Klassen et al. 2011: Supplemental Information (SI)

Supplementary Figure Legends

Supplementary Figure 1: Scattergrams of SNP totals Vs target gene length relating to Table 1 and Figure 1.
A. Histogram of all candidate genes showing the total number of validated discovered SNPs in the gene plotted against the size of the gene.
B. Histogram of all validated nsSNPs in the candidate gene plotted against the size of the gene. In both A and B the largest genes (RYR1-RYR3) contained the most variants, however the slightly smaller KCNMA1 gene had relatively low SNP count thus the number of total SNPs and the number of nsSNPs in the gene is not readily predictable by gene size alone.

Supplementary Figure 2: Histograms of SNP Blosum80 scores as identified in Supplementary Table X and used in permutation analysis described in Supplementary Experimental Procedures.
A. Histogram of all coding SNPs in each cohort (sSNP; nsSNP) arranged by Blosum80 score. Coding SNPs in both affected and control groups are found across the range of Blosum80 scores. **B.** Enlargement of the most unfavorable amino acid substitutions (Blosum80 -6 to 0) from A. The proportion of nsSNP severity is similar across both groups.

Supplemental Table 3: Summary of all[1] single nucleotide polymorphisms discovered using Sanger sequencing on 237 target ion channel genes in 139 controls and 151 individuals with IE, only those SNPs validated through secondary means are enumerated in Table 1.

Type/Location of SNP	*Number of SNPs*	*Percent of Total Dataset*	*Number of Novel SNPs*
Promoter[2]	458	4.3	333
5' UTR	382	3.4	257
3' UTR	2752	24.8	1858
Synonymous (sSNP)	1947	17.5	993
Nonsynonymous (nsSNP)	1673	15.1	1246
Nonsense /Stop Codon	30	<1	30
Splice Site SNP[3]	39	<1	36
Splice Region SNP[4]	358	3.2	246
Intron SNP	3325	29.9	2213
Undefined[5]	138	1.2	90
TOTALS	**11102**	**100.0**	**7303**

1. All discovered SNPs include those that were validated and used in the main study, as well as unvalidated SNPs detected using the automated detection and annotation program SNPdetector V. 3. This discovery set has been previously reported online at http://www.hgsc.bcm.tmc.edu/ionchannel-snpList.xsp
2. SNPs in promoter regions are reported as such if the promoter for the gene is known and defined in our gene models.
3. splice site (+2 to -2 bp from defined exon boundary at/near splice junction).
4. splice region (-2 to -15 bp from defined exon boundary, located in the intron, from splice junction).

Supplementary Experimental Procedures

Study population structure – We evaluated all self reported white Caucasian and white Hispanic individuals regardless of the age or gender presenting to the Baylor College of Medicine affiliated hospitals in accordance with the approved study protocol and accepted policies and guidelines of the Institutional Review Board at Baylor College of Medicine. Race and ethnicity were established according to Risch et al. (Tang et al., 2005) based on a structured questionnaire that examined the racial and ethnic origin of an individual back three generations on either side of a family tree. Demographic characteristics of the cohort are detailed in Table 1 below. DNA from individuals recruited into the study (both cases and controls) was sent to the Coriell Institute for Medical Research for archiving in their cell line repository and is accessible upon request (http://ccr.coriell.org/Sections/Collections/NINDS/Epilepsy). Please note that the samples we have provided Coriell are not differentiated from the other "epilepsy" samples in the repository.

Cases:

Patients presenting to the Baylor College of Medicine affiliated hospitals who met the accepted criteria for either idiopathic or cryptogenic epilepsy (1989) were recruited in accordance with the approved study protocol . In the proposed new classification the terms "idiopathic" and "cryptogenic" have been revisited in describing the underlying cause (etiology) of the disorder (Berg et al., 2010). The recommendation is for the term "genetic epilepsy" is used in the place of "idiopathic epilepsy" and is reserved for seizures occurring in epilepsy patients with presumed genetic origin, such as benign rolandic epilepsy and childhood absence epilepsy. Similarly, the usage of "epilepsy of unknown cause" is recommended to replace "cryptogenic epilepsy" because the underlying cause is yet unknown. This proposed classification system allows that cryptogenic epilepsy may result from a genetic defect or some yet unidentified mechanism.

Clinical assessment: Study subjects were evaluated for epilepsy by board- certified neurologists following a standard protocol for the diagnostic evaluation of their epilepsy (Mizrahi et al., 1990; York et al., 2003). The neurological assessment was classified as abnormal if deficits such as hemiparesis, ataxia or cognitive impairment were present. The presence of clumsiness or other "soft" neurological signs was insufficient to classify a patient as having neurological

abnormalities. A structured questionnaire was used to collect information from the patient and any witnesses to the seizures to complement data recorded in the clinic chart.

Electroencephalographic (EEG) evaluations: All study patients underwent at least one of the following; (a) routine sleep deprived EEG, (b) prolonged ambulatory (A-EEG), or (c) video EEG (V-EEG) monitoring. EEG studies were reviewed and interpreted according to ILAE standards by neurologists board certified in clinical neurophysiology at the Baylor Comprehensive Epilepsy Center.

Imaging studies: Brain magnetic resonance imaging (MRI) was performed in all seizure subjects in accordance with the standard protocol at our center. Visual analysis of oblique and coronal T1, T2, and FLAIR images was done at the Baylor College of Medicine affiliated hospitals. Epileptic seizures and syndromes were classified using information obtained from the history, physical evaluation, and above specified ancillary diagnostic investigations according to the ILAE guidelines that were accepted clinical practice at the time of the design and execution of this study (1989).

Controls:

Control subjects were recruited in accordance with the approved study protocol among unrelated visitors, spouses, significant others, and friends of patients presenting to the Baylor College of Medicine affiliated hospitals. The clinical assessment of controls was limited to answers to a structured general medical questionnaire with specific emphasis on any neurological or cardiac symptoms, including recurrent spells or seizure like episodes. Only individuals with an entirely negative past and present history were included. Individuals with a family history positive for any neurological disorder outside of cerebrovascular disease and undocumented migraines were also excluded.

Study Cohort

	CASES			**CONTROLS**		
	Males	*Females*	***All subjects***	*Males*	*Females*	***All subjects***
White Caucasian	49	71	**120**	60	75	**135**
White Hispanic	17	14	**31**	2	2	**4**
Age range (years)	1-75	1-65	**1-75**	10-58	13-59	**10-59**
Average age (years)	29.64	25.55	**27.32**	45.48	45.89	**45.71**

Ion Channel Gene Targets and Gene Models

We selected all known ion channelopathy genes, their family members and functionally similar ion channels (Harmar et al., 2009). This list includes members of both the voltage-gated, ligand-gated and background "leak" channels, and a small number of ion channel interacting proteins like ANK2 (ankyrin-B). Ion channel gene models were defined using the annotations available through the UCSC Genome Browser for the hg17 assembly from March 2004. Exon boundaries were defined using Refseq coordinates where available. All reported splice isoforms were included for sequencing. Protein sequences from *in silico* translation of each transcript model were aligned to the reported protein sequence in PubMed. Manual curation of a handful of gene models was required to reconcile discrepancies. Gene models are available on request.

Amplicon Definition and Primer Design

Amplicons selected for PCR amplification and downstream sequencing were single exons smaller than 600 bp in size. For exons larger than 500 bp, multiple amplicons were subdivided to allow for optimal product lengths but still span the entire exon with overlapping ends. When two or more exons with the spanning intron were collectively less than 500 bp in size, the total region was defined as a single amplicon for primer design. This consolidation reduced the number of targets by 10%. For ion channel gene targets where alternative isoforms exist, all undefined exons (alternatively spliced exons) were targeted as amplicons and reference cDNA sequences from NCBI for the alternative isoform were used for annotation. Primers were designed as 20mers and were placed 50 bp from either end of an amplicon to ensure accurate reads at the exon/intron boundaries. Primers were generated automatically using Primer3, and primer sequences are available on request.

Extraction and Archiving of Genomic DNA

After counselling and informed consent, 2 vials of 8 mls of blood were drawn from each individual. Genomic DNA was extracted using the Gentra Puregene Blood Kit (Qiagen) and stored in de-identified bar coded tubes prior to use in amplicon generation. Samples of high quality, high molecular weight genomic DNA were also sent to the Coriell Sample Repository for cell line archiving and are openly accessible (http://www.coriell.org/ ; Home > Collections > NINDS > Epilepsy).

Highthrough-put Parallel Sanger Sequencing

A high-throughput, parallel Sanger sequencing pipeline was designed to move multiple samples through the amplification, sequencing and subsequent SNP detection at one time. All PCR reactions were performed in a 25 ul reaction volume containing 20-40 ng of individual genomic DNA, 10 pmol of each primer and 12.5 ul of Qiagen muliplex PCR master mix. General PCR conditions were initial denaturation and activation step 15 minutes at 95C, followed by 30 cycles of the following; denaturation 94C for 30 s, annealing 50C-65C (depending on target) for 90 s followed by a 1 min extension step at 72C. A final 10 minute completion step at 72C followed amplification. Sequencing was performed using Big Dye Terminator Cycle Sequencing v3.1 and visualized using an ABI3700 (PE Applied Biosystems). Sequencing chromatograms were compared to reference gene models and single nucleotide polymorphisms were detected and annotated using SNPdetector v3.0 (Zhang et al., 2005)and individual channotypes were output for analysis. Validation of SNPs included in this study was performed by visual confirmation of the chromatogram data, the presence of the SNP in dbSNP, the generation of a custom MIP chip, and a combination of Biotage and/or 454 sequencing.

Permutation Bioinformatic Analysis

To asseses the enrichment and predictive power of rare variants within our study population we performed a permutation-based rare variant enrichment test for case-control studies that accounts for amplicon-specific group oversampling. After filtering for the desired SNPs and corresponding genotypes, we observed the number of minor alleles in each group. Any observed enrichment of rare variants in either group was compared against the null distribution generated by 10,000 permutations (random shuffling) of the group labels (case vs control) to determine the false discovery rate.

Multiple Alignments for Schematic Models

For every nucleotide reference sequence (NM_number) used in this study, we obtained the corresponding protein sequences (NP_number) in NCBI. Protein alignments for all members of a gene family that were structurally related (eg. SCN1A-SCN11A or GABRG1-GABRG3) were performed using the MUSCLE webserver (Edgar, 2004a, b). For the human epilepsy (hEP) genes, structural models with reference comparative alignments or annotated amino acid sequence were obtained from the literature. These models were used to position individual nonsense or missense mutations on schematic representations of channel structure.

Supplementary References

(1989). Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia *30*, 389-399.
Berg, A., Berkovic, S., Brodie, M., Buchhalter, J., Cross, J., van Emde Boas, W., Engel, J., French, J., Glauser, T., Mathern, G.*, et al.* (2010). Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. Epilepsia *51*, 676-685.

Edgar, R.C. (2004a). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Edgar, R.C. (2004b). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797.

Harmar, A., Hills, R., Rosser, E., Jones, M., Buneman, O., Dunbar, D., Greenhill, S., Hale, V., Sharman, J., Bonner, T.*, et al.* (2009). IUPHAR-DB: the IUPHAR database of G protein-coupled receptors and ion channels. Nucleic Acids Res *37*, D680-685.

Mizrahi, E.M., Kellaway, P., Grossman, R.G., Rutecki, P.A., Armstrong, D., Rettig, G., and Loewen, S. (1990). Anterior temporal lobectomy and medically refractory temporal lobe epilepsy of childhood. Epilepsia *31*, 302-312.

Tang, H., Quertermous, T., Rodriguez, B., Kardia, S.L., Zhu, X., Brown, A., Pankow, J.S., Province, M.A., Hunt, S.C., Boerwinkle, E.*, et al.* (2005). Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. Am J Hum Genet *76*, 268-275.

York, M.K., Rettig, G.M., Grossman, R.G., Hamilton, W.J., Armstrong, D.D., Levin, H.S., and Mizrahi, E.M. (2003). Seizure control and cognitive outcome after temporal lobectomy: a comparison of classic Ammon's horn sclerosis, atypical mesial temporal sclerosis, and tumoral pathologies. Epilepsia *44*, 387-398.

Zhang, J., Wheeler, D., Yakub, I., Wei, S., Sood, R., Rowe, W., Liu, P., Gibbs, R., and Buetow, K. (2005). SNPdetector: a software tool for sensitive and accurate SNP detection. PLoS Comput Biol *1*, e53.

Supplemental Figure

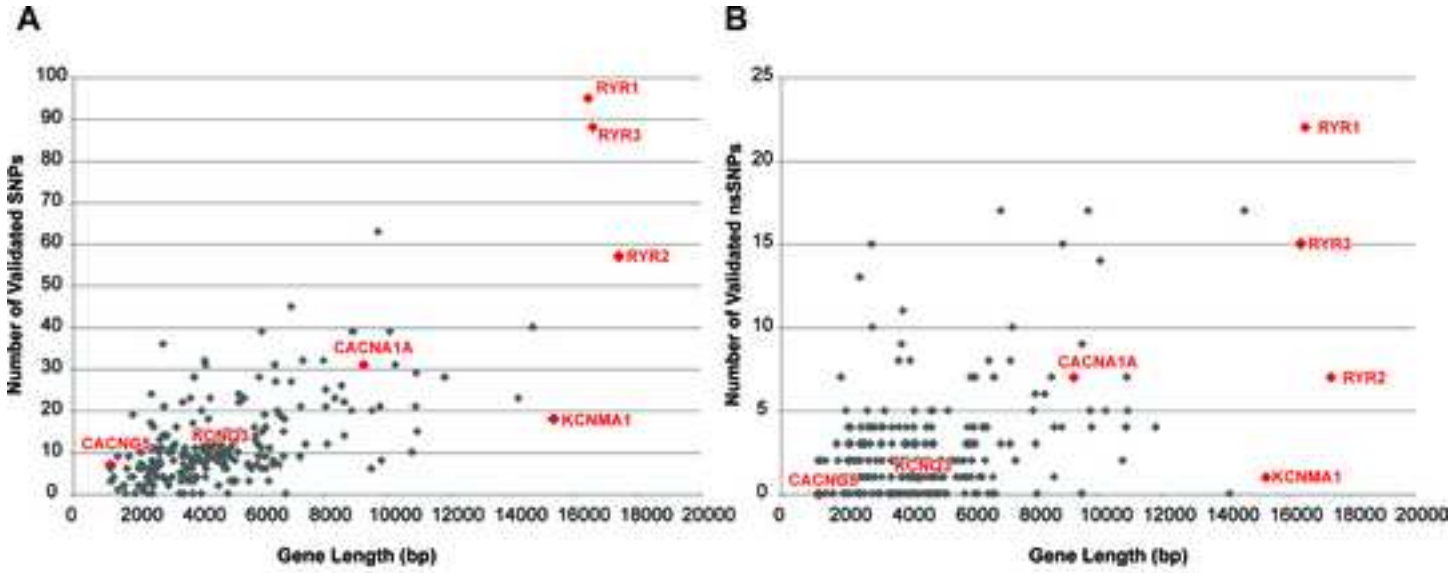

Figure S1

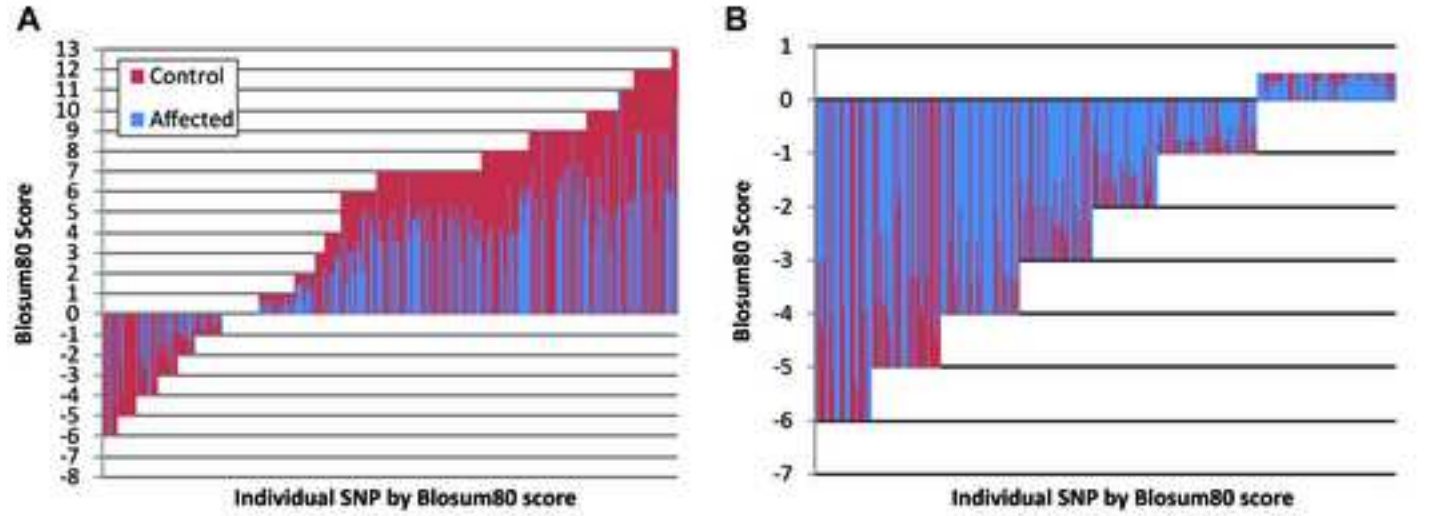

Figure S2