



# Whole Genome Sequence Data From Captive Baboons Implicate *RBFOX1* in Epileptic Seizure Risk

Mark Z. Kos<sup>1\*†</sup>, Melanie A. Carless<sup>2,3†</sup>, Lucy Blondell<sup>1</sup>, M. Michelle Leland<sup>4</sup>,  
Koyle D. Knappe<sup>5</sup>, Harald H. H. Göring<sup>1†</sup> and Charles Ákos Szabó<sup>5,6†</sup>

## OPEN ACCESS

### Edited by:

Lawrence Todd Reiter,  
University of Tennessee Health  
Science Center (UTHSC),  
United States

### Reviewed by:

James S. Sutcliffe,  
Vanderbilt University, United States  
Olivia J. Veatch,  
University of Kansas Medical Center,  
United States

### \*Correspondence:

Mark Z. Kos  
mark.kos@utrgv.edu

### †ORCID:

Mark Z. Kos  
orcid.org/0000-0003-3129-7057  
Melanie A. Carless  
orcid.org/0000-0001-8350-2302  
Harald H. H. Göring  
orcid.org/0000-0002-0507-088X  
Charles Ákos Szabó  
orcid.org/0000-0003-2058-3357

### Specialty section:

This article was submitted to  
Genetics of Common and Rare  
Diseases,  
a section of the journal  
Frontiers in Genetics

**Received:** 25 May 2021

**Accepted:** 27 July 2021

**Published:** 20 August 2021

### Citation:

Kos MZ, Carless MA, Blondell L,  
Leland MM, Knappe KD, Göring HHH  
and Szabó CÁ (2021) Whole Genome  
Sequence Data From Captive  
Baboons Implicate *RBFOX1*  
in Epileptic Seizure Risk.  
Front. Genet. 12:714282.  
doi: 10.3389/fgene.2021.714282

<sup>1</sup> Department of Human Genetics, South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Edinburg, TX, United States, <sup>2</sup> Department of Biology, The University of Texas at San Antonio, San Antonio, TX, United States, <sup>3</sup> Brain Health Consortium, The University of Texas at San Antonio, San Antonio, TX, United States, <sup>4</sup> Laboratory Animal Research, UT Health San Antonio, San Antonio, TX, United States, <sup>5</sup> Department of Neurology, UT Health San Antonio, San Antonio, TX, United States, <sup>6</sup> South Texas Comprehensive Epilepsy Center, San Antonio, TX, United States

In this study, we investigate the genetic determinants that underlie epilepsy in a captive baboon pedigree and evaluate the potential suitability of this non-human primate model for understanding the genetic etiology of human epilepsy. Archived whole-genome sequence data were analyzed using both a candidate gene approach that targeted variants in baboon homologs of 19 genes ( $n = 20,881$  SNPs) previously implicated in genetic generalized epilepsy (GGE) and a more agnostic approach that examined protein-altering mutations genome-wide as assessed by snpEff ( $n = 36,169$ ). Measured genotype association tests for baboon cases of epileptic seizure were performed using SOLAR, as well as gene set enrichment analyses (GSEA) and protein-protein interaction (PPI) network construction of top association hits genome-wide ( $p < 0.01$ ;  $n = 441$  genes). The maximum likelihood estimate of heritability for epileptic seizure in the pedigreed baboon sample is 0.76 (SE = 0.77;  $p = 0.07$ ). Among candidate genes for GGE, a significant association was detected for an intronic SNP in *RBFOX1* ( $p = 5.92 \times 10^{-6}$ ; adjusted  $p = 0.016$ ). For protein-altering variants, no genome-wide significant results were observed for epilepsy status. However, GSEA revealed significant positive enrichment for genes involved in the extracellular matrix structure (ECM; FDR = 0.0072) and collagen formation (FDR = 0.017), which was reflected in a major PPI network cluster. This preliminary study highlights the potential role of *RBFOX1* in the epileptic baboon, a protein involved in transcriptomic regulation of multiple epilepsy candidate genes in humans and itself previously implicated in human epilepsy, both focal and generalized. Moreover, protein-damaging variants from across the genome exhibit a pattern of association that links collagen-containing ECM to epilepsy risk. These findings suggest a shared genetic etiology between baboon and human forms of GGE and lay the foundation for follow-up research.

**Keywords:** whole-genome sequence, association test, *Papio*, baboon, idiopathic generalized epilepsy, epilepsy, genetic generalized epilepsy, *RBFOX1*

## INTRODUCTION

Epilepsy is a chronic, highly heterogeneous neurological disorder, with recurrent seizures being the diagnostic hallmark. Genetic generalized epilepsies (GGEs) account for at least 40% of United States cases, (Durón et al., 2005) with heritability estimates from twin studies ranging from 25 to 69% (Kjeldsen et al., 2001). With the rise of high-throughput genomic sequencing, a multitude of epilepsy-associated genetic variants have been identified in recent years, including rare variants, such as recurrent microdeletions of large effect (Epi 4K Consortium and Epilepsy Phenome/Genome Project, 2017). Currently, epilepsies are suspected of having complex polygenic architectures, involving variants of many genes from across the frequency spectrum. However, segregation analyses in some families have shown that rare pathogenic variants convey most of the risk, (Myers and Mefford, 2016) with a potentially important role played by *de novo* mutagenesis, as established in patients with focal and generalized epilepsies as well as epileptic encephalopathies (Jones et al., 2011).

Toward elucidating the various genes and molecular mechanisms underlying epileptogenesis and seizure generation, animal models have played a critical role, as clinical studies of human patients preclude the possibility of invasive, and often more insightful, methods. For decades now, rodents have been the most common animal models used in epilepsy research, contributing greatly toward our understanding of acquired epilepsy and acutely induced seizure events. For chronic phenotypes that more closely resemble human epilepsy, a number of well-described mouse models have been developed that exhibit spontaneous absence-like seizures with spike-wave EEG discharges analogous to GGE, but with important drawbacks, most notably their monogenic basis and other neurological abnormalities. More promising are two rat models with presumed polygenic inheritance, Genetic Absence Rats from Strasbourg (GAERS) and Wistar Albino Glaxo rat from Rijswijk (WAG/Rij), (Gauguier et al., 2004; Rudolf et al., 2004) that have shown electroclinical and pathophysiological similarities to human GGE with absence epilepsy. Despite these successes, translatability of these models is limited due to differences in neurodevelopment and the natural evolutionary histories of the species, as well as response to treatment, (Löscher, 2017) suggesting the need for alternate animal models that can further characterize the complex genetic architecture of GGE in humans and better identify new etiological pathways with novel therapeutic targets.

Baboons are suitable genetic models for humans, as their genomes are more similar to humans than those of less closely related model organisms, and share neuroanatomical, biochemical, and physiological features stemming from their close phylogenetic relationship. Baboons have been used as animal models in a broad range of human diseases, including diabetes, heart disease, osteoporosis, and chronic infectious diseases (Rogers and Hixson, 1997). Baboons also represent a natural model for GGE (Szabó et al., 2012, 2013; Croll et al., 2019). For the pedigreed baboon colony housed at the Southwest National Primate Research Center (SNPRC; San Antonio, TX,

United States), absence, myoclonic, and generalized tonic-clonic seizures have been reported to occur spontaneously or triggered by intermittent light stimulation, with electroclinical findings suggesting strong similarities with human epilepsy, in particular juvenile myoclonic epilepsy (Szabó et al., 2013). In a preliminary analysis of 1,400 animals, (Croll et al., 2019) we previously estimated moderate heritabilities for seizures ( $h^2 = 0.33$ ,  $p < 1 \times 10^{-7}$ ) and interictal epileptic discharges ( $h^2 = 0.19$ ,  $p < 0.002$ ), although the underlying genetic variants are unknown. With high-throughput whole-genome sequencing (WGS) now performed on a subset of these animals and the potential enrichment of rare pathogenic variants within large multigenerational pedigrees that improve statistical power for detecting their risk effects, the SNPRC baboon colony represents a largely untapped genetic resource for epilepsy research (Jun et al., 2018).

This study provides an initial assessment of the baboon as a genetic model for epilepsy. Archived WGS data from the SNPRC colony were tested for genetic associations with epileptic seizure, targeting single nucleotide polymorphisms (SNPs) within GGE-associated genes, as well as protein-altering variants genome-wide. Based on these results, as well as pathway enrichment analyses of top association hits, a significant association was observed for a common variant in the gene *RBFOX1* (RNA Binding Fox-1 Homolog 1), providing insight into the etiological underpinnings of the disorder.

## MATERIALS AND METHODS

### Baboon Colony

There are currently about 1,000 baboons in the pedigreed colony housed at the SNPRC, representing a mixture of species, with extensive veterinary records generated since 1980. The animals have been treated in strict accordance with United States Public Health Service policy, (Institute for Laboratory Animal Research, 1996) with regular inspections by the United States Department of Agriculture (latest inspection in 2019). This study was approved by the IACUC of the University of Texas Health Science Center at San Antonio and the Texas Biomedical Research Institute.

We performed a retrospective survey of seizures from veterinary records available for 1,528 baboons, both living and deceased. This included both witnessed seizures, either spontaneous or provoked by ketamine administration or handling, and suspected seizures, assessed by treatments for head injuries, in particular craniofacial trauma that are often related to falls during seizures, as well as peri-ictal behaviors such as confusion and lethargy. As previously reported, (Szabó et al., 2012) a total of 3,389 seizure events were detected in the records of 1,098 baboons, of which 1,537 (45%) were witnessed in 404 animals, with 1,267 (82%) being unprovoked. For this study, archived WGS data were available for 90 baboons with clinical histories [*Papio hamadryas anubis* (*P.h. anubis*) and *P.h. anubis/cynocephalus* hybrids], including scalp EEG recordings of interictal epileptic discharges for 53 of these animals (Szabó et al., 2013). Diagnosis of epilepsy was adapted from criteria established

by the International League Against Epilepsy (ILAE), (Fisher et al., 2014) including: (1) evidence of at least one witnessed seizure or craniofacial trauma with a diagnostic EEG; (2) two or more witnessed seizures or craniofacial trauma without a diagnostic EEG; (3) or recording of seizures, spontaneous or photoepileptic, by video-EEG (Szabó et al., 2013, 2014). Baboons without a history of witnessed seizures and/or craniofacial trauma and normal EEG measures constituted our control group. Using these criteria, 42 baboons were diagnosed with epilepsy, 19 as healthy controls, and 29 as indeterminate (single witnessed seizure or acute craniofacial trauma and without a diagnostic EEG).

## Processing of Archived WGS Data

Archived WGS data for the SNPRC baboon colony are available in the NCBI Sequence Read Archive (SRA) repository. These data have been submitted under BioProject PRJNA433868, described as a “mixture of low-coverage and high-coverage whole-genome sequence data” generated on Illumina HiSeq X Ten and HiSeq 2500 instruments. Using a submission date cut-off of 8/2/2018, a total of 791 WGS runs representing 788 different baboons were downloaded as compressed SRA-formatted files onto the STDOI Genomics Computing Center cluster at the University of Texas Rio Grande Valley using the SRA FTP platform (command *prefetch*, followed by *fastq-dump*), representing approximately 11 TB of data.

For the 90 baboons that were assessed for epileptic seizure, FASTQ files were extracted from the compressed WGS files, representing about  $6.8 \times$  genomic coverage on average, and then aligned against the reference *P.h. anubis* genome assembly “Panu\_3.0” (NCBI BioProject PRJNA54005) with the software BWA v. 0.7.12, (Li and Durbin, 2010) using the faster and more accurate BWA-MEM algorithm (Li, 2013) and compressed into the binary BAM format using the *SamFormatConverter* Java executable in the software package Picard v. 1.121<sup>1</sup> Using the Picard toolkit, the aligned sequences were then (1) soft-clipped for beyond-end-of-reference alignments with *CleanSam*; (2) sorted based on genomic coordinates with *AddOrReplaceReadGroups*, with meta information added to the reads; (3) tagged for duplicate reads with *MarkDuplicates*; (4) confirmed for their consistency to the SAM format specifications with *ValidateSamFile*; and (5) used to generate BAM index files with *BuildBamIndex* to allow for faster processing of the data downstream.

For the next steps in the WGS pipeline, the GATK genomic analysis toolkit (v. 3.5 and v. 4.1.3.0) (McKenna et al., 2010) was employed as follows: (1) potential mapping errors due to indel events were identified in the tagged BAM files and realigned with *RealignerTargetCreator* and *IndelRealigner*; (2) quality scores for individual base calls were examined and adjusted for systematic technical errors stemming from the original sequencing work using base quality score recalibration (BQSR), a machine-learning approach to empirically build a covariation model of these errors using a set of known variants [in this case, previously published sequence variation of six species

of *Papio* baboon and *Theropithecus gelada* (Rogers et al., 2019)]; (3) call SNPs from realigned and recalibrated sequence data via local *de novo* assembly using *HaplotypeCaller* (using -ERC GVCF mode); (4) aggregate the resulting 90 GVCF files into multi-sample “datastores” per chromosome using *GenomicsDBImport*; and (5) perform joint genotyping on each datastore with *GenotypeGVCFs*, creating raw SNP VCF files.

Lastly, based on recommendations made by the GATK development team,<sup>2</sup> low-quality SNPs were excluded from the genotyped VCF files using *SelectVariants* for the following hard thresholds: Phred-scaled quality score (QUAL) < 30; quality score by depth (QD) < 2; symmetric odds ratio (SOR) > 3; measure of strand bias by Fisher’s exact test (FS) > 60; root mean square of mapping quality (MQ) < 40; rank sum test for mapping qualities of reference and alternate allele reads (MQRankSum) < -12.5; rank sum test for relative positioning of reference and alternate alleles within reads (ReadPosRankSum) < -8; and allele frequency (AF) < 0.05.

## Genetic Analyses

The WGS data were examined using both candidate gene and genome-wide approaches to improve the statistical power of the study. For the former, genes implicated in GGE or the synonymous idiopathic generalized epilepsy (IGE) were reviewed as positional candidate genes by searching the PubMed database with the following terms:

(GENE NAME) AND [(“genetic generalized epilepsy”) OR (“idiopathic generalized epilepsy”) OR (“childhood absence epilepsy”) OR (“juvenile absence epilepsy”) OR (“juvenile myoclonic epilepsy”)].

The search was not exhaustive but limited to genes with *a priori* evidence for GGE susceptibility in the databases *Online Mendelian Inheritance in Man* (OMIM) and *EpilepsyGene*, (Ran et al., 2015) including genetic epilepsy with febrile seizures plus, childhood absence epilepsy, juvenile absence epilepsy, and juvenile myoclonic epilepsy. Positional candidate genes were identified as such by evaluating the collective *in silico* evidence yielded by the PubMed search, including the seminal GWAS mega-analysis conducted by ILAE, (The International League Against Epilepsy Consortium on Complex Epilepsies, 2018) with particular emphasis paid to studies that *independently* reported significant associations or familial segregation of gene variants and any of the various forms of GGE (thus disqualifying novel, unreplicated findings). In the end, 19 positional candidate genes for GGE were selected for our study, with each having a homologous gene or homolog in the baboon genome (confirmed using the chain/net alignments for genomes GRCh38 and Panu\_3.0 from the UCSC Genome Browser): *BRD2*, *CACNA1A*, *CACNA1H*, *CACNB4*, *CASR*, *CHRNA7*, *EFHC1*, *GABRA1*, *GABRA6*, *GABRB3*, *GABRD*, *GABRG2*, *ICK*, *ME2*, *NIPA2*, *PNPO*, *RBF0X1*, *SCN1A*, and *SLC2A1*.

Heritability ( $h^2$ ) estimation and genetic association testing were performed using the software SOLAR, (Almasy and Blangero, 1998) based on maximum likelihood variance

<sup>1</sup><http://broadinstitute.github.io/picard>

<sup>2</sup><https://gatk.broadinstitute.org/hc/en-us/articles/360035531112--How-to-Filter-variants-either-with-VQSR-or-by-hard-filtering>



decomposition. Genetic relationships between the baboons were inferred from a previously reported, single six-generation pedigree comprising 2,455 animals, including 38 cases of epileptic seizure and 16 controls with processed WGS data (total  $n = 54$ ), which were the focus of the genetic analyses. Associations between SNPs and seizure status were computed using measured genotype association (MGA) analysis under a liability threshold in which kinship is treated as a random effect and SNPs as fixed effects, with sex and the first principal component (PC) as covariates. The likelihoods of the MGA models, each maximized for the effects of kinship and SNP genotypes, were compared to null models with SNP effects constrained to zero. For the baboon homologs of the 19 positional candidate genes for GGE, 20,881 SNPs were identified in the WGS data (including variants 5 kb upstream and downstream of the leading and trailing UTR sequences) and tested for association. A more agnostic, genome-wide approach was also taken, in which protein-altering variants of “moderate” impact (e.g., missense variants) and “high” impact (e.g., start and stop codons gained or lost, splice acceptor and donor site mutations) on protein function were tested in a separate run ( $n = 36,169$  SNPs). These variants were predicted by the annotation tool SnpEff (Cingolani et al., 2012b) using a SNP prediction library built for assembly Panu\_3.0/Panu 4 (gene annotation GTF files downloaded from the UCSC database) and filtered with SnpSift (Cingolani et al., 2012a). Association  $p$ -values were adjusted for the effective number of independent SNPs tested, which was estimated per chromosome in SOLAR using the Li & Ji method (Li and Ji, 2005) based on the computation of eigenvalues of a genotypic correlation matrix.

For the association results of genome-wide deleterious SNPs, the genes impacted by these variants (and with known human homologs) were tested for enrichment for genes from biological pathways and molecular ontologies using the program Gene Set Enrichment Analysis (GSEA v. 4.1) (Subramanian et al., 2005) for the Molecular Signatures Database (MSigDB v. 7.2). GSEA represents a modified Kolmogorov–Smirnov test, with significance of the unweighted enrichment score established *via* 10,000 permutations. Additionally, protein–protein interaction (PPI) networks were constructed for these results using the online tool STRING v. 11.0, (Szklarczyk et al., 2019) limited to interactions of “high confidence” (score  $\geq 0.7$ ) from all available sources, with functional enrichment analysis performed on PPI nodes for the entire network, as well as targeted analysis of the largest observed cluster, for reactome pathways and GO-terms.

## Assessment of Inflation of the Association Test Statistic

Potential inflation of the MGA test statistic was assessed by calculating the genomic inflation factor  $\lambda$  (ratio of the median of the empirically observed distribution of  $\chi^2$  to the expected median of 0.456). This was done for 10 random subsets of 20,881 intragenic SNPs from the processed WGS data (mirroring the number of SNPs tested in the 19 candidate genes). For the MGA model with sex as the sole covariate, some inflation was observed among association results generated for these random subsets, with a mean  $\lambda$  of 1.13 (SD = 0.02), suggesting potential

confounding due to population substructure. To investigate this, principal components analysis (PCA) was performed on WGS data from autosomes 1–20 ( $n = 5,542,179$  SNPs) with the EIGENSTRAT method, using the statistical package AssocTests v. 1.0-1 (Wang et al., 2020) in R v. 4.0.3. (R Core Team, 2020). A plot of the first two PCs (accounting for 9.0 and 4.6% of the genotypic variation, respectively; see **Supplementary Figure 1**) revealed stratification within the baboon sample, especially relative to the first PC. Hence, PC1 was included as a covariate in the MGA model and association testing was re-run on the random SNP subsets, effectively controlling the genome-wide inflation, with a mean  $\lambda$  of 1.01 (SD = 0.01). Thus, sex and PC1 were used as covariates in all MGA tests performed in the study.

## RESULTS

### SNP Association Testing of GGE Candidate Gene Homologs

After realignment and recalibration of the base quality scores of the baboon WGS data, a total of 46,947,610 variants, both SNPs and indels, were successfully called. These variants were then filtered for various QC thresholds, targeting SNPs within known genic regions, yielding 5,634,214 SNPs for 54 case-control animals. For this subset, variant confidence scores normalized by sequence coverage (QualByDepth or QD) range from 2.0 to 42.9, with a mean of 17.5 and median of 16.0 (SD = 6.0). Most (~79%) of the variants are intronic, with a transition-to-transversion (Ts/Tv) ratio of 2.48. Over 25 million effects on gene products were identified with SnpEff, with 110,771 (0.4%) and 2,394 (0.009%) characterized as “MODERATE” and “HIGH,” respectively, including ones arising from missense and nonsense mutations and splice site variants.

Using a maximum likelihood, variance components approach to our genetic analyses, heritability of epileptic seizure in the baboon cohort was estimated at 0.77 (SE = 0.73;  $p = 0.07$ ). Targeting baboon homologs of genes implicated in GGE and related epilepsies in humans, 20,881 QC filtered SNPs from 19 genes were tested for association with epileptic seizure (see **Table 1** and **Supplementary Table 1**;  $\lambda = 1.10$ ). The effective number of independent SNPs tested for association (i.e., accounting for LD between variants) was estimated at 2,686 using the Li & Ji method, (Li and Ji, 2005) which was used in the adjustment of  $p$ -values for multiple testing *via* the conservative Bonferroni method. The top association result is for a common intronic SNP in the gene *RBFox1* [chromosome 20 (NC\_018171), base pair (bp) position 5,642,021 in intron 3 (ENSPANT00000033712.2); minor allele frequency (MAF) = 0.48], with an estimated  $\beta$  coefficient of 1.37 (SE = 0.30) and Bonferroni-corrected  $p = 0.016$  (raw  $p = 5.92 \times 10^{-6}$ ; see **Figure 1**). The SNP was successfully genotyped for 51 of the 54 animals, with genotype ratios (AA:AG:GG) of 0:3:12 for healthy controls and 17:12:7 for cases of epileptic seizure, revealing a significant enrichment of the A allele among cases (29 carriers; AF = 0.64) relative to controls (three carriers, all heterozygotes; AF = 0.1). The QC sequencing metrics for the called SNP appear

**TABLE 1** | Top ten candidate genes based on SNP association results for epileptic seizure in baboons.

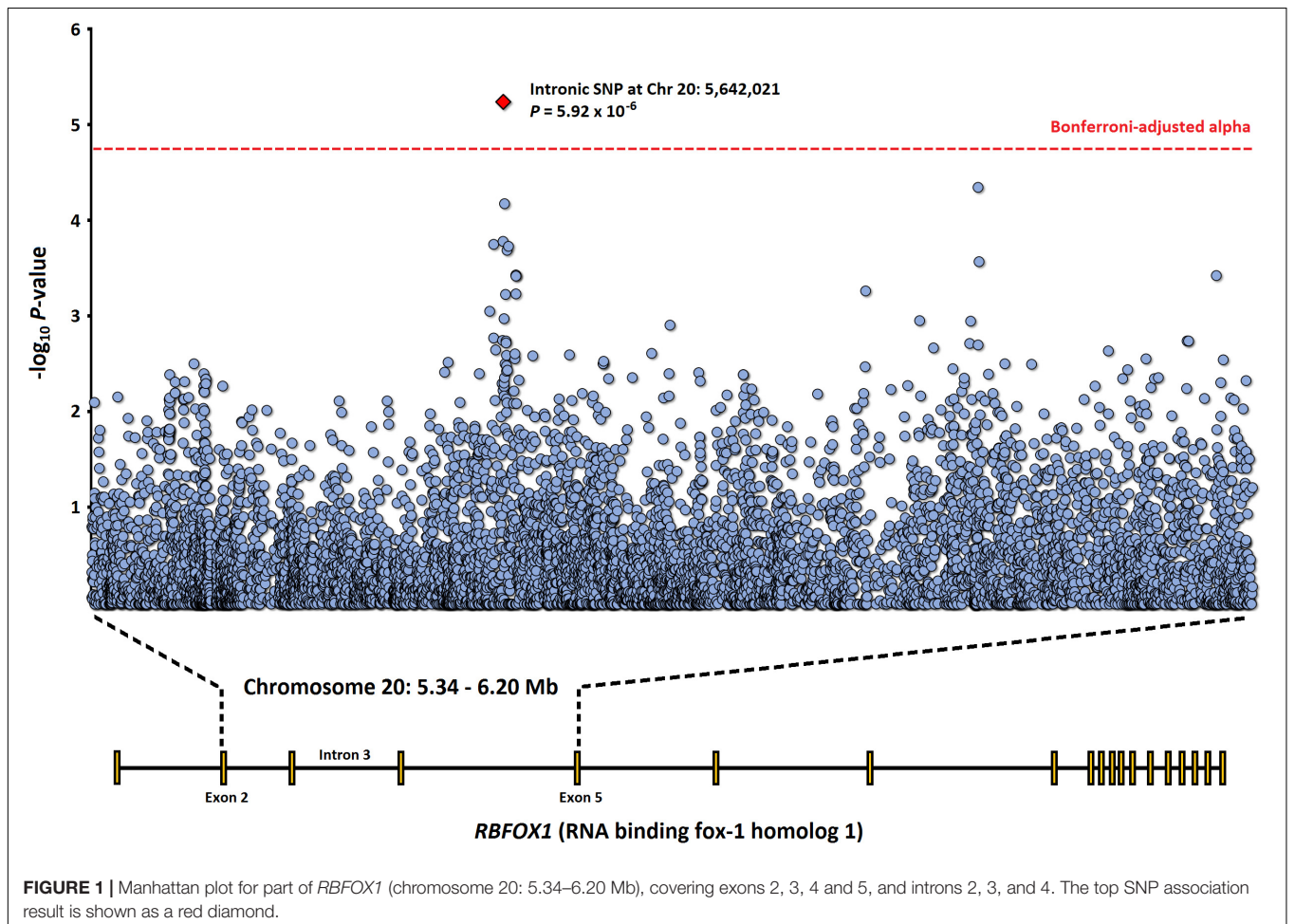
Gene	Chr	Position (bp)	MA	MAF	Variant type	Epilepsy seizure	
						Beta (SE)	p-Value
<i>RBFOX1</i>	20	5,642,021	A	0.48	Intron	1.37 (0.30)	$5.92 \times 10^{-6}$ *
<i>GABRB3</i>	7	2,517,399	G	0.10	Intron	5.64 (1.58)	$3.67 \times 10^{-4}$
<i>SCN1A</i>	12	51,750,877	A	0.22	Intron	3.94 (1.24)	$1.50 \times 10^{-3}$
<i>GABRG2</i>	6	156,497,930	C	0.09	Intron	1.42 (0.45)	$1.73 \times 10^{-3}$
<i>CASR</i>	2	41,913,861	A	0.08	Intron	4.65 (1.52)	$2.29 \times 10^{-3}$
<i>CACNA1A</i>	19	12,102,543	T	0.19	Synonymous	3.56 (1.19)	$2.84 \times 10^{-3}$
<i>GABRA1</i>	6	156,221,558	C	0.15	Intron	3.54 (1.25)	$3.42 \times 10^{-3}$
<i>CHRNA7</i>	7	6,949,236	C	0.09	Intron	5.19 (1.77)	$3.43 \times 10^{-3}$
<i>BRD2</i>	4	32,316,440	G	0.03	Upstream	4.38 (1.50)	$3.49 \times 10^{-3}$
<i>ICK</i>	4	52,249,073	A	0.09	Intron	4.78 (1.68)	$4.55 \times 10^{-3}$

The table lists the top ten gene homologs for human candidate genes for GGE and related epilepsies based on their top SNP association results for epileptic seizure in baboons.

SNP positions relative to reference *P.h. anubis* genome assembly Panu\_3.0 (NCBI BioProject PRJNA54005).

Variant type determined with the annotation tool SnpEff.

\*Significant after Bonferroni correction for multiple testing ( $\alpha = 0.05/2,686 = 1.86 \times 10^{-5}$ ).



**TABLE 2** | Top ten genome-wide association results for protein-altering SNPs.

Gene	Chr	Position (bp)	MA	MAF	Variant type	Epilepsy seizure	
						Beta (SE)	p-Value
SYNE1	4	110,510,700	C	0.06	Missense	4.35 (0.97)	$7.52 \times 10^{-6}$
PBXIP1	1	125,308,470	G	0.32	Missense	4.09 (0.96)	$2.02 \times 10^{-5}$
SYCP1	1	112,209,853	A	0.18	Missense	5.63 (1.33)	$2.21 \times 10^{-5}$
LOC108584956	3	43,438,092	A	0.06	Missense	5.93 (1.42)	$3.02 \times 10^{-5}$
MLLT6	16	45,549,650	G	0.20	Missense	5.61 (1.38)	$4.51 \times 10^{-5}$
ZNF692	1	217,379,527	G	0.24	Missense	5.44 (1.33)	$4.61 \times 10^{-5}$
GLYAT	14	14,601,373	G	0.42	Missense	1.37 (0.34)	$5.16 \times 10^{-5}$
RRM2B	8	97,957,365	T	0.44	Missense	1.15 (0.30)	$1.04 \times 10^{-4}$
UHRF1BP1	4	34,215,535	A	0.07	Missense	5.31 (1.39)	$1.31 \times 10^{-4}$
SELENOP	6	42,584,897	C	0.19	Missense	5.02 (1.34)	$1.82 \times 10^{-4}$

The table lists the top ten SNP association results for epilepsy seizure in baboons for genome-wide variants with deleterious effects on gene products ( $n = 36,169$ ), as predicted by SnpEff (variants deemed as having either “moderate” or “high” deleterious effect).

SNP positions relative to reference *P.h. anubis* genome assembly Panu\_3.0 (NCBI BioProject PRJNA54005).

Mutation types analyzed included missense, nonsense, stop codon lost, start codon lost, splice acceptor, and splice donor variants.

satisfactory, including: QD = 31.3; average genotype quality score (GQ) = 32.3; and approximately  $5.9 \times$  depth of coverage.

## Genome-Wide Association Testing of Protein-Altering Variants

Given the limited sample size, any comprehensive examination of the millions of called SNPs would be severely underpowered. To achieve a genome-wide perspective of the WGS data, protein-altering SNPs with predicted deleterious effects ( $n = 36,169$ ) were selected for a second round of association testing. It is likely that many of the genetic determinants that underlie seizures in baboons impact protein structure, and thus, targeting this mutation class may allow for broader, functionally meaningful association patterns to emerge. The top association (see **Table 2** and **Supplementary Table 2**;  $\lambda = 1.07$ ) is for a missense variant (p.Glu2066Gln) in the gene *SYNE1* ( $\beta = 4.35$ , SE = 0.97;  $p = 7.52 \times 10^{-6}$ ), which is statistically non-significant after multiple-testing correction. Although the gene has no reported links to epilepsy, *SYNE1* has been identified as the causal gene underlying spinocerebellar ataxia 1 (Gros-Louis et al., 2007).

For SNP associations with  $p$ -values less than 0.01 ( $n = 610$  SNPs), representing 441 different genes, PPI relationships were examined with STRING v. 11.0 (see **Figure 2**). Restricted to curated interaction scores with high confidence levels ( $\geq 0.7$ ), 34 PPI clusters were generated, ranging in size from 2 to 22 nodes, comprising a total of 211 edges, with 264 of the input genes isolated from the network (i.e., unconnected). Given the number of genes examined, the expected number of edges is 172, which yielded a PPI enrichment  $p$ -value of  $2.31 \times 10^{-3}$ , thus suggesting broad patterns of biomolecular connectivity among the gene products impacted by these variants. Enrichment analyses were performed on the PPI network (including unconnected nodes), with the following top-enriched GO-terms (FDR = 0.0062): GO:0062023 “Collagen-containing extracellular matrix” (13 of 144 genes belonging to the GO-term category were counted); GO:0005929 “Cilium” (31/570 count); and GO:0005604 “Basement membrane” (11/91

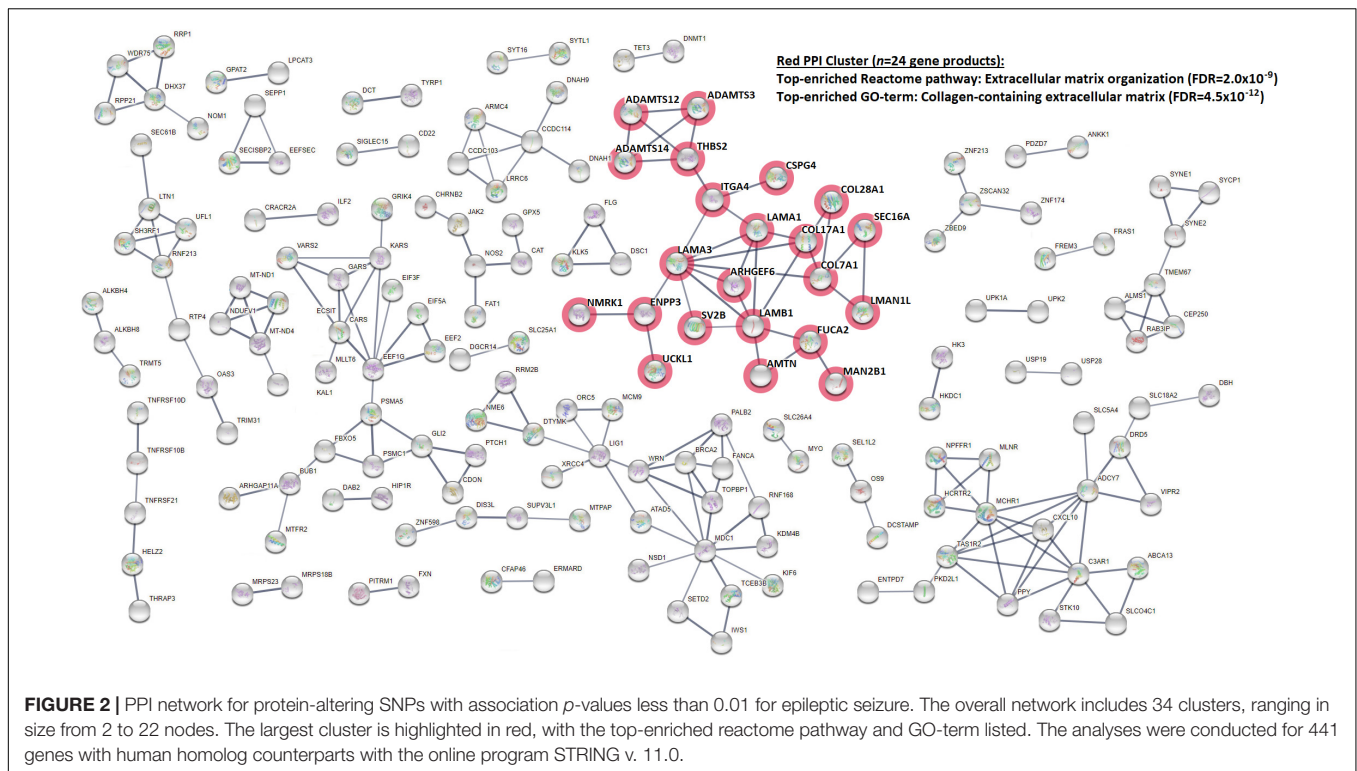
count). Reactome pathways were also tested, but none achieved statistical significance. Narrowing the enrichment analysis to the largest connected cluster in the network, color-coded in **Figure 2** as red ( $n = 22$  nodes), the top-enriched GO-term and Reactome pathway are GO:0062023 “Collagen-containing extracellular matrix” (9/144 count; FDR =  $4.48 \times 10^{-12}$ ) and HSA-1474244 “Extracellular matrix organization” (9/298 count; FDR =  $2.01 \times 10^{-9}$ ), respectively.

GSEA was also performed on the entire set of protein-altering SNPs, with each of the impacted genes ( $n = 8,802$  with human homologs) pre-ranked based on the intragenic SNP with the best association result (and thus not relying on an arbitrarily set  $p$ -value threshold that partitions the association results). Three gene sets from the Reactome pathways and GO-terms (**Table 3**) were found to exhibit significant positive enrichment (i.e., overrepresentation among highly ranked genes). Similar to the PPI network, enriched gene sets relate to collagen and the extracellular matrix (ECM). The enriched Reactome pathway is HSA-1474290 “Collagen formation” (FDR = 0.017; see **Figure 3A**), with 63 genes in the pathway observed in the pre-ranked list, with the top-ranked gene being *ADAMTS3* (73rd gene; max enrichment score (ES) at 2,832nd gene). For the GO-terms, significant enrichment was detected for a pair of gene sets related to the ECM, with the top result belonging to GO:0005201 “Extracellular matrix structural constituent” (FDR = 0.0072; see **Figure 3B**), with 110 genes in the pre-ranked list, with the top-ranked gene being *MXRA5* (146th gene; max ES at 2,955th gene).

## DISCUSSION

In this study, genetic association testing was performed for epileptic seizure in a pedigreed baboon colony using available WGS data from the NCBI SRA database. Targeting homologs of genes previously implicated in human epilepsy, we observed a significant association for an intronic SNP in the gene *RBFox1*, encoding an RNA-binding protein that regulates splicing,





including transcripts of other epilepsy candidate genes (e.g., *GABRG2*, *KCNQ2*, *SCN8A*, *SLC12A5*, and *SYN1*, among others), and plays a key role in neuronal excitation in the mammalian brain (Gehman et al., 2011). A more agnostic approach was also used in our analyses, as protein-altering SNPs were identified throughout the baboon genome and tested for association with epilepsy status. Although no significant associations were observed, genes involved in the collagen-containing ECM showed significant overrepresentation among top results. To the best of our knowledge, these findings represent the first

characterization of potential variants underlying the genetic architecture of epilepsy in baboons, pointing to commonalities with what is known about the human genetic etiology, thus suggesting the suitability of baboons as an animal model for the disorder.

The *RBFox1* gene has repeatedly been linked to epilepsy, with the earliest reports based on *de novo* structural gene variants observed in patients comorbid for autism, (Martin et al., 2007) intellectual disability, (Bhalla et al., 2004) or pontocerebellar hypoplasia (Gallant et al., 2011). These findings were followed up by a series of association studies by Lal et al. that targeted exon-disrupting microdeletions in *RBFox1*, revealing a significant excess among GGE patients compared to population controls, (Lal et al., 2013b) as well as among cases of Rolandic epilepsy (Lal et al., 2013a) and sporadic focal epilepsy, (Lal et al., 2015) suggesting that *RBFox1* deletions are involved in a wide spectrum of both focal and generalized epilepsies. In mice, brain-specific *Rbfox1* knockouts have displayed spontaneous seizures and heightened epileptogenic response to kainic acid (Gehman et al., 2011). Although the pathophysiological mechanisms of *RBFox1* remain to be clarified, more recent *in vivo* and *in vitro* studies have yielded insights, including regulation of cerebral cortex development, (Hamada et al., 2015) changes to transcriptomic expression and splicing patterns of neuronal genes, (Fogel et al., 2012) and crosstalk with microRNA miR-129-5p that impacts homeostatic downscaling of excitatory synapses (Rajman et al., 2017).

As one of the longest genes in the human genome, *RBFox1* has an extended 5' region that harbors several transcription start sites at varying first exons, enabling complex isoform

**TABLE 3 |** Gene set enrichment analyses (GSEA) of association results for protein-damaging SNPs.

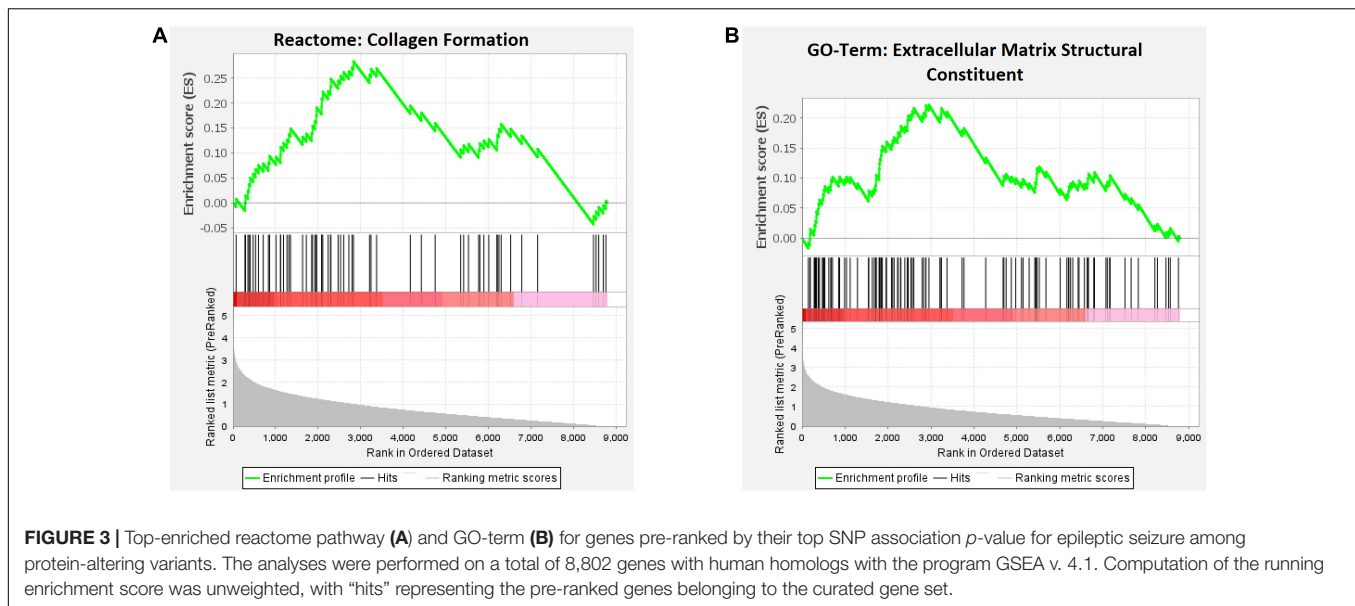
Gene set	Size	NES	FDR <i>q</i> -value
Reactome pathways			
Collagen formation	63	2.62	1.67 × 10 <sup>-2</sup>
GO-terms			
ECM structural constituent	110	2.71	7.16 × 10 <sup>-3</sup>
ECM structural constituent conferring tensile strength	32	2.48	2.88 × 10 <sup>-2</sup>

The table lists significant GSEA results for Reactome pathways and GO-terms (MSigDB v. 7.2).

The gene list input is based on the association results for epileptic seizure in baboons for 36,169 protein-damaging SNPs, pre-ranked according to the top associations in each gene (negative logarithmic transformation of unadjusted *p*-values).

Size is the number of pathway/GO-term genes observed in the gene list.

NES is the normalized enrichment score, which represents the degree to which a gene set is overrepresented at the top (or bottom) of the pre-ranked list of genes.



versatility to produce differential patterns of subcellular and tissue-specific localization (Conboy, 2017). The key regulatory feature of the RBF0X1 protein, however, is located in the downstream coding region, a single RNA recognition motif (RRM) that preferentially binds the sequence (U)GCAUG in pre-mRNA introns, mRNA 3' UTRs, and microRNA hairpins. Alternative splicing modulates the inclusion of the various exons in this region, generating active and inactive forms of RRM, as well as directing isoform localization. Overall, the gene is highly conserved in most primates, including its homolog in *P.h. anubis* that is approximately 92.4% identical to its human counterpart, with a Gene Order Conservation score of 100 (Ensembl Release 102 data).

For the significant *RBF0X1* association reported here, the SNP is positioned in intron 3 (of 19 in total), which is part of the extended 5' region described above. In the human genome (GRCh38), the SNP and its flanking 50 bp region aligns within *RBF0X1* at chromosome 16: 5,767,827–5,767,900 bp (cytogenetic band 16p13.3) via the UCSC BLAST-like alignment tool (score = 58; span = 74 bp; identity = 89.2%), and appears to have limited functional and clinical relevance, besides a few nearby *cis*-regulatory elements (DNase-H3K4me3, proximal enhancer-like signatures). The intron itself harbors SNPs listed in the NHGRI-EBI GWAS catalog ( $p < 1.0 \times 10^{-5}$ ), with perhaps the most noteworthy being rs13332522, a variant significantly associated with ADHD-related variation in intracranial volume (Klein et al., 2019). For GWAS of epilepsy, the recent mega-analysis by ILAE is the most substantive dataset to-date, (The International League Against Epilepsy Consortium on Complex Epilepsies, 2018) revealing 16 genome-wide significant loci. Although this does not include *RBF0X1*, a suggestive multi-variant association signal is evident in intron 7 for GGE, with a peak at rs11639540 ( $p = 1.84 \times 10^{-6}$ ).

Further, our analysis of genome-wide protein-damaging variants in baboon indicates a potential etiological role for

collagen-containing ECM, with links to epilepsy reported in literature. ECM has long been known to regulate various aspects of neural development, providing structural support (e.g., basement membrane scaffolded by collagen fibrils), and stimulating various pathways that drive progenitor proliferation, neuronal differentiation and migration, and synaptogenesis (Long and Huttner, 2019). During epileptogenesis, the perineuronal net (PNN), a condensed ECM structure that envelops fast-spiking interneurons and that is responsible for synaptic plasticity and excitation-inhibition stabilization, has shown evidence of degradation by endogenous proteases, resulting in impaired GABAergic inhibition (Tewari et al., 2018). Matrix metalloproteinase-9 (MMP-9) is one such PNN protease, which has exhibited increased mRNA and protein levels after seizure in both animal models and patients, including generalized tonic-clonic events, (Cudna et al., 2017) with inhibitors viewed as potential therapeutic drug targets (Broekaart et al., 2021).

These findings provide a compelling first glimpse into the genetic underpinnings of epilepsy in baboons, targeting common, less penetrant risk variants in the allele frequency spectrum. A significant burden of rare, deleterious microdeletions at recurrent hotspots among neurodevelopmental genes and subject to negative selection has been reported in GGE patients, including *RBF0X1* as noted above, and thus, analyses of genome-wide indels in the baboon sequence data, including the delineation of haplotype structures at key loci, could be highly informative. Moreover, the sample size of this preliminary investigation is relatively small and could be boosted by additional baboons from the pedigreed colony, providing the requisite power to detect novel risk variants that are rarer and of smaller effect size. With its extensive archived WGS data and veterinary records, the SNPRC colony could provide a unique resource for epilepsy research, namely, a primate model with a shared genetic etiology, which can garner new insights and lead to much-needed alternative approaches to therapeutic treatment.



## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The animal study was reviewed and approved by IACUC of the University of Texas Health Science Center at San Antonio and the Texas Biomedical Research Institute.

## AUTHOR CONTRIBUTIONS

MK, MC, and CS devised the project. MK processed the whole-genome sequence data and performed the genetic analyses. ML, KK, and CS conducted the retrospective survey of veterinary records. CS performed the diagnoses of epilepsy in the baboon colony. MK and CS took the lead in writing the manuscript. MC, LB, and HG provided critical feedback that helped shape the analysis and manuscript. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Almasy, L., and Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* 62, 1198–1211. doi: 10.1086/301844
- Bhalla, K., Phillips, H. A., Crawford, J., McKenzie, O. L. D., Mulley, J. C., Eyre, H., et al. (2004). The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J. Hum. Genet.* 49, 308–311. doi: 10.1007/s10038-004-0145-4
- Broekaart, D. W. M., Bertran, A., Jia, S., Korotkov, A., Senkov, O., Bongaarts, A., et al. (2021). The matrix metalloproteinase inhibitor IPR-179 has antiseizure and antiepileptogenic effects. *J. Clin. Invest.* 131:e138332.
- Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., et al. (2012a). Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front. Genet.* 3:35. doi: 10.3389/fgene.2012.00035
- Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012b). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6, 80–92. doi: 10.4161/fly.19695
- Conboy, J. G. (2017). Developmental regulation of RNA processing by Rbfox proteins. *Wiley Interdiscip. Rev. RNA* 8. doi: 10.1002/wrna.1398
- Croll, A., Szabo, C. A., Abou-Madi, N., and Devinsky, O. (2019). Epilepsy in nonhuman primates. *Epilepsia* 60, 1526–1538. doi: 10.1111/epi.16089
- Cudna, A., Jopowicz, A., Mierzejewski, P., and Kurkowska-Jastrzębska, I. (2017). Serum metalloproteinase 9 levels increase after generalized tonic-clonic seizures. *Epilepsy Res.* 129, 33–36. doi: 10.1016/j.eplepsyres.2016.11.006
- Durón, R. M., Medina, M. T., Martínez-Juárez, I. E., Bailey, J. N., Perez-Gosíengfiao, K. T., Ramos-Ramírez, R., et al. (2005). Seizures of idiopathic generalized epilepsies. *Epilepsia* 46(Suppl. 9), 34–47.
- Epi 4K Consortium and Epilepsy Phenome/Genome Project (2017). Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol.* 16, 135–143.
- Fisher, R. S., Acevedo, C., Arzimanoglu, A., Bogacz, A., Cross, J. H., Elger, C. E., et al. (2014). A practical definition of epilepsy. *Epilepsia* 55, 475–482.
- Fogel, B. L., Wexler, E., Wahnich, A., Friedrich, T., Vijayendran, C., Gao, F., et al. (2012). RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum. Mol. Genet.* 21, 4171–4186. doi: 10.1093/hmg/dd240

## FUNDING

This study was supported by NIH/NINDS [Grant Nos: R01 NS047755 (L Blondell) and R21 NS084198 (CA Szabó)], NIH/NCRR (Grant No: P51 RR013986), and NIH/ORIP (Grant No: P51 OD011133); and conducted in facilities constructed with support from NIH Research Facilities Improvement grants C06 RR013556, C06 RR014578, and C06 RR015456.

## ACKNOWLEDGMENTS

Baboon pedigree data used in the heritability estimation and MGA analyses were kindly provided by MC Mahaney.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.714282/full#supplementary-material>

- Gallant, N. M., Baldwin, E., Salamon, N., Dipple, K. M., and Quintero-Rivera, F. (2011). Pontocerebellar hypoplasia in association with de novo 19p13.1p13.12 microdeletion. *Am. J. Med. Genet. A* 155A, 2871–2878. doi: 10.1002/ajmg.a.34286
- Gauguier, D., van Luijckelaar, G., Bihoreau, M. T., Wilder, S. P., Godfrey, R. F., Vossen, J., et al. (2004). Chromosomal mapping of genetic loci controlling absence epilepsy phenotypes in the WAG/Rij rat. *Epilepsia* 45, 908–915. doi: 10.1111/j.0013-9580.2004.13104.x
- Gelman, L. T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., et al. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat. Genet.* 43, 706–711. doi: 10.1038/ng.841
- Gros-Louis, F., Dupre, N., Dion, P., Fox, M. A., Laurent, S., Verreault, S., et al. (2007). Mutations in SYNE1 lead to a newly discovered form of autosomal recessive cerebellar ataxia. *Nat. Genet.* 39, 80–85. doi: 10.1038/ng1927
- Hamada, N., Ito, H., Iwamoto, I., Morishita, R., Tabata, H., and Nagata, K. (2015). Role of the cytoplasmic isoform of RBFOX1/A2BP1 in establishing the architecture of the developing cerebral cortex. *Mol. Autism* 6:56.
- Institute for Laboratory Animal Research. (1996). *Guide for the Care and Use of Laboratory Animals*. Washington, DC: National Academies Press.
- Jones, N. C., O'Brien, T. J., and Powell, K. L. (2011). Morphometric changes and molecular mechanisms in rat models of idiopathic generalized epilepsy with absence seizures. *Neurosci. Lett.* 497, 1185–1193.
- Jun, G., Manning, A., Almeida, M., Zawistowski, M., Wood, A. R., Teslovich, T. M., et al. (2018). Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees. *Proc. Natl. Acad. Sci. U. S. A.* 115, 379–384.
- Kjeldsen, M. J., Kyvik, K. O., Christensen, K., and Friis, M. L. (2001). Genetic and environmental factors in epilepsy: a population-based study of 11,900 Danish twin pairs. *Epilepsy Res.* 44, 167–178. doi: 10.1016/s0920-1211(01)00196-6
- Klein, M., Walters, R. K., Demontis, D., Stein, J. L., Hibar, D. P., Adams, H. H., et al. (2019). Genetic markers of ADHD-related variations in intracranial volume. *Am. J. Psychiatry* 176, 228–238. doi: 10.1176/appi.ajp.2018.18020149
- Lal, D., Pernhorst, K., Klein, K. M., Reif, P., Tozzi, R., Toliat, M. R., et al. (2015). Extending the phenotypic spectrum of RBFOX1 deletions: sporadic focal epilepsy. *Epilepsia* 56, e129–e133.
- Lal, D., Reinthaler, E. M., Altmüller, J., Toliat, M. R., Thiele, H., Nürnberg, P., et al. (2013a). RBFOX1 and RBFOX3 mutations in rolandic epilepsy. *PLoS One* 8:e73323. doi: 10.1371/journal.pone.0073323

- Lal, D., Trucks, H., Möller, R. S., Hjalgrim, H., Koeleman, B. P. C., de Kovel, C. G. F., et al. (2013b). Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. *Epilepsia* 54, 265–271. doi: 10.1111/epi.12084
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv[Preprint]* arXiv:1303.3997,
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698
- Li, J., and Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95, 221–227. doi: 10.1038/sj.hdy.6800717
- Long, K. R., and Huttner, W. B. (2019). How the extracellular matrix shapes neural development. *Open Biol.* 9:180216. doi: 10.1098/rsob.180216
- Löscher, W. (2017). Animal models of seizures and epilepsy: past, present, and future role for the discovery of antiseizure drugs. *Neurochem. Res.* 42, 1873–1888. doi: 10.1007/s11064-017-2222-z
- Martin, C. L., Duvall, J. A., Ilkin, Y., Simon, J. S., Arreaza, M. G., Wilkes, K., et al. (2007). Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 869–876. doi: 10.1002/ajmg.b.30530
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- Myers, C. T., and Mefford, H. C. (2016). Genetic investigations of the epileptic encephalopathies: recent advances. *Prog. Brain Res.* 226, 35–60. doi: 10.1016/bs.pbr.2016.04.006
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rajman, M., Metge, F., Fiore, R., Khudayberdiev, S., Aksoy-Aksel, A., Bicker, S., et al. (2017). A microRNA-129-5p/Rbfox crosstalk coordinates homeostatic downscaling of excitatory synapses. *EMBO J.* 36, 1770–1787. doi: 10.15252/emboj.201695748
- Ran, X., Jinchen, L., Shao, Q., Chen, H., Lin, Z., Sun, Z. S., et al. (2015). EpilepsyGene: a genetic resource for genes and mutations related to epilepsy. *Nucleic Acids Res.* 43, D893–D899.
- Rogers, J., and Hixson, J. E. (1997). Insights from model systems: baboons as an animal model for genetic studies of common human disease. *Am. J. Hum. Genet.* 61, 489–493. doi: 10.1086/515527
- Rogers, J., Raveendran, M., Harris, R. A., Mailund, T., Leppälä, K., Athanasiadis, G., et al. (2019). The comparative genomics and complex population history of *Papio* baboons. *Sci. Adv.* 5:eaau6947.
- Rudolf, G., Bihoreau, M. T., Godfrey, R. F., Wilder, S. P., Cox, R. D., Lathrop, M., et al. (2004). Polygenic control of idiopathic generalized epilepsy phenotypes in the genetic absence rats from Strasbourg (GAERS). *Epilepsia* 45, 301–308. doi: 10.1111/j.0013-9580.2004.50303.x
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Szabó, C. Á., Knape, K. D., Leland, M. M., Bauer, C., and Williams, J. T. (2014). Craniofacial trauma as a clinical marker of seizures in baboon colony. *Comp. Med.* 64, 135–139.
- Szabó, C. Á., Knape, K. D., Leland, M. M., Cwikla, D. J., Williams-Blangero, S., and Williams, J. T. (2012). Epidemiology and characterization of seizures in pedigreed baboon colony. *Comp. Med.* 62, 535–538.
- Szabó, C. Á., Knape, K. D., Leland, M. M., and Williams, J. T. (2013). Electroclinical phenotypes in a pedigreed baboon colony. *Epilepsy Res.* 105, 77–85. doi: 10.1016/j.eplepsyres.2013.02.012
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613.
- Tewari, B. P., Chaunsali, L., Campbell, S. L., Patel, D. C., Goode, A. E., and Sontheimer, H. (2018). Perineuronal nets decrease membrane capacitance of peritumoral fast spiking interneurons in a model of epilepsy. *Nat. Commun.* 9:4724.
- The International League Against Epilepsy Consortium on Complex Epilepsies. (2018). Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat. Commun.* 9:5269.
- Wang, L., Zhang, W., and Li, Q. (2020). AssocTests: an R package for genetic association studies. *J. Stat. Softw.* 94, 1–26.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kos, Carless, Blondell, Leland, Knape, Göring and Szabó. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.