Genetics inMedicine | SYSTEMATIC REVIEW

EFHC1 variants in juvenile myoclonic epilepsy: reanalysis according to NHGRI and ACMG guidelines for assigning disease causality

Julia N. Bailey, PhD[1,2,3], Christopher Patterson, BA[1,2], Laurence de Nijs, PhD[2,4], Reyna M. Durón, MD[1,2,5], Viet-Huong Nguyen, PharmD, MPH[1,2,6], Miyabi Tanaka, MD[1,2,7], Marco T. Medina, MD[2,8], Aurelio Jara-Prado, PhD[2,9], Iris E. Martínez-Juárez, MD, MSc[2,9], Adriana Ochoa, MSc[2,9], Yolli Molina, MSc[2,8], Toshimitsu Suzuki, PhD[2,10], María E. Alonso, MD[2,9], Jenny E. Wight, MPH[1,2], Yu-Chen Lin, MPH[1,2], Laura Guilhoto, MD[2,11], Elza Marcia Targas Yacubian, MD[2,11], Jesús Machado-Salas, MD, PhD[1,2], Andrea Daga, PhD[2,12], Kazuhiro Yamakawa, PhD[2,10], Thierry M. Grisar, MD, PhD[2,4], Bernard Lakaye, PhD[2,4] and Antonio V. Delgado-Escueta, MD[1,2,7]

Purpose: *EFHC1* variants are the most common mutations in inherited myoclonic and grand mal clonic-tonic-clonic (CTC) convulsions of juvenile myoclonic epilepsy (JME). We reanalyzed 54 *EFHC1* variants associated with epilepsy from 17 cohorts based on National Human Genome Research Institute (NHGRI) and American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of sequence variants.

Methods: We calculated Bayesian LOD scores for variants in coinheritance, unconditional exact tests and odds ratios (OR) in case–control associations, allele frequencies in genome databases, and predictions for conservation/pathogenicity. We reviewed whether variants damage *EFHC1* functions, whether *efhc1*$^{-/-}$ KO mice recapitulate CTC convulsions and "microdysgenesis" neuropathology, and whether supernumerary synaptic and dendritic phenotypes can be rescued in the fly model when *EFHC1* is overexpressed. We rated strengths of evidence and applied ACMG combinatorial criteria for classifying variants.

Results: Nine variants were classified as "pathogenic," 14 as "likely pathogenic," 9 as "benign," and 2 as "likely benign." Twenty variants of unknown significance had an insufficient number of ancestry-matched controls, but ORs exceeded 5 when compared with racial/ethnic-matched Exome Aggregation Consortium (ExAC) controls.

Conclusions: NHGRI gene-level evidence and variant-level evidence establish *EFHC1* as the first non–ion channel microtubule–associated protein whose mutations disturb R-type VDCC and TRPM2 calcium currents in overgrown synapses and dendrites within abnormally migrated dislocated neurons, thus explaining CTC convulsions and "microdysgenesis" neuropathology of JME.

Genet Med advance online publication 28 July 2016

Key Words: causality; *EFHC1*; juvenile myoclonic epilepsy; whole-exome sequencing

INTRODUCTION

On 12 September 2011, the US National Human Genome Research Institute (NHGRI) convened an expert working group to address the challenges of assigning disease causality to sequence variants. Clear guidelines for distinguishing disease-causing sequence variants from false-positive reports of causality were provided.[1] The US Centers for Disease Control and Prevention[2] in the same year and the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG) in 2013, concerned about accuracy of clinical laboratory reports to clinical practitioners, also convened a workgroup consisting of clinical service providers.[3]

The NHGRI working group cautioned that the vast majority of genes reported as causally linked to monogenic diseases are true positives, but 27% of 406 published severe disease mutations in 104 sequenced individuals either were common polymorphisms or lacked direct evidence for pathogenicity.[4–8] The NHGRI working group defined rare germ-line variants with minor allele frequencies of <0.01 that have relatively large

The first two authors contributed equally to this work.
[1]Epilepsy Genetics/Genomics Lab, Neurology and Research Services, VA GLAHS/UCLA, Los Angeles, California, USA; [2]GENESS International Consortium; [3]Department of Epidemiology, Fielding School of Public Health, UCLA, Los Angeles, California, USA; [4]GIGA-Neurosciences, University of Liège, Liège, Belgium; [5]Facultad de Ciencias de la Salud, Universidad Tecnológica Centroamericana (UNITEC), Tegucigalpa, Honduras; [6]Chapman University School of Pharmacy, Irvine, California, USA; [7]Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, California, USA; [8]National Autonomous University of Honduras, Tegucigalpa, Honduras; [9]National Institute of Neurology and Neurosurgery, Mexico City, Mexico; [10]Neurogenetics Lab, RIKEN Brain Science Institute, Saitama, Japan; [11]Unidade de Pesquisa e Tratamento das Epilepsias (UNIPETE), Universidade Federal de Sao Paulo (UNIFESP-EPM), Sao Paulo, Brazil; [12]Eugenio Medea Scientific Institute, Conegliano and Dulbecco Telethon Institute, Italy.
Correspondence: Antonio V. Delgado-Escueta (AEscueta@mednet.ucla.edu) or Julia N. Bailey (JBailey@mednet.ucla.edu)

Submitted 23 December 2015; accepted 9 May 2016; advance online publication 28 July 2016. doi:10.1038/gim.2016.86

effects on disease risk and variants that have been implicated in severe monogenic diseases and complex diseases.[9] The working groups' intended scope excluded common small-effects variants typically identified by genome-wide association studies of complex traits.[10,11] Because "unambiguous assignment of disease causality for sequence variants is often impossible," the NHGRI working group introduced the concept of "implication"—implicating by genetic evidence sequence variant(s) of a gene that is in the process of integrating and assessing experimental evidence for pathogenicity.[1]

For "implication," the NHGRI workgroup emphasized the "critical primacy of strong robust statistical genetic support," such as coinheritance in family studies, case–control association studies, allele frequencies in public genome databases, and predictions for conservation and pathogenicity. Strong statistical genetic support for "implication" is then supplemented by variant specific experimental studies that demonstrate that a gene product is functionally disrupted by variants. The NHGRI workgroup values disease models that recapitulate the relevant pathology of human disease and allow rescue of the phenotype when the molecular disease pathway is knocked out or eliminated.[1] The ACMG workgroup, concerned more with reports of clinical genomic testing that impact medical decision making, took these evidentiary data and rated their strengths as "very strong," "strong," "moderate," and "supportive." They set rules for combining the strengths of evidentiary data when classifying sequence variants into "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign"[3] (**Figure 1**).

The NHGRI core guidelines,[1] the ACMG consensus recommendations for interpretation of sequence variants,[3] and large genome databases representing different racial and continental populations, such as the Genome 1000 (ref. 12), the Exome Variant Server 6500 (ref. 13), and the Exome Aggregation Consortium (ExAC) (ref. 14), were not available in 2004 when variants of the EF-hand domain (C-terminal) containing 1 gene (*EFHC1*) were reported as disease-causing mutations in myoclonic and grand mal clonic-tonic-clonic (CTC) convulsions produced by juvenile myoclonic epilepsy (JME). Consequently, all *EFHC1* variants discovered in the first decade of this millennium and reported with respect to epilepsy or not[15–28] have not been "vetted" through NHGRI and ACMG guidelines. More importantly, both NHGRI and ACMG guidelines advise that "with evidence on variants evolving" and the "content of sequencing tests expanding," "rigorous evaluation" and "reanalysis of variants are encouraged" to prevent misannotation of the pathogenicity of variants in public databases. For all these reasons, we applied NHGRI guidelines and ACMG rules (**Figure 1**) for combining evidentiary criteria in reanalyzing 54 *EFHC1* variants, of which 33 were originally published as mutations.

MATERIALS AND METHODS

We gathered 54 *EFHC1* variants reported in regard to epilepsy from scientific and medical literature, bibliographic resources from the NCBI PubMed literature server, the ClinVar database,[15] and personal communications with authors of abstracts and posters published during neurological, epilepsy, and genetic meetings. The 33 purported *EFHC1* mutations are scattered across the 640 amino acid protein of Myoclonin1 (**Figure 2**; **Supplementary Table S1** online places all 54 variants in the GRCh37/hg19 coordinate system, provides rsID number if identified in dbSNP 142, and translates the cDNA and protein nomenclature for all four coding transcripts identified by Ensembl).

Coinheritance

Twelve families were reported with *EFHC1* variants cosegregating identically by descent with all disease-affected members[16,17,19,23,25] and with two variants that did not cosegregate (**Table 1**). We calculated Bayesian factor linkage likelihood to evaluate the significance of sequence variants[29] based on the pedigrees as published. Eleven different genotyped *EFHC1* variants that were found in these families were used as marker alleles. JME was assumed to be in linkage disequilibrium with the markers and at = 0, with a standard penetrance model for JME (pn_{p2} = 0.001; pn_{pq} = 0.7; pn_{q2} = 0.7). We corrected for ascertainment bias as proposed by Thompson et al.[29]

Case–control association studies

Supplementary Table S2 online lists the study design of all published case–control studies, their population groups, their specific racial/ethnic groups and countries of their residencies, the number of index cases and controls, and the targets used in screening for *EFHC1* mutations in 12 cohorts from 9 countries.

We studied the association of *EFHC1* variants with JME or genetic generalized epilepsy index cases versus the association of *EFHC1* variants with ancestry/race-matched controls as originally published. **Table 2** summarizes the actual results of the unconditional exact homogeneity/independence test (Z-pooled method, one-tailed),[30,31] which assessed whether the proportion of variants associated between the two groups reflected a statistically significant difference. **Supplementary Table S3** online provides all the details of the case–control studies. Because many of the published studies on *EFHC1* had insufficient control sample sizes, we also calculated odds ratios (OR) and statistical significance (*P* values) for both the study as published and the allele counts available in race-matched population groups from the ExAC database.[14] An OR of 1.0 means that the variant does not affect the odds of having the disease; values higher than 1.0 indicate that there is an association between the variant and the risk for the disease.

Allele frequencies

We extracted the minor allele frequencies for all putative pathogenic *EFHC1* variants in exomes from race-matched and ancestry-matched presumably normal populations of 6,503 persons stored in the 2013 Exome Variant Server[13] (ESP6500SI-V2), in genome sequences of 2,504 persons of the 1000 Genomes Project[12] (phase 3, release 16), and in 60,706, exomes collected by the ExAC consortium (**Supplementary Table S4a**,**b** online).[14]

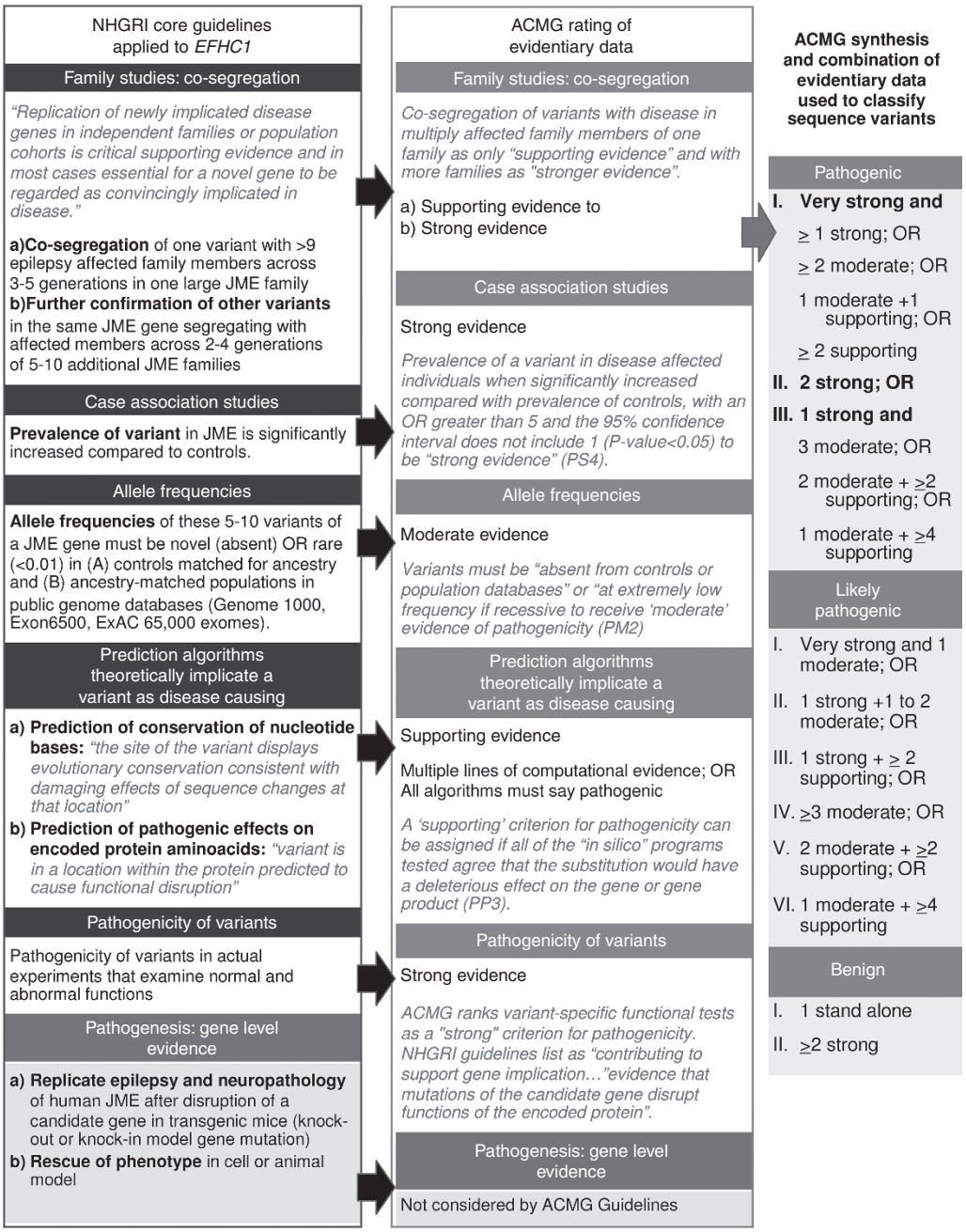

Figure 1 Assigning disease causality to sequence variants according to the National Human Genome Research Institute and American College of Medical Genetics and Genomics Guidelines.

Studies of epilepsy prevalence were selected primarily on the basis of whether they classified electroclinical syndromes such as genetic generalized epilepsies, JME, and childhood absence epilepsy in observed cases. Population prevalence for Hispanics was determined by a door-to-door study performed in rural Bolivia in 1999.[32] Prevalence among Caucasians was determined by a study of the Norwegian National registry in 2015,[33] and East Asian prevalence by a study of the regional registry of patients older than 15 years in Hong Kong, China.[34] Another estimate of Southeast Asian prevalence was produced via a random cluster survey of Cambodian villages.[35] Differing in methodology and ascertainment and not ideally race-matched to JME index cases studied, these studies represent the only data we could use to compare allele frequencies of *EFHC1* variants with JME disease prevalence.

Algorithms predicting conservation and pathogenicity

We analyzed the theoretical pathogenicity of all 54 variants by applying: (i) four algorithms (PhyloP,[36,37] SiPhy,[38,39] GERP^{++},[40] and PHASTCons[41]) that measure evolutionary

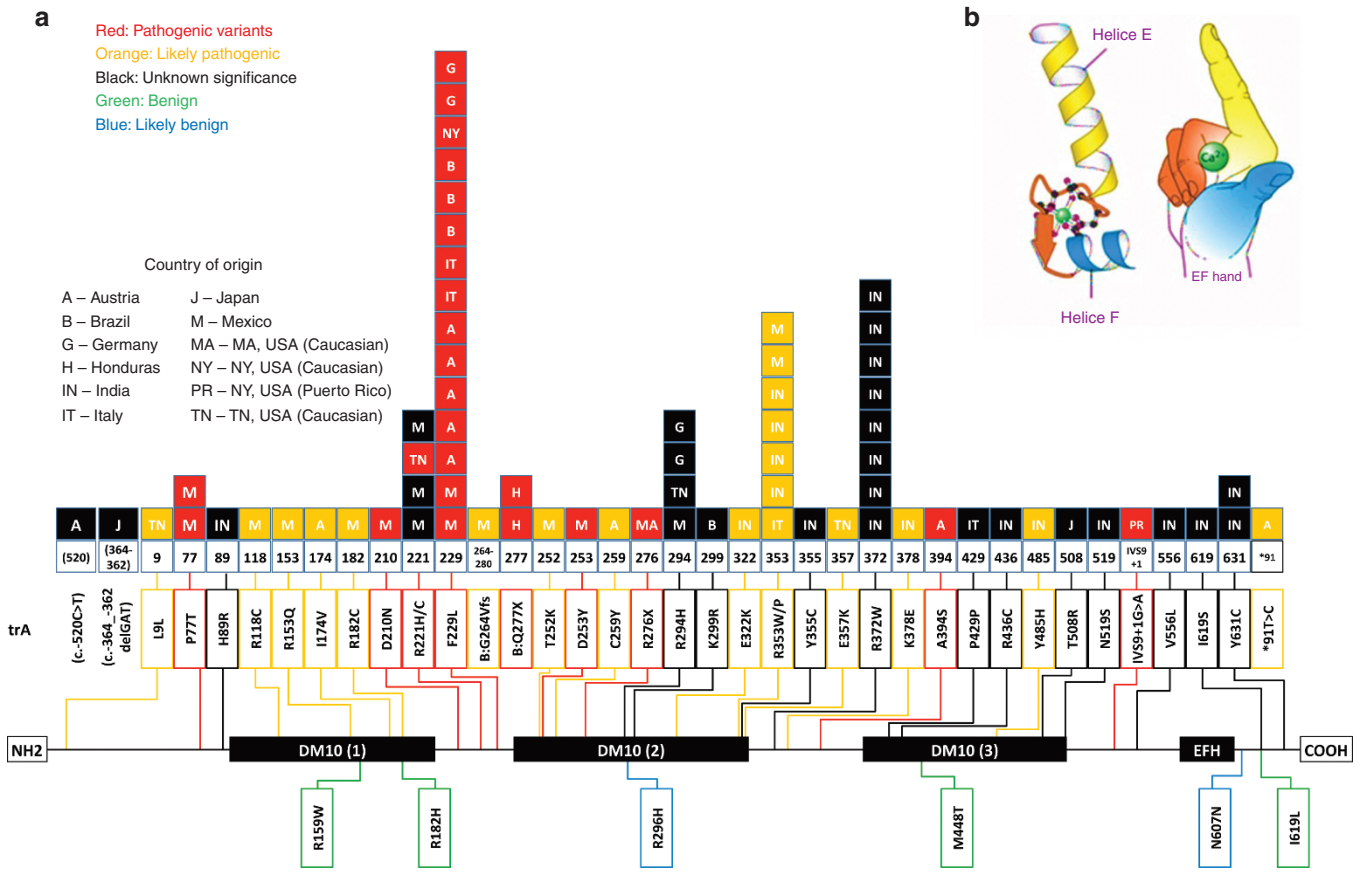

Figure 2 An EF hand–containing calcium-binding gene (*EFHC1*) spans 72 kb, has 11 exons, three DM10 domains (DM refers to *Drosophila melanogaster* sequences), and one EF-hand motif that is calcium binding (HGNC: 16406). (**a**) The *EFHC1* gene encodes a 640-amino-acid protein called myoclonin1. The EF-hand motif is located at the C-terminal between amino acids 578 and 606 and is encoded by a nucleotide sequence that is present in exon 10. Because the only motif of *EFHC1* whose function was known consisted of the EF hand, the gene was first called *EFHC1* for "EF hand containing one."[15] The diagram shows the domain organization of *EFHC1* protein, the positions of various mutations found in JME families, and their frequency (number of independent families with a given mutation). Mutation numbering is based on the GenBank reference protein sequence with accession number 608816. A, Africa; B, Brazil; DM10, Domain 10; H, Honduras; IN*, India; IS, Israel; IT, Italy; J, Japan; M, Mexico; NY, New York. (**b**) Schematic representation of an EF-hand motif comprising two helixes—E and F—linked by a calcium-binding sequence. Symbolic representation of the same motif as a right hand in which the E helix corresponds to the index and the F helix to the thumb. Note: Part b of Figure 2 (EFH domain) has been reproduced with permission (publicly available) from: "http://www.ncbi.nlm.nih.gov/books/NBK98188/" Myoclonin1/*EFHC1* in cell division, neuroblast migration, synapse/dendrite formation in juvenile myoclonic epilepsy. Jasper's Basic Mechanisms of the Epilepsies [Internet]. 4th edition. Noebels JL, Avoli M, Rogawski MA, Olsen RW, Delgado-Escueta AV, editors. Bethesda (MD): "http://www.ncbi.nlm.nih.gov/" National Center for Biotechnology Information (US); 2012.

conservation at the level of DNA base pairs and (ii) seven algorithms (SIFT,[42,43] PolyPhen2-HVAR,[44,45] LRT,[46] Mutation Taster,[47] Mutation Assessor[48] and FATHMM,[49] and CADD[50]) that calculate amino acid conservation and the likelihood of deleteriously altering the encoded amino acid function (**Supplementary Table S5a–c** online). These algorithms were chosen because they were included in the database of Nonsynonymous Functional Predictions (dbNSFP v2.6),[51,52] which is available for download from the ANNOVAR[53] website. Variants found in non-RefSeq-defined transcripts were annotated manually when possible. To determine whether variants might affect the splicing consensus, all were run through the online algorithms provided by Human Splicing Finder.[54] To assess the ACMG criterion for benign status (BP4), we used the Multiz alignment of 62 mammalian species available on the UCSC Genome Browser to determine whether the amino acid substitution would compromise function.

Variant-specific experimental functional studies on pathogenicity

Only 20 variants have undergone variant-level experimental functional studies (**Supplementary Table S6a–c** online). These consist of five *EFHC1* mutations and three polymorphisms, which we originally reported in 2004 (ref. 16) and 12 *EFHC1* mutations reported from India (ref. 28). We summarize the results of functional studies in three tables: **Supplementary Table S6a** online, which presents the molecular and cellular models, **Table 6b**, which displays *in vivo* models of neurodevelopment, and **Table 6b**, which covers the protein–protein interactions (PPIs). We also report the level of statistical significance between each variant and the wild-type allele published in original articles.[55–58]

Table 1 Cosegregation analysis of *EFHC1* variants found in juvenile myoclonic epilepsy families

Variant	Family	Clinically affected	SW/PSW-affected	Asymptomatic carriers	Non carriers	Clinical penetrance	Clinical and EEG penetrance	Bayesian factor LOD score	Pathogenic odds ratio
Y33H	Index4a[25]	3	0	0	0	1.0000	1.0000	0.3106	2.04
P77T/R221H	Mexico1[16]	2	3	1	4	0.3333	0.8333	1.9492	88.97
	Mexico2[16]	2	1	1	0	0.5000	0.7500	0.3016	2.00
	Total	4	4	2	4	0.4000	0.8000	2.2509	178.19
R118C	MexicoA[18]	3	0	0	1	1.0000	1.0000	0.4874	3.07
D210N	Mexico5[16]	2	1	1	0	0.5000	0.7500	–0.0350	0.92
R221H	TN1[23]	2	0	2	1	0.5000	0.5000	0.7791	6.01
	Total[a]	6	4	4	5	0.4286	0.7143	3.0300	1,071.43
F229L	Mexico3[16]	2	1	0	0	0.6667	1.0000	0.3004	2.00
	Mexico4[16]	3	2	2	4	0.4286	0.7143	0.5998	3.98
	Italy13[19]	2	0	1	0	0.6667	0.6667	1.0155	10.36
	Total	7	3	3	4	0.5385	0.7692	1.9157	82.35
D253Y	Mexico6[16]	2	0	0	1	1.0000	1.0000	0.4874	3.07
Q277X	Honduras15[17]	4	5	5	14	0.2857	0.6429	2.4434	277.59
R353W	Italy25[19]	3	0	0	0	1.0000	1.0000	–0.0489	0.89
R353Q	Index4b[25]	1[b]	0	3	3	–	–	–4.9308	1.17E-05
P429P	Italy4[19]	2[b]	0		0	–	–	–2.2438	0.01
Total (across all IBD cosegregating variants)		30	13	13	25	0.5357	0.7679		

The pedigrees are summarized here as they were originally published with counts of clinically affected individuals, clinically asymptomatic individuals with SW/PSW EEG traits, asymptomatic carriers of *EFHC1* variants with no observed EEG traits, and noncarrier family members. Estimates of "clinical" and "clinical and EEG trait" penetrances were calculated for each family and across each variant.

[a]The total data for the R221H variant also include the Mexico1 and Mexico2 families in which the variant was found in a double heterozygous haplotype with P77T/R221H.
[b]R353Q cosegregates with only one of the four affected members in this family. P429P segregated with only two of the three affected members in this family.

Paradigm V—gene level functional studies on pathogenesis

EFHC1 function has been studied at the gene level by knocking out *EFHC1* orthologs in both mice[59,60] and flies (**Supplementary Table S7** online).[61] Both mouse and fly models presented seizure-related and electroclinical phenotypes and neuroanatomical measures similar to those in the variant-level functional studies. These measures and their level of statistical significance with respect to the wild type as published are summarized in **Supplementary Table S7** online.

RESULTS

EFHC1

Ensembl identifies a total of four alternative coding transcripts and two noncoding transcripts for *EFHC1* (see **Supplementary Table S1** online). Most reported *EFHC1* variants associated with epilepsies cause amino acid substitutions in two of the isoforms: transcripts A (length: 640aa) and B (length: 278aa). The two transcripts share the first 241-amino-acid sequence, and transcript B translates two additional variants, causing a frameshift and nonsense change in the protein, but it is prematurely truncated at the end of exon 4. Transcript B retains the first DM10 domain (see **Figure 2** for an illustration of *EFHC1*). Transcript B is not expressed in the mouse brain, but it is expressed in human and chimpanzee neural tissue. Transcripts C and D have not been evaluated for their role in epilepsy or their expression in other mammalian species.

Cosegregation

Of 33 putative pathogenic variants, eight were reported to cosegregate with 40 clinically and EEG polyspike wave–affected family members across two to four generations of 12 JME families in four separate cohorts from Mexico, Honduras, Italy, and Tennessee (USA).[16,17,19,23] The remaining variants were detected only in singletons (**Table 1**).

Table 1 summarizes the pedigrees studied and calculates the estimated penetrances within and across all families. We reanalyzed the patterns of cosegregation with JME and the EEG polyspike wave trait in affected carrier families and calculated a Bayesian method for evaluating causality of variants[29] using eight *EFHC1* variants as markers (**Table 1**). Our reanalysis identified the following variants as 3.07- to 1,000-times more likely (LOD: 0.4874 to 3.0300) to have cosegregation occur not by chance. P77T/R221H (as a double heterozygous variant), R221H, R118C, R221H, D253Y, F229L, and Q277X were identified as single autosomal dominant heterozygous variants.

The families in which R353W and D210N were found were too small and hence underpowered to detect significant linkage. P77T/R221H and trB:Q277X both had LOD scores >2.0

Table 2 OR or risk for JME or GGE derived from case–control studies on putative pathogenic *EFHC1* variants as originally published

		Cases						Study-specific			
cDNA	Protein	JME		GGEs		Controls			OR	*P* value	
Pathogenic variants											
229C>A	**P77T**										
	Latino	2	64	–	–	0	504		∞	0.0043	**
	Latino	0	46	0	92	1	120		0	1	
	African	–	–	0	34	2	46	GGE	0	1	
628G>A	**D210N**										
	Latino	1	64	–	–	0	504		∞	0.0467	*
662G>A	**R221H**										
	Latino	2	64	–	–	0	504		∞	0.0043	**
	Latino	0	46	0	92	1	120		0	1	
	African	–	–	0	34	2	46	GGE	0	1	
	Caucasian	1	108	–	–	0	240		∞	0.1396	
685T>C	**F229L**										
	Latino	2	64	–	–	0	504		∞	0.0043	**
	Caucasian	–	–	1	184	0	206	GGE	∞	0.2425	
	Caucasian	2	54	–	–	0	100		∞	0.0469	*
	Caucasian	‡	48	‡	122	4	736		‡	‡	
	South American	3	204	–	–	0	100		∞	0.1827	
	Caucasian	2	76	–	–	–	–		–	–	
829C>T	**Q277X**										
	Latino	2	88	–	–	0	1,246		∞	0.0013	**
757G>T	**D253Y**										
	Latino	1	64	–	–	0	504		∞	0.0467	*

OR of "∞" occurred when the variant was not detected in controls. *P* values for each were calculated using an unconditional exact test (Z-pooled, one-tailed) for results within the study only.

‡Syndrome-specific allele counts were not provided in the referenced paper. Moreover, other epilepsy patients (e.g., temporal lobe epilepsy) were included in the cohort. For these reasons, specific risk for JME or GGE could not be calculated.

**P* value <0.05. ***P* value <0.01.

GGE, genetic generalized epilepsy; JME, juvenile myoclonic epilepsy; OR, odds ratios.

(pathogenic OR >100×), which is suggestive of linkage. R221H (including in the two families where it was found with P77T) had a LOD score >3.0, which is significant for linkage. Our reanalysis does not show whether the combination of P77T and R221H produces a disruption of *EFHC1* function or whether R221H by itself is causal.

One reported variant, P429P, segregated with two out of three affected individuals in a family from Italy, and presumptively in trans with the nonpathogenic variant R182H, which segregated with all three affected individuals.[19] The Bayesian LOD score for this variant was −2.2438. Another variant, R353Q, cosegregated with one of the four clinically affected members in the family[25] with a Bayesian LOD score of −4.9308. These variants meet the "strong" criterion (BS4) for benign status.

Our reanalysis of cosegregating families indicated that *EFHC1* variants were transmitted in an autosomal dominant manner and suggested the implication of *EFHC1* variants in JME. In weighing the value of cosegregation as ACMG evidentiary data, we first considered it as only "supporting" evidence for pathogenicity because it is not clear from ACMG guidelines how to weight evidence of cosegregation in one large multigenerational family. However, because there was increasing segregation data in at least seven families from diverse ethnic backgrounds, these cosegregation data could be weighed by ACMG as moderate to strong evidence.[3]

Case association studies

The 12 case–control association studies yielded 32 putative pathogenic variants (note: R276X was not studied in case–controls) (**Table 2** and **Supplementary Table S3** online). An additional four variants (L9L, R353P, E357K, and P429P) were originally published as polymorphisms, because they either were found in one control or produced a synonymous change in the protein product. These four variants were found to be

absent in their race-matched population in ExAC.[14] The final variant included in **Supplementary Table S3** online is R159W, which has been reported as benign in nearly all studies but reached statistical significance during case association within two studies. It should be noted that the variant was no longer significant when compared with the ExAC population data. For each variant, all studies that genotyped the variant in a case or a control are summarized in **Table 2** and all details are presented in **Supplementary Table S3** online.

Within the scope of their specific study design, nine variants reached statistical significance and had an OR >5: P77T, D210N, R221H, F229L, trB: G264Vfs*280, trB:Q277X, T252K, D253Y, and c.*91T>C. When using the race-matched ExAC population data, all of the variants listed here still reached statistical significance within their own study, and an additional 18 variants also met the ACMG criterion for strong evidence (PS4). Three variants were replicated and met the criterion in two or more studies: R221H, F229L, and R353W.

Allele frequencies

Eight variants were completely absent across all populations in the ExAC database: C259Y, R276X, E322K, K378E, A394S, P429P, c.1640+1G>A, and c.*91T>C (**Supplementary Table S4a**,**b** online). Another eight variants were absent in the race-matched subpopulation in ExAC: L9L, R118C, R152Q, I174V, T252K, R353P, E357K, and Y485H. Two additional variants were absent in the race-matched European American subpopulation of the ESP6500si database: R221H and R353W.

Of the 54 variants we examined, 17 were found at allele frequencies greater than the expected population prevalence: P77T, H89R, R182C, D210N, R221H (only in the Latino subpopulation), F229L, trB:G264Vfs, trB:Q277X, R294H, R353W, R436C, M448T, T508R, N607N, I619S, and Y631C. The ESP6500si database corroborated only three of these variants as being greater than the population prevalence: F229L, M448T, and R294H. Finally, six variants met the stand-alone BA1 criteria based on their allele frequency in ExAC: c.-148_147delGC, R159W, R182H, c.573+10A>G, I619L, and c.*121C>A. ExAC and ESP6500si do not target intronic regions, so certain variants do not have allele frequencies. 1000 Genomes identifies two additional intronic variants, c.1492+175_176delTT and c.1851+59C>T, that were at allele frequencies >0.05.

In silico analysis for conservation and damaging effects

All exonic variants, except the two transcript B variants, were predicted to be evolutionarily constrained across a phylogeny of species by at least one nucleotide conservation measure (**Supplementary Table S5a–c** online). None of the conservation scores predicted either of the transcript B variants to be conserved across species. This is predictable because transcript B was not found to be expressed in the mouse brain, but it was expressed in humans and chimpanzees; therefore, it may not be under evolutionary constraint across the entirety of the vertebrate or mammalian clades. V556L was not found to be conserved by any of the nucleotide-based calculations, but it was determined to be part of a conserved element by GERP++ spanning exon 10.

Eight exonic variants were predicted to be benign by all eight pathogenicity algorithms: P77T, R221H, R353P, E357K, A394S, M448T, K378E, and V556L. Of these variants, only P77T and R221H underwent experimental functional testing and were found to have a significant effect on several measures of calcium channel–dependent activities and neurodevelopment (see Results and Discussion, and **Supplementary Table S6a,b** online). The other variants that showed similar significant differences in functional experiments were D253Y (predicted to be pathogenic by five of the algorithms), D210N (predicted to be pathogenic by seven), and F229L (predicted to be pathogenic by four algorithms). The only variant that was predicted to be pathogenic by all eight measures—the ACMG requirement for "supporting" pathogenic evidence (PP3)—was R436C. It should be noted that the FATHMM algorithm predicted all but one of the variants to be benign. Five variants were found as the reference allele in two or more mammalian species: R221H, R296H, M448T, and Y355C, which meets the requirement for "supporting" benign evidence (BP4).

Human Splice Finder results are summarized in **Supplementary Table S5c** online. Ten variants were predicted to create new donor or acceptor sites, and one variant (c.1640+1G>A) disrupted the wild-type donor sites. Thirty-one variants were predicted to disrupt splicing enhancer motifs, and 17 variant were predicted to create new splicing silencer motifs. Eight variants were predicted to not alter the splicing consensus by any of the algorithms. We applied the PP3 and BP4 criteria only to synonymous variants and those affecting canonical splicing sites.

Variant-specific experimental evidence for pathogenicity

Only a few variants of *EFHC1* have undergone functional testing (**Supplementary Table S6a–c** online). Five *EFHC1* variants were originally reported in JME patients from Mexico[16] (P77T, D210N, R221H, F229L, and D253Y), in reverse TRPM2-induced (transient receptor potential calcium permeable M2 channel) apoptosis, and in current densities.[57] Four of these variants (the exception was P77T, which was not tested) produce severe mitotic spindle defects during cell division[55] and impair early radial and tangential migration of neuroblasts,[56] thus providing experimental evidence that *EFHC1* variants are damaging to gene function. **Supplementary Table S6a**,**b** online show the published statistical results of 14 experimental measures demonstrating a significant difference between the tested variants and the wild-type protein. Three variants (R159W, R182H, and I619L) classified as benign polymorphism did not produce statistically significant results in almost all of the measures in comparison to the wild type.[16,55–57] R182H showed a small, but significant difference in apoptotic activity in primary mouse hippocampal neurons in culture.[16]

Most recently, Sahni et al.[58] demonstrated that wild-type *EFHC1* proteins interacted with products of 16 genes. Of the *EFHC1* disease alleles tested, R221H and A394S perturbed

Table 3 Combination of evidentiary data, their weight, and strength of evidence used to classify variants into the five-tier American College of Medical Genetics and Genomics classification of sequence variants

Nine pathogenic variants	
c.229C>A	p.P77T
Strong	• Experimental functional studies show damaging effect on function or dominant/negative effect on calcium signaling and apoptosis. Perturbed 15 of 16 protein–protein-interactions (PPIs) detected in the *EFHC1*-WT (Edgetic). • Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0043$) and ExAC population rates (OR: 19.61; $P = 0.0140$)
Support	• Found to cosegregate with four clinically affected and four asymptomatic SW-affected individuals in two families (Bayesian LOD score: 2.2509)
Benign	• MAF in ExAC is greater than expected for the disorder
c.628G>A	p.D210N
Strong	• Experimental functional studies show damaging effect on function or dominant/negative effect on calcium signaling, apoptosis, and neuroblast migration • Prevalence significantly increased in JME over both study controls (OR: ∞; $p = 0.0467$) and ExAC population rates (OR: 26.23; $P = 0.0244$)
Support	• Found to cosegregate with two clinically affected and one asymptomatic SW-affected individuals in one family (Bayesian LOD score: –0.0350)
Benign	• MAF in ExAC is greater than expected for the disorder
c.662G>A	p.R221H
Strong	• Experimental functional studies show damaging effect on function or dominant/negative effect on calcium signaling, apoptosis, and neuroblast migration. Perturbed all 16 protein–protein interactions detected in the *EFHC1*-WT (quasi-null) • Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0043$) and ExAC population rates in two studies (OR: 18.63, $P = 0.0036$; OR: 311.84, $P = 0.0006$)
Moderate	• Absent in European Americans in ESP6500si
Support	• Found to cosegregate with six clinically affected and four asymptomatic SW-affected individuals in three families (Bayesian LOD score: 3.0300)
Benign	• MAF in ExAC is greater than expected for the disorder (for Latinos) • Found as reference allele in 4 out of 62 mammalian species
c.685T>C	p.F229L
Strong	• Experimental functional studies show damaging effect on function or dominant/negative effect on calcium signaling and apoptosis • Prevalence significantly increased in JME over both study controls in two studies (OR: ∞, $P = 0.0043$; OR: ∞, $P = 0.0469$) and ExAC population rates in three studies (OR: 16.94, $P = 0.0041$; OR: 8.43, $P = 0.0074$; OR: 5.89, $P = 0.0159$)
Support	• Found to cosegregate with seven clinically affected and three asymptomatic SW-affected individuals in three families (Bayesian LOD score: 1.9157)
Benign	• MAF in ExAC and ESP6500si is greater than expected for the disorder
trB: c.829C>T	trB: p.Q277X
Strong	• Found de novo in clinically affected patient with paternity and maternity confirmed • Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0043$) and ExAC population rates (OR: 19.61; $P = 0.0140$)
Moderate	• Stop-gain mutation truncating the final residue in the transcript
Support	• Found to cosegregate with four clinically affected and five asymptomatic SW-affected individuals in one family (Bayesian LOD score: 2.4434)
Benign	• MAF in ExAC is greater than expected for the disorder
c.757G>T	p.D253Y
Strong	• Found de novo in clinically affected patient with paternity and maternity confirmed • Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0043$) and ExAC population rates (OR: 19.61; $P = 0.0140$)
Support	• Found to cosegregate with four clinically affected and five asymptomatic SW-affected individuals in one family (Bayesian LOD score: 2.4434)
c.826C>T	p.R276X
Very strong	• Nonsense mutation causing the deletion of six exons of the protein
Moderate	• Absent in all populations in the ExAC and ESP6500si • Protein length changes due to in-frame deletions/insertions in a nonrepeat region or stop-loss variants
c.1180G>T	p.A394S
Strong	• Perturbed all 16 protein–protein-interactions detected in the *EFHC1*-WT (quasi-null) • Prevalence significantly increased in GGEs over ExAC population rates (OR: ∞; $P = 2.59E-05$)
Moderate	• Absent in all populations in the ExAC and ESP6500si

ExAC, Exome Aggregation Consortium; GGE, genetic generalized epilepsy; JME, juvenile myoclonic epilepsy; OR, odds ratios.

Table 3 Continued

Table 3 Continued on next page

c.1640+1G>A	
Very strong	• Null variant disrupting the canonical +1 splice site at the end of exon 9
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 0.0015$)
Moderate	• Absent in all populations in the ExAC and ESP6500si
Support	• Predicted to break the WT donor site by Human Splicing Finder
14 likely pathogenic variants	
c.25T>C	p.L9L
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 1.10E-07$)
Moderate	• Absent in Europeans in ExAC and European Americans in ESP6500si
Support	• Human Splice Finder predicted the variant to alter the WT splicing consensus
c.352C>T	p.R118C
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 0.0027$)
Moderate	• Absent in Latinos in ExAC.
Support	• Found to cosegregate with three clinically affected individuals in a small family (Bayesian LOD score: 0.4874)
c.458G>A	p.R153Q
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: 70.83; $P = 0.0097$)
Moderate	• Absent in Latinos in ExAC.
c.520A>G	p.I174V
Strong	• Prevalence significantly increased in GGEs over ExAC population rates (OR: ∞; $P = 2.61E-05$)
Moderate	• Absent in Europeans in ExAC and all populations in ESP6500si
trB: c.786delA	trB: p.G264Vfs
Strong	• Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0260$) and ExAC population rates (OR: 14.57; $P = 0.0174$)
Moderate	• Frameshift deletion altering the final 16 residues in the transcript
Benign	• MAF in ExAC is greater than expected for the disorder
c.755C>A	p.T252K
Strong	• Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0260$) and ExAC population rates (OR: ∞; $P = 0.0028$)
Moderate	• Absent in Latinos in ExAC
c.776G>A	p.C259Y
Strong	• Prevalence significantly increased in GGEs over ExAC population rates (OR: ∞; $P = 2.61E-05$)
Moderate	• Absent in all populations in the ExAC and ESP6500si
c.964G>A	p.E322K
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 2.61E-05$)
Moderate	• Absent in all populations in ExAC
c.1057C>T	p.R353W
Strong	• Prevalence significantly increased in JME over ExAC population rates in two studies (OR: 50.33, $P = 0.0021$; OR: 6.27, $P = 0.207$)
Moderate	• Absent in European Americans in ESP6500si
Support	• Found to cosegregate with three clinically affected individuals in a family (Bayesian LOD score: –0.0489)
Benign	• MAF in ExAC is greater than expected for the disorder
c.1058G>C	p.R353P
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 1.23E-05$)
Moderate	• Absent in Latinos in ExAC
c.1069G>A	p.E357K
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 2.29E-05$)
Moderate	• Absent in Europeans in ExAC and European Americans in ESP6500si
c.1132A>G	p.K378E
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 0.0212$)
Moderate	• Absent in all populations in ExAC
c.1453T>C	p.Y485H
Strong	• Prevalence significantly increased in JME over ExAC population rates (OR: ∞; $P = 0.0198$)
Moderate	• Absent in South Asians in ExAC
c.*91T>C	
Strong	• Prevalence significantly increased in JME over both study controls (OR: ∞; $P = 0.0231$) and ExAC population rates (OR: ∞; $P = 0.0009$)
Moderate	• Absent in all populations in ExAC

ExAC, Exome Aggregation Consortium; GGE, genetic generalized epilepsy; JME, juvenile myoclonic epilepsy; OR, odds ratios.

EFHC1 interaction with all 16 genes. P77T disrupted interaction with all but one protein (TEX11, which is expressed exclusively in male germ cells and therefore may not play a role in epileptogenesis), and T508R disrupted interaction with four proteins. The PPI profiles of the two polymorphisms, R159W and M448T, did not show any significant perturbation of the wild-type network (see **Supplementary Table S6c** online for summary of these findings). Interactions with five gene products—CCDC36 (Coiled-coil Domain Containing 36), EIF4ENIF1 (Eukaryotic Translation Initiation Factor 4E Nuclear Import Factor 1), REL (V-Rel Avian Reticuloendotheliosis Viral Oncogene Homolog), TCF4 (Transcription Factor 4), and ZBED1 (Zinc-Finger, BED containing 1)—were interrupted by all four of the *EFHC1* disease alleles. Four of these proteins—EIF4ENIF1, REL, TCF4, and TRAF2—play important roles in the regulation of neuron-specific differentiation or apoptosis.[62–65] Two of the interactors, GOLGA2 and ZBED1, have been implicated in cell-cycle control and cell proliferation.[66,67] TRIP6 is a positive regulator of lysophosphatidic acid (LPA)-induced cell migration.[68] Finally, three of the proteins whose interactions were perturbed—REL, TRAF2, and TRIP6—play roles in the NF-κB signaling pathway and have been implicated in the processes of hippocampal synaptic plasticity and memory.[69]

Gene-level experimental evidence for pathogenesis

Supplementary Table S7 online summarizes 20 gene-level experimental studies of *efhc1*-deficient mouse[59,60] and fly models.[61] The most impressive experimental evidence for disease causality seeing the epileptic disorder manifest in a knockout animal model. **Supplementary Video S1** online shows an *efhc1*KO mouse (*efhc1*$^{-/-}$) having several massive myoclonias and a CTC convulsion.[60] The massive myoclonic seizures and CTC convulsions shown in the video can occur in homozygous *efhc1*$^{-/-}$ and heterozygous *efhc1*$^{+/-}$ mutant mice.

Heterozygous (*efhc1*$^{+/-}$) and null (*efhc1*$^{-/-}$) mutants exhibit more spontaneous positive myoclonias than wild-type mice and quick, high-amplitude polyspikes on their EMG.[59] Two measures of seizure susceptibility—the percentage of animals exhibiting generalized seizures within 600 s after treatment with pentylenetretrazole (PTZ) and latency to clonic seizures after PTZ treatment—are significantly increased in the same measures of wild-type mice, with the greatest significance reached in 9- to 12-month-old mice. Both *efhc1*$^{+/-}$ and *efhc1*$^{-/-}$ mutant mice show ependymal cilia in the lateral ventricles, with abnormally decreased movements at 3 months and slightly enlarged lateral ventricles and decreased hippocampal volume at 12 months.[59] In mice with massive myoclonias and grand mal CTC convulsions, cell death occurs in ependymal cells along periventricular zones (both lateral and fourth ventricles) and in striatal cells, while disorganization of cell layering in paraventricular nucleus of hypothalamus, thalamus, hippocampus, and neocortex is present.[60]

The notion of overgrown and overexcitable neurons and neurites was further examined *in vivo* in *Drosophila melanogaster*.[61] Knocking out the *Drosophila DEFHC1.1* gene, a homolog of human *EFHC1*, resulted in supernumerary synaptic boutons at the neuromuscular junction synapse and increased terminal branching of dendritic arborization, along with increased spontaneous neurotransmitter release. The notion of overgrown and overexcitable neurons and neurites was further solidified when *DEFHC1.1* overexpression rescued and markedly reduced dendrite branching and complexity.[61] These rescue experiments strongly recommended by NHGRI core guidelines argue convincingly that *EFHC1*'s main function is to restrain excessive synaptic and dendritic growth and arborization.

DISCUSSION

In research laboratories, the primary purpose of the search for disease-associated variants is to identify molecular disease mechanisms that can lead to a quest for a curative molecule.[1] During clinical laboratory testing, however, the primary purpose of searching for disease-causing sequence variants is to support medical decision making.[3] Here, in the reanalysis of *EFHC1* variants, we search for disease mechanisms that produce the most feared and most neurologically damaging seizure phenotype—the grand mal CTC convulsions of JME—while we weigh evidence for or against pathogenicity of a given *EFHC1* variant that can be used in medical decision making.

Applying NHGRI guidelines and ACMG classification to all 54 *EFHC1* variants in literature, we show that *EFHC1* is definitely implicated in JME. **Table 2** provides evidence used to classify all the variants as "pathogenic" and "likely pathogenic." **Supplementary Table S8a** online summarizes the ACMG criteria used for classifying all 54 *EFHC1* variants included in this study. Using ACMG combinatorial criteria, we classified 9 *EFHC1* variants as "pathogenic," 14 variants as "likely pathogenic," and 20 variants "of unknown significance" (**Table 3**). Eight *EFHC1* variants were benign and 3 were likely benign.

Of the "pathogenic" variants, the five original *EFHC1* variants discovered in Mexican families[16] met two "strong" criteria for pathogenicity; they were found to be statistically increased in JME cases in comparison to controls in at least one study and also demonstrated a significant difference in 4 to 10 experimental measures of neuron function and neurodevelopment.[55–57] trB:Q277X was found to be statistically increased in disease cases compared with controls, and it was found de novo in a singleton whose parentage was confirmed.[17] R276X, a nonsense variant that truncates the final six exons of the primary transcript, was associated with a JME patient and reported in the ClinVar database.[15] It was also found as a de novo mutation in a single case of epileptic encephalopathy (personal communication during presentation of a poster by S. Jamuar et al. during 2013 American Epilepsy Society Meeting.), thereby meeting the ACMG "very strong" criteria for pathogenicity (PVS1).

Within the scope of their originally reported individual studies, 20 *EFHC1* variants did not reach statistical significance during case–control association because these studies, having an insufficient number of racial/ethnic- and ancestry-matched controls, were statistically underpowered. However, when

compared with racial/ethnic-matched ExAC population controls, these same *EFHC1* variants have ORs >5 and reach statistical significance.

Fourteen variants classified as "likely pathogenic" would have been classified as "unknown" if only the results within their specific case–control study were used. Although most of these variants were absent in the study controls run, two variants—L9L and E357K—were found in the study ancestry-matched controls but were completely absent from Europeans in ExAC.[14] These variants may be examples of population-specific benign variations whose case association would have resulted in false positives if the ancestry-matched population controls were not included in the study and only the public genome databases were used as reference panels. These observations demonstrate the necessity of using both (i) study controls, matched for ancestry and country of residence, who are well screened to be free of epilepsy and febrile convulsions in their families and (ii) large public genome databases that have not been screened for epilepsy.

In the framework of the larger population groups used by ExAC,[14] we attempted to quantify the statistical effect size that pathogenic variants would have on JME patients. These estimates were calculated using the number of JME index cases in each population group who were screened for variants in all exons of *EFHC1* (see **Supplementary Table S2** online). Among Latinos, variants meeting the ACMG standard of "pathogenic" were identified as heterozygous mutations in 4.10% of individuals with JME, and "likely pathogenic" variants were found in another 3.59%. In Caucasians, both "pathogenic" and "likely pathogenic" variants were discovered in 2.59% (each) of individuals with JME. Currently, the variants discovered in India only meet the standard for "likely pathogenic" and account for 1.46% of all its screened JME patients.[28] However, we are still awaiting publication of experimental functional studies of the *EFHC1* variants discovered in JME patients in India. These studies will probably change their classification. Finally, the JME cohort from Brazil found variants meeting the "pathogenic" standard in only 2.94% of their JME cases (unpublished observations).

Further evidence for a large effect on the phenotype of JME is provided by a nonconsanguineous Moroccan-Jewish family in which three of their seven children afflicted with intractable epilepsy during infancy and who died at 18–36 months.[24] Whole-exome sequencing of the family revealed a homozygous mutation of F229L in two of the three affected children (the third child could not be tested). These children began experiencing seizures 6–12 h after birth and subsequently developed severe psychomotor retardation and microcephaly. Brain MRI of one of the children at 2 years of age exhibited decreased cerebellar volume, hypomyelination, and enlarged lateral and third ventricles, consistent with both our variant-specific experiments and gene-level experiments in knockout models of *EFHC1* (*efhc1*$^{-/-}$ KO mice and the loss-of-function fly model).

When a variant's allele frequency and its population prevalence are greater than the disease prevalence, ACMG recommendations consider this result "strong" criterion for the benign status for the variant. This criterion needs to be revisited by the AMG workgroup and rediscussed in the context of nonmonogenic disorders because of the uncertainty of the prevalence of specific diseases such as a specific epilepsy syndrome like JME. Although this criterion did not change the pathogenic classification of *EFHC1* variants in question according to the combinatorial rules of ACMG, 5 of the 7 "pathogenic" variants and 2 of the 16 "likely pathogenic" variants were found to have allele frequencies higher than the expected prevalence of the JME in their respective populations. There are three possible explanations for this observation:

1. The population prevalence studies are poorly matched to the populations in which the variants were found. Most epidemiology studies focused only on the prevalence of "active" epilepsies, a measure that is important for public health policy and estimation of economic impact. However, the lifetime prevalence, which would attempt to also capture individuals with a history of epilepsy prior to the date of ascertainment but who may be in remission or no longer seeking treatment, would be a better measure for comparison in genetic studies. Furthermore, studies of epilepsy prevalence frequently do not capture further information of seizure types or classifications of electroclinical syndromes. When they do, different diagnostic and ascertainment criteria make it difficult to compare results between studies.
2. The ExAC database, our primary reference panel for estimating minor allele frequencies, includes exomes of several disease populations. Notably, these three studies were targeted toward the identification of neurological diseases such as schizophrenia and Tourette syndrome. Although *EFHC1* has not been specifically implicated in either of these conditions, other genes implicated in JME and childhood absence epilepsy show overlap with those implicated in schizophrenia, specifically *GABRA1*, *GABRB3*, *GABRG2* (refs. 70,71), and *CHRNA7* (ref. 72), indicating that the ExAC database may be enriched with these alleles.
3. Like variants associated with other diseases that have a complex genetic architecture, some *EFHC1* variants may not be sufficient by themselves to cause epilepsy; however, they may have an additive effect toward the pathogenesis of JME disease in conjunction with other alleles associated with epilepsy. This would also explain the discovery of JME disease alleles in screened study controls[20,21] as well as in the case of multiple disease alleles in linkage disequilibrium and the autosomal dominant transmission with incomplete penetrance that we see in our large families.

In the case of a fully penetrant monogenic disease with a large statistical effect size (the disease model used when creating both the NHGRI and ACMG standards), the comparison

of allele frequencies to population prevalence may provide an appropriate guideline for classifying diseases. However, genetic cases of JME exhibit a high degree of phenotypic heterogeneity, and even within the affected families only 50.94% of all heterozygous carriers actually develop clinical epilepsy. Of all the implicated JME genes, *EFHC1* has been replicated by more studies than any other, but "pathogenic" and "likely pathogenic" variants still only account for approximately 3% of JME cases in Caucasians and 8% of those in Hispanics. JME does not perfectly fit the disease model for which these standards were created. However, even with this limitation, the ACMG and NHGRI guidelines enabled us to classify the purported disease-causing variants in *EFHC1*.

In conclusion, we found the NHGRI and ACMG guidelines to be useful in quantifying the amounts and types of evidence that implicate sequence variants of *EFHC1* as disease-causing in JME. Vetting *EFHC1* variants through NHGRI guidelines[1] definitely implicates these *EFHC1* variants in JME. Using ACMG recommendations, scoring rules, and combinatorial criteria to choose a classification from the five-tier system,[3] our reanalysis showed that 9 *EFHC1* variants are "pathogenic," 14 are "likely pathogenic," and 20 are "variants of unknown significance." (See **Supplementary Table S8b** online for criteria and classification of all variants.) NHGRI gene-level evidence and variant-level evidence establish *EFHC1* as the first non–ion channel microtubule–associated protein[53] whose mutations disturb R-type VDCC[16] and TRPM2 calcium currents[57] in overgrown synapses and dendrites[61] within abnormally migrated dislocated neurons,[56] thus explaining myoclonic and grand mal CTC convulsions and "microdysgenesis"[73,74] neuropathology of JME.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/gim

ACKNOWLEDGMENTS

This project was supported by the National Institutes of Health (1R01NS055057), a VACO Merit Review Grant (5IO1CS000743 to A.V.D-.E.), and a CONACYT grant (to A.J-.P.). Some genotype services were provided by the Center for Inherited Disease Research (CIDR) (1X01HG007145 to A.V.D.-E.). CIDR is fully funded through a federal contract from the National Institutes of Health to Johns Hopkins University (contract HHSN2682012000081).

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature* 2014;508:469–476.
2. Gargis AS, Kalman L, Berry MW, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol* 2012;30:1033–1036.
3. Richards S, Aziz N, Bale S, et al.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–424.
4. Bell CJ, Dinwiddie DL, Miller NA, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci Transl Med* 2011;3:65ra4.
5. Xue Y, Chen Y, Ayub Q, et al.; 1000 Genomes Project Consortium. Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet* 2012;91:1022–1032.
6. Norton N, Robertson PD, Rieder MJ, et al.; National Heart, Lung and Blood Institute GO Exome Sequencing Project. Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. *Circ Cardiovasc Genet* 2012;5:167–174.
7. Weng L, Kavaslar N, Ustaszewska A, et al. Lack of MEF2A mutations in coronary artery disease. *J Clin Invest* 2005;115:1016–1020.
8. Hunt KA, Smyth DJ, Balschun T, et al.; Type 1 Diabetes Genetics Consortium; UK Inflammatory Bowel Disease (IBD) Genetics Consortium; Wellcome Trust Case Control Consortium. Rare and functional SIAE variants are not associated with autoimmune disease risk in up to 66,924 individuals of European ancestry. *Nat Genet* 2012;44:3–5.
9. Epi4K and EPGP Investigators. *De novo* mutations in epileptic encephalopathies. *Nature* 2013;501:217–221.
10. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–753.
11. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013;369:1502–1511.
12. Abecasis GR, Altshuler D, et al. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010;467:1061–1073.
13. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA. http://evs.gs.washington.edu/EVS/. Accessed 2 April 2014.
14. Exome Aggregation Consortium (ExAC), Cambridge, MA. http;//exac.broadinstitute.org. Accessed date 29 November 2015.
15. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016;44(D1):D862–D868.
16. Suzuki T, Delgado-Escueta AV, Aguan K, et al. Mutations in EFHC1 cause juvenile myoclonic epilepsy. *Nat Genet* 2004;36:842–849.
17. Medina MT, Suzuki T, Alonso ME, et al. Novel mutations in Myoclonin1/EFHC1 in sporadic and familial juvenile myoclonic epilepsy. *Neurology* 2008;70(22 Pt 2):2137–2144.
18. Jara-Prado A, Martínez-Juárez IE, Ochoa A, et al. Novel Myoclonin1/EFHC1 mutations in Mexican patients with juvenile myoclonic epilepsy. *Seizure* 2012;21:550–554.
19. Annesi F, Gambardella A, Michelucci R, et al. Mutational analysis of EFHC1 gene in Italian families with juvenile myoclonic epilepsy. *Epilepsia* 2007;48:1686–1690.
20. Stogmann E, Lichtner P, Baumgartner C, et al. Idiopathic generalized epilepsy phenotypes associated with different EFHC1 mutations. *Neurology* 2006;67:2029–2031.
21. Subaran RL, Conte JM, Stewart WC, Greenberg DA. Pathogenic EFHC1 mutations are tolerated in healthy individuals dependent on reported ancestry. *Epilepsia* 2015;56:188–194.
22. von Podewils F, Kowoll V, Schroeder W, et al. Predictive value of EFHC1 variants for the long-term seizure outcome in juvenile myoclonic epilepsy. *Epilepsy Behav* 2015;44:61–66.
23. Ma S, Blair MA, Abou-Khalil B, Lagrange AH, Gurnett CA, Hedera P. Mutations in the GABRA1 and EFHC1 genes are rare in familial juvenile myoclonic epilepsy. *Epilepsy Res* 2006;71:129–134.
24. Berger I, Dor T, Halvardson J, et al. Intractable epilepsy of infancy due to homozygous mutation in the EFHC1 gene. *Epilepsia* 2012;53:1436–1440.
25. Coll M, Allegue C, Partemi S, et al. Genetic investigation of sudden unexpected death in epilepsy cohort by panel target resequencing. *Int J Legal Med* 2016;130:331–339.
26. Bai D, Bailey JN, Durón RM, et al. DNA variants in coding region of EFHC1: SNPs do not associate with juvenile myoclonic epilepsy. *Epilepsia* 2009;50:1184–1190.
27. Pinto D, Louwaars S, Westland B, et al. Heterogeneity at the JME 6p11-12 locus: absence of mutations in the EFHC1 gene in linked Dutch families. *Epilepsia* 2006;47:1743–1746.
28. Raju Pedabaliyarasimhuni PK. Connecting the paralogs: contribution of EFHC1 and EFHC2 in juvenile myoclonic epilepsy. Thesis, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India, 2014.
29. Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 2003;73:652–655.

30. Galili T, Calhoun P. *Barnard's exact test—a powerful alternative for Fisher's exact test (implemented in R). R Statistics Blog, 2 February 2010*. Accessed 28 October 2015. http://www.r-statistics.com/2010/02/barnards-exact-test-a-powerful-alternative-for-fishers-exact-test-implemented-in-r/.
31. Mehrotra DV, Chan IS, Berger RL. A cautionary note on exact unconditional inference for a difference between two independent binomial proportions. *Biometrics* 2003;59:441–450.
32. Nicoletti A, Reggio A, Bartoloni A, et al. Prevalence of epilepsy in rural Bolivia: a door-to-door survey. *Neurology* 1999;53:2064–2069.
33. Syvertsen M, Nakken KO, Edland A, Hansen G, Hellum MK, Koht J. Prevalence and etiology of epilepsy in a Norwegian county-A population based study. *Epilepsia* 2015;56:699–706.
34. Fong GC, Mak W, Cheng TS, Chan KH, Fong JK, Ho SL. A prevalence study of epilepsy in Hong Kong. *Hong Kong Med J* 2003;9:252–257.
35. Bhalla D, Chea K, Hun C, et al. Epilepsy in Cambodia-treatment aspects and policy implications: a population-based representative survey. *PLoS One* 2013;8:e74817.
36. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res* 2010;20:110–121.
37. Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A; NISC Comparative Sequencing Program. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* 2005;15:901–913.
38. Lindblad-Toh K, Garber M, Zuk O, et al. A high-resolution map of human evolutionary constraint using 20 mammals. *Nature* 2011;478:476–482.
39. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics* 2009;25:i54–i62.
40. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* 2010;6:e1001025.
41. Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005;15:1034–1050.
42. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003;31:3812–3814.
43. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–1081.
44. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–249.
45. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* 2013;Chapter 7:Unit7.20.
46. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. *Genome Res* 2009;19:1553–1561.
47. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 2010;7:575–576.
48. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* 2011;39:e118.
49. Shihab HA, Gough J, Cooper DN, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat* 2013;34:57–65.
50. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–315.
51. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat* 2011;32:894–899.
52. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human nonsynonymous SNVs and their functional predictions and annotations. *Human Mutat* 2013;34:E2393–402.
53. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
54. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009;37:e67.
55. de Nijs L, Léon C, Nguyen L, et al. EFHC1 interacts with microtubules to regulate cell division and cortical development. *Nat Neurosci* 2009;12:1266–1274.
56. de Nijs L, Wolkoff N, Coumans B, Delgado-Escueta AV, Grisar T, Lakaye B. Mutations of EFHC1, linked to juvenile myoclonic epilepsy, disrupt radial and tangential migrations during brain development. *Hum Mol Genet* 2012;21:5106–5117.
57. Katano M, Numata T, Aguan K, et al. The juvenile myoclonic epilepsy-related protein EFHC1 interacts with the redox-sensitive TRPM2 channel linked to cell death. *Cell Calcium* 2012;51:179–185.
58. Sahni N, Yi S, Taipale M, et al. Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* 2015;161:647–660.
59. Suzuki T, Miyamoto H, Nakahari T, et al. EFHC1 deficiency causes spontaneous myoclonus and increased seizure susceptibility. *Hum Mol Genet* 2009;18:1099–1109.
60. Machado-Salas J, Tanaka M, Avila Costa M, et al. Neuronal migration arrest in a juvenile myoclonic epilepsy KO-mice [abstr]. *Epilepsy Curr* 2012: 472. http://dx.doi.org/10.5698/1535-7511-13.s1.1.
61. Rossetto MG, Zanarella E, Orso G, et al. DEFHC1.1, a homologue of the juvenile myoclonic gene EFHC1, modulates architecture and basal activity of the neuromuscular junction in Drosophila. *Hum Mol Genet* 2011;20:4248–4257.
62. Yang G, Smibert CA, Kaplan DR, Miller FD. An eIF4E1/4E-T complex determines the genesis of neurons from precursors by translationally repressing a proneurogenic transcription program. *Neuron* 2014;84:723–739.
63. Ahn HJ, Hernandez CM, Levenson JM, Lubin FD, Liou HC, Sweatt JD. c-Rel, an NF-kappaB family transcription factor, is required for hippocampal long-term synaptic plasticity and memory formation. *Learn Mem* 2008;15:539–549.
64. Flora A, Garcia JJ, Thaller C, Zoghbi HY. The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc Natl Acad Sci USA* 2007;104:15382–15387.
65. Shinoda S, Skradski SL, Araki T, et al. Formation of a tumour necrosis factor receptor 1 molecular scaffolding complex and activation of apoptosis signal-regulating kinase 1 during seizure-induced neuronal death. *Eur J Neurosci* 2003;17:2065–2076.
66. Shamseldin HE, Bennett AH, Alfadhel M, Gupta V, Alkuraya FS. GOLGA2, encoding a master regulator of golgi apparatus, is mutated in a patient with a neuromuscular disorder. *Hum Genet* 2016;135:245–251.
67. Yamashita D, Sano Y, Adachi Y, et al. hDREF regulates cell proliferation and expression of ribosomal protein genes. *Mol Cell Biol* 2007;27:2003–2013.
68. Xu J, Lai YJ, Lin WC, Lin FT. TRIP6 enhances lysophosphatidic acid-induced cell migration by interacting with the lysophosphatidic acid 2 receptor. *J Biol Chem* 2004;279:10459–10468.
69. Salles A, Romano A, Freudenthal R. Synaptic NF-kappa B pathway in neuronal plasticity and memory. *J Physiol Paris* 2014;108:256–262.
70. Sun J, Kuo PH, Riley BP, Kendler KS, Zhao Z. Candidate genes for schizophrenia: a survey of association studies and gene ranking. *Am J Med Genet B Neuropsychiatr Genet* 2008;147B:1173–1181.
71. Ross CA, Margolis RL, Reading SA, Pletnikov M, Coyle JT. Neurobiology of schizophrenia. *Neuron* 2006;52:139–153.
72. Allen NC, Bagade S, McQueen MB, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008;40:827–834.
73. Meencke HJ, Janz D. Neuropathological findings in primary generalized epilepsy: a study of eight cases. *Epilepsia* 1984;25:8–21.
74. Meencke HJ, Janz D. The significance of microdysgenesia in primary generalized epilepsy: an answer to the considerations of Lyon and Gastaut. *Epilepsia* 1985;26:368–371.