

Appendix S1. Materials & methods

Monogenic analysis

Single Nucleotide Variants (SNVs) in the 555 gene panel were analyzed in all patients. Variants were filtered according to quality criteria (using GATK guidelines (<https://gatk.broadinstitute.org/hc/en-us>)), allelic frequency ($\leq 1\%$ in GnomADv3.1 (<https://gnomad.broadinstitute.org/>) and ≤ 1 allele in our patient cohort) and predicted pathogenicity (CADD score ≥ 30 (1)). The 140 remaining SNVs were examined individually for inheritance, phenotype compatibility and reported pathogenicity in the ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>). An independent search for monogenic variants in selected GGE and NAFE patients was conducted in parallel by the Human Genetic Variation Research Group (Royal College of Surgeons in Ireland (RCSI), Dublin). The findings were provided as a personal communication.

A subset of 56 patients with a phenotype of rare and complex epilepsy were reanalyzed in the ERN EpiCARE (<https://epi-care.eu/>) – Solve-RD project (<https://solve-rd.eu/>). Several variant filtering pipelines were applied in search of SNVs, Structural Variants (SVs), Copy Number variants (CNVs), mitochondrial genome variants and variants involving splice sites in the 1342 gene panel (2,3) (manuscript in preparation). The Solve-RD genetic analysis platform GPAP was also used as a collaborative platform to identify additional patients with pathogenic variants in new genes (4).

Twelve patients with a very severe phenotype, of which 10 were also included in the Solve-RD project, were further analyzed in search for homozygous or compound heterozygous variants in the whole exome. In addition, an analysis of all heterozygous variants included in a previously identified chromosome 12 locus was performed in a patient from a large family with autosomal dominant temporal lobe epilepsy and febrile seizures (5,6). A Southern Blot test in search of a *CSTB* dodecamere expansion was performed in another patient because of a clinical evolution suggestive of progressive myoclonus epilepsy.

All reported variants were classified according to ACMG guidelines (7). Whenever possible, the variants were confirmed with Sanger sequencing, and familial segregation was studied.

Oligogenic analysis – cases and controls selection

Out of 422 cases and 5308 controls, we selected 386 cases and 800 controls that were sequenced using the same Illumina bait set. All the cases were recruited in Belgian centers. Their ethnicity, based on the ethnic origin of the four grandparents, was available in 372 out of 386 patients: 252 originated from Belgium, 37 from North Western Europe outside Belgium, 34 fully or in part from Southern Europe, 10 fully or in part from Eastern Europe, 28 fully or in part from Africa, 9 fully or in part from Asia and 2 fully or in part from America. The 800 controls were referred from different centers: 383

from Italy, 194 from Japan, 124 from Taiwan, 39 from Turkey, 31 from Lebanon, and 29 from Cyprus. A Principal Component Analysis was performed using genotypic data from cases and controls (Figure S2). The outliers were selected based on visual evaluation, and were removed. 34 cases were excluded, and 352 remained. 326 controls were excluded, and 474 remained. All the controls from the Japanese and Taiwanese centers were removed (318 controls). The final Principal Component Analysis plot (Figure S2) showed that the remaining cases and controls, although not perfectly overlapping, were close to each other. Finally, 112 cases with structural epilepsies were excluded. 240 cases and 474 controls remained for the oligogenic analysis.

Oligogenic analysis – variant filtration

SNVs from the whole exome were filtered according to quality criteria (using GATK hard filtering criteria: $FS \leq 60$, $MQ \geq 40$, $MQRankSum \geq -12.5$, $ReadPosRankSum \geq -8$, $QD > 2.0$ (<https://gatk.broadinstitute.org/hc/en-us>)) and allelic frequency ($< 1\%$ in GnomADv3.1 (<https://gnomad.broadinstitute.org/>)). Furthermore, the variants were filtered according to their allelic frequency in the cohort, and only the variants with a frequency $< 3\%$ within the cases and within each of the control populations from the different referring centers (Italy, Turkey, Lebanon, and Cyprus) were kept for analysis, in order to remove probable sequencing errors, and population specific variants. All indels, synonymous variants and intronic variants were excluded from the analysis.

Oligogenic analysis – Hop tool score (8)

The final score from Hop is established by combining a pathogenicity score (PS) and a disease relevance score (DS). The PS is the prediction score from the VarCoPP2.0 model. In this study, only variant combinations with prediction scores included in the 99%-confidence zone of the predictor, corresponding to a prediction score ≥ 0.647 , were considered for analysis. This 99%-confidence zone was established by testing neutral combinations from the 1000 Genome Project and obtaining the minimal prediction resulting in 1% false positives (9). The DS is calculated within the Hop tool and scores the biological proximity of the two genes to the gene panel provided as disease information (8).

Oligogenic analysis – individual approach

The individual approach concerned six selected patients with a severe phenotype. The top 50 variant combinations from the patient were obtained with the Hop tool (8). A second, more restrictive variant filtering was performed in order to reduce the number of gene combinations. Individual variants from the combinations were filtered according to allelic frequency ($< 0.5\%$ in GnomADv3.1 (<https://gnomad.broadinstitute.org/>)) and < 4 allele in our patient cohort) and predicted pathogenicity

(CADD score >15 (1)). The variant combinations containing the excluded individual variants were removed. The remaining variant pairs were combined into small networks of variant combinations to try to detect potential oligogenic signatures.

References

1. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014 Mar;46(3):310–5. doi:10.1038/ng.2892 PubMed PMID: 24487276; PubMed Central PMCID: PMC3992975.
2. Matalonga L, Hernández-Ferrer C, Piscia D, Solve-RD SNV-indel working group, Schüle R, Synofzik M, et al. Solving patients with rare diseases through programmatic reanalysis of genome-phenome data. *Eur J Hum Genet EJHG.* 2021 Sep;29(9):1337–47. doi:10.1038/s41431-021-00852-7 PubMed PMID: 34075210; PubMed Central PMCID: PMC8440686.
3. Denommé-Pichon AS, Matalonga L, de Boer E, Jackson A, Benetti E, Banka S, et al. A Solve-RD ClinVar-based reanalysis of 1522 index cases from ERN-ITHACA reveals common pitfalls and misinterpretations in exome sequencing. *Genet Med Off J Am Coll Med Genet.* 2023 Apr;25(4):100018. doi:10.1016/j.gim.2023.100018 PubMed PMID: 36681873.
4. Laurie S, Piscia D, Matalonga L, Corvó A, Fernández-Callejo M, Garcia-Linares C, et al. The RD-Connect Genome-Phenome Analysis Platform: Accelerating diagnosis, research, and gene discovery for rare diseases. *Hum Mutat.* 2022 Jun;43(6):717–33. doi:10.1002/humu.24353 PubMed PMID: 35178824; PubMed Central PMCID: PMC9324157.
5. Claes L, Audenaert D, Deprez L, Van Paesschen W, Depondt C, Goossens D, et al. Novel locus on chromosome 12q22-q23.3 responsible for familial temporal lobe epilepsy associated with febrile seizures. *J Med Genet.* 2004 Sep;41(9):710–4. doi:10.1136/jmg.2004.019257 PubMed PMID: 15342703; PubMed Central PMCID: PMC1735896.
6. Depondt C, Van Paesschen W, Matthijs G, Legius E, Martens K, Demaerel P, et al. Familial temporal lobe epilepsy with febrile seizures. *Neurology.* 2002 May 14;58(9):1429–33. doi:10.1212/wnl.58.9.1429 PubMed PMID: 12011300.
7. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med Off J Am Coll Med Genet.* 2015 May;17(5):405–24. doi:10.1038/gim.2015.30 PubMed PMID: 25741868; PubMed Central PMCID: PMC4544753.
8. Gravel B, Renaux A, Papadimitriou S, Smits G, Nowé A, Lenaerts T. Prioritization of oligogenic variant combinations in whole exomes. *Bioinforma Oxf Engl.* 2024 Mar 29;40(4):btac184. doi:10.1093/bioinformatics/btac184 PubMed PMID: 38603604; PubMed Central PMCID: PMC11037482.
9. Versbraegen N, Gravel B, Nachtegaele C, Renaux A, Verkinderen E, Nowé A, et al. Faster and more accurate pathogenic combination predictions with VarCoPP2.0. *BMC Bioinformatics.* 2023 May 1;24(1):179. doi:10.1186/s12859-023-05291-3 PubMed PMID: 37127601; PubMed Central PMCID: PMC10152795.