

# UCSF

## UC San Francisco Previously Published Works

### Title

Application of rare variant transmission disequilibrium tests to epileptic encephalopathy trio sequence data.

### Permalink

<https://escholarship.org/uc/item/82j6m40z>

### Journal

European journal of human genetics : EJHG, 25(7)

### ISSN

1018-4813

### Authors

Epi4K Consortium  
EuroEPINOMICS-RES Consortium  
Epilepsy Phenome Genome Project

### Publication Date

2017-06-01

### DOI

10.1038/ejhg.2017.61

Peer reviewed

## SHORT REPORT

# Application of rare variant transmission disequilibrium tests to epileptic encephalopathy trio sequence data

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

The classic epileptic encephalopathies, including infantile spasms (IS) and Lennox–Gastaut syndrome (LGS), are severe seizure disorders that usually arise sporadically. *De novo* variants in genes mainly encoding ion channel and synaptic proteins have been found to account for over 15% of patients with IS or LGS. The contribution of autosomal recessive genetic variation, however, is less well understood. We implemented a rare variant transmission disequilibrium test (TDT) to search for autosomal recessive epileptic encephalopathy genes in a cohort of 320 outbred patient–parent trios that were generally prescreened for rare metabolic disorders. In the current sample, our rare variant transmission disequilibrium test did not identify individual genes with significantly distorted transmission over expectation after correcting for the multiple tests. While the rare variant transmission disequilibrium test did not find evidence of a role for individual autosomal recessive genes, our current sample is insufficiently powered to assess the overall role of autosomal recessive genotypes in an outbred epileptic encephalopathy population.

*European Journal of Human Genetics* (2017) 25, 894–899; doi:10.1038/ejhg.2017.61; published online 17 May 2017

## INTRODUCTION

Epileptic encephalopathies are severe and therapy-resistant epilepsies of childhood, which frequently lead to developmental delay and multiple associated medical issues. Infantile spasms (IS) and Lennox–Gastaut syndrome (LGS) represent two of the more common broad subtypes of epileptic encephalopathies. Many novel genes for epileptic encephalopathies have been discovered in the last 5 years, fueled by the access to whole-exome sequencing. In particular, exome sequencing has highlighted the important role of *de novo* variants with current estimates suggesting that over 15% of non-Dravet epileptic encephalopathy cases are explained by a disease-causing *de novo* variant in an established epileptic encephalopathy gene<sup>1,2</sup> with this estimate increased to over 80% among individuals diagnosed with Dravet syndrome.<sup>3</sup> Up to a further 3% have been reported to be explained by likely clinically relevant *de novo* copy-number variants.<sup>4</sup>

While the role of *de novo* genetic variation in epileptic encephalopathies is increasingly understood, the role of recessive genetic variation, outside of recessive neurometabolic disorders such as lysosomal disorders, amino acid or organic acid imbalances, congenital disorders of glycosylation, and some mitochondrial diseases, remains unclear. In our current study, we systematically assessed autosomal recessive inheritance in 320 IS or LGS patient–parent trios who did not have a likely disease-causing *de novo* variant among one of the established dominant epileptic encephalopathy genes.<sup>1,2</sup> In general, the 320 cases studied here had already been intensively studied for neurometabolic disorders using biochemical assessments.

## SUBJECTS AND METHODS

### Cohort

Three-hundred and twenty epileptic encephalopathy trios were recruited through multiple international consortia, including 57 IS or LGS trios unpublished in our earlier studies.<sup>1,2</sup> Patients did not have a clearly identified metabolic or genetic cause for their epilepsy based on clinically available testing, which varied across institutions. This collection of 320 trios did not include: (a)

patients previously found to have a disease-causing *de novo* variant in an established dominant epileptic encephalopathy gene, and (b) trios where exome sequencing was based on a lymphoblastoid cell line source for at least one of the three family members. The overall cohort was not enriched for consanguineous parents. Only two parent pairs showed an identity-by-descent  $>0.125$ , both  $<0.15$ , which is approximately equivalent to third degree relatives.<sup>5</sup>

Among the 320 trios; two families reported multiple affected children. For one of these families both the proband and affected sibling were investigated through exome sequencing, while for the second family only the proband and parents were studied. Sequencing methods used to generate the sequence data have been previously described.<sup>1,2</sup>

### Transmission disequilibrium tests

For the transmission test, we used two approaches that we have previously introduced.<sup>6,7</sup> First, we tested for an autosomal homozygous or compound heterozygous effect using core TDT.<sup>7</sup> In computing the test, we selected loss-of-function and missense single-nucleotide substitution variants found at a global population minor allele frequency  $<5\%$  ( $MAF < 0.05$ ). The loss-of-function variants were defined as stop gain, stop lost, start lost, and canonical splice acceptor and donor site variants. For the missense variants, we used our in-house Analysis Tool for Annotated Variants platform to identify the possibly and probably damaging variants based on a maximum Polyphen-2 HumDiv and HumVar prediction score<sup>8</sup> of  $>0.4333$ . This test was applied to each autosomal gene individually as well as collectively across a set of 99 autosomal recessive neurometabolic genes published by van Karnebeek *et al.*<sup>9</sup> The recessive neurometabolic gene-set analysis allowed us to assess whether there was evidence for elevated rate of recessive genotypes among recessive neurometabolic disorder genes beyond what had already been screened out by the conventional biochemical assessments performed on this patient sample.

Second, we tested for a general effect of inherited autosomal variation by using a rare variant TDT that uses information from an independent collection of population controls (6503 EVS<sup>10</sup> plus 1303 IGM sequenced controls) to weigh the contribution of variants to the final test statistic.<sup>6</sup> In this analysis, qualifying variants were defined using the same PolyPhen-2 thresholds as above and were again required to have a global  $MAF < 5\%$ . Given that population

A full list of consortium members appears after the References section.

\*Correspondence: Epi4K Consortium, 701 West 168th Street, New York, NY 10032, USA. E-mail: Epi4K@columbia.edu

Received 27 September 2016; revised 8 March 2017; accepted 27 March 2017; published online 17 May 2017

stratification can impact the power of the test but not the type I error, we restricted this second analysis to trios with European ancestry ( $n = 286$  trios).

### Power simulation

As we had not previously performed power simulations for the type of gene-set application conducted in our neurometabolic analysis, we conduct a new power simulation to evaluate the types of effects that we could exclude based on this analysis (Figure 1). In these simulations, we conditioned on the parental genotype information contained in this IS/LGS population sample and characterized the distribution of offspring genotypes, given this information and the fact that the offspring is affected. This distribution is a function of the number of causal genes, for which the family is informative, which is related to the density of causal genes within the actual gene set, and the relative risk of the offspring developing disease, given that they have two affected gene copies. We give the details of this procedure here.

Let  $G_f$ ,  $G_m$ ,  $G_o$  be the number of gene copies harboring a qualifying variant in the trio's father, mother, and offspring, respectively. We condition our power analysis on the observed parental genotype and study our ability to identify signal given a differing proportion of causal genes (out of the total number of genes considered),  $\gamma$  and differing relative risks,  $R$ , of being diseased, given two gene copies (of a causal disease gene) are affected *versus* less than two copies are affected. Since the analysis is conditional on the observed parental data, only a subset of genes and families are informative.<sup>7</sup> Specifically, only 20 genes across 54 families can have compound genotypes that lead to informative transmissions, that is,  $G_f = G_m = 1$ ,  $G_f = 1$ ,  $G_m = 2$  or  $G_f = 2$ ,  $G_m = 1$ . A total of 46 families are informative for only one gene and eight families are informative for two genes. In each of these eight families, the two genes are located on different chromosomes, so we assume that the transmissions of each gene are independent.

Let  $D_o = 1$  indicate the fact that the offspring is affected. Let  $C$  be an indicator of whether the gene whose transmission is being considered is among the set of disease causal genes or not. When a family is informative for two genes disease causal indicators are given for each gene by  $C_1$  and  $C_2$ . Note, we assume the disease risk for samples with multiple affected disease genes are the same with those with only one affected disease gene.

To simulate trios under the alternative, we first randomly select  $20\gamma$  genes as disease causal and then generate offspring as follows.

If the family is informative for only one gene, the distribution of both offspring's gene copies being affected is given by:

$$\begin{aligned} \Pr(G_o = 2 | G_f = 1, G_m = 1, D_o = 1, C = 1) &= \frac{R}{R+3} \\ \Pr(G_o = 2 | G_f + G_m = 3, D_o = 1, C = 1) &= \frac{R}{R+1} \\ \Pr(G_o = 2 | G_f = 1, G_m = 1, D_o = 1, C = 0) &= 0.75 \\ \Pr(G_o = 2 | G_f + G_m = 3, D_o = 1, C = 0) &= 0.5. \end{aligned}$$

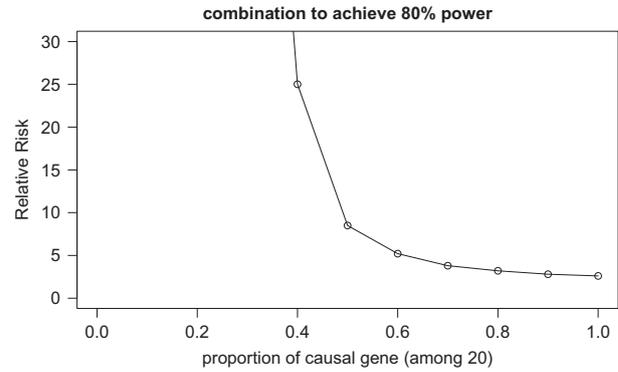
If the family is informative for two genes and no more than one of them are disease causal, the compound genotype, the two genes, can be computed independently of one another using the equation above. When both genes are disease causal, their transmissions are not independent given the offspring is affected. In this case, the compound genotypes of the offspring, for the two genes, can be given by,

$$\begin{aligned} \Pr(G_{o1} = 2, G_{o2} = 2 | G_{f1} = G_{m1} = G_{f2} = G_{m2} = 1, D_o = 1, C_1 = C_2 = 1) &= \frac{R}{R+9} \\ \Pr(G_{o1} = 2, G_{o2} \neq 2 | G_{f1} = G_{m1} = G_{f2} = G_{m2} = 1, D_o = 1, C_1 = C_2 = 1) &= \frac{6R}{7R+9} \\ \Pr(G_{o1} = 2, G_{o2} = 2 | G_{f1} = G_{m1} = 1, G_{f2} = 1, G_{m2} = 2, D_o = 1, C_1 = C_2 = 1) &= \frac{R}{5R+3} \\ \Pr(G_{o1} = 2, G_{o2} \neq 2 | G_{f1} = G_{m1} = 1, G_{f2} = 1, G_{m2} = 2, D_o = 1, C_1 = C_2 = 1) &= \frac{R}{5R+3} \\ \Pr(G_{o1} \neq 2, G_{o2} = 2 | G_{f1} = G_{m1} = 1, G_{f2} = 1, G_{m2} = 2, D_o = 1, C_1 = C_2 = 1) &= \frac{3R}{5R+3}, \end{aligned}$$

where  $G_{o1}$ ,  $G_{m1}$ ,  $G_{f1}$  and  $G_{o2}$ ,  $G_{m2}$ ,  $G_{f2}$  denotes the trio's compound genotypes at the first and second gene, respectively. We apply coreTDT to each simulated data set, and for each combination of  $\gamma$  and  $R$ , we use 1000 replicates to estimate the power. The combination of  $\gamma$  and  $R$  that obtains 80% power are presented in Figure 1.

### Accession numbers

The exome-sequencing data reported in this paper are deposited in the Database of Genotypes and Phenotypes (dbGaP) with the accession number phs000654.v2.p1. The EuroEPINOMICS-RES data are deposited in the



**Figure 1** coreTDT power simulation conditional on the parental genotype of 54 informative families and 20 informative genes in the compound heterozygous analysis. Presents the combination of the relative risk and the proportion of disease causal genes among these 20 informative genes as a single unit, under which the tests can achieve 80% power.

European Genome-phenome Archive with the accession numbers: EGAS00001000190, EGAS00001000386, and EGAS00001000048.

### RESULTS

We assessed the role of inherited rare variation using the population control-weighted rare variant TDT.<sup>6</sup> This test was applied to each autosomal gene across 320 eligible trios. No gene reached exome-wide significance after correcting for the 17 816 consensus coding sequence (CCDS release 14) autosomal genes (adjusted  $\alpha = 2.81 \times 10^{-6}$ , Table 1). Though population stratification cannot affect the false positive rate of the test, it can affect the power.<sup>6</sup> We also conducted an analysis that was restricted to the 286 trios of European ancestry. Again, no gene reached the exome-wide significance level (Table 1).

We then tested for the presence of a recessive effect in each autosomal gene across the 320 trios. After quality control, only 3472 autosomal genes were found to have at least one informative family, that is, contain qualifying variants within the gene and that could, potentially, lead to homozygous or compound heterozygous offspring. None of these 3472 genes achieved significance after correcting for the number of genes tested (adjusted  $\alpha = 1.44 \times 10^{-5}$ ). The 10 most significant genes are listed in Table 2. To investigate whether there is any evidence of recessive neurometabolic involvement in this sample, we also applied the coreTDT to the set of 99 autosomal recessive neurometabolic genes,<sup>9</sup> looking for an enrichment of homozygous or compound heterozygous offspring across the entire gene set as a single unit. No enrichment was found ( $P = 0.51$ ).

We investigated the power of this analysis. Since only 54 families are informative for at least one of the 99 autosomal recessive neurometabolic genes, and only 20 genes have at least one informative family, our analyses are effectively restricted to these 54 families and 20 genes. We vary the proportion of informative genes that are actually disease causal and the relative risk and identify combinations of these parameters that attain at least 80% power (see 'Power simulation' for details). The results of this analysis can be found in Figure 1. As can be seen, even when the compound heterozygous or homozygous qualifying variants are fully penetrant, the causal gene proportion must be  $>40\%$  to attain 80% power. When the proportion of causal genes is larger, for example, 80%, we will have high power to detect an effect even with a relatively low relative risk.

Using established standards to identify clinically relevant recessive genotypes,<sup>11,12</sup> one trio was found to have inherited two clinically relevant *SPATA5* variants in a compound heterozygous

**Table 1 Top 10 genes from the analysis of rvTDT with 320 and subsequently with the subset of 286 European ancestry trios**

Test	Gene	Number of SNVs <sup>a</sup>	Uncorrected P-value
Analysis of rvTDT with 320 trios	<i>ABCA13</i>	52	1.17E-05
	<i>CENPO</i>	4	0.00236
	<i>DST</i>	51	0.00297
	<i>IPPK</i>	5	0.00198
	<i>PCF11</i>	10	0.00032
	<i>SCAPER</i>	5	0.00137
	<i>SLC46A3</i>	6	0.00185
	<i>STEAP4</i>	4	0.00291
	<i>TRAF3IP1</i>	7	0.00292
	<i>ZNF878</i>	7	0.00252
Analysis of rvTDT with 286 trios	<i>ABCA13</i>	52	0.00025
	<i>SCAPER</i>	5	0.00046
	<i>ANKZF1</i>	11	0.00079
	<i>DST</i>	51	0.00123
	<i>TRAF3IP1</i>	7	0.00172
	<i>STEAP4</i>	4	0.00247
	<i>SBSN</i>	12	0.00295
	<i>OR2B2</i>	5	0.00298
	<i>SAP130</i>	8	0.00326
	<i>SCARF1</i>	11	0.00344

P-value is based on the linear combination test with population controls. Adjusted  $\alpha$  correcting for the number of genes is  $P < 2.81 \times 10^{-6}$ .

<sup>a</sup>Representing the number of qualifying variants found in this gene across all families.

**Table 2 Top 10 autosomal genes from the analysis of coreTDT with 320 trios**

Gene	Number of SNVs <sup>a</sup>	Uncorrected P-value
<i>PGM2L1</i>	4	0.00195
<i>CEP120</i>	8	0.00391
<i>CR1</i>	25	0.00756
<i>C14ORF177</i>	4	0.01288
<i>CNTRL</i>	24	0.01563
<i>DACT1</i>	8	0.01563
<i>KATNB1</i>	10	0.01563
<i>SYNJ1</i>	17	0.01563
<i>ZNF677</i>	11	0.01563
<i>KIAA1614</i>	15	0.01973

Adjusted  $\alpha$  correcting for the number of genes is  $P < 2.81 \times 10^{-6}$ .

<sup>a</sup>Representing the number of qualifying variants found in this gene across all families.

manner.<sup>13</sup> The proband's phenotype is consistent with the *SPATA5* disease literature, and both variants (NM\_145207.2; c.1677C>A (p.(Tyr559\*)) and c.251G>A (p.(Arg84Gln)) have previously been described as clinically relevant among other patients with *SPATA5* encephalopathy.<sup>13</sup>

## DISCUSSION

A number of rare recessive disorders can present with an epileptic encephalopathy, particularly neurometabolic disorders; the latter are generally identified by biochemical analyses of blood, urine or CSF. We performed a global, hypothesis-free test to assess the role of autosomal recessive genetic variation in 320 patients with classic epileptic encephalopathies undiagnosed with standard clinical

workups. Our sample of patient–parent trios did not identify a genome-wide significant departure in the observed number of offspring with recessive genotypes from that expected for any specific gene, or among 99 genes compiled for autosomal recessive neurometabolic disorders.

Many classical recessive metabolic disorders are routinely identified through biochemical screening prior to research study enrollment. Within our sample of 320 trios, we did not find any genetic neurometabolic disorders that were missed through the conventional biochemical screening. From a clinical perspective, this emphasizes that conventional biochemical screening for these treatable causes should continue to be pursued. We did identify a single case among the 320 with a clinically relevant recessive genotype in *SPATA5*,<sup>13</sup> a recently described gene for a recessive condition characterized by seizures, microcephaly, intellectual disability, and hearing loss.

The role of various dominant epilepsy genes including *ALG13*, *CDKL5*, *DNM1*, *GABRB3*, *SCN1A*, *SCN2A*, and *STXBP1*, for epileptic encephalopathies was securely established through exome sequencing of 356 trios and subsequent genome-wide assessments for excess *de novo* variants identified in individual genes.<sup>1,2</sup> No single gene passes a comparable threshold among the 320 trios studied here when assessing autosomal recessive genotypes. We demonstrate that the current sample of 320 trios is insufficiently powered to appropriately estimate what overall contribution autosomal recessive epilepsy genes have on the epileptic encephalopathies; however, our power analyses show that we do have sufficient power to rule out a large role from known recessive neurometabolic genes among this patient sample that has been previously screened for such factors using conventional biochemical assessments. Using a similar approach, a recent study on 4125 patient–parent trios with various developmental disorders identified two novel autosomal recessive disease genes exceeding genome-wide significance,<sup>14</sup> emphasizing the importance of acquiring larger numbers to more confidently interpret the current lack of signal for very rare genetic epilepsies with recessive inheritance. Large-scale collaborative initiatives like the epilepsy genetic initiative and the Epi25 effort will aid the efforts to analyze genomic data on this scale.

## CONFLICT OF INTEREST

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## ACKNOWLEDGEMENTS

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 (C Freyer, K Fox, R Fahlstrom, S Cristofaro, and K McGovern), Bioinformatics Core (G Nesbitt, K McKenna, and V Mays), staff at Coriell Institute – NINDS Genetics Repository, and members of the Institute for Genomic Medicine, Columbia University (P Cansler, J Charoensri, B Copeland, S Kamalakaran, J Keebler, B Krueger, C Malone, C Mebane, and M Cook) for their dedication and commitment to this work. We also thank R Stewart, K Gwinn, R Corriveau, B Fureman, and V Whittemore from the National Institute of Neurological Disorders and Stroke for their careful oversight and guidance of both EPGP and Epi4K. We thank the following organizations for assistance in publicizing EPGP; enabling us to recruit participants effectively: AED Pregnancy Registry, American Epilepsy Society, Association of Child Neurology Nurses, California School Nurses Organization, Child Neurology Society, Citizens United for Research in Epilepsy, Dravet Syndrome Foundation, Epilepsy Alliance of Orange County, Epilepsy Foundation, Epilepsy Therapy Project, Finding a Cure for Epilepsy and Seizures, IDEA League, InfantileSpasms.com, Lennox–Gastaut Syndrome Foundation, PatientsLikeMe, People Against Childhood Epilepsy, PVNH Support & Awareness, and Seizures

& Epilepsy Education. We would like to acknowledge the following individuals or groups for the contributions of control samples: D Daskalakis; P Lugar; J Milner; T Young and K Whisenhunt; Z Farfel, D Lancet, and E Pras; W Lowe; R Gbadegesin and M Winn; K Schmader, S McDonald, HK White, and M Yanamadala; A Holden; E Behr; C Moylan; AM Diehl and M Abdelmalek; S Palmer; G Nestadt; J Samuels; Y Wang; M Carrington; M Harms; T Miller; A Pestronk; R Bedlack; R Brown; N Shneider; S Gibson; J Ravits; A Gilter; J Glass; F Baas; E Simpson; and G Rouleau; K Welsh-Bomer, C Hulette, J Burke; The ALS Sequencing Consortium; The Murdoch Study Community Registry and Biorepository; M Connors, L Morris, and the CHAVI investigators; the Carol Woods and Crosdaile Retirement Communities; and DUHS (Duke University Health System) Nonalcoholic Fatty Liver Disease Research Database and Specimen Repository. The collection of control samples and data was funded in part by: Biogen Idec.; The Duke Chancellor's Discovery Program Research Fund 2014; Bill and Melinda Gates Foundation; The Division of Intramural Research; B57 SAIC-Fredrick Inc M11-074; Bryan ADRC NIA P30 AG028377; The Ellison Medical Foundation New Scholar award AG-NS-0441-08; National Institute of Mental Health (K01MH098126, R01MH097993); National Institute of Allergy and Infectious Diseases (1R56AI098588-01A1); and National Institute of Allergy and Infectious Diseases Center (U19-AI067854, UMI-AI100645). We thank the NHLBI GO Exome Sequencing Project and its ongoing studies that produced and provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926), and the Heart GO Sequencing Project (HL-103010). This work was supported by grants from the National Institute of Neurological Disorders and Stroke (The Epilepsy Phenome/Genome Project NS053998; Epi4K NS077364, NS077274, NS077303, and NS077276), The Andrew's Foundation, Finding a Cure for Epilepsy and Seizures, the Richard Thalheimer Philanthropic Fund, and the Eurocores program EuroEPINOMICS-RES of the European Science Foundation. The project received further support through grants from the Fund for Scientific Research Flanders (FWO); the Academy of Finland (141549); the Folkhälsan Research Foundation; the program 'Investissements d'avenir' ANR-10-IAIHU-06; the Federal Ministry for Education and Research (IonNeuONet: 01GM1105), the German Research Foundation (DFG: HE5415/3-1; Le1030/11-1; RO3396/2-1), the German Society for Epileptology (DGfE), the Foundation noepilep.; the Swiss National Science Foundation (SNF: 32EP30\_136042/1); the Wellcome Trust (09805); intramural funds of the University of Kiel; the Popgen 2.0 network (P2N) through the German Ministry for Education and Research (01EY1103); and the European Union through Seventh Framework Programme (FP7) under the project DESIRE (N602531). The project also received infrastructural support through the Institute of Clinical Molecular Biology in Kiel, supported in part by DFG Cluster of Excellence 'Inflammation at Interfaces' and 'Future Ocean'. Orrin Devinsky, David B Goldstein, Steve Petrou and Slavé Petrovski have interests in companies related to epilepsy precision medicine.

## AUTHOR CONTRIBUTIONS

Sequence analysis and statistical interpretation: YJ (lead analyst), ASA, and DBG. Bioinformatics processing: JBr, ELH, YJ, SIP, ZR, and QW. Clinical expert panel: IH, HCM, AP, and SW. Writing of manuscript: ASA, SFB, PDJ, DBG, ELH, IH, YJ, DHL, and SIP. Epi4K Steering Committee: ASA, SFB, PC, ND, DD, EEE, MPE, TG, DBG, ELH, MRJ, RuKu, DHL, AGM, HCM, TJO, RO, StP, SIP, AP, IES, and ES. EuroEPINOMICS-RES consortium leadership and study coordination: IH, PDJ, and SW. EuroEPINOMICS study design: A-EL, AS, BK, HL, IH, PG, PDJ, and SW. EPGP study design: BKA, OD, DD, MPE, RuKu, DHL, RO, ES, and MRW. Epi4K epileptic encephalopathy phenotyping: SFB, PC, DD, RuKu, DHL, HCM, RO, AP, IES, and ES. EuroEPINOMICS-RES proband recruitment and phenotyping: A-EL, AR, CM, CD, DC, DP, DH, EL, FZ, FRos, HC, HH, HM, HL, IH, JJ, JL, JS, KSe, KMK, KSt, NB, PS, RG, RSM, SvS, SW, SB, TL, TT, US, VK, and YW. EPGP proband recruitment and phenotyping: BA, EA, FA, DA, JFB, SFB, GDC, DCo, PCr, OD, DD, MEF, NBF, DF, EBG, TG, SG, SRH, JH, KH, SLH, HEK, RCKn, EK, RaKu, RuKu, DHL, SMM, PVM, EJN, JMP, JP, KP, AP, IES, JJS, RAS, JSi, LS, MS, LLT, AV, EPGV, GKV, JW, and PW. EPGP phenotype data analysis: BA,

BKA, AB, JB, GDC, OD, DD, MPE, JF, TG, SJ, AK, RCKn, RuKu, DHL, RO, JMP, AP, IES, RAS, RS, ES, JJS, JSu, PW, and MRW.

## CONSORTIUM MEMBERSHIP

Andrew S Allen<sup>1</sup>, Samuel F Berkovic<sup>2</sup>, Joshua Bridgers<sup>3</sup>, Patrick Cossette<sup>4</sup>, Dennis Dlugos<sup>5</sup>, Michael P Epstein<sup>6</sup>, Tracy Glauser<sup>7</sup>, David B Goldstein<sup>3</sup>, Erin L Heinzen<sup>3</sup>, Yu Jiang<sup>1</sup>, Michael R Johnson<sup>8</sup>, Ruben Kuzniecky<sup>9</sup>, Daniel H Lowenstein<sup>10</sup>, Anthony G Marson<sup>11</sup>, Heather C Mefford<sup>12</sup>, Terence J O'Brien<sup>13</sup>, Ruth Ottman<sup>14,15</sup>, Steven Petrou<sup>13,16,17</sup>, Slavé Petrovski<sup>2,3,14</sup>, Annapurna Poduri<sup>18</sup>, Zhong Ren<sup>3</sup>, Ingrid E Scheffer<sup>19,20</sup>, Elliott Sherr<sup>21</sup>, Quanli Wang<sup>3</sup>

<sup>1</sup>Department of Biostatistics and Bioinformatics, Duke Clinical Research Institute, Durham, NC, USA; <sup>2</sup>Department of Medicine, Epilepsy Research Centre, University of Melbourne (Austin Health), Heidelberg, Victoria, Australia; <sup>3</sup>Institute for Genomic Medicine, Columbia University Medical Center, New York, NY, USA; <sup>4</sup>Department of Neurosciences and CHUM Research Center, Université de Montréal, CHUM-Hôpital Notre-Dam, Montréal, Quebec, Canada; <sup>5</sup>Department of Neurology and Pediatrics, The Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA; <sup>6</sup>Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA; <sup>7</sup>Division of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; <sup>8</sup>Centre for Clinical Translation, Division of Brain Sciences, Imperial College London, London, UK; <sup>9</sup>NYU School of Medicine, New York University, New York, NY, USA; <sup>10</sup>Department of Neurology, University of California, San Francisco, San Francisco, CA, USA; <sup>11</sup>Department of Molecular and Clinical Pharmacology, University of Liverpool, Clinical Sciences Centre, Liverpool, UK; <sup>12</sup>Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, WA, USA; <sup>13</sup>Department of Medicine and Neurology, The Royal Melbourne Hospital, Parkville, Victoria, Australia; <sup>14</sup>Department of Epidemiology and Neurology, The G.H. Sergievsky Center, Columbia University, New York, NY, USA; <sup>15</sup>Division of Epidemiology, New York State Psychiatric Institute, New York, NY, USA; <sup>16</sup>Florey Institute for Neuroscience and Mental Health, The University of Melbourne, Melbourne, Victoria, Australia; <sup>17</sup>Centre for Neural Engineering, The University of Melbourne, Melbourne, Victoria, Australia; <sup>18</sup>Department of Neurology, Division of Epilepsy and Clinical Neurophysiology, Boston Children's Hospital, Boston, MA, USA; <sup>19</sup>Department of Medicine, Epilepsy Research Centre, University of Melbourne (Austin Health), Heidelberg, Victoria, Australia; <sup>20</sup>Department of Pediatrics, Florey Institute, Royal Children's Hospital, University of Melbourne, Melbourne, Victoria, Australia; <sup>21</sup>Departments of Neurology, Pediatrics and Institute of Human Genetics, University of California, San Francisco, San Francisco, CA, USA

### EuroEPINOMICS-RES

Rudi Balling<sup>22</sup>, Nina Barisic<sup>23</sup>, Stéphanie Baulac<sup>24,25,26,27,28</sup>, Hande Caglayan<sup>29</sup>, Dana Craiu<sup>30,31</sup>, Peter De Jonghe<sup>32,33</sup>, Christel Depienne<sup>34,35</sup>, Renzo Guerrini<sup>36</sup>, Ingo Helbig<sup>37,38</sup>, Helle Hjalgrim<sup>39,40</sup>, Dorota Hoffman-Zacharska<sup>41</sup>, Johanna Jahn<sup>37</sup>, Karl Martin Klein<sup>42,43</sup>, Bobby Koelman<sup>44</sup>, Vladimir Komarek<sup>45</sup>, Roland Krause<sup>22</sup>, Eric Leguern<sup>24,25,26,27,28</sup>, Anna-Elina Lehesjoki<sup>46</sup>, Johannes R Lemke<sup>47</sup>, Holger Lerche<sup>48</sup>, Tarja Linnankivi<sup>49</sup>, Carla Marini<sup>36</sup>, Patrick May<sup>22</sup>, Rikke S Møller<sup>39,40</sup>, Hiltrud Muhle<sup>37</sup>, Deb Pal<sup>50</sup>, Aarno Palotie<sup>51,52,53</sup>, Felix Rosenow<sup>42,43</sup>, Kaja Selmer<sup>54</sup>, Jose M Serratos<sup>55</sup>, Sanjay Sisodiya<sup>56,57</sup>, Ulrich Stephani<sup>37</sup>, Katalin Sterbova<sup>45</sup>, Pasquale Striano<sup>58</sup>, Arvid Suls<sup>32,33,59</sup>, Tiina Talvik<sup>60,61</sup>, Sarah von Spiczak<sup>37</sup>, Yvonne Weber<sup>48</sup>, Sarah Weckhusen<sup>32,33</sup>, Federico Zara<sup>62</sup>

<sup>22</sup>Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg, UK; <sup>23</sup>Department of Paediatrics, University of Zagreb, Medical School, University Hospital Centre Zagreb, Zagreb, Croatia; <sup>24</sup>Sorbonne Universités, UPMC Univ Paris 06, UM 75, Paris, France; <sup>25</sup>INSERM, U1127, Paris, France; <sup>26</sup>CNRS, UMR 7225, Paris, France; <sup>27</sup>ICM, Paris, France; <sup>28</sup>Department of Genetic and Cytogenetic, AP-HP, Hôpital Pitié-Salpêtrière, Paris, France; <sup>29</sup>Department of Molecular Biology and Genetics, Bogaziçi University, Istanbul, Turkey; <sup>30</sup>Department of Neurology, Pediatric Neurology, Psychiatry, Neurosurgery, Pediatric Neurology Clinic II, 'Carol Davila' University of Medicine, Bucharest, Romania; <sup>31</sup>Pediatric Neurology Clinic, 'Professor Doctor Alexandru Obregia' Clinical Hospital, Sector 4, Bucharest, Romania; <sup>32</sup>Department of Molecular Genetics, Neurogenetics group, VIB,

Antwerp, Belgium; <sup>33</sup>Laboratory of Neurogenetics, Institute Born-Bunge, University of Antwerp, Antwerp, Belgium; <sup>34</sup>Département de Médecine translationnelle et Neurogénétique, IGBMC, CNRS UMR 7104/INSERM U964/ Université de Strasbourg, Illkirch, France; <sup>35</sup>Laboratoire de cytogénétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; <sup>36</sup>Pediatric Neurology Unit and Laboratories, Children's Hospital A. Meyer—University of Florence, Florence, Italy; <sup>37</sup>Department of Neuropediatrics, University Medical Center Schleswig-Holstein, Christian-Albrechts University, Kiel, Germany; <sup>38</sup>Division of Neurology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; <sup>39</sup>Danish Epilepsy Centre, Dianalund, Denmark; <sup>40</sup>Institute for Regional Health research, University of Southern Denmark, Odense, Denmark; <sup>41</sup>Department of Medical Genetics, Institute of Mother and Child, Warsaw, Poland; <sup>42</sup>Epilepsy Center Rhine-Main, Center of Neurology and Neurosurgery, Goethe-University Frankfurt, Frankfurt, Germany; <sup>43</sup>Epilepsy Center Hessen, Philipps-University Marburg, Marburg, Germany; <sup>44</sup>Department of Medical Genetics, University Medical Center, Utrecht, The Netherlands; <sup>45</sup>Department of Child Neurology, Second Medical Faculty and University Hospital Motol, Praha, Czech Republic; <sup>46</sup>Folkhälsan Institute of Genetics, Research Programs Unit, Molecular Neurology and Neuroscience Center, University of Helsinki, Helsinki, Finland; <sup>47</sup>Institute of Human Genetics, University of Leipzig Hospitals and Clinics, Leipzig, Germany; <sup>48</sup>Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany; <sup>49</sup>Children's Hospital, Pediatric Neurology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; <sup>50</sup>Department of Clinical Neuroscience, Institute of Psychiatry, King's College London, London, UK; <sup>51</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridgeshire, UK; <sup>52</sup>Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; <sup>53</sup>Program in Medical and Population Genetics and Genetic Analysis Platform, The Broad Institute of MIT and Harvard, Cambridge, MA, USA; <sup>54</sup>Department of Medical Genetics, Oslo University Hospital, Oslo, Norway; <sup>55</sup>Department of Neurology, Epilepsy Unit, Institute of Medical Research at Fundación Jiménez Díaz University Hospital and Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER) Unit 744, Madrid, Spain; <sup>56</sup>Department of Clinical and Experimental Epilepsy, UCL Institute of Neurology, London, UK; <sup>57</sup>Epilepsy Society, Buckinghamshire, UK; <sup>58</sup>Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, Pediatric Neurology and Muscular Diseases Unit, University of Genova and Gaslini Institute, Genova, Italy; <sup>59</sup>GENOMED, Center of Medical Genetics, University of Antwerp and Antwerp University Hospital, Antwerp, Belgium; <sup>60</sup>Department of Pediatrics, University of Tartu, Tartu, Estonia; <sup>61</sup>Department of Neurology and Neurorehabilitation, Children's Clinic, Tartu University Hospital, Tartu, Estonia; <sup>62</sup>Department of Neurosciences, Laboratory of Neurogenetics, Pediatric Neurology and Muscular Diseases Unit, University of Genova and Gaslini Institute, Genova, Italy

## EPGP

Bassel Abou-Khalil<sup>63</sup>, Brian K Alldredge<sup>64,65</sup>, Dina Amrom<sup>66,67</sup>, Eva Andermann<sup>68,69</sup>, Frederick Andermann<sup>70,71</sup>, Jocelyn F Bautista<sup>72</sup>, Samuel F Berkovic<sup>2</sup>, Judith Bluvstein<sup>9</sup>, Gregory D Cascino<sup>73</sup>, Damian Consalvo<sup>74</sup>, Patricia Crumrine<sup>75</sup>, Orrin Devinsky<sup>76</sup>, Dennis Dlugos<sup>5</sup>, Michael P Epstein<sup>6</sup>, Miguel E Fiol<sup>77</sup>, Nathan B Fountain<sup>78</sup>, Jacqueline French<sup>79</sup>, Daniel Friedman<sup>80</sup>, Tracy Glauser<sup>7</sup>, Kevin Haas<sup>81</sup>, Sheryl R Haut<sup>82</sup>, Jean Hayward<sup>83</sup>, Sucheta Joshi<sup>84</sup>, Andres Kanner<sup>85</sup>, Heidi E Kirsch<sup>86</sup>, Eric H Kossoff<sup>87</sup>, Rachel Kuperman<sup>88</sup>, Ruben Kuzniecky<sup>9</sup>, Daniel H Lowenstein<sup>10</sup>, Shannon M McGuire<sup>89</sup>, Paul V Motika<sup>90</sup>, Edward J Novotny<sup>91</sup>, Ruth Ottman<sup>14,15</sup>, Juliann M Paolicchi<sup>92</sup>, Jack Parent<sup>93,94</sup>, Kristen Park<sup>95,96</sup>, Annapurna Poduri<sup>18</sup>, Ingrid E Scheffer<sup>19,20</sup>, Renée A Shellhaas<sup>97</sup>, Elliott Sherr<sup>21</sup>, Joseph Sirven<sup>98</sup>, Michael C Smith<sup>99</sup>, Joseph Sullivan<sup>10</sup>, Liu Lin Thio<sup>100</sup>, Anu Venkat<sup>101</sup>, Eileen PG Vining<sup>102</sup>, Gretchen K Von Allmen<sup>103</sup>, Judith L Weisenberg<sup>104</sup>, Peter Widdess-Walsh<sup>105</sup>, Melodie R Winawer<sup>106</sup>

<sup>63</sup>Department of Neurology, Vanderbilt University Medical Center, Nashville, TN, USA; <sup>64</sup>Department of Clinical Pharmacy, UCSF School of Pharmacy, San Francisco, CA, USA; <sup>65</sup>Department of Neurology, UCSF School of Medicine, San Francisco, CA, USA; <sup>66</sup>Neurogenetics Unit, Montreal Neurological Hospital and Institute, Montreal, Québec, Canada; <sup>67</sup>Departments of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada; <sup>68</sup>Neurogenetics Unit and Epilepsy Research Group, Montreal Neurological

Hospital and Institute, Montreal, Québec, Canada; <sup>69</sup>Departments of Neurology and Neurosurgery and Human Genetics, McGill University, Montreal Quebec, Canada; <sup>70</sup>Epilepsy Research Group, Montreal Neurological Hospital and Institute, Montreal, Québec, Canada; <sup>71</sup>Departments of Neurology and Neurosurgery and Pediatrics, McGill University, Montreal, Quebec, Canada; <sup>72</sup>Department of Neurology, Cleveland Clinic Lerner College of Medicine and Epilepsy Center of the Cleveland Clinic Neurological Institute, Cleveland, OH USA; <sup>73</sup>Division of Epilepsy, Mayo Clinic, Rochester, MN, USA; <sup>74</sup>Epilepsy Center, Neurology Division, Ramos Mejia Hospital, Buenos Aires, Argentina; <sup>75</sup>Medical Epilepsy Program and EEG and Child Neurology, Children's Hospital of Pittsburgh of UPMC, Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; <sup>76</sup>NYU and Saint Barnabas Epilepsy Centers, NYU School of Medicine, New York, NY, USA; <sup>77</sup>Department of Neurology, Epilepsy Care Center, University of Minnesota Medical School, Minneapolis, MN, USA; <sup>78</sup>FE Drefuss Comprehensive Epilepsy Program, University of Virginia, Charlottesville, VA, USA; <sup>79</sup>NYU Comprehensive Epilepsy Center, New York, NY, USA; <sup>80</sup>Department of Neurology, NYU School of Medicine, New York, NY, USA; <sup>81</sup>Vanderbilt University Medical Center, Nashville, TN, USA; <sup>82</sup>Comprehensive Epilepsy Center, Montefiore Medical Center, Bronx, NY, USA; <sup>83</sup>The Kaiser Permanente Group, Oakland, CA, USA; <sup>84</sup>Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, MI, USA; <sup>85</sup>Department of Neurology, Comprehensive Epilepsy Center, Miller School of Medicine, University of Miami, Miami, FL, USA; <sup>86</sup>Departments of Neurology and Radiology, University of California, San Francisco, CA, USA; <sup>87</sup>Neurology and Pediatrics, Child Neurology, Pediatric Neurology Residency Program, Johns Hopkins Hospital, Baltimore, MD, USA; <sup>88</sup>Epilepsy Program, Children's Hospital and Research Center Oakland, Oakland, CA, USA; <sup>89</sup>Clinical Neurology, Children's Hospital Epilepsy Center of New Orleans, New Orleans, LA, USA; <sup>90</sup>Comprehensive Epilepsy Center, Oregon Health and Science University, Portland, OR, USA; <sup>91</sup>Departments of Neurology and Pediatrics, University of Washington School of Medicine, Seattle Children's Hospital, Seattle, WA, USA; <sup>92</sup>Weill Cornell Medical Center, New York, NY, USA; <sup>93</sup>Department of Neurology and Neuroscience Graduate Program, University of Michigan Medical Center, Ann Arbor, MI, USA; <sup>94</sup>Ann Arbor Veterans Administration Healthcare System, Ann Arbor, MI, USA; <sup>95</sup>University of Colorado School of Medicine, Aurora, CO, USA; <sup>96</sup>Department of Pediatrics, Division of Neurology, Children's Hospital Colorado, Aurora, CO, USA; <sup>97</sup>University of Michigan, Pediatric Neurology, Ann Arbor, MI, USA; <sup>98</sup>Department of Neurology, Mayo Clinic, Scottsdale, AZ, USA; <sup>99</sup>Department of Neurological Sciences, Rush Epilepsy Center, Rush University Medical Center, Chicago, IL, USA; <sup>100</sup>Department of Neurology, Washington University School of Medicine, St Louis, MO, USA; <sup>101</sup>Department of Pediatrics, Saint Peter's University Hospital, New Brunswick, NJ, USA; <sup>102</sup>Department of Neurology, Johns Hopkins Hospital, Baltimore, MD, USA; <sup>103</sup>Division of Child and Adolescent Neurology, Department of Pediatrics, University of Texas Medical School, Houston, TX, USA; <sup>104</sup>Department of Neurology, Division of Pediatric Neurology, Washington University School of Medicine, St Louis, MO, USA; <sup>105</sup>Department of Neurology, Beaumont Hospital and Royal College of Surgeons, Dublin, Ireland; <sup>106</sup>Department of Neurology, The G.H. Sergievsky Center, Columbia University, New York, NY, USA

- 1 Epi4K Consortium. Epilepsy Phenome/Genome Project, Allen AS *et al*: *De novo* mutations in epileptic encephalopathies. *Nature* 2013; **501**: 217–221.
- 2 EuroEPINOMICS-RES Consortium. Epilepsy Phenome/Genome Project. Epi4K Consortium: *De novo* mutations in synaptic transmission genes including DNMT1 cause epileptic encephalopathies. *Am J Hum Genet* 2014; **95**: 360–370.
- 3 Harkin LA, McMahon JM, Iona X *et al*: The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain* 2007; **130** (Pt 3): 843–852.
- 4 Epilepsy Phenome/Genome Project Epi4K Consortium: Copy number variant analysis from exome data in 349 patients with epileptic encephalopathy. *Ann Neurol* 2015; **78**: 323–328.
- 5 Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 6 Jiang Y, Satten GA, Han Y *et al*: Utilizing population controls in rare-variant case-parent association tests. *Am J Hum Genet* 2014; **94**: 845–853.

- 7 Jiang Y, McCarthy JM, Allen AS: Testing the effect of rare compound-heterozygous and recessive mutations in case-parent sequencing studies. *Genet Epidemiol* 2015; **39**: 166–172.
- 8 Adzhubei IA, Schmidt S, Peshkin L *et al*: A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–249.
- 9 van Karnebeek CD, Shevell M, Zschocke J, Moeschler JB, Stockler S: The metabolic evaluation of the child with an intellectual developmental disorder: diagnostic algorithm for identification of treatable causes and new digital resource. *Mol Genet Metab* 2014; **111**: 428–438.
- 10 Exome Variant Server (EVS), NHLBI GO Exome Sequencing Project (ESP), Seattle, WA, USA.
- 11 Zhu X, Petrovski S, Xie P *et al*: Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med* 2015; **17**: 774–781.
- 12 Need AC, Shashi V, Hitomi Y *et al*: Clinical application of exome sequencing in undiagnosed genetic conditions. *J Med Genet* 2012; **49**: 353–361.
- 13 Tanaka AJ, Cho MT, Millan F *et al*: Mutations in SPATA5 are associated with microcephaly, intellectual disability, seizures, and hearing loss. *Am J Hum Genet* 2015; **97**: 457–464.
- 14 Akawi N, McRae J, Ansari M *et al*: Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. *Nat Genet* 2015; **47**: 1363–1369.