerance to LOF variants, including constraint indexes such as the probability of being LOF intolerant (pLI),[14] the LOF observed/expected upper bound fraction (LOEUF),[14] and the probability of haploinsufficiency (pHaplo).[29] The *ZFHX3* gene has pLI of 0.997 (> 0.9), LOEUF of 0.147 (< 0.35), and pHaplo of 1(≥ 0.84), all indicating that *ZFHX3* is intolerant to heterozygous LOF variants. Considering the recurrent lethal and neurodevelopment/epileptic phenotypes in *Zfhx3* knockout/knockdown animals, it is possible that severe *ZFHX3* deficiency, such as biallelic null variants, will cause early lethality, while moderate *ZFHX3* deficiency, such as heterozygous null variants or biallelic missense variants with moderate damaging effect, was potentially associated with epilepsy and/or neurodevelopmental abnormities. In this study, most of the variants were located outside the functional domains and predicted to be without hydrogen bond changes with surrounding residues (**Figure 2**), indicating mild damaging effects, which potentially explain the pathogenesis of compound heterozygous variants. Three patients with mild phenotype had an identical variant p.Pro3618Gln, which was located in the last part of the C-terminus and presented a low MAF (0.000301) (instead of not present) in the general population, indicating a potentially mild damaging effect. Another patient with mild phenotype had a frameshift variant with truncation at the C-terminus, which was inherited from the asymptomatic father and also presented with low MAF, thus potentially being of mild damaging. These findings suggested a potential genotype-phenotype correlation.

The ZFHX3 protein, a transcription factor containing 23 C_2H_2-type zinc fingers and 4 homeodomains, belongs to the C_2H_2-type zinc finger protein family, which plays an essential role in neurodevelopment.[30] Previously, *ZFHX3 de novo* variants were occasionally detected in patients with neurodevelopmental disorders, including four variants in autism spectrum disorder,[31-33] three variants in developmental disorder,[34] and one variant in development and epileptic encephalopathy (**Table S3**),[35] suggesting

a potential role in neurodevelopmental disorders. In the presented study, neurodevelopment abnormalities were observed in three patients. These data indicated an possible association between *ZFHX3* and neurodevelopmental disorders. Further studies are needed to determine the whole phenotypic spectrum of *ZFHX3*.

Infantile spasms is generally a severe form of epilepsy with poor outcome.[36, 37] Clinically, approximate 34% of patients with spasms presented favorable outcomes,[25] of which the underlying mechanism is unknown. Our recent study showed that GDS is associated with the natural course and outcomes of diseases.[26-28] In this study, the two patients with earlier spasms presented favorable outcomes. Studies on the temporal expression stage showed that *ZFHX3* orthologs were highly expressed in the embryonic stage and decreased dramatically in childhood, which is potentially one of the explanations for the favorable outcomes. Considering that the expression of *ZFHX3* was slightly increased around 30 years old, long-term follow-up is required to observe the future evolution of disease.

In conclusion, this study suggested that *ZFHX3* is potentially a novel causative gene of partial epilepsy and infant spasms. The correlation between the outcome and gene expression stage provided insight into the underlying mechanism of the natural course of illness, potentially being helpful in management of the patients.

Acknowledgements

We thank the patients and their families for participating in this study. We thank Prof. Liu Ji-Yong for donating *tub-Gal4* driver line and *Canton-s* line flies; and Schrödinger (New York, NY, USA) for supporting open-source PyMOL software.

Funding

This study was supported by grants from the National Natural Science Foundation of China (No. 82271505 to W.-P.L. No. 81870903 to Y.-H.Y); Science and Technology Project of Guangzhou (Grant Nos 202201020106 to Liu Xiaorong, 33121030, and

33121173), Multi-center Clinical Research Fund Project of the Second Affiliated Hospital of Guangzhou Medical University (Grant Nos 010G271099, 2020-LCYJ-DZX-03 to Liao Weiping, and 2021-LCYJ-DZX-02 to Liu Xiaorong), and Scientific Research Project of Guangzhou Education Bureau (Grant No. 202235395 to Liu Xiaorong). The funders had no role in study design, data collection and analysis and decision to publish or in manuscript preparation.

Competing interests

The authors declare no competing interests.

Supplementary material

Supplementary table 1. Primer sequences used in this study.

Supplementary table 2. Genetic feature of *ZFHX3* variants identified in this study.

Supplementary table 3. Previously reported *ZFHX3* variants and associated phenotypes.

Supplementary reference (only for review). *PCDH15* in self-limited partial epilepsy and the mechanism of distinct phenotypes (unpublished).

References

1. Jung CG, Kim HJ, Kawaguchi M, et al. Homeotic factor ATBF1 induces the cell cycle arrest associated with neuronal differentiation. *Development*. Dec 2005;132(23):5137-45. doi:10.1242/dev.02098
2. Ishii Y, Kawaguchi M, Takagawa K, et al. ATBF1-A protein, but not ATBF1-B, is preferentially expressed in developing rat brain. *J Comp Neurol*. Oct 6 2003;465(1):57-71. doi:10.1002/cne.10807
3. Jung CG, Uhm KO, Miura Y, et al. Beta-amyloid increases the expression level of ATBF1 responsible for death in cultured cortical neurons. *Molecular neurodegeneration*. Jul 5 2011;6:47. doi:10.1186/1750-1326-6-47
4. Kim TS, Kawaguchi M, Suzuki M, et al. The ZFHX3 (ATBF1) transcription factor induces PDGFRB, which activates ATM in the cytoplasm to protect cerebellar neurons from oxidative stress. *Dis Model Mech*. Nov-Dec 2010;3(11-12):752-62.

doi:10.1242/dmm.004689
5. Sun X, Fu X, Li J, et al. Heterozygous deletion of Atbf1 by the Cre-loxP system in mice causes preweaning mortality. *Genesis*. Nov 2012;50(11):819-27. doi:10.1002/dvg.22041
6. Fuller TD, Westfall TA, Das T, Dawson DV, Slusarski DC. High-throughput behavioral assay to investigate seizure sensitivity in zebrafish implicates ZFHX3 in epilepsy. *J Neurogenet*. Mar - Jun 2018;32(2):92-105. doi:10.1080/01677063.2018.1445247
7. Wang JY, Wang J, Lu XG, et al. Recessive PKD1 Mutations Are Associated With Febrile Seizures and Epilepsy With Antecedent Febrile Seizures and the Genotype-Phenotype Correlation. *Front Mol Neurosci*. 2022;15:861159. doi:10.3389/fnmol.2022.861159
8. Chen Z, Luo S, Liu ZG, et al. CELSR1 variants are associated with partial epilepsy of childhood. *Am J Med Genet B Neuropsychiatr Genet*. Oct 2022;189(7-8):247-256. doi:10.1002/ajmg.b.32916
9. Yan HJ, He YY, Jin L, Guo Q, Zhou JH, Luo S. Expanding the phenotypic spectrum of KCNK4: From syndromic neurodevelopmental disorder to rolandic epilepsy. *Front Mol Neurosci*. 2022;15:1081097. doi:10.3389/fnmol.2022.1081097
10. Luo S, Liu ZG, Wang J, et al. Recessive LAMA5 Variants Associated With Partial Epilepsy and Spasms in Infancy. *Front Mol Neurosci*. 2022;15:825390. doi:10.3389/fnmol.2022.825390
11. Tian Y, Zhai QX, Li XJ, et al. ATP6V0C Is Associated With Febrile Seizures and Epilepsy With Febrile Seizures Plus. *Front Mol Neurosci*. 2022;15:889534. doi:10.3389/fnmol.2022.889534
12. Martin HC, Jones WD, McIntyre R, et al. Quantifying the contribution of recessive coding variation to developmental disorders. *Science*. Dec 7 2018;362(6419):1161-1164. doi:10.1126/science.aar6731
13. Converge c. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature*. Jul 30 2015;523(7562):588-91. doi:10.1038/nature14659
14. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. May 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7
15. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res*. Jan 4 2017;45(D1):D840-D845. doi:10.1093/nar/gkw971
16. Epi25 Collaborative. Electronic address sbuea, Epi C. Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. *Am J Hum Genet*. Aug 1 2019;105(2):267-282. doi:10.1016/j.ajhg.2019.05.020
17. Liu CQ, Qu XC, He MF, et al. Efficient strategies based on behavioral and electrophysiological methods for epilepsy-related gene screening in the Drosophila model. *Front Mol Neurosci*. 2023;16:1121877. doi:10.3389/fnmol.2023.1121877
18. Parker L, Howlett IC, Rusan ZM, Tanouye MA. Seizure and epilepsy: studies of

seizure disorders in Drosophila. *Int Rev Neurobiol*. 2011;99:1-21. doi:10.1016/B978-0-12-387003-2.00001-X
19. Bayat A, Liu Z, Luo S, et al. A new neurodevelopmental disorder linked to heterozygous variants in UNC79. *Genet Med*. May 11 2023:100894. doi:10.1016/j.gim.2023.100894
20. Stone B, Burke B, Pathakamuri J, Coleman J, Kuebler D. A low-cost method for analyzing seizure-like activity and movement in Drosophila. *J Vis Exp*. Feb 19 2014;(84):e51460. doi:10.3791/51460
21. Gu H, O'Dowd DK. Cholinergic synaptic transmission in adult Drosophila Kenyon cells in situ. *J Neurosci*. Jan 4 2006;26(1):265-72. doi:10.1523/JNEUROSCI.4109-05.2006
22. Qiao J, Yang S, Geng H, Yung WH, Ke Y. Input-timing-dependent plasticity at incoming synapses of the mushroom body facilitates olfactory learning in Drosophila. *Curr Biol*. Nov 21 2022;32(22):4869-4880 e4. doi:10.1016/j.cub.2022.09.054
23. Wang J, Lin ZJ, Liu L, et al. Epilepsy-associated genes. *Seizure*. Jan 2017;44:11-20. doi:10.1016/j.seizure.2016.11.030
24. Ran D, Cai S, Wu H, Gu H. Di (2-ethylhexyl) phthalate modulates cholinergic mini-presynaptic transmission of projection neurons in Drosophila antennal lobe. *Food Chem Toxicol*. Sep 2012;50(9):3291-7. doi:10.1016/j.fct.2012.03.070
25. Chourasia N, Yuskaitis CJ, Libenson MH, et al. Infantile spasms: Assessing the diagnostic yield of an institutional guideline and the impact of etiology on long-term treatment response. *Epilepsia*. May 2022;63(5):1164-1176. doi:10.1111/epi.17209
26. Zou DF, Li XY, Lu XG, et al. Association of FAT1 with focal epilepsy and correlation between seizure relapse and gene expression stage. *Seizure*. Mar 11 2023;doi:10.1016/j.seizure.2023.03.003
27. He N, Li B, Lin ZJ, Zhou P, Su T, Liao WP. Common genetic epilepsies, pathogenicity of genes/variants, and genetic dependence. *Seizure*. Jul 2023;109:38-39. doi:10.1016/j.seizure.2023.05.008
28. Lin Si-Mei ST, Qiao Jing-Da, et al. PCDH15 in self-limited partial epilepsy and the mechanism of distinct phenotypes (unpublished results).
29. Collins RL, Glessner JT, Porcu E, et al. A cross-disorder dosage sensitivity map of the human genome. *Cell*. Aug 4 2022;185(16):3041-3055 e25. doi:10.1016/j.cell.2022.06.036
30. Al-Naama N, Mackeh R, Kino T. C(2)H(2)-Type Zinc Finger Proteins in Brain Development, Neurodevelopmental, and Other Neuropsychiatric Disorders: Systematic Literature-Based Analysis. *Front Neurol*. 2020;11:32. doi:10.3389/fneur.2020.00032
31. Wang T, Guo H, Xiong B, et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nat Commun*. Nov 8 2016;7:13316. doi:10.1038/ncomms13316
32. Hashimoto R, Nakazawa T, Tsurusaki Y, et al. Whole-exome sequencing and neurite outgrowth analysis in autism spectrum disorder. *J Hum Genet*. Mar

2016;61(3):199-206. doi:10.1038/jhg.2015.141
33. Kosmicki JA, Samocha KE, Howrigan DP, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nat Genet*. Apr 2017;49(4):504-510. doi:10.1038/ng.3789
34. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. Feb 23 2017;542(7642):433-438. doi:10.1038/nature21062
35. Epi KC, Epilepsy Phenome/Genome P, Allen AS, et al. De novo mutations in epileptic encephalopathies. *Nature*. Sep 12 2013;501(7466):217-21. doi:10.1038/nature12439
36. Riikonen R. Infantile Spasms: Outcome in Clinical Studies. *Pediatr Neurol*. Jul 2020;108:54-64. doi:10.1016/j.pediatrneurol.2020.01.015
37. Zuberi SM, Wirrell E, Yozawitz E, et al. ILAE classification and definition of epilepsy syndromes with onset in neonates and infants: Position statement by the ILAE Task Force on Nosology and Definitions. *Epilepsia*. Jun 2022;63(6):1349-1397. doi:10.1111/epi.17239
38. Turner TN, Wilfert AB, Bakken TE, et al. Sex-Based Analysis of De Novo Variants in Neurodevelopmental Disorders. *Am J Hum Genet*. Dec 5 2019;105(6):1274-1285. doi:10.1016/j.ajhg.2019.11.003

Table 1. Clinical feature of the individuals with *ZFHX3* variants

Case	Variants (NM_006885)	Sex	Age	Onset age	Sz course	Sz-free duration	Effective AEDs	EEG	Brain MRI	Development	Diagnosis
Case 1	c.314C>T/p.Pro105Leu c.2282G>C/p.Gly761Ala	M	juvenile	childhood	sGTCS, 1-2 times/yr	1 yr	-	Bilateral spike-slow waves with frontal dominance	Normal	Normal	PE
Case 2	c.2419G>A/p.Glu807Lys c.10853C>A/p.Pro3618Gln	F	childhood	childhood	sGTCS, 1-5 times/mo	1.5 yr	LTG	Right frontal-centro-temporal sharp-slow waves	Normal	Normal	PE
Case 3	c.2671T>C/p.Phe891Leu c.9583_9584insT/ p.Pro3195LeufsTer44	M	childhood	toddler	CPS, 2-3 times/yr	1 yr	OXC	Spike-slow waves in bilateral frontal-central area with right dominance	Normal	Normal	PE
Case 4	c.2686G>A/p.Ala896Thr c.10439C>T/p.Ala3480Val	M	toddler	infant	Spasms, CPS >10 times/day	2 yr	LEV, LTG	Poly-spike-slow waves	Normal	GDD	Spasms, PE
Case 5	c.4125_4127del/p.Ala1376del c.10853C>A/p.Pro3618Gln	M	childhood	childhood	sGTCS, 3 times; CPS, twice	5 yr	OXC, VPA	Right temporal spike-slow waves	Normal	Normal	PE
Case 6	c.5152A>C/p.Met1718Leu c.10510G>C/p.Val3504Leu	M	childhood	toddler	Spasms 5-7 times/day for 4 mo; then CPS 0-10 times/day	9 yr	VPA, LEV, LTG	Right frontal spike-slow waves with tendency of generalization	Normal	GDD	Spasms, PE
Case 7	c.6161C>T/p.Ala2054Val c.10445G>T/p.Ser3482Ile	M	childhood	infant	FS, 3 times in infant; sGTCS and CPS twice/week; NCSE 2-3 times in childhood	4 mo	VPA, PER, LEV	Left frontal, temporal, and occipital 1.5-2Hz spike-slow waves; Diffused 1.5-2.5Hz spike-slow waves	Normal	GDD	PE
Case 8	c.10445G>T/p.Ser3482Ile c.10853C>A/p.Pro3618Gln	F	childhood	childhood	sGTCS, 2 times for 2 yr	1.5 yr	TPM	Bilateral occipital-temporal sharp-slow waves	Normal	Normal	PE

Abbreviations: AEDs, antiepileptic drugs; CPS, complex partial seizures; GDD, globally developmental delay; EEG, electroencephalogram; FS, febrile seizures; LEV, levetiracetam; LTG, lamotrigine; NCSE, nonconvulsive status epilepticus; OXC, oxcarbazepine; PE, partial epilepsy; PER, Perampanel; Sz, seizures; sGTCS, secondary generalized tonic-clonic seizures; TPM, topamax; VPA, valproate; mo, month; yr, year

Table 2. Analysis of the aggregate frequency of *ZFHX3* variants identified in this study

Identified *ZFHX3* variants	Allele count/number in this study	Allele count/number in controls of gnomAD-all populations	Allele count/number in controls of gnomAD-East Asian populations	Allele count/number in ExAC populations	Allele count/number in controls of Epi25 WES browser	Number of Homozygotes in controls of gnomAD
c.2419G>A/p.Glu807Lys	1/756	10/109406	0/9046	6/121404	1/66888	0
c.10853C>A/p.Pro3618Gln	3/756	35/116270	35/9678	23/112466	5/66408	0
c.5152A>C/p.Met1718Leu	1/756	2/109258	2/9022	1/119740	1/66886	0
c.10510G>C/p.Val3504Leu	1/756	−/−	−/−	−/−	−/−	-
c.10445G>T/p.Ser3482Ile	2/756	30/119978	30/9938	25/117840	9/66884	0
c.6161C>T/p.Ala2054Val	1/756	2/105822	1/8870	2/115778	−/−	0
c.2686G>A/p.Ala896Thr	1/756	24/108908	1/9046	17/119382	2/66868	0
c.10439C>T/p.Ala3480Val	1/756	7/119956	7/9942	5/118166	0/66884	0
c.4125_4127del/p.Ala1376del	1/756	8/120262	8/9962	6/121388	0/66888	0
c.314C>T/p.Pro105Leu	1/756	34/112352	25/9804	23/104844	15/66264	0
c.2282G>C/p.Gly761Ala	1/756	22/109332	11/9042	16/119178	10/66888	0
c.2671T>C/p.Phe891Leu	1/756	−/−	−/−	−/−	−/−	-
c.9583_9584insT/p.Pro3195LeufsTer44	1/756	8/119370	2/9686	60/118664	−/−	0
Total	16/756	182/120262	122/9962	184/121404	43/66888	0
P value[†]		2.3×10^{-13}	4.36×10^{-2}	2.34×10^{-13}	3.56×10^{-18}	
OR (95% CI)		14.27 (7.94-23.95)	1.74 (0.96-2.97)	14.25 (7.93-23.91)	33.61 (18.85-59.94)	

[†]P values and odds ratio were estimated with a 2-sided Fisher's exact test.

Abbreviations: CI, confidence interval; gnomAD, Genome Aggregation Database; ExAC, Exome Aggregation Consortium; Epi25, a whole-exome sequencing case-control study of epilepsy; OR, odds ratio

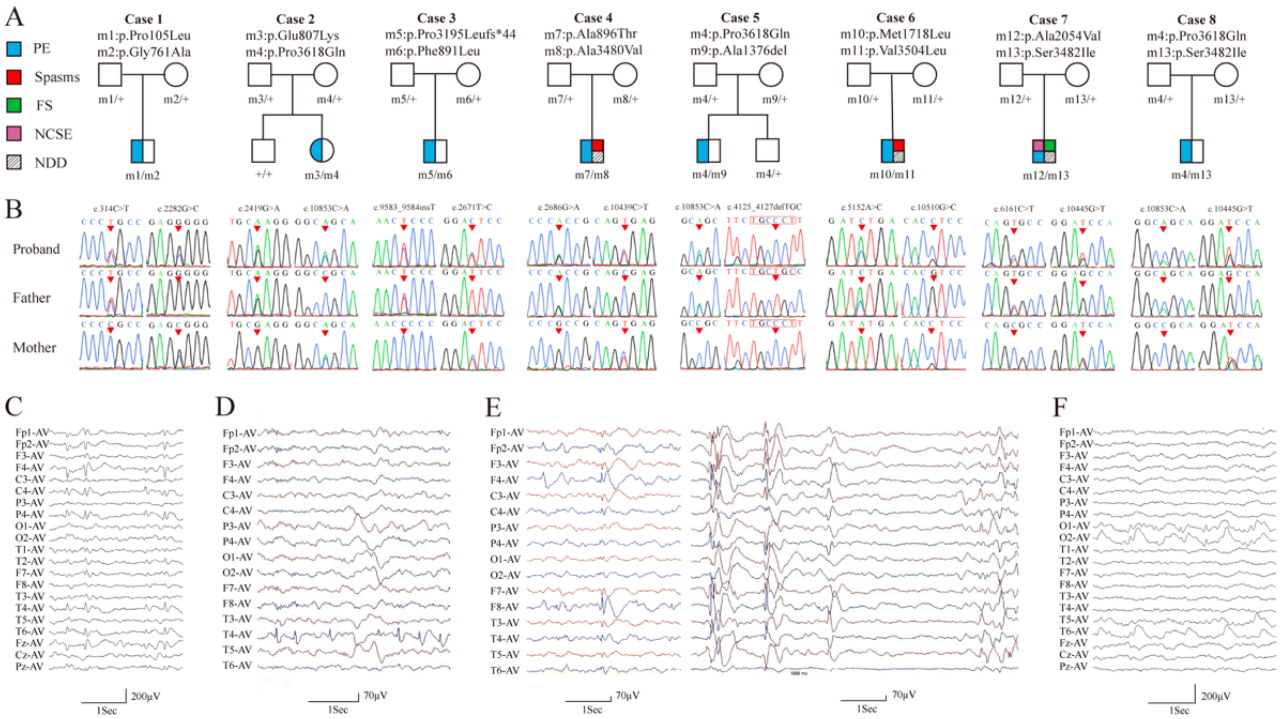

Figure 1 Genetic data and representative EEG recordings of the cases with *ZFHX3* variants. (A) Pedigrees of eight cases with compound heterozygous *ZFHX3* variants and their corresponding phenotypes. PE, partial epilepsy; FS, febrile seizures; NCSE, nonconvulsive status epilepticus; NDD, neurodevelopmental delay. (B) DNA sequencing chromatogram of *ZFHX3* variants. Arrows indicate the position of the mutations. (C) EEG of case 2 showed right frontal-centro-temporal sharp-slow waves. (D) EEG of case 5 showed right temporal spike-slow waves. (E) EEG of case 6 showed right frontal spike-slow waves (left) and generalized poly-spike-slow waves (right). (F) EEG of case 8 showed bilateral occipital-temporal sharp-slow waves.

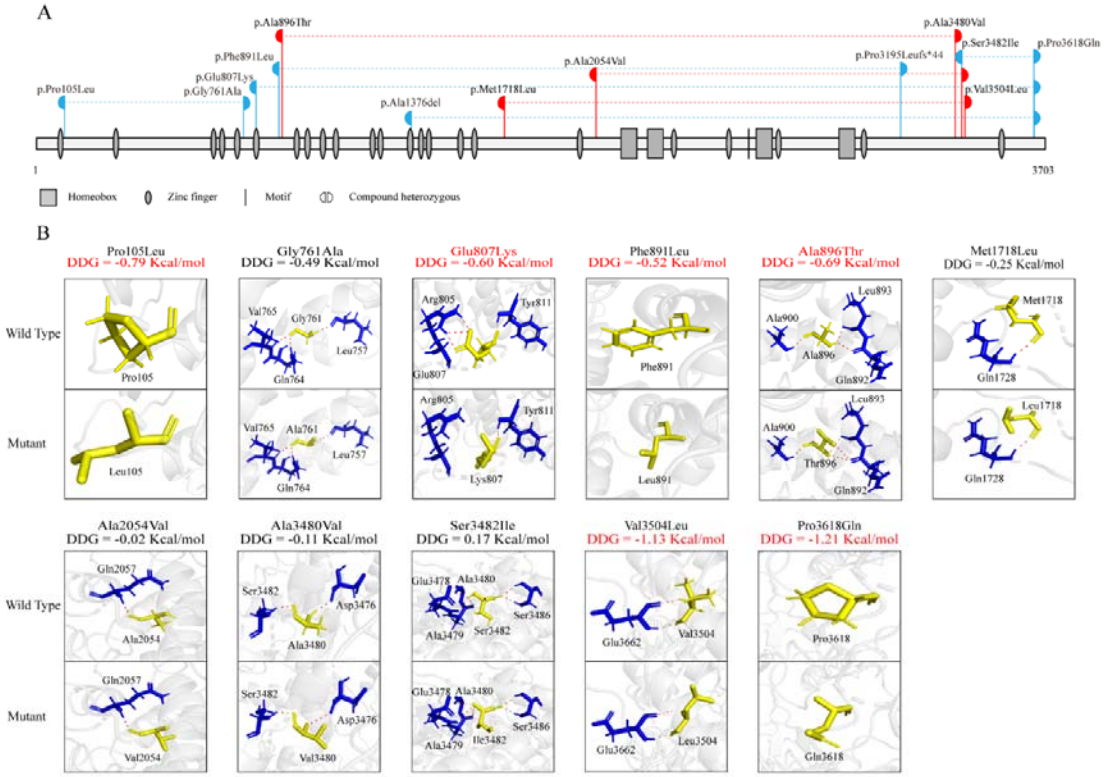

Figure 2 Schematic illustration of *ZFHX3* variants. (A) Schematic illustration of the Zinc finger homeobox protein 3 and locations of *ZFHX3* variants identified in this study. Two variants of the same height were a pair of biallelic variants. Red color represented cases with severe partial epilepsy with neurodevelopmental abnormities. Bule color represented cases with mild partial epilepsy. (B) Hydrogen bond changes (Red) and free energy stability changes (△△G, Kcal/mol) value of the variants from the present study.

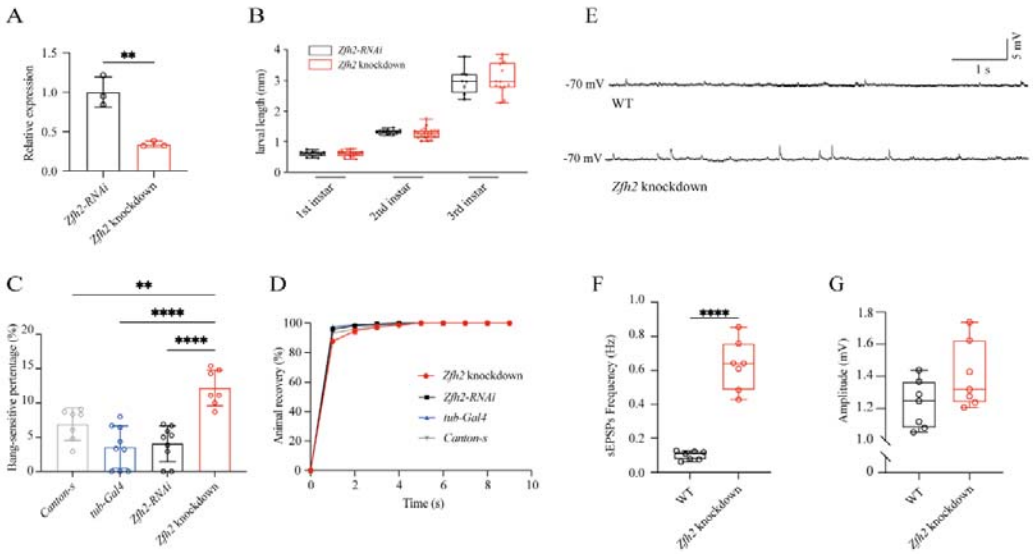

Figure 3 Functional studies on *Zfh2* knockdown flies (A) Relative *Zfh2* mRNA expression levels of the knockdown flies and the *UAS-Zfh2-RNAi* controls. (B) The body length of *Zfh2* knockdown and *UAS-Zfh2-RNAi* control larvae. (C) Seizure-like behaviors occurred at a higher rate in *Zfh2* knockdown flies (*tub-Gal4* > *UAS-Zfh2-RNAi*) than in *UAS-Zfh2-RNAi* line [12.20 ± 0.99% (n = 7) vs 4.06 ± 0.87% (n = 9); P < 0.0001], in *Canton-s* line[12.20 ± 0.99% (n = 7) vs 6.88 ± 0.89% (n = 7); P = 0.0017], and in *tub-Gal4* line[12.20 ± 0.99% (n = 7) vs 3.56 ± 1.03% (n = 7); P < 0.0001]. (D) The recovery time from seizure-like behaviors of *Zfh2* knockdown flies was longer than that of the controls. (E) Representative traces of projection neurons' sEPSPs in *Canton-s* wild-type (WT) flies and *Zfh2* knockdown flies. (F) The frequency of sEPSPs was significantly higher in *Zfh2* knockdown flies than that in wild-type flies [0.6312 ± 0.06 Hz (n = 7) vs 0.09907 ± 0.01 Hz (n = 7); P < 0.0001]. (G) There was no difference in amplitude of sEPSPs between *Zfh2* knockdown flies and wild-type flies [1.405 ± 0.2039 pA (n = 7) vs 1.228 ± 0.1486 pA (n = 7), P = 0.0888].

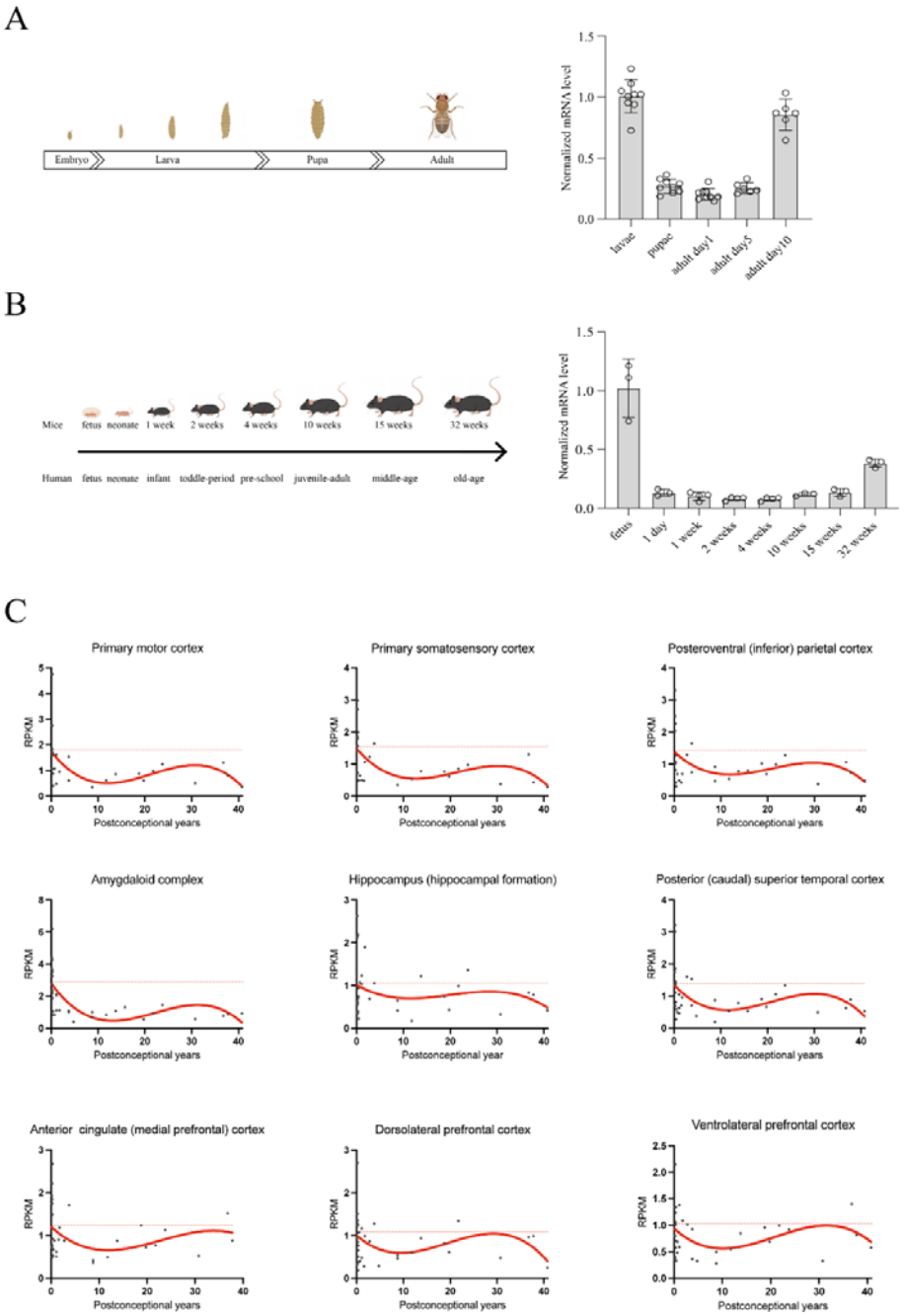

Figure 4 The temporal expression profile of *ZFHX3* orthologs. (A) Schematic illustration developmental stages of *Drosophila melanogaster* (left) and the temporal expressions of *Zfh2* in flies (right). The mRNA levels of *Zfh2* were examined by RT-qPCR. The *Zfh2* mRNA levels were normalized to the mRNA levels of the third instar larvae. (B) Schematic illustration development stages of mice (left) and the temporal expressions of *Zfhx3* in frontal cortex of mice (right). The *Zfhx3* mRNA levels were normalized to the mRNA levels of fetal mice. (C) Temporal expression pattern of *ZFHX3* in the human brain. RPKM, Reads Per Kilobase per Million mapped reads.

Table S1. Primer sequences of RT-qPCR

Primers for *Drosophila melanogaster*	
GAPDH-F:	5’-CGTCAACGATCCCTTCATCGATGTC-3’
GAPDH-R:	5’-CAGCACTGGCCCAGTTGATGTTG-3’
Zfh2-F:	5’-GCGATAACAAGAGGCAGGAGG-3’
Zfh2-R:	5’-CGGACACTACACGAAAAGAATGG-3’
Primers for mice	
GAPDH-F:	5’-TGTTCCTACCCCCAATGTGTC-3’
GAPDH-R:	5’-AAGTCGCAGGAGACAACCTG-3’
Zfhx3-F:	5’-GACAAACCCAGTAGCATGGAG-3’
Zfhx3-R:	5’-GTAGGTCTGGAGGCTGGAAAAG-3’

Table S2. Genetic feature of the individuals with *ZFHX3* variants

Case	Coordinate	cDNA change (NM_006885)	Protein change	Inheritance	MAF	SIFT	PP2_Var	Mutation-Taster	CADD	GERP++	phyloP	phastCons	SiPhy
Case 1	chr16:72993731	c.314C>T	p.Pro105Leu	paternal	2.74e-04	T (0.056)	B (0.001)	D (1.000)	T (17.01)	C (5.11)	NC (0.874)	NC (0.534)	C (12.922)
	chr16:72991763	c.2282G>C	p.Gly761Ala	maternal	1.54e-04	D (0.020)	P (0.719)	D (1.000)	T (19.17)	C (4.51)	C (3.856)	NC (0.964)	C (13.864)
Case 2	chr16:72991626	c.2419G>A	p.Glu807Lys	paternal	5.20e-05	T (0.153)	P (0.994)	D (1.000)	D (24.70)	C (5.52)	C (7.858)	C (1.000)	C (19.474)
	chr16:72821322	c.10853C>A	p.Pro3618Gln	maternal	2.39e-04	D (0.001)	P (0.669)	D (1.000)	T (17.82)	C (4.39)	C (6.706)	C (1.000)	C (16.590)
Case 3	chr16:72991374	c.2671T>C	p.Phe891Leu	maternal	-	T (0.064)	P (0.977)	D (1.000)	D (25.30)	C (5.52)	C (9.268)	C (1.000)	C (15.677)
	chr16:72822591	c.9583_9584insT	p.Pro3195_LeufsTer44	paternal	3.93e-05	-	-	-	-	-	-	-	-
Case 4	chr16:72991359	c.2686G>A	p.Ala896Thr	paternal	1.52e-04	T (0.352)	B (0.015)	D (1.000)	T (19.79)	NC (-4.08)	NC (0.020)	NC (0.933)	NC (8.169)
	chr16:72821736	c.10439C>T	p.Ala3480Val	maternal	5.38e-05	T (0.114)	P (0.970)	D (1.000)	D (22.80)	C (4.24)	C (7.999)	C (1.000)	C (16.983)
Case 5	chr16:72832453	c.4125_4127del	p.Ala1376del	maternal	3.18e-05	-	-	-	-	-	-	-	-
	chr16:72821322	c.10853C>A	p.Pro3618Gln	paternal	2.39e-04	D (0.001)	P (0.669)	D (1.000)	T (17.82)	C (4.39)	C (6.706)	C (1.000)	C (16.590)
Case 6	chr16:72831429	c.5152A>C	p.Met1718Leu	paternal	1.99e-05	T (0.665)	B (0.009)	D (1.000)	T (13.16)	C (5.74)	C (4.166)	C (1.000)	C (16.331)
	chr16:72821665	c.10510G>C	p.Val3504Leu	maternal	-	D (0.022)	P (0.537)	D (1.000)	D (21.00)	C (4.22)	C (4.744)	C (1.000)	C (16.949)
Case 7	chr16:72830420	c.6161C>T	p.Ala2054Val	paternal	1.85e-05	D (0.000)	P (0.998)	D (1.000)	D (23.10)	C (5.29)	C (9.940)	C (1.000)	C (19.309)
	chr16:72821730	c.10445G>T	p.Ser3482Ile	maternal	2.47e-04	D (0.015)	B (0.052)	D (0.950)	D (20.20)	C (4.24)	NC (1.133)	NC (0.797)	NC (8.770)
Case 8	chr16:72821730	c.10445G>T	p.Ser3482Ile	maternal	2.47e-04	D (0.015)	B (0.052)	D (0.950)	D (20.20)	C (4.24)	NC (1.133)	NC (0.797)	NC (8.770)
	chr16:72821322	c.10853C>A	p.Pro3618Gln	paternal	2.39e-04	D (0.001)	P (0.669)	D (1.000)	T (17.82)	C (4.39)	C (6.706)	C (1.000)	C (16.590)

Abbreviation: B, benign; C, conserved; CADD, Combined Annotation Dependent Depletion; D, damaging; GERP, Genomic Evolutionary Rate Profiling; MAF, minor allele frequency from general population of gnomAD; NC, nonconserved; P, probably_damaging; PP2, Polyphen2_HVAR; phastCons, Phylogenetic Analysis with Space/Time models conservation scoring and identification of conserved elements; phyloP, Computation of p-values for conservation or acceleration, either lineage-specific or across all branches; SIFT, Sorting Intolerant From Tolerant; T, tolerable.

Table S3. Previously reported *ZFHX3* variants and associated phenotypes

Case	Variants (NM_006885)	Original	MAF	Phenotypes	References
Case 1	c.2532C>A/His844Gln	*De novo*	-	ASD	Wang (2016)[31]
Case 2	c.3808C>T/His1270Tyr	*De novo*	-	ASD	Hashimoto (2016)[32]
Case 3	c.4857dupA/His1620Thrfs*42	*De novo*	-	ASD	Wang (2016)[31]
Case 4	c.9622dupC/Gln3208Profs*31	*De novo*	-	ASD	Kosmicki (2017)[33]
Case 5	c.6880C>T/Arg2294*	*De novo*	-	DD	McRae (2017)[34]
Case 6	c.8993G>A/Arg2998Gln	*De novo*	-	DD	McRae (2017)[34]
Case 7	c.3035_3036delAC/Tyr1012Serfs*34	*De novo*	-	DD	Turner (2019)[38]
Case 8	c.7562C>T/Ala2521Val	*De novo*	-	DEE	Epi4K (2013)[35]