De Novo Mutations in Synaptic Transmission Genes Including DNM1 Cause Epileptic Encephalopathies

EuroEPINOMICS-RES Consortium,* Epilepsy Phenome/Genome Project, and Epi4K Consortium

Emerging evidence indicates that epileptic encephalopathies are genetically highly heterogeneous, underscoring the need for large cohorts of well-characterized individuals to further define the genetic landscape. Through a collaboration between two consortia (EuroEPINOMICS and Epi4K/EPGP), we analyzed exome-sequencing data of 356 trios with the “classical” epileptic encephalopathies, infantile spasms and Lennox Gastaut syndrome, including 264 trios previously analyzed by the Epi4K/EPGP consortium. In this expanded cohort, we find 429 de novo mutations, including de novo mutations in DNM1 in five individuals and de novo mutations in GABBR2, FASN, and RYR3 in two individuals each. Unlike previous studies, this cohort is sufficiently large to show a significant excess of de novo mutations in epileptic encephalopathy probands compared to the general population using a likelihood analysis (p = 8.2 × 10^-5), supporting a prominent role for de novo mutations in epileptic encephalopathies. We bring statistical evidence that mutations in DNM1 cause epileptic encephalopathy, find suggestive evidence for a role of three additional genes, and show that at least 12% of analyzed individuals have an identifiable causal de novo mutation. Strikingly, 75% of mutations in these probands are predicted to disrupt a protein involved in regulating synaptic transmission, and there is a significant enrichment of de novo mutations in genes in this pathway in the entire cohort as well. These findings emphasize an important role for synaptic dysregulation in epileptic encephalopathies, above and beyond that caused by ion channel dysfunction.

Introduction

Epileptic encephalopathies (MIM 308350) comprise a range of severe epilepsy syndromes determined by age of onset, seizure types, developmental and clinical course, and electroencephalographic findings.1 Infantile spasms are one of the most common seizure types seen in epileptic encephalopathies, and in combination with hypsarrhythmia on EEG and developmental regression define West syndrome. In some individuals, West syndrome evolves to Lennox-Gastaut syndrome, characterized by multiple seizure types including atonic and tonic seizures and generalized slow-spike-wave activity on EEG. Both entities are considered prototypic epileptic encephalopathies.

An underlying genetic cause has been shown in a growing proportion of persons with epileptic encephalopathies. Many causal mutations arise de novo.2,7 Our earlier exome-sequencing study of 264 trios identified de novo mutations in seven genes reported at the time in the Mendelian Inheritance in Man (OMIM) database as linked to epileptic encephalopathy (CDKL5 [MIM 300203], KCNQ2 [MIM 602235], KCNT1 [MIM 608167], SCN1A [MIM 182389], SCN2A [MIM 182390], SCN8A [MIM 600702], and STXBP1 [MIM 602926]) and provided clear evidence of pathogenicity for de novo mutations in two not previously implicated genes, ALG13 (MIM 300776) and GABBR3 (MIM 137192).2 Although we found multiple additional candidate genes affected by de novo mutations, statistical power was insufficient to implicate a pathogenic role for these genes.7 By combining samples from two international consortia, we now analyze a joint cohort of 356 proband-parent trios with infantile spasms or Lennox-Gastaut syndrome to search for additional causal de novo mutations.

Subjects and Methods

Study Subjects and Procedure of Exome Sequencing

Three epileptic encephalopathy cohorts were evaluated in this study: (1) Epilepsy Phenome/Genome Project (EPGP) cohort 1 (n = 264 trios) used in the previously published paper,2 (2) new EPGP cohort 2 (n = 73 trios), and (3) EuroEPINOMICS-RES cohort (n = 19 trios). The study was approved by the local ethics committee of each participating center. Parents or the legal guardian of each proband signed an informed consent form for participation. The EPGP cohort inclusion criteria have been reported previously.2 Probands recruited through the EuroEPINOMICS-RES consortium for the project on nonlesional epileptic spasms (NLES) were enrolled by clinical sites in 18 different European countries. Genomic DNA of the individuals was extracted from peripheral blood according to standard procedures. All phenotypic data were entered in the web-based platform Cartagenia-BENCH, which was adapted for this study. Inclusion criteria included documented infantile spasms with onset in the first 2 years of life, normal routine metabolic screening, and exclusion of causal structural abnormalities on MRI of the brain. Most probands had undergone testing for known genetic causes of infantile spasms, but this was not required for study inclusion. Individuals with a family history of epilepsy in first-degree relatives were excluded. Detailed exome-sequencing and data analysis methods are provided in the legend of Table S1 available online.

Trio-Specific Callable Real-Estate

Since the ability to call a de novo mutation requires that all three individuals comprising the trio are well sequenced, we calculated the percent of the exome and individual genes, as defined by the consensus coding sequence (CCDS release 9, GRCh37.p5) and the 2 bp flanking splice sites, where each base was sequenced at least 10-fold with a multisampling (GATK) raw phred-scaled confidence score of ≥ 20 in the presence or absence of a variant. These

*Correspondence: euroepinomics-RES@ua.ac.be
http://dx.doi.org/10.1016/j.ajhg.2014.08.013. ©2014 by The American Society of Human Genetics. All rights reserved.
estimates for callable real-estate were incorporated into the architecture and gene-enrichment analyses.

Genetic Architecture Likelihood Analyses

We conducted a likelihood analysis of parameters describing the genetic architecture of epileptic encephalopathies, including the relative risk ($\gamma$) and proportion of the exome related to epileptic encephalopathies ($\eta$). Because there were substantial differences in sequenced regions across the different exome-sequencing methods used in this study, we adapted our previously proposed likelihood model to incorporate the sequence-specific mutation rate for the individual proband's callable real-estate. This likelihood can be written as

\[
L(\gamma, \eta) = \prod \left\{ \frac{\lambda_i^y e^{-\lambda_i}}{y!} \left[ \gamma + (1 - \gamma)(1 - c_i) \right]^y \gamma + (1 - \gamma)e^{-\lambda_i} \right\}
\]


\[
(1)
\]

where $x_i$ and $\lambda_i$ are the de novo counts and mutation rate, respectively, for the $i$th trio. In previous analyses, $C$ is assumed to be a known constant taking the value of 0.25. Point estimates were obtained by optimizing Equation 1 and likelihood ratio tests were computed by computing the log-likelihood at this optimum to the value obtained under the null hypothesis. Marginal confidence intervals for $\eta$ and $\gamma$ were derived based on a profile likelihood calculated from the equation above.

“Hot-Zone” Analysis

To further assess the likelihood of pathogenicity of mutations, we considered two scores for each de novo single-nucleotide variant (SNV): the gene-level residual variation intolerance score (RVIS) assesses the intolerance of genes to functional genetic variation and the variant-level score assesses the probability that a given variant damages protein function. For the variant-level score, we use PolyPhen-2 (HumVar) to score missense variants. Synonymous and loss-of-function (nonsense and splice acceptor/donor) mutations were scored 0 and 1, respectively. For comparison, we also assessed de novo mutations in previously published control trios that reported at least one protein-coding de novo mutation (n = 411). The de novo mutations previously reported in controls were subjected to the same filtering criteria we used in this study, including presence in the CCDS protein-coding regions or in the 2 bp flanking splice donor and acceptor sites and absence from control cohort and ESP6500SI database. We also reassessed variant function using the same annotation pipeline used for the epileptic encephalopathy probands. Overlapping control samples, which were reported across multiple studies, were only considered once.

For each case and control sample that reported multiple de novo mutations, we considered only the single most damaging de novo mutation based on the shortest Euclidian distance from the most damaging coordinate [1,0] in the 2D space that is constructed based on the variant-level vector (PolyPhen-2 score) along the x axis and the gene-level vector (RVIS percentile score) along the y axis (Figure 1).

A two-tailed Fisher’s exact was used to test whether the single most damaging de novo mutation found in cases preferentially lay in the “hot zone,” defined by a PolyPhen-2 score of $\geq$0.95 and RVIS $\leq$25th percentile, in comparison to the single most damaging de novo mutation found in previously published control trios.

Calculation of the Gene-Specific Mutation Rate

The probabilities of getting greater than or equal de novo point mutations by chance when considering the observed numbers of de novo point mutations were calculated using procedures similar to the one that we introduced previously, with one minor modification: for the X chromosome, instead of calculating the average mutation rate and multiplying it by the number of gene copies, the mutation rate was calculated as the sum of the diploid rate in each female trio and the haploid rate in each male trio (Table 1).

Determination of “Solved” Epileptic Encephalopathy Trios

To determine what proportion of the 356 epileptic encephalopathy probands are currently explained by an identified de novo mutation, we first generated a list of genes in which mutations with high confidence cause an epileptic encephalopathy (further referred to as “epileptic encephalopathy genes”). We searched the OMIM database (January 2014) as well as several recent publications to obtain a comprehensive list of putative epileptic encephalopathy genes. We considered five categories of genes: (1) OMIM “epileptic encephalopathy (EIEE),” n = 18; (2) OMIM “Dravet Syndrome,” n = 1; (3) OMIM Neurological Clinical Synopsis containing “epileptic encephalopathy” in disorders with a “known molecular basis,” n = 8; (4) genes implicated in OMIM due to “epileptic encephalopathy” in the gene summaries, n = 19; and (5) genes implicated in epileptic encephalopathy in the recent literature, n = 6 (ALG13, DNMT1 [MIM 602377], GABRA1 [MIM 137160], GABRB3 [MIM 138252], and SLC35A2 [MIM 314375]). For the first two categories, if the mode of inheritance was recessive or if the first reported allelic variant in OMIM was published before 2010, we considered these definitive EE genes (n = 15; ARHGEF9 [MIM 300429], ARX [MIM 300382], CDKL5, KCNQ2 [MIM 602235], CDK6, PLCB1 [MIM 607120], PNKP [MIM 605610], PNPO [MIM 603287], SCN1A, SCN2A, SLC25A22 [MIM 609302], ST3GAL3 [MIM 606494], STXBP1, SST2 [MIM 615463], and TBC1D24 [MIM 613577]). For all remaining genes on this comprehensive list of putative epileptic encephalopathy genes, a panel of clinicians reviewed the specific phenotypic details to determine the relevance to epileptic encephalopathy specifically. This highlighted 13 genes of interest, including eight genes from category 1, two from category 3, one from category 4, and six from category 5 (Table S4). We then applied the test for a significant excess of de novo mutations in these genes to statistically analyze their association with epileptic encephalopathy (Table S4). Of the 13 tested genes, 11 showed significant association with epileptic encephalopathy (ALG13, CHD2 [MIM 602119], DNMT1, GABRA1, GABRB3, GNA01 [MIM 139311], GRIN2A [MIM 138253], KCN1, SCN8A, SLC35A2, and SYNGAP1 [MIM 603384]), giving us a total of 26 high-confidence epileptic encephalopathy genes (11 newly tested genes + 15 original “definitive epileptic encephalopathy genes”). A full list of contributing variants and corresponding literature references is provided in Table S4.

Gene List Enrichment Analyses

To determine whether our list of de novo mutations was preferentially located in genes contained in a set of relevant gene lists, we used a previously published method. In brief, a binomial probability calculation was adopted to determine whether the de novo mutations identified in this cohort of epileptic encephalopathy
probands were selectively enriched within the well-sequenced coding sequence of genes within a particular gene list. This process corrects each gene list not only for the protein-coding real estate (size) of genes, but also for the average “callable real estate” across all trios.

We performed two additional tests to confirm the relationship of the observed gene list enrichments among genes harboring a de novo mutation in epileptic encephalopathy probands. First, for each gene list, we calculated the matching exact binomial test based on de novo single-nucleotide mutations identified among controls, corrected for corresponding CCDS protein-coding sizes. Then, for each gene list, we also directly compared the rate of overlapping single-nucleotide de novo mutations between our epileptic encephalopathy cohort and a control cohort.

For this direct comparison we used a two-tail Fisher’s exact test to determine whether there was a significantly increased frequency of de novo mutations overlapping a gene list among either the cases or controls.

Protein-Protein Interaction Network Analyses

Ingenuity Pathway Analysis (Ingenuity Systems) was used to assess connectivity of proteins harboring a de novo substitution. The requirements for assessing protein-protein interconnectivity were experimentally observed interactions. The permitted interaction types were: protein-protein, protein-DNA, activation, inhibition, phosphorylation, and ubiquitination. As described previously, preferential gene list enrichment among the network, compared

---

**Figure 1. Hot-Zone Analysis**

Illustrates where de novo single-nucleotide mutations reside along the variant-level vector (x axis: PolyPhen-2 HumVar score) and the gene-level vector (y axis: RVIS percentile score). We identified 411 controls9–14,16 (A, adapted from Zhu et al.34) and 232 epileptic encephalopathy cases (B) that reported at least one single-nucleotide de novo mutation. For each of the cases and controls, we plot the single most damaging single-nucleotide de novo mutation, as defined by the shortest Euclidian distance from the most damaging coordinate [1,0]. The hot zone (red shading) is defined as the region that reflects a PolyPhen-2 score ≥0.95 and a RVIS percentile score ≤25%.

Details for all hot-zone de novo mutations are available in Table S1. Among the case hot-zone insert we flag with red the hot-zone de novo mutations that occur in the epileptic encephalopathy gene list as described in the panel.
### Results

#### Distribution of De Novo Mutations in the Epileptic Encephalopathy Cohort

In the 356 trios, 29.18 ± 1.07 Mb (range 25.50–31.06 Mb) of the CCDS and associated splice sites were sufficiently sequenced in each family to detect a de novo mutation. These estimates were similar between the two cohorts (Epi4K, 29.18 ± 1.10 Mb [range 25.50–31.06 Mb]; EuroEPINOMICS-RES, 29.20 ± 0.25 Mb [range 28.77–29.54 Mb]). We identified a total of 429 de novo mutations in the 356 individuals analyzed (mean 1.2 per person). Of these, 58 (13.5%) are predicted loss-of-function (nonsense, frame-shift insertion-deletion, splice site) mutations and 281 (65.5%) are missense or in-frame deletions (Table S1). The frequency of loss-of-function mutations is significantly higher than the 7.1% observed in exome-sequenced control trios,9–14,16 (exact binomial test, p = 3.0 × 10^-6).

Using a likelihood analysis, we compared the distribution of the number of de novo mutations per individual in probands with epileptic encephalopathies (Figure S1) to the expected distribution in a control population. We find that individuals with epileptic encephalopathies have significantly more exonic de novo mutations compared to controls (p = 1.9 × 10^-4, Figure S2).

We show that by taking the single most damaging de novo mutation per individual in a population of controls,9–14,16 12% of de novo SNVs fall in the hot zone (Figure 1A). In contrast, among individuals with epileptic encephalopathies, the proportion of all single most damaging de novo mutations in the hot zone is 32%. This translates into an excess of 46 de novo mutations in the hot zone in our sample of probands with epileptic encephalopathies compared to control expectations.

| Gene          | Chr | Average Effectively Sequenced Length (bp) | Average Mutation Rate | Number of De Novo Mutations | p Value  
|---------------|-----|------------------------------------------|----------------------|-----------------------------|---------
| ALG13         | X   | 1,134.24                                 | 1.57 × 10^-3         | 2                           | 8.07 × 10^-12*** |
| CDKL5         | X   | 2,860.67                                 | 3.97 × 10^-3         | 3                           | 1.33 × 10^-6***  |
| CHD2          | 15  | 5,369.04                                 | 7.34 × 10^-3         | 2                           | 1.32 × 10^-3     |
| DNM1          | 9   | 2,334.48                                 | 4.58 × 10^-3         | 5                           | 2.99 × 10^-10*** |
| FASN          | 17  | 6,139.81                                 | 1.30 × 10^-4         | 2                           | 4.05 × 10^-3     |
| GABBR2        | 9   | 2,560.29                                 | 4.43 × 10^-3         | 2                           | 4.86 × 10^-4     |
| GABBR3        | 15  | 1,263.05                                 | 2.02 × 10^-3         | 4                           | 1.75 × 10^-9***  |
| GNAO1         | 16  | 1,319.68                                 | 2.49 × 10^-3         | 2                           | 1.55 × 10^-4     |
| HDAC4         | 2   | 2,769.92                                 | 6.04 × 10^-3         | 2                           | 9.00 × 10^-4     |
| KCNQ2         | 20  | 1,708.42                                 | 3.84 × 10^-3         | 2                           | 3.66 × 10^-4     |
| PIK3AP1       | 10  | 2,431.02                                 | 4.01 × 10^-3         | 2                           | 3.99 × 10^-4     |
| RANGAP1       | 22  | 1,557.37                                 | 2.48 × 10^-3         | 2                           | 1.54 × 10^-4     |
| RYR3          | 15  | 14,777.56                                | 2.18 × 10^-4         | 2                           | N/A  
| SCN1A         | 2   | 6,067.97                                 | 8.07 × 10^-3         | 7                           | 3.91 × 10^-3***  |
| SCN2A         | 2   | 5,899.52                                 | 7.69 × 10^-3         | 3                           | 2.07 × 10^-9***  |
| SCN8A         | 12  | 5,858.57                                 | 8.35 × 10^-3         | 2                           | 1.70 × 10^-3     |
| SLC35A2       | X   | 1,107.91                                 | 1.95 × 10^-3         | 2                           | 4.86 × 10^-5     |
| STXBP1        | 9   | 1,922.03                                 | 3.22 × 10^-3         | 5                           | 5.20 × 10^-11*** |
| TTN           | 2   | 104,698.88                               | 1.35 × 10^-3         | 2                           | 2.50 × 10^-1     |

*aAdjusted α is equivalent to 0.05 / 18,091 = 2.76 × 10^-6, indicated by single asterisk; 0.01 / 18,091 = 5.53 × 10^-7, indicated by double asterisk; or 0.001 / 18,091 = 5.53 × 10^-8, indicated by triple asterisk.

*bExact same mutation was observed in multiple unique individuals in this gene; in these instances, a p value reflecting probability for site-specific recurrent mutations was also calculated. The smaller p value is shown in the table.

*cMIM 602119.
*dMIM 605314.
*eMIM 607942.
*fMIM 602362.

gMutation rates for small insertion-deletion de novo mutations are not able to be estimated, therefore probability could not be estimated.

hMIM 314375.

iMIM 188840.
Considering only intolerant genes in the likelihood analysis (RVIS \( \leq 25\text{th} \) percentile), we likewise see an excess of de novo mutations in intolerant genes (\( p = 4.1 \times 10^{-4} \), Figure S2). From this analysis, we estimate that 2.8% (\( \chi^2 = 0.028 \); CI = 0–1) of the \(-4,000\) intolerant genes (RVIS \( \leq 25\% \)) in the genome contribute to epileptic encephalopathy risk. Because the confidence interval limits reliability of the parameter estimate, we simulated the sample size needed to reduce the boundaries of the CI. We found that sample sizes approximately four times the size of the current cohort would be expected to reduce the CI to half its current size.

**Screen for Genes Not Previously Implicated in Epileptic Encephalopathy**

We found 19 genes with a de novo mutation in two or more probands (Table 1). Seven genes carried a significant excess of de novo mutations, independently supporting their roles in epileptic encephalopathies (Table 1, Figure S3), including six genes known to carry epileptic encephalopathy-causing mutations (ALG13, \[^2\] Cdk5, \[^17\] GABRB3, \[^2\] SCN1A, \[^4\] SCN2A, \[^18\] and STXBP1 \[^19\]) and one, DNM1, not previously linked to disease. Clinical features of persons with a de novo mutation in DNM1 are summarized in Table 2. We note that our method analyzing the likelihood of seeing multiple de novo mutations in the same gene in the aggregate epileptic encephalopathy cohort does not take into account that some of the genes with mutations in more than one individual might be particularly associated with specific epileptic encephalopathy subphenotypes.

**Protein-Protein Interaction Network Analyses and Gene Enrichment Analysis**

The addition of de novo mutation data from 92 samples to our original investigation on \( 264 \) probands \[^4\] shows an extension of the interconnected network to \( 139 \) interacting proteins, and we continue to see an enrichment of known epileptic encephalopathy genes (Table S3). Strikingly, among the \( 139 \) genes forming this protein network, \( 42 \) (30.2%) encode proteins annotated to the “synaptic junction transmission” gene list, as defined by Ingenuity Pathway Analysis (IPA, Ingenuity Systems), including DNM1 and two other genes with multiple de novo mutations in this study (GABBR2 [MIM 607340] and RYR3 [MIM 180903]) (Figures 2 and S4). Furthermore, of the \( 26 \) definitive epileptic encephalopathy genes, as defined in Supplemental Methods, \( 13 \) genes (ARHGEF9, DNM1, GABRA1, GABRB3, GNAO1, GRIN2A, KCNQ2, PLCB1, SCN1A, SCN2A, SCN8A, STXBP1, SYNGAP1) encode proteins that are part of IPA’s “synaptic junction transmission” gene list. Given these patterns, we first assessed whether the single nucleotide de novo mutations found in the \( 356 \) probands were preferentially occurring within the \( 1,085 \) genes annotated to synaptic junction transmission by IPA and found significant enrichment (\( p = 7.3 \times 10^{-14} \)) (Table S5). Interestingly, we find enrichment for synaptic transmission genes even discounting all genes encoding ion channels (\( p = 3.1 \times 10^{-6} \)), but the reverse is not true (\( p = 0.38 \)).

Consistent with earlier studies, we also find that de novo mutations are preferentially drawn from a number of other groups of genes, including fragile X mental retardation protein (FMRP)-regulated genes (\( p = 3.9 \times 10^{-12} \))\[^2\] \[^20\] and Mouse Genome Informatics (MGI) seizure ortholog genes (\( p = 3.7 \times 10^{-19} \)).\[^21\] We see no similar preferential enrichment of de novo single-nucleotide mutations among any of the above gene lists when looking at de novo mutations observed in control trios (Table S5).

**Discussion**

Infantile spasms and Lennox-Gastaut syndrome represent classical phenotypes of epileptic encephalopathy. Analysis of exome data of \( 356 \) trios with these severe epileptic disorders shows an excess of exonic de novo mutations compared to control trios. While a number of studies report de novo mutations in neuropsychiatric diseases,\[^10\]–\[^13\],\[^22\],\[^23\] they did not show a statistically significant overall excess (genome-wide) of de novo mutations. Furthermore, we show that there is an excess of de novo mutations in mutation-intolerant genes and in a hot zone defined by a PolyPhen-2 score of \( \geq 0.95 \) and RVIS \( \leq 25\text{th} \) percentile in probands. These findings underscore the prominent role of de novo mutations in the etiology of epileptic encephalopathies. Pathway analysis further shows that there is a strong enrichment of de novo mutations in genes annotated to synaptic junction transmission, even when excluding ion channel genes. Disturbance of synaptic transmission thus seems to be a key factor in the pathogenesis of epileptic encephalopathies.

In the current data set, we find five individuals with a de novo missense mutation in DNM1, one of the genes involved in synaptic transmission. In our earlier study, de novo mutations in DNM1 were identified in two persons, but statistical evidence was insufficient to support pathogenicity.\[^2\] In this larger cohort we now securely implicate DNM1 as a gene in which mutations cause epileptic encephalopathy. All five probands with de novo mutations in DNM1 had infantile spasms with onset between \( 2 \) and \( 13 \) months. Four of the five persons evolved to Lennox-Gastaut syndrome. All individuals had severe to profound intellectual disability with pronounced hypotonia and absence of speech. The presence of some developmental delay prior to epilepsy onset in two probands may suggest an influence of DNM1 on neurodevelopment independent of seizure activity.

Dynamin-1 (DNM1) is an exclusively brain-expressed GTpase localizing to the presynaptic terminal. It is involved in activity-dependent synaptic vesicle endocytosis and membrane recycling.\[^24\] More precisely, it provides the mechanical force necessary to pinch off budding vesicles.

---

364 The American Journal of Human Genetics 95, 360–370, October 2, 2014
<table>
<thead>
<tr>
<th>Trio</th>
<th>LGSkj</th>
<th>ISg</th>
<th>LGSaix</th>
<th>NLES16</th>
<th>NLES7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, age</td>
<td>F, 15 years</td>
<td>M, 8 years</td>
<td>M, 6 years</td>
<td>E, 13 years</td>
<td>M, 6 years</td>
</tr>
<tr>
<td>Exam at birth</td>
<td>normal</td>
<td>normal</td>
<td>normal</td>
<td>normal</td>
<td>normal</td>
</tr>
<tr>
<td>Development prior to epilepsy onset</td>
<td>probably normal (lost skills between 9 and 11 months)</td>
<td>head control 2–3 months, at 6 months some delay noted</td>
<td>normal</td>
<td>all milestones delayed</td>
<td>all milestones delayed</td>
</tr>
<tr>
<td>Seizure onset</td>
<td>7 months</td>
<td>6 months</td>
<td>2 months</td>
<td>12 months</td>
<td>13 months</td>
</tr>
<tr>
<td>Seizure type at onset</td>
<td>atypical spasms</td>
<td>epileptic spasms</td>
<td>epileptic spasms</td>
<td>epileptic spasms</td>
<td>epileptic spasms</td>
</tr>
<tr>
<td>Other seizure types</td>
<td>atypical absences with eyelid fluttering, drop attacks, generalized tonic clonic seizures</td>
<td>atonic and tonic seizures</td>
<td>none</td>
<td>myoclonic, atypical absences, tonic, focal dyscognitive seizures, generalized tonic clonic seizures, obtundation status</td>
<td>atypical absences, tonic, focal dyscognitive seizures, obtundation status</td>
</tr>
<tr>
<td>Antiepileptic drug response</td>
<td>therapy resistant, longer periods of seizure freedom on vigabatrin and valproic acid</td>
<td>therapy resistant, some response to ketogenic diet</td>
<td>seizure free on ketogenic diet since age 3.5 years</td>
<td>therapy resistant</td>
<td>therapy resistant, no effect of ketogenic diet</td>
</tr>
<tr>
<td>Seizure outcome</td>
<td>seizure free between 3 and 8 years, then relapse</td>
<td>on-going frequent seizures</td>
<td>seizure free on ketogenic, off antiepileptic drugs</td>
<td>ongoing frequent seizures</td>
<td>ongoing frequent seizures</td>
</tr>
<tr>
<td>EEG at onset</td>
<td>slow background, multifocal discharges</td>
<td>hypsarrhythmia</td>
<td>high voltage bilateral slow spike-wave discharges</td>
<td>modified hypsarrhythmia</td>
<td>hypsarrhythmia</td>
</tr>
<tr>
<td>Course of EEG</td>
<td>slow background, slow generalized spike-wave discharges and multifocal (poly)speaks</td>
<td>slow background, left temporal slowing, slow generalized spike-wave discharges, diffuse (poly)speaks</td>
<td>not available</td>
<td>slow background, diffuse slow spike-wave discharges, sharp waves-slow waves; (poly)speaks; paroxysmal fast activity</td>
<td>slow background, diffuse multifocal sharp waves and sharp waves-slow waves; paroxysmal fast activity</td>
</tr>
<tr>
<td>Neurological examination</td>
<td>mild diffuse hypotonia, mild ataxia with wide based gait, mild tremor</td>
<td>general hypotonia</td>
<td>general hypotonia</td>
<td>axial hypotonia, secondary microcephaly</td>
<td>axial hypotonia</td>
</tr>
<tr>
<td>Development at last follow up</td>
<td>severe intellectual disability; no speech; autism spectrum disorder, behavioral problems with self-injurious behavior</td>
<td>severe intellectual disability; no speech; does not walk</td>
<td>severe intellectual disability; no speech; does not walk; behavioral problems with self-injurious behavior</td>
<td>profound intellectual disability; no speech; no visual fixation; does not sit or walk</td>
<td>profound intellectual disability; no speech; no visual fixation; does not sit or walk</td>
</tr>
<tr>
<td>MRI</td>
<td>normal</td>
<td>normal</td>
<td>normal</td>
<td>generalized cerebral atrophy</td>
<td>generalized cerebral atrophy</td>
</tr>
</tbody>
</table>

*Annotated based on NCBI IDs: NM_004408.3 and NP_004399.2.
from the synaptic membrane, and becomes essential during high levels of neuronal activity.24 The expression of DNM1 is upregulated during postnatal brain development and peaks during neurite and synapse formation.25 Functionally, DNM1 has five domains: (1) a GTPase domain that binds and hydrolyzes GTP and contains four GTP binding motifs (G1–G4), (2) a middle domain that is important for oligomerization, (3) a pleckstrin homology domain that binds lipids, (4) a GTPase effector domain that fulfills an assembly function and stimulates GTPase activity, and (5) a proline-rich domain that interacts with SH3 domains in other proteins (Figure 3).

All five DNM1 substitutions identified in this study affect highly conserved residues (Figure 3). Four substitutions are located in the GTPase domain, and two of these lie in a G motif (c.194C>A [p.Thr65Asn] in G1 and c.618G>C [p.Lys206Asn] in G4 [RefSeq accession numbers NM_004408.3, NP_004399.2]). The fifth substitution lies in the middle domain (Figure 3).

Several previous studies have introduced different dynamin mutant constructs into mammalian cells.26,27 The overall effect was an impairment of endocytosis, as was also shown for a mutant affecting the same amino acid in G1, p.Thr65Ala, which blocked hydrolysis of GTP.26 Several other mutants have been shown to exert a dominant-negative effect. Of particular interest in this regard is the spontaneous substitution p.Ala408Thr (DNM1ftfl), located in the middle domain of DNM1, in the so-called fitful mouse.28 Mice heterozygous for the substitution present with recurrent, often intractable, seizures whereas...
homozygous mice have a more severe phenotype with ataxia and often lethal seizures. DNM1<sup>flfl</sup> showed impaired oligomerization stalling endocytosis, and binding to wild-type DNM1, thereby exerting a dominant-negative effect. Because the loss of DNM1 preferentially affects inhibitory synapses, it is possible that the stalled endocytosis leads to inefficient recycling of synaptic vesicles with impaired tonic firing at inhibitory synapses and thus seizures as a result. It remains to be shown whether the five mutations identified in this study exert a similar dominant-negative effect.

While no other novel genes reached statistical significance, the excess of mutations among intolerant genes occurs only partially in either known or newly established genes for epileptic encephalopathies. This indicates that further genetic risk factors are present in our cohort. In particular, we identify multiple de novo mutations in GABBR2, FASN (MIM 600212), and RYR3 (Tables S1 and S2), making them strong candidate epileptic encephalopathy genes. All three genes are among the genes most intolerant to functional variation (RVIS: GABBR2, 5.08%; FASN, 0.38%; RYR3, 0.06%). GABBR2 encodes a subunit of the brain-expressed G protein coupled GABAB receptor. Knockout mice exhibit spontaneous seizures and severe memory impairment. Fatty acid synthase, a multifunctional protein encoded by FASN, plays a key role in de novo lipogenesis and is highly active in adult neural stem cells. Deletion of the gene in mice results in impairment of adult neurogenesis. Finally, RYR3 is a brain-expressed ryanodine receptor, responsible for calcium release from intracellular stores. RYR3 plays a role in synaptic plasticity and Ryr3 knockout mice exhibit impaired spatial learning.

After establishing a list of 26 definite epileptic encephalopathy genes, we estimate that at least 12% (42/356) of persons with infantile spasms or Lennox-Gastaut syndrome have definitive disease-causing de novo mutations in protein-coding regions (Table S1). We emphasize that the method used to define this list of epileptic encephalopathy genes is highly conservative and relies only on statistical evidence supporting an excess of reported de novo SNVs in epileptic encephalopathy probands across multiple studies. It does not incorporate any assessments of the biological consequences of mutations. As a result, GRIN2B did not reach our threshold in this analysis, although it is clearly linked to epileptic encephalopathy when considering evidence of both the presence of mutations in selected proband cohorts and functional and electrophysiological effects of mutations.

The likelihood and hot-zone analysis suggest additional causal genes beyond those securely implicated in our study. Evidence for these genes will accumulate with increasing sample sizes. This continuous increase of causally implicated genes is already illustrated by the identification of six additional genes with epileptic encephalopathy-causing mutations since the publication of the initial screen for de novo mutations in epileptic encephalopathies (CHD2, DNM1, GABRA1, GNAO1, SLC35A2, and GRIN2A). In all six genes, our initial study had identified at least one de novo mutation, but we did not have definite statistical evidence for their role at the time. In the current cohort of 356 individuals, we find additional de novo mutations in four of these new genes (CHD2, DNM1, GNAO1, and SLC35A2, Table S1; clinical data of not previously reported persons, Table S2).

In conclusion, by combining the data from two international consortia, we demonstrate the important role of damaging de novo mutations, confirm the role of five recently described genes with epileptic encephalopathy-causing mutations, and find secure evidence for a causative role of DNM1 mutations in severe epileptic...
encephalopathies. Moreover, we provide suggestive evidence for three other genes (GABBR2, FASN, and RYR3). The extent of the genetic heterogeneity associated with epileptic encephalopathies strongly motivates genome-wide screening in the clinical setting, rather than more targeted approaches. We advocate the establishment of a centralized genome data repository on epileptic encephalopathy probands to increase the speed of gene discovery in these devastating neurological disorders. Collectively, our data present evidence for a predominant role of synaptic dysregulation in epileptic encephalopathies, complementing the prevailing channelopathy paradigm in epilepsy.

Supplemental Data

Supplemental Data include four figures, six tables, consortia member affiliations, and Supplemental Acknowledgments and Author Contributions and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2014.08.013.

Consortia


Received: June 3, 2014
Accepted: August 29, 2014
Published: September 25, 2014

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ClustalW2, http://www.ebi.ac.uk/Tools/msa/clustalw2/
DeNovoGear, http://sourceforge.net/projects/denovogear/
European Genome-phenome Archive (EGA), https://www.ebi.ac.uk/ega
GATK, http://www.broadinstitute.org/gatk/
Genic Intolerance, http://chgv.org/GenicIntolerance/
GenomeComb, http://genomcomb.sourceforge.net/
Mouse Genome Informatics, http://www.informatics.jax.org/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
PolyPhen-2, http://www.genetics.bwh.harvard.edu/pph2/

Acknowledgments

We are deeply grateful to the probands, their families, clinical research coordinators, and referring physicians for their participation and provision of phenotype data and DNA samples used in this study.

We thank the EPGP Administrative Core (C. Freyer, K. Schärdein, R. Fahlström, S. Cristofoo, and K. McGovern), EPGP Bioinformatics Core (G. Neshitt, K. McKenna, and V. Mays), staff at the Coriell Institute – NINDS Genetics Repository (C. Tarn and A. Scuttii), and members of the Duke Center for Human Genome Variation (J. Bridgers, J. Kebbielli, H.S. Kim, S. Gewalt, J. Maia, E. Campbell, K. Cronin, L. Hong, S. Sliwa, J. Quinn, C. Mebane, E. Kazura, R. Perez-Morro, and M. Cook) for their extraordinary dedication and commitment to this work. We also thank S. Shinnar (Albert Einstein College of Medicine) and N. Risch (University of California, San Francisco) for valuable input into the creation of EPGP and Ep4K, and R. Stewart, K. Gwinn, and R. Corriveau from the National Institute of Neurological Disorders and Stroke for their careful oversight and guidance of both EPGP and Ep4K.

This work was supported by grants from the National Institute of Neurological Disorders and Stroke (The Epilepsy Phenome/Genome Project NS053998; Ep4K NS077364, NS077274, NS077303, and NS077276), The Andrew's Foundation, Finding a Cure for Epilepsy and Seizures, the Richard Thalheimer Philanthropic Fund, and the Eucoros program EuroEPINOMICS-RES of the European Science Foundation. Additional funding sources are summarized in the Supplemental Acknowledgments. The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.
Accession Numbers

The Epi4K exome-sequencing data reported in this paper are deposited in the Database of Genotypes and Phenotypes (dbGaP), accession number phs000654.v2.p1. The EuroEPINOMICS-RES exome-sequencing data are deposited in the European Genome-phenome Archive, accession numbers EGAS00001000190, EGAS000001000386, and EGAS000001000048.

References


