Yang et al. *Zool. Res.* 2023, 44(5): 837−847
https://doi.org/10.24272/j.issn.2095-8137.2022.514

ZR | Zoological Research

Article Open Access

Population genetics of marmosets in Asian primate research centers and loci associated with epileptic risk revealed by whole-genome sequencing

XiangYu Yang[1,2,3], YaFei Mao[1,5], Xuan-Kai Wang[1], Dong-Ni Ma[1,3], Zhen Xu[6], Neng Gong[6], Barbara Henning[5], Xu Zhang[1,3], Guang He[1], Yong-Yong Shi[1,7], Evan E. Eichler[5,8], Zhi-Qiang Li[1,7,*], Eiki Takahashi[1,9,*], Wei-Dong Li[1,2,3,4,*]

[1] *Bio-X Institutes*, *Key Laboratory for the Genetics of Development and Neuropsychiatric Disorders (Ministry of Education)*, *Shanghai Key Laboratory of Psychotic Disorders*, *and Brain Science and Technology Research Center*, *Shanghai Jiao Tong University*, *Shanghai* 200240, *China*

[2] *Institute of Psychology and Behavioral Science*, *Shanghai Jiao Tong University*, *Shanghai* 200030, *China*

[3] *WLA Laboratories*, *World Laureates Association*, *Shanghai* 201203, *China*

[4] *Global Institute of Future Technology*, *Shanghai Jiao Tong University*, *Shanghai* 200240, *China*

[5] *Department of Genome Sciences*, *University of Washington School of Medicine*, *Seattle WA* 98195, *USA*

[6] *Institute of Neuroscience*, *Key Laboratory of Primate Neurobiology*, *CAS Center for Excellence in Brain Science and Intelligence Technology*, *Chinese Academy of Sciences*, *Shanghai* 200031, *China*

[7] *Affiliated Hospital of Qingdao University* & *Biomedical Sciences Institute of Qingdao University*, *Qingdao Branch of SJTU Bio-X Institutes*, *Qingdao University*, *Qingdao*, *Shandong* 266003, *China*

[8] *Howard Hughes Medical Institute*, *University of Washington*, *Seattle WA* 98195, *USA*

[9] *Department of Biomedicine*, *Graduate School of Medical Sciences*, *Kyushu University*, *Fukuoka* 812-8582, *Japan*

ABSTRACT

The common marmoset (*Callithrix jacchus*) has emerged as a valuable nonhuman primate model in biomedical research with the recent release of high-quality reference genome assemblies. Epileptic marmosets have been independently reported in two Asian primate research centers. Nevertheless, the population genetics within these primate centers and the specific genetic variants associated with epilepsy in marmosets have not yet been elucidated. Here, we characterized the genetic relationships and risk variants for epilepsy in 41 samples from two epileptic marmoset pedigrees using whole-genome sequencing. We identified 14 558 184 single nucleotide polymorphisms (SNPs) from the 41 samples and found higher chimerism levels in blood samples than in fingernail samples. Genetic analysis showed fourth-degree of relatedness among marmosets at the primate centers. In addition, SNP and copy number variation (CNV) analyses suggested that the WW domain-containing oxidoreductase (*WWOX*) and Tyrosine-protein phosphatase nonreceptor type 21 (*PTPN21*) genes may be associated with epilepsy in marmosets. Notably, *KCTD18-like* gene deletion was more common in epileptic marmosets than control marmosets. This study provides valuable population genomic resources for marmosets in two Asian primate centers. Genetic analyses identified a reasonable breeding strategy for genetic diversity maintenance in the two centers, while the case-control study revealed potential risk genes/variants associated with epilepsy in marmosets.

Keywords: Common marmoset (*Callithrix jacchus*); Population genetics; Whole-genome sequencing; Genetic chimerism; Epilepsy; Risk locus

INTRODUCTION

The common marmoset (*Callithrix jacchus*) is a small New World monkey native to the Atlantic Forest and the semiarid Caatinga biome of northeastern Brazil. Given its anatomical, physiological, and genetic similarities to humans, the common marmoset serves as a valuable nonhuman primate (NHP)

Received: 23 April 2023; Accepted: 26 July 2023; Online: 27 July 2023

Foundation items: This work was supported by the National Natural Science Foundation of China (82001372), National Key Research and Development Program of China (2018YFE0126700), Shanghai Jiao Tong University 2030 Initiative (WH510363001-7), Shanghai Municipal Commission of Science and Technology Program (21dz2210100), and Shanghai Education Commission Research and Innovation Program (2019-01-07-00-02-E00037), as well as a National Institutes of Health (NIH) grant (5R01HG002385) to E.E.E.

*Corresponding authors, E-mail: lizqsjtu@163.com; eiki.takahashi.sjt@gmail.com; liwd@sjtu.edu.cn

model for biomedical studies (Kishi et al., 2014; Malukiewicz et al., 2020; Shimogori et al., 2018; Youlatos, 2009). In addition, marmosets offer several advantages over other NHP models, including small body size (250–450 g), shorter generation time (approximately 2 years), and high fecundity (biannual births of twins or triplets, with triplets being more common in captivity) (Homman-Ludiye & Bourne, 2020). However, genetic chimerism remains a considerable obstacle in genetic studies of marmosets. While initially identified in marmoset hematopoietic cells and subsequently observed in somatic and germline tissues (Benirschke et al., 1962; Ross et al., 2007), the degree of DNA chimerism in different tissues remains a subject of debate (Sweeney et al., 2012).

Based on public resource data, more than 7 000 marmosets are currently housed in research facilities around the world (National Academies of Sciences, Engineering, and Medicine, 2019). Approximately 1 000 marmosets, the largest captive population in China, are housed at the Center for Excellence in Brain Science and Intelligence Technology of the Chinese Academy of Sciences (CAS) (Poo et al., 2016). In Japan, a substantial population of marmosets are maintained within the laboratory facilities of RIKEN, Central Institute for Experimental Animals (CIEA), and CLEA Japan, Inc. (Nishijima et al., 2012; Okano & Mitra, 2015; Okano et al., 2016). The first short-read genome assembly for marmosets was reported in 2014 (Marmoset Genome Sequencing and Analysis Consortium, 2014). Recently, Yang et al. used long-read technology with a trio-binning approach to produce a comprehensive chromosome-level, fully haplotype-resolved diploid genome assembly of the common marmoset (Yang et al., 2021). These assemblies offer considerable insights into the genomics of New World monkeys. Nevertheless, despite these advancements, there remains a notable lack of population genomics studies focusing on marmosets from distinct colonies (Sato et al., 2015).

During routine breeding procedures at the CAS and CLEA primate centers, researchers discovered that handling could trigger epileptic seizures in certain marmosets (Yang et al., 2022). In an earlier study, we characterized the phenotypes of epileptic marmosets maintained at the CLEA facility (Yang et al., 2022). The observed behavioral phenotype of these marmosets resembled the generalized tonic-clonic seizures typically seen in human epileptic patients. However, the collection of ictal phase EEG data, necessary for determining specific epilepsy types, proved challenging due to the limited duration and variability of the seizures. The epileptic marmosets from the CAS facility exhibited comparable behavioral phenotypes to their counterparts at CLEA, suggesting a stable inheritance pattern. Based on phenotypic mapping, we speculate that this form of epilepsy likely follows an autosomal dominant inheritance pattern in marmosets. However, genetic investigations on the pedigrees of marmosets with epilepsy at the CLEA and CAS facilities have not yet been conducted.

Epilepsy is a complex brain disorder characterized by recurrent seizures with diverse etiologies (Scharfman, 2007; Shorvon, 2011). Previous studies have shown that approximately one-third of human epilepsy cases result from a genetic predisposition (Mullen et al., 2018; Thomas & Berkovic, 2014). Genetic epilepsies are commonly categorized into three main groups: i.e., chromosomal disorders, Mendelian disorders, and non-Mendelian/complex disorders (Battaglia & Guerrini, 2005; Gardiner, 2000; Hardies et al., 2016; Robinson & Gardiner, 2004). Both inherited and *de novo* mutations have been reported in humans with epilepsy (Thijs et al., 2019), affecting various genes and loci, such as *SCN1A*, *STXBP1*, *CHD2*, *RBFOX1*, and 15q13.3, associated with ion channel regulation, receptor function, synaptic function, cell growth and proliferation, chromatin remodeling, and other cellular functions (Ellis et al., 2020; Helbig et al., 2009). These mutations have also been implicated in epilepsy phenotypes observed in animal models, including mice, dogs, and baboons (Ekenstedt et al., 2012; Kos et al., 2021; Marshall et al., 2021; Wielaender et al., 2017). Therefore, further study on the genetic risk variants associated with epileptic marmosets in the CLEA and CAS colonies is warranted.

In the current study, we utilized high-coverage whole-genome sequencing (WGS) to compare genetic chimerism in blood and fingernail tissues. We further investigated the population genetics of marmosets residing in the Chinese and Japanese captive colonies. In addition, we explored the potential risk variants associated with epilepsy in marmosets from the two colonies. Our study provides a valuable dataset on the population genetics of marmosets and offers insights into the presence of genetic variants potentially associated with epilepsy in captive marmoset colonies.

MATERIALS AND METHODS

Animals

The common marmosets (*Callithrix jacchus*) in this study were reared at the CLEA Marmoset Breeding Facility (CLEA Japan Inc., Gifu, Japan) under approval of the Institutional Animal Care and Use Committee (IACUC) of CLEA Japan Inc. (No.55-019CJ), and the CAS Marmoset Breeding Facility under approval of the Key Laboratory of Primate Neurobiology, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, China (CEBSIT-2022007). All animal experiments were approved by the IACUC of Shanghai Jiao Tong University (No.10645). Marmosets that exhibited at least one witnessed seizure induced by handling were classified as epileptic marmosets, and the epileptic phenotype was further verified through EEG analysis (Supplementary Figure S1). These marmosets displayed symptoms resembling the generalized tonic-clonic seizures observed in humans with epilepsy.

Sampling

We conducted a retrospective survey of seizures based on family history of marmosets that had experienced handling-evoked seizures (Yang et al., 2022). Samples, including peripheral blood, were collected from living marmosets, while fingernail tissue samples were collected from living or euthanized marmosets, depending on experimental and veterinary care considerations. A total of 41 samples were collected from 38 marmosets (16 euthanized marmosets and 22 living marmosets; 32 fingernail samples and nine blood samples). Peripheral blood samples were taken from the femoral/caudal vein without anesthesia using a heparinized syringe. For fingernail collection, 2-mm long fragments were clipped from the fingers and wiped with ethanol-soaked cotton, with large debris carefully removed using forceps. The collected samples were then stored on ice (4 ℃) until DNA extraction.

Whole-genome sequencing (WGS)

Whole-genome DNA was collected using a TIANam/Genomic DNA Kit (DP304-02, Tiangen Biotech, China). Qualified genomic DNA (OD260/280=1.8–2.0 and concentration≥20 ng/μL) was fragmented into ~350 bp segments, which were then end-polished, A-tailed, and ligated with adapters for Illumina sequencing, followed by polymerase chain reaction (PCR) amplification with the P5 and P7 primers. The final PCR products for library construction were purified (AMPure XP system, USA), followed by quality control assessment, including size distribution evaluation using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and molarity measurement using real-time PCR. Finally, the library was subjected to WGS using a PE150 sequencer (Illumina, USA) at Novogene (China). All individuals were sequenced at greater than 17.2× coverage (26.94±8.037), and sequences with a read depth>5 were used for single nucleotide polymorphism (SNP) calling.

Single-nucleotide variant (SNV) calling

The long-read genome assembly (Callithrix_jacchus_cj1700_1.1/calJac4) published by Washington University on the NCBI database and UCSC Genome Browser was used in this study. The BWAsoftare package (v0.7.5) (Li, 2013) was applied to map the WGS data to the marmoset reference genome (cj1700_1.1/calJac4). SAMtools (v1.9) (Li, 2011; Li et al., 2009) was then used to fix mate information and sort the mapped SAM files. All SNV calling pipelines were based on our previously published paper (Mao et al., 2018). Variants with an allele balance below 0.25 were excluded to minimize the influence of chimerism. HaplotypeCaller in GATK (v3.7.0) (McKenna et al., 2010) was used to call the SNVs and filter low-quality SNVs using the parameters QD<20.0; ReadPosRankSum<-8.0; filterName LowQualFilter_RPRS\DP<5; DP>200; QUAL<30; FS>60.0; MQ<50.0; MQRankSum<-12.5; HaplotypeScore>13.0; -rf UnmappedRead -rf DuplicateRead. SAMtools was also used to calculate the mapping coverage of each individual.

Nucleotide diversity (π), genetic divergence (F_{ST}), and absolute divergence (D_{XY}) of each chromosome were estimated using pixy (v1.2.7) in 100 kb windows. Analysis focused on fingernail samples, after three unrelated individuals were excluded, to compare genetic diversity between the two populations. The combined genotyping VCF file was used as input in Helmsman tools (Carlson et al., 2018) for mutation spectrum analysis. The ggplot2 program in R was used to plot the figures. Each mutation class was tested for each pair using the Wilcoxon rank-sum test.

Next, PLINK (v1.9) (Purcell et al., 2007) was used to calculate heterozygosity and perform principal component analysis (PCA). Allele frequency distribution was determined using R (v3.6.1). PLINK (v1.9) and Kinship-based INference for Genome-wide association studies (KING) (Manichaikul et al., 2010) were used to calculate identity by descent (IBD) within/between each marmoset group/pair of groups. After excluding nine blood samples and mis-sampled JF303, 31 fingernail samples were used for GWAS analysis (CLEA: 11 control marmosets/nine epileptic marmosets; CAS: six control marmosets/five epileptic marmosets). As some important nodes in the pedigrees were unavailable for typical linkage analysis, we performed association analysis using Linkage and Association Modeling in Pedigrees (LAMP), which uses a maximum-likelihood model to extract information on genetic linkage and associations from pedigree samples (Li et al., 2005, 2006). Case-control-based analysis was conducted. Unaffected siblings and unrelated unaffected individuals were treated as controls according to the LAMP manual, and the logarithm of odds (LOD) score was calculated to test for association (Li et al., 2006). We compared the likelihood maximized under the general model ($0\leq r^2\leq 1$) (L_{GM}) with the likelihood maximized under the null model ($r^2=0$) (L_{LE}) using the likelihood-ratio statistic $T_{LE}=2[\ln(L_{GM})-\ln(L_{LE})]$. Significance of the test statistic was assessed empirically by simulating marker genotypes under the null hypothesis and comparing the observed statistic with the simulated null distribution. In this analysis, we only used SNPs (including coding and noncoding regions) with a call rate>0.90. A total of 8 898 345 SNPs were retained, with a *P*-value of 1.0e-5 taken as the threshold for suggestive evidence of association. Short-read sequencing data were used to phase the heterozygous SNVs with WhatsHap (v1.4) (Martin et al., 2016; Patterson et al., 2015).

Copy number variation (CNV) calling

We compared the frequency of genes disrupted by CNV events between the epileptic and control marmosets. Mapped BAM files were used to call CNVs with both Manta (v1.5.0) (Chen et al., 2016) and Smoove (v0.2.5, https://github.com/brentp/smoove) callers (Pedersen, 2018). The caller VCFs were filtered to select: (1) deletions, duplications, and insertions; (2) events in autosomes or X chromosome; (3) events with more than 75% nonsegmental duplications (SegDups) and more than 75% nonshort tandem repeat bp; and (4) events overlapping exons by at least 1 bp. We merged events of the same type, genotype, and sample that overlapped with each other if they had 50% reciprocal overlap and breakpoints<10 kb apart. The CJ1700_v1.1 annotation files from the UCSC table browsers were used for gene annotation. For each gene and type of CNV call, case and control samples with at least one CNV event intersecting the gene by at least 1 bp were counted. Estimation of the frequency of CNV events per gene was performed for the callset from each CNV caller independently. Due to the different numbers of case and control samples in each population, Fisher's exact test was applied to examine whether the proportion of CNV events was higher in case samples than in control samples in each population. We used raw *P*<0.05 to select candidate genes enriched in CNV events in case samples. To validate the CNV events in candidate genes, we reconstructed a track hub (https://eichlerlab.gs.washington.edu/public/track_hubs/dgordon/marmoset_read_depth/hub.txt) with read depth tracks for each sample. Given the phylogenetic distance of marmosets from humans, we could not take the human genome assembly as a reference to run whole-genome shotgun detection (WSSD) for validation filtering. We manually validated the CNV calls that contained genes using read depth information in the track hub (available from the UCSC Genome Browser).

SNP and CNV validation

A total of 312 SNPs in 5–17 CLEA marmosets were validated by PCR amplification and Sanger sequencing. The primers used for PCR were designed with Primer-BLAST (Ye et al., 2012). The PCR analyses were performed in a 40 μL reaction containing 1.0 μL of genomic DNA, 2 μL of reverse and forward primers (10 μmol/L), 20 μL of 2×T8 High-Fidelity Master Mix (Tsingke, China), and 15 μL of ddH_2O. The PCR

amplification protocol included pre-denaturation at 98 °C for 1 min, followed by 35 cycles of denaturation for 10 s at 98 °C, annealing for 10 s at 60 °C, elongation at 72 °C for 40 s, and a final extension at 72 °C for 5 min, with subsequent cooling to 4 °C. The PCR products were then sequenced using an ABI 3730xl DNA Analyzer (Thermo Fisher Scientific, USA), and sequence alignment was conducted using DNAMAN (Lynnon Biosoft, Canada) and Chromas (Technelysium Pty, Australia).

For CNV validation, whole-genome DNA collected from the fingernails of each marmoset was used as a template. PCR amplification was performed using the ProFlex™ PCR system (Thermo Fisher Scientific, USA). The thermocycling conditions were: 94 °C for 2 min, followed by 32 cycles of 94 °C for 30 s, 58 °C for 30 s, 72 °C for 2 min, and a final extension of 72 °C for 2 min. DNA-free water was used as a negative control template following the same protocol described above. The PCR primers used for *KCTD18-like* deletion validation were 5'-GCTTTGTTCCCAGAGTT-3' and 5'-CAGTGCAAGTA CAATGAT-3'. The PCR products were analyzed by agarose gel electrophoresis. After electrophoresis, the images were recorded using a Tanon 2500 gel image system (Tanon, China).

RESULTS

Sequence coverage and heterozygosity in marmosets

We first collected data on research colonies of marmosets from a public database (Figure 1A; Supplementary Table S1), then investigated the population genetics of 27 individuals from the Japanese marmoset colony (CLEA, Yaotsu, Gifu; Figure 1B) and 11 individuals from the Chinese marmoset colony (CAS, Shanghai; Figure 1B). Blood and fingernail tissue samples were used to extract DNA for WGS. In total, we sequenced 2 982 Gb of data from all samples, with 33 individuals sequenced at over 20× coverage and three sequenced at over 40× (Figure 1C; Supplementary Table S2).

The Genome Analysis Toolkit (GATK, v3.7.0) was used to call SNPs for each sample and for all 41 genotyped samples, for a total of 14 558 184 SNPs. The heterozygosity of blood and fingernail samples was then compared to investigate which tissue had a lower level of genetic chimerism. Analysis revealed that the heterozygosity of blood samples was significantly higher than that of fingernail samples (Figure 1D, Wilcoxon rank sum test; Blood_CLEA vs. Fingernail_CAS: P=0.0007978; Blood_CLEA vs. Fingernail_CLEA: P=0.02456; Fingernail_CAS vs. Fingernail_CLEA: P=0.3267). In addition, a greater frequency of sites with an allele balance less than 0.3 was observed in blood samples than in fingernail samples (Supplementary Figure S2). Based on these results, we suggest that fingernail tissues are a better material than blood for population genetic studies of marmosets due to their lower chimerism.

Genetic difference between Japanese and Chinese marmoset pedigrees

Based on allele frequency analysis, 9 084 947 biallelic SNPs had a frequency exceeding 5%. These SNPs also exhibited a maximum missing genotyping rate of 25% (Figure 2A). To reduce sequencing bias from blood samples and low-allele frequency SNPs, we excluded SNPs with an allele frequency less than 5% and removed nine individuals that were only blood sampled from downstream analysis. Linkage pruning was performed on the remaining biallelic SNPs, resulting in 986 260 unlinked SNPs (10.86% of the original set), which were used for PCA (Figure 2B). Population structure analysis indicated that individuals from CAS clustered together, while individuals from CLEA clustered together based on the first component (5.55%) but not the second (3.29%, Supplementary Figure S3). This suggest that marmosets from CLEA have higher genetic diversity than those from CAS. We also calculated nucleotide diversity (π) for each colony (autosome: π_{CAS}=0.19±0.02, π_{CLEA}=0.19±0.02; Chr X: π_{CAS}=0.13±0.08, π_{CLEA}=0.18±0.07) as well as genetic and absolute divergence (F_{ST} and D_{XY}, respectively) between the colonies (Supplementary Figures S4, S5 and Table S3). In addition, mutation spectrum analysis was performed on the CAS and CLEA colonies, which showed that the spectra were nearly identical (Wilcoxon rank-sum test, Supplementary Figure S6A and Table S4).

Next, we determined the degree of relatedness between the CAS and CLEA colonies. To better estimate IBD and kinship coefficients, we focused on the 814 362 biallelic SNPs on autosomes without any missing genotypes. The PI-HAT calculated by IBD was significantly correlated with the kinship coefficient (R^2=0.88, P<1×10^{-16}), and the pairwise samples with first-degree of relatedness based on pedigree information had the largest PI-HAT and kinship coefficient values (Figure 2C, Supplementary Table S5), except for the JF303-related pairs. One possible explanation is that JF303 was mis-sampled during our preparation procedure, and thus, we excluded this sample from subsequent downstream analysis. As expected, the pairs of samples between the CAS and CLEA colonies had the lowest PI-HAT and kinship coefficient values. Based on kinship coefficients, CAS and CLEA colonies were predicted to have at least fourth-degree of relatedness.

Analysis of epileptic marmosets and controls

The epilepsy phenotypes of marmosets from the two pedigrees were independently observed. We next focused on whether the epileptic marmosets and controls exhibited the same mutation spectrum and significant genetic differentiation at certain loci. Therefore, we examined 9 084 947 biallelic SNPs and performed 96-trinucleotide mutation spectrum analysis on the epileptic marmosets and controls. Notably, results showed nearly identical 96-mutation spectra between the datasets (Wilcoxon rank-sum test, Supplementary Figure S6B and Table S4). Mutations (ACG→ATG/ATC→ACC) accounted for the largest portion due to the higher mutation rate on the CpG island (Supplementary Figures S6, S7).

As genetic variation (SNPs/CNVs) can contribute to epilepsy, we performed SNP- and CNV-based candidate gene screening. We also performed association analysis to search for SNPs that may contribute to epilepsy status. Because key nodes in the pedigrees were not available for sequencing, we performed association analysis using LAMP (Li et al., 2005, 2006). We identified 35 SNPs with a LOD score>5.0 and P<1×10^{-5} (LAMP likelihood-ratio test, Table 1; Figure 3A; Supplementary Figure S8 and Table S6). The sites that exceeded the threshold were mainly located in two protein-coding genes, WW domain-containing oxidoreductase (*WWOX)* and *Tyrosine-protein phosphatase nonreceptor type 21 (PTPN21)* (Figure 3B). Some SNPs in the CLEA-pedigree marmosets were validated through Sanger sequencing (Supplementary Table S6 and Figure S9). We also reconstructed the haplotypes through phasing information of

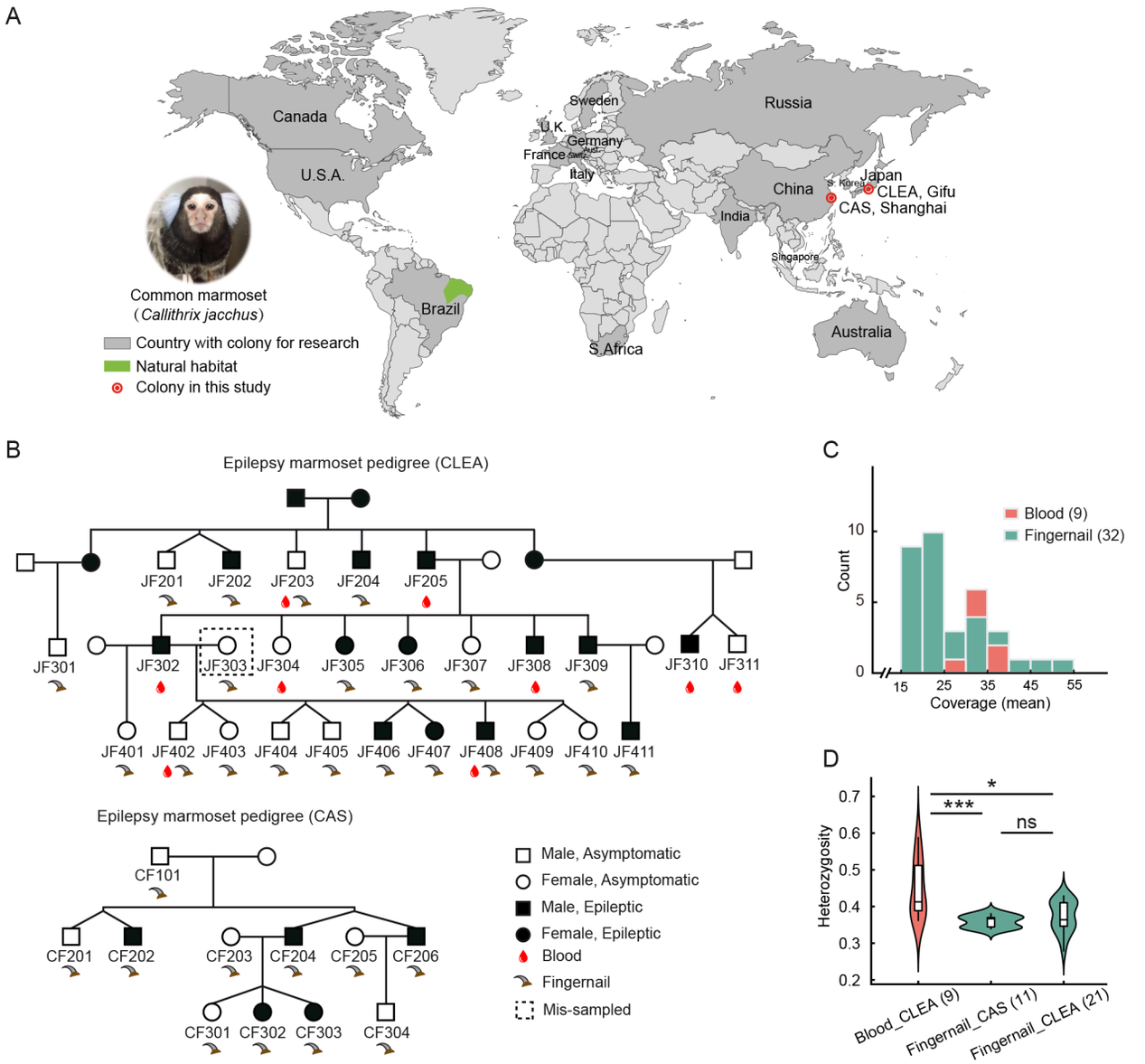

Figure 1 Research colony distribution of common marmosets and sequencing information of epileptic marmoset pedigrees

A: Worldwide geographic distribution of marmoset (*Callithrix jacchus*) research colonies based on the literature. B: Pedigrees with sequencing information for the Japanese (CLEA) and Chinese colonies (CAS). Mislabeled individual JF303 is marked with a dashed rectangle. C: Sequencing coverage plots of 41 whole-genome DNA samples. D: Violin plot of blood and fingernail sample heterozygosity. Wilcoxon rank-sum test. ns: Not significant; *: P<0.05; ***: P<0.001. Source data are provided in Supplementary Table S2.

SNVs. Several SNPs in the *WWOX* gene were confirmed in the same haplotype (Figure 3B, C; Supplementary Table S6).

We also examined CNVs in the samples using two independent methods (Manta v1.5.0 and Smoove v0.2.5). In total, 497 genes were disrupted by CNVs, with read depth (implemented in the marmoset genome browser track hub https://eichlerlab.gs.washington.edu/public/track_hubs/dgordon/marmoset_read_depth/hub.txt) then applied to validate 51 events for 100 randomly chosen genes (validation rate: 51%). The validation rate was similar to that reported in previous human CNV analyses (Coe et al., 2014; Cooper et al., 2011). Fisher's exact test was performed between the control and epileptic marmosets in different the pedigrees. We identified nine and six risk candidate genes in the CLEA and CAS pedigree groups, respectively. In the combined CLEA and CAS dataset, 16 candidate genes were called (Supplementary Figure S10 and Table S7). The integration of all results yielded a total of 16 risk candidate genes disrupted by deletions or duplications (Figure 4A; Supplementary Table S7).

We manually validated the 16 CNV-disrupted risk genes (Table 2). The successfully validated genes included three protein-coding genes. Notably, the *KCTD18-like* deletion was enriched in epileptic marmosets of the CLEA pedigree (Figure 4B, Fisher's exact test, P=0.012). KCTD18 is considered a highly conserved protein in primates (Supplementary Figure S11). The *KCTD18-like* gene deletion was experimentally validated in the marmosets. The tentative genotype was inferred based on the ratio of CNV region depth to average depth. PCR analysis determined the validation rates for the controls (100%, 11/11) and cases (85.7%, 6/7) (Figure 4D; Supplementary Figure S12 and Table S8). In addition, whole-genome shotgun detection showed that *KCTD18* has two alleles in humans and other great apes but

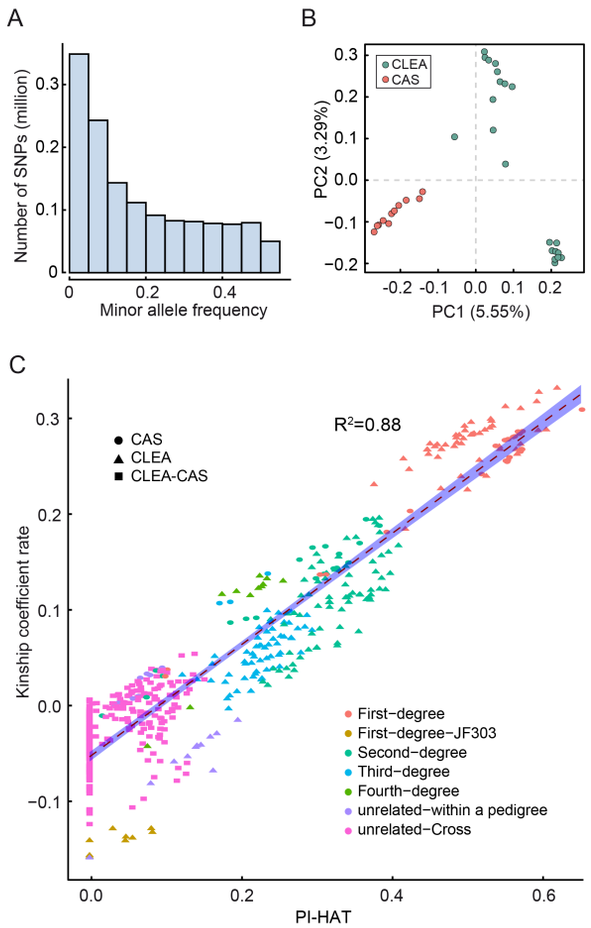

Figure 2 Population genetic analysis of CLEA and CAS colonies
A: Minor allele frequency distribution of SNPs in all marmoset samples (bins=0.05, *n*=41). B: Population structure assessed using PCA. Red: CAS population; Green: CLEA population; C: Relationship plot between 32 marmosets based on kinship coefficients measured using PI-HAT; y-axis and x-axis describe kinship coefficient and PI-HAT values, respectively. Kinship coefficient: IBD probability of a randomly selected allele at a particular locus in two individuals; PI-HAT (Proportion IBD): coefficient of relatedness among all possible pairwise comparisons of marmoset individuals. Source data are provided in Supplementary Table S5.

three in marmosets (Figure 4C; Supplementary Figure S13).

DISCUSSION

Understanding the population genetics of marmosets living in colonies is crucial for biomedical research (Warren et al., 2020). This information forms the basis for appropriate breeding strategies to ensure adequate population diversity in research populations, which is critical for maintaining immunity and reproductive success of captive marmosets. Here, we described two captive marmoset pedigrees and conducted high-coverage WGS of a total of 41 samples. First, the sequencing data showed that blood samples were more likely to contain DNA information from two individuals due to a higher level of DNA chimerism. This increased heterozygosity suggests that blood samples may not be an ideal sampling method for WGS in future marmoset genetic studies, such as *de novo* mutation rate estimation and trio-based genome assembly. Additionally, we analyzed the population structure and degree of relatedness between the two colonies. Our findings indicated that the individuals in the two colonies shared at least fourth-degree of relatedness. In addition, individuals in the CLEA colony exhibited higher genetic diversity than those in the CAS colony. Given the impacts of long-term inbreeding on genetic diversity within closed colonies, these results suggest that interbreeding the two colonies may be a viable strategy to increase the genetic diversity of captive research colonies and reduce genetic disease caused by inbreeding (National Academies of Sciences, Engineering, and Medicine, 2019). Finally, we conducted a preliminary analysis to estimate the genetic differences between the epileptic and control marmosets in the two pedigrees.

The field of marmoset research is witnessing a gradual increase worldwide, as evidenced by our survey showing that nearly 20 countries maintain captive marmosets for research purposes, with even more countries involved in collaborative marmoset-based research. However, a long-standing question in marmoset studies revolves around the patterns of chimerism observed in different tissues. Initially, hematopoietic chimerism was discovered in marmoset twins (Benirschke et al., 1962), with later studies also reporting the presence of chimerism in both somatic and germline tissues (Ross et al., 2007). However, Sweeney et al. (2012) proposed that chimerism exists only in hematopoietic cells, suggesting that the chimerism found in other tissues is the result of blood contamination. Therefore, the extent of DNA chimerism across different tissues remains uncertain. In the present study, we demonstrated that the level of heterozygosity in fingernails was much lower than that in blood, indicating potential variability in the level of chimerism among different tissues in marmosets. The fingernail samples used for whole-genome DNA extraction emerged as a preferable source material for genetic analyses in marmosets. Nevertheless, it should be noted that the current data and samples are insufficient for conducting a quantitative analysis of chimerism. Thus, future studies should explore quantitative analysis of chimerism in different tissues of marmosets using more advanced technology, such as single-cell or long-read sequencing.

After ensuring a lower level of chimerism in the fingernail samples, we used these data to analyze population genetics in the two colonies. Our results indicated that the CLEA colony had greater diversity than the CAS colony. Inbreeding may lead to a reduction in individual adaptability, fecundity, and immunity in a colony with low genetic diversity (National Academies of Sciences, Engineering, and Medicine, 2019). Thus, given the prolonged inbreeding in each Asian marmoset center, we propose that an exchange of genetic information between the two populations would be beneficial for maintaining diversity, although attention must also be paid to the risk of genetic disease.

We also conducted a preliminary search for candidate genes responsible for epilepsy in the two colonies. Although we did not find any signals that reached genome-wide significance ($P<5\times10^{-8}$) in the SNP association study, results suggested that *WWOX*, *PTPN21*, and *KCTD18-like* may be candidates with pathogenic mutations in epileptic marmosets. *WWOX* is a putative tumor suppressor that encodes a member of the short-chain dehydrogenases/reductase (SDR) protein family, which is also related to epileptic encephalopathy (Johannsen et al., 2018; Repudi et al., 2021). *WWOX* plays an essential role in balancing neocortical excitability, and its deletion can lead to spontaneous seizures

Table 1 Single-nucleotide variant (SNV) candidate loci associated with epilepsy in marmosets

Chr	NCBI ID	Gene	BP	REF	ALT	LOD	*P* value	CLEA			CAS		
								R/R	R/A	A/A	R/R	R/A	A/A
chr10	NC_048392.1	*ADCK1*	109317561	A	G	5.4	3.90E-06	0/2	3/7	8/0	0/0	1/5	5/0
chr10	NC_048392.1	*PTPN21*	119874705	C	G	5.18	6.60E-06	6/0	3/9	2/0	2/0	1/5	3/0
chr10	NC_048392.1	*PTPN21*	119890620	C	T	5.18	6.60E-06	6/0	3/9	2/0	2/0	1/5	3/0
chr10	NC_048392.1	*PTPN21*	119897794	C	T	5.89	1.30E-06	6/0	3/9	2/0	2/0	0/5	4/0
chr10	NC_048392.1	*PTPN21*	119901287	C	T	5.89	1.30E-06	6/0	3/9	2/0	2/0	0/5	4/0
chr10	NC_048392.1	*PTPN21*	119901688	C	T	5.89	1.30E-06	6/0	3/9	2/0	2/0	0/5	4/0
chr10	NC_048392.1	*PTPN21*	119901711	A	G	5.89	1.30E-06	6/0	3/9	2/0	2/0	0/5	4/0
chr10	NC_048392.1	*PTPN21*	119907345	C	T	5.18	6.60E-06	6/0	3/9	2/0	2/0	1/5	3/0
chr10	NC_048392.1	*PTPN21*	119916320	T	C	5.18	6.60E-06	6/0	3/9	2/0	2/0	1/5	3/0
chr10	NC_048392.1	*EML5*	120220737	T	C	5.89	1.30E-06	6/0	3/9	2/0	2/0	0/5	4/0
chr20	NC_048402.1	*WWOX*	34231005	T	A	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34231014	A	G	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34231295	C	T	5.2	6.30E-06	10/0	1/9	0/0	3/0	3/5	0/0
chr20	NC_048402.1	*WWOX*	34231314	T	C	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34231612	C	T	5.2	6.30E-06	10/0	1/9	0/0	3/0	3/5	0/0
chr20	NC_048402.1	*WWOX*	34232290	A	T	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34234379	C	T	5.63	2.40E-06	10/0	1/6	0/3	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34234735	G	A	5.2	6.30E-06	10/0	1/9	0/0	3/0	3/5	0/0
chr20	NC_048402.1	*WWOX*	34236189	T	A	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34238010	T	G	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34238067	C	T	5.3	5.00E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34243849	C	T	5.04	9.20E-06	10/0	1/8	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34243891	G	A	5.63	2.40E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34249323	G	A	5.63	2.40E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34250685	G	A	5.51	3.10E-06	10/0	1/9	0/0	3/0	3/5	0/0
chr20	NC_048402.1	*WWOX*	34257180	C	T	5.13	7.40E-06	9/0	2/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*WWOX*	34262270	C	T	5.63	2.40E-06	10/0	1/9	0/0	3/0	3/4	0/1
chr20	NC_048402.1	*MPHOSPH6*	37343079	T	C	5.48	3.30E-06	0/1	10/0	1/8	3/5	3/0	0/0
chr4	NC_048386.1	*LOC118153086*	10041131	T	C	5.06	8.70E-06	8/0	2/9	0/0	4/0	2/4	0/1

Chr: chromosome; BP: base pair; REF (R): reference; ALT (A): alteration; LOD: −log10(*P*-value); R/R: Wild type; R/A: heterozygote; A/A: homozygote; The numbers indicate the number of each genotype in Control group/Case group; Details are provided in Supplementary Table S6); The annotation (synonymous, nonsynonymous, intronic, etc.) of SNPs are provided in Supplementary Table S6.

in mice (Breton et al., 2021). During neurogenesis and early brain development, loss of *WWOX* can affect different cytoskeletal components and alter prenatal cortical development (Iacomino et al., 2020). WWOX also binds with Tau via its C-terminal SDR domain and interacts with Tau-phosphorylating enzymes ERK, JNK, and GSK-3β, which may limit Alzheimer's disease (AD) progression (Hsu et al., 2021). *PTPN21* is involved in hereditary human diseases such as schizophrenia (Chen et al., 2011) and positively influences cortical neuronal survival and neuritic elongation via ErbB4/NRG3 signaling (Plani-Lam et al., 2015). However, evidence related to its potential role in epilepsy is limited, and further exploration is required to understand its involvement in epileptogenesis.

KCTD18-like deletion was likely enriched in the epileptic marmosets in the CLEA colony (Fisher's exact test, P=0.012). *KCTD18* belongs to the *KCTD* gene family, which is highly expressed in the brain (Usui et al., 2013). Human genetic analysis has shown that neurodevelopmental, neuropsychiatric, and neurodegenerative disorders are associated with the *KCTD* gene family (Liu et al., 2013; Teng et al., 2019). Notably, a patient with a *KCTD18* duplication presented with epilepsy (Usui et al., 2013). While recent study reported that *KCTD18* may serve as a regulator of fat cell proliferation (Kulyté et al., 2022), few studies have reported on the function of *KCTD18* in the neural system. *KCTD18-like* is a duplicated gene in marmosets as well as a single-exon gene. Interestingly, other single-exon functional genes exist in the *KCTD* gene family (e.g., *KCTD11* and *KCTD12*) (Grzybowska, 2012). The gene structure and protein functional domain similarity of the *KCTD* genes suggest that the *KCTD18-like* gene may be functional. However, we note that the Fisher's test *P*-value for *KCTD18-like* was no longer significant after multiple testing correction (P>0.05) due to the limited sample size in this study. Given the potential function of the *KCTD18-like* gene and its association with human epilepsy, it would be worth investigating whether deletion of the *KCTD18-like* gene is functionally associated with epilepsy in marmosets. Despite our efforts to identify pathogenic CNV mutations in these marmosets, we do not suggest that the identified risk candidate loci are the sole cause of epilepsy in marmosets. Additionally, it is important to note that short-read sequencing has limitations in fully characterizing CNVs, and more accurate detection may necessitate the utilization of third-generation long-read sequencing technology. In addition to genetic factors, we are also interested in exploring the potential involvement of environmental factors in the onset of epilepsy in marmosets, as well as the contribution of polygenic variants (Leu et al., 2019; Moreau et al., 2020; Stewart, 2010).

There are several limitations that should be acknowledged

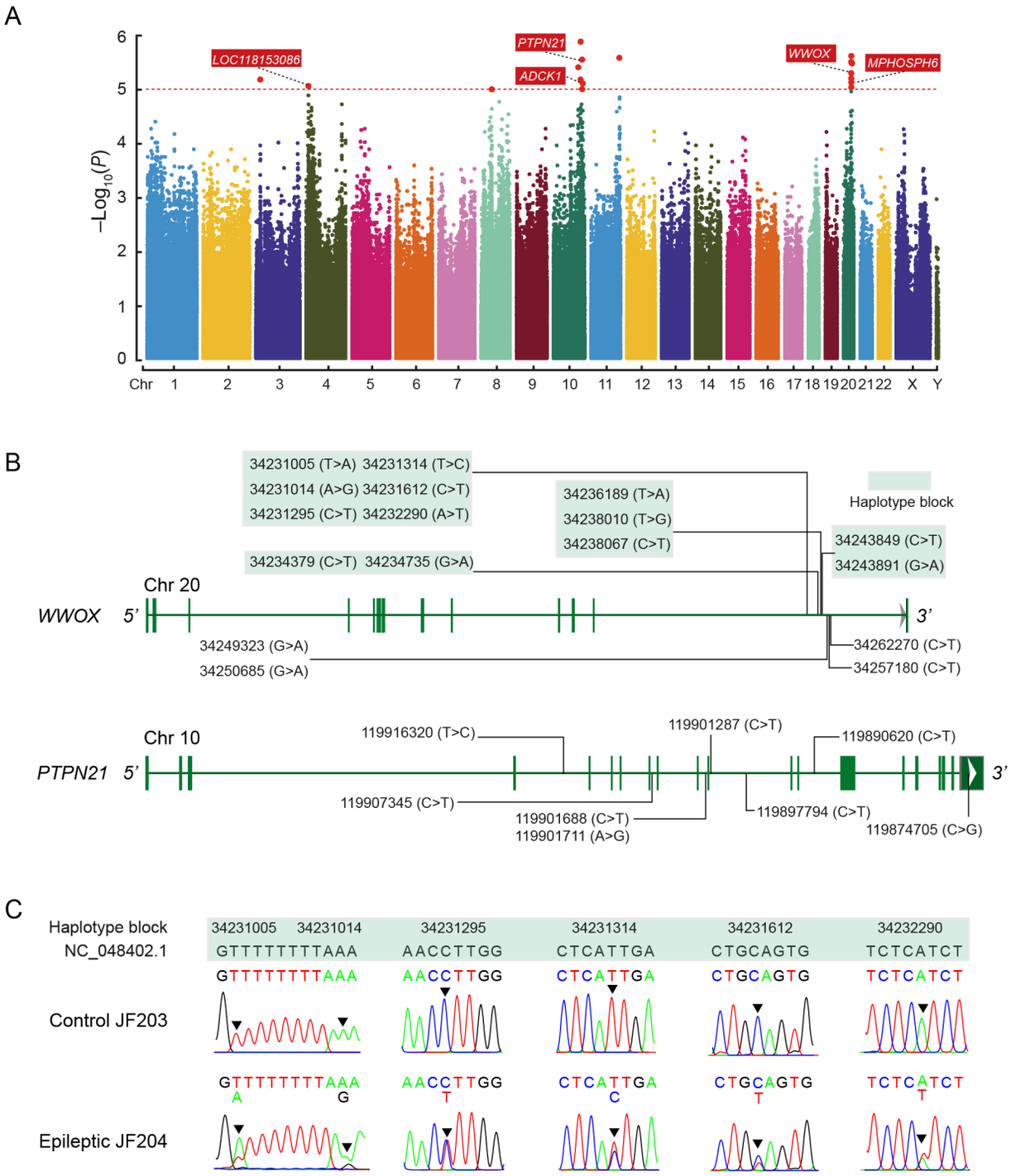

Figure 3 Genome-wide association analysis of marmoset variations associated with epilepsy

A: Manhattan plot showing top associations with epilepsy variation. Significance level (–log 10 transformed *P*-values) for SNPs is plotted along the genome in chromosomal order. Horizontal line indicates significance threshold $P=1.00 \times 10^{-5}$. B: Distribution of SNPs in *WWOX* and *PTPN21* genes in marmoset genome. Each SNP is marked with its genomic position. C: Representative Sanger sequencing validation results for control and epilepsy samples in the CLEA pedigree. SNPs in *WWOX*, confirmed within the same haplotype, are marked with light green haplotype blocks. Source data are provided in Supplementary Table S6.

in our study. Firstly, the relatively small sample size of available marmosets may limit the reliability of our findings. Increasing the sample size and employing other targeted approaches for genetic association testing should provide more insights into the genetic factors contributing to the development of epilepsy in marmosets. Secondly, due to the limited duration and variability of the seizures, we were unable to collect ictal EEG data for identifying the specific type of epilepsy. Therefore, further studies aimed at collecting such data would be required to refine this epileptic marmoset model, thereby enabling a more comprehensive understanding of the phenotypic characteristics associated with epileptic marmosets.

In summary, our study provides a valuable dataset on

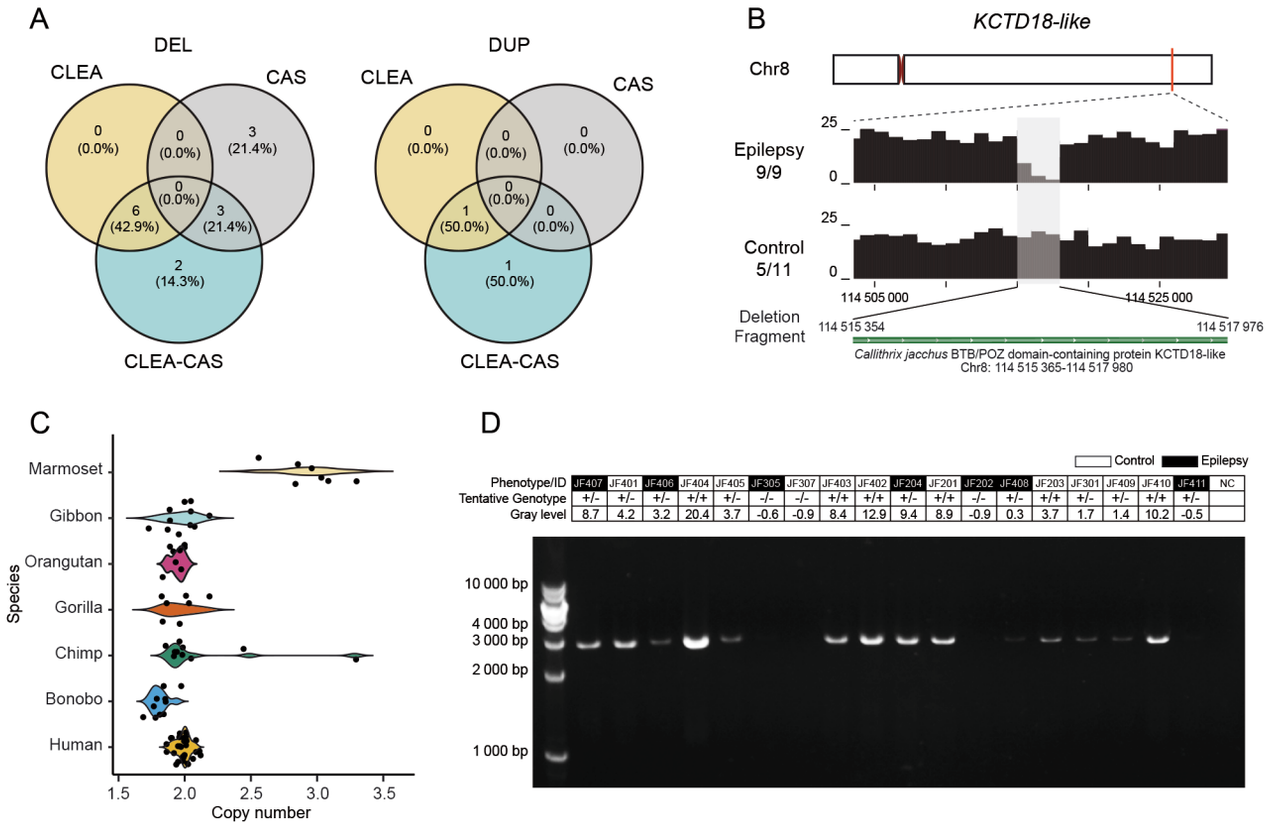

Figure 4 Candidate susceptibility gene screening and validation

A: Venn diagram of deleted or duplicated gene numbers in different datasets. B: Read depth of *KCTD18-like* genome segment in epileptic and control marmosets from CLEA. Decreased read depth was found in all nine epileptic marmosets and five control marmosets. C: Copy number of *KCTD18* in different species populations. D: *KCTD18-like* gene depletion was validated by PCR. Cells with a black background correspond to epileptic marmosets. Tentative genotype was inferred from the CNV region depth to average depth ratio based on read depth analysis. At ratios of ≤ 0.25, 0.25–0.75, and ≥0.75, the loci were defined as homozygous deletion (−/−), heterozygous deletion (+/−), and nondeletion (+/+), respectively. Source data are provided in Supplementary Table S9.

Table 2 Candidate susceptibility genes screened by CNV calling

Gene ID	Gene Symbol	Gene Type	Supporting Caller	Population	Type
XR_004732517.1	*LOC118146658*	Nonprotein coding	Manta	CLEA, CLEA&CAS	DEL
XM_035254712.1	*KCTD18-like*	Protein coding	Manta, Smoove	CLEA, CLEA&CAS	DEL
XM_002758028.4	*C14H2orf50*	Protein coding	Manta, Smoove	CLEA&CAS	DEL
XM_017967938.2	*ZNF331*	Protein coding	Manta	CLEA&CAS	DUP
XR_004743798.1	*LOC118153954*	Nonprotein coding	Manta, Smoove	CAS	DEL

marmoset population genetics and illustrates the genetic differences between two closed marmoset colonies in Asia. In addition, our study provides preliminary insights into potential candidate variants/genes associated with epilepsy in marmosets. These efforts represent a fundamental step towards studying captive marmoset populations and developing an epileptic marmoset disease model.

DATA AVAILABILITY

The raw genomic sequencing reads can be downloaded from the NCBI (PRJNA807054), China National Center for Bioinformation (PRJCA016608), and Science Data Bank databases (DOI: 10.57760/sciencedb.08059).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

E.E.E. is a scientific advisory board (SAB) member of Variant Bio, Inc. All other authors declare no competing interests.

AUTHORS' CONTRIBUTIONS

X.Y.Y., Y.F.M., and W.D.L. designed the experiment. X.Y.Y., Z.X., E.T., N.G., and W.D.L. contributed to sample collection and sample processing. X.Y.Y., Y.F.M., X.K.W., Z.Q.L., B.H., G.H., and Y.Y.S. contributed to data analysis. D.N.M. and X.Z. contributed to SNP and CNV validation. E.E.E. provided computational resources. X.Y.Y., Y.F.M., X.K.W., W.D.L., E.T., and Z.Q.L. generated tables and figures and drafted the manuscript. All authors contributed to editing and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Tomoyuki Araki (NIPPON Genetics Co., Ltd. Japan) for helping to purify DNA samples and T. Brown for manuscript proofreading and editing. E.E.E. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Battaglia A, Guerrini R. 2005. Chromosomal disorders associated with epilepsy. *Epileptic Disorders*, **7**(3): 181−192.

Benirschke K, Anderson JM, Brownhill LE. 1962. Marrow chimerism in marmosets. *Science*, **138**(3539): 513−515.

Breton VL, Aquilino MS, Repudi S, et al. 2021. Altered neocortical oscillations and cellular excitability in an *in vitro Wwox* knockout mouse model of epileptic encephalopathy. *Neurobiology of Disease*, **160**: 105529.

Carlson J, Li JZ, Zöllner S. 2018. Helmsman: fast and efficient mutation signature analysis for massive sequencing datasets. *BMC Genomics*, **19**(1): 845.

Chen JC, Lee G, Fanous AH, et al. 2011. Two non-synonymous markers in *PTPN21*, identified by genome-wide association study data-mining and replication, are associated with schizophrenia. *Schizophrenia Research*, **131**(1–3): 43–51.

Chen XY, Schulz-Trieglaff O, Shaw R, et al. 2016. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*, **32**(8): 1220−1222.

Coe BP, Witherspoon K, Rosenfeld JA, et al. 2014. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics*, **46**(10): 1063−1071.

Cooper GM, Coe BP, Girirajan S, et al. 2011. A copy number variation morbidity map of developmental delay. *Nature Genetics*, **43**(9): 838−846.

Ekenstedt KJ, Patterson EE, Mickelson JR. 2012. Canine epilepsy genetics. *Mammalian Genome*, **23**(1–2): 28–39.

Ellis CA, Petrovski S, Berkovic SF. 2020. Epilepsy genetics: clinical impacts and biological insights. *The Lancet Neurology*, **19**(1): 93−100.

Gardiner RM. 2000. Impact of our understanding of the genetic aetiology of epilepsy. *Journal of Neurology*, **247**(5): 327−334.

Grzybowska EA. 2012. Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochemical and Biophysical Research Communications*, **424**(1): 1−6.

Hardies K, Weckhuysen S, De Jonghe P, et al. 2016. Lessons learned from gene identification studies in Mendelian epilepsy disorders. *European Journal of Human Genetics*, **24**(7): 961−967.

Helbig I, Mefford HC, Sharp AJ, et al. 2009. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nature Genetics*, **41**(2): 160−162.

Homman-Ludiye J, Bourne JA. 2020. The marmoset: the next frontier in understanding the development of the human brain. *ILAR Journal*, **61**(2–3): 248–259.

Hsu CY, Lee KT, Sun TY, et al. 2021. WWOX and its binding proteins in neurodegeneration. *Cells*, **10**(7): 1781.

Iacomino M, Baldassari S, Tochigi Y, et al. 2020. Loss of Wwox perturbs neuronal migration and impairs early cortical development. *Frontiers in Neuroscience*, **14**: 644.

Johannsen J, Kortüm F, Rosenberger G, et al. 2018. A novel missense variant in the SDR domain of the WWOX gene leads to complete loss of WWOX protein with early-onset epileptic encephalopathy and severe developmental delay. *Neurogenetics*, **19**(3): 151−156.

Kishi N, Sato K, Sasaki E, et al. 2014. Common marmoset as a new model animal for neuroscience research and genome editing technology. *Development, Growth & Differentiation*, **56**(1): 53−62.

Kos MZ, Carless MA, Blondell L, et al. 2021. Whole genome sequence data from captive baboons implicate *RBFOX1* in epileptic seizure risk. *Frontiers in Genetics*, **12**: 714282.

Kulyté A, Aman A, Strawbridge RJ, et al. 2022. Genome-wide association study identifies genetic loci associated with fat cell number and overlap with genetic risk loci for type 2 diabetes. *Diabetes*, **71**(6): 1350−1362.

Leu C, Stevelink R, Smith AW, et al. 2019. Polygenic burden in focal and generalized epilepsies. *Brain*, **142**(11): 3473−3481.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**(21): 2987−2993.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*,doi: arXiv:1303.3997.

Li H, Handsaker B, Wysoker A, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**(16): 2078−2079.

Li MY, Boehnke M, Abecasis GR. 2005. Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *The American Journal of Human Genetics*, **76**(6): 934−949.

Li MY, Boehnke M, Abecasis GR. 2006. Efficient study designs for test of genetic association using sibship data and unrelated cases and controls. *The American Journal of Human Genetics*, **78**(5): 778−792.

Liu ZP, Xiang YQ, Sun GH. 2013. The KCTD family of proteins: structure, function, disease relevance. *Cell & Bioscience*, **3**(1): 45.

Malukiewicz J, Boere V, De Oliveira MAB, et al. 2020. An introduction to the *Callithrix* genus and overview of recent advances in marmoset research. *ILAR Journal*, **61**(2–3): 110–138.

Manichaikul A, Mychaleckyj JC, Rich SS, et al. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*, **26**(22): 2867−2873.

Mao YF, Economo EP, Satoh N. 2018. The roles of introgression and climate change in the rise to dominance of *Acropora* corals. *Current Biology*, **28**(21): 3373−3382.e5.

Marshall GF, Gonzalez-Sulser A, Abbott CM. 2021. Modelling epilepsy in the mouse: challenges and solutions. *Disease Models & Mechanisms*, **14**(3): dmm047449.

Martin M, Patterson M, Garg S, et al. 2016. WhatsHap: fast and accurate read-based phasing. *BioRxiv*, doi:https://doi.org/10.1101/085050.

McKenna A, Hanna M, Banks E, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297−1303.

Moreau C, Rébillard RM, Wolking S, et al. 2020. Polygenic risk scores of several subtypes of epilepsies in a founder population. *Neurology Genetics*, **6**(3): e416.

Mullen SA, Berkovic SF, the ILAE Genetics Commission. 2018. Genetic generalized epilepsies. *Epilepsia*, **59**(6): 1148−1153.

National Academies of Sciences, Engineering, and Medicine. 2019. Care, use, and welfare of marmosets as animal models for gene editing-based biomedical research. *In*: Proceedings of a Workshop. Washington: The National Academies Press.

Nishijima K, Saitoh R, Tanaka S, et al. 2012. Life span of common marmoset (*Callithrix jacchus*) at CLEA Japan breeding colony. *Biogerontology*, **13**(4): 439−443.

Okano H, Mitra P. 2015. Brain-mapping projects using the common marmoset. *Neuroscience Research*, **93**: 3−7.

Okano H, Sasaki E, Yamamori T, et al. 2016. Brain/MINDS: a Japanese national brain project for marmoset neuroscience. *Neuron*, **92**(3): 582−590.

Patterson M, Marschall T, Pisanti N, et al. 2015. WhatsHap: weighted haplotype assembly for future-generation sequencing reads. *Journal of Computational Biology*, **22**(6): 498−509.

Pedersen B. 2018. Smoove: structural variant calling and genotyping with existing tools, but, smoothly.https://github.com/brentp/smoove.

Plani-Lam JHC, Chow TC, Siu KL, et al. 2015. PTPN21 exerts pro-neuronal survival and neuritic elongation *via* ErbB4/NRG3 signaling. *The International Journal of Biochemistry & Cell Biology*, **61**: 53−62.

Poo MM, Du JL, Ip NY, et al. 2016. China brain project: basic neuroscience, brain diseases, and brain-inspired computing. *Neuron*, **92**(3): 591−596.

Purcell S, Neale B, Todd-Brown K, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, **81**(3): 559−575.

Repudi S, Steinberg DJ, Elazar N, et al. 2021. Neuronal deletion of *Wwox*, associated with WOREE syndrome, causes epilepsy and myelin defects. *Brain*, **144**(10): 3061−3077.

Robinson R, Gardiner M. 2004. Molecular basis of Mendelian idiopathic epilepsies. *Annals of Medicine*, **36**(2): 89−97.

Ross CN, French JA, Ortí G. 2007. Germ-line chimerism and paternal care in marmosets (*Callithrix kuhlii*). *Proceedings of the National Academy of Sciences of the United States of America*, **104**(15): 6278−6282.

Sato K, Kuroki Y, Kumita W, et al. 2015. Resequencing of the common marmoset genome improves genome assemblies and gene-coding sequence analysis. *Scientific Reports*, **5**: 16894.

Scharfman HE. 2007. The neurobiology of epilepsy. *Current Neurology and Neuroscience Reports*, **7**(4): 348−354.

Shimogori T, Abe A, Go Y, et al. 2018. Digital gene atlas of neonate common marmoset brain. *Neuroscience Research*, **128**: 1−13.

Shorvon SD. 2011. The etiologic classification of epilepsy. *Epilepsia*, **52**(6): 1052−1057.

Stewart I. 2010. Environmental risk factors for temporal lobe epilepsy–Is prenatal exposure to the marine algal neurotoxin domoic acid a potentially preventable cause?. *Medical Hypotheses*, **74**(3): 466−481.

Sweeney CG, Curran E, Westmoreland SV, et al. 2012. Quantitative molecular assessment of chimerism across tissues in marmosets and tamarins. *BMC Genomics*, **13**: 98.

Teng XC, Aouacheria A, Lionnard L, et al. 2019. *KCTD*: a new gene family involved in neurodevelopmental and neuropsychiatric disorders. *CNS Neuroscience & Therapeutics*, **25**(7): 887−902.

The Marmoset Genome Sequencing and Analysis Consortium. 2014. The common marmoset genome provides insight into primate biology and evolution. *Nature Genetics*, **46**(8): 850−857.

Thijs RD, Surges R, O'Brien TJ, et al. 2019. Epilepsy in adults. *The Lancet*, **393**(10172): 689−701.

Thomas RH, Berkovic SF. 2014. The hidden genetics of epilepsy —a clinically important new paradigm. *Nature Reviews Neurology*, **10**(5): 283−292.

Usui D, Shimada S, Shimojima K, et al. 2013. Interstitial duplication of 2q32.1–q33.3 in a patient with epilepsy, developmental delay, and autistic behavior. *American Journal of Medical Genetics Part A*, **161**(5): 1078−1084.

Warren WC, Harris RA, Haukness M, et al. 2020. Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science*, **370**(6523): eabc6617.

Wielaender F, Sarviaho R, James F, et al. 2017. Generalized myoclonic epilepsy with photosensitivity in juvenile dogs caused by a defective DIRAS family GTPase 1. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(10): 2669−2674.

Yang CT, Zhou Y, Marcus S, et al. 2021. Evolutionary and biomedical insights from a marmoset diploid genome assembly. *Nature*, **594**(7862): 227−233.

Yang XY, Chen ZT, Wang ZY, et al. 2022. A natural marmoset model of genetic generalized epilepsy. *Molecular Brain*, **15**(1): 16.

Ye J, Coulouris G, Zaretskaya I, et al. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, **13**: 134.

Youlatos D. 2009. The smallest anthropoids. *Journal of Long-Term Effects of Medical Implants*, **18**(2): 175−179.