Advancing the phenome alongside the genome in epilepsy studies

Ingo Helbig, MD
Dick Lindhout, MD

Correspondence to
Dr. Helbig:
helbigi@email.chop.edu

Neurology® 2017;89:14–15

The development of high-throughput sequencing technologies has led to remarkable progress in understanding the genetic basis of human epilepsies. Over the last 5 years, more than 20 genes have been identified through genome-first approaches, identifying novel genes first and then working backwards to understand the associated phenotypes. This approach has been useful in epileptic encephalopathies, where family-based exome studies have identified a growing list of genes with causative de novo mutations.[1]

The reliance on genetic rather than phenotypic data has led to an unusual schism in the epilepsy field. Historically, precise and detailed phenotyping was the key component to successful gene identification. For example, the discovery of *SCN1A* in genetic epilepsy with febrile seizures plus relied on phenotypic subtleties between febrile seizures and febrile seizures plus, and the identification of *CHRNA4* as the first gene in human epilepsy required a careful distinction between nocturnal frontal lobe seizures and parasomnias.[2,3] However, in 2017, precise phenotyping is often perceived as a bottleneck in large-scale studies that is dependent on human effort.

In the current issue of *Neurology*®, Tobochnik and collaborators[4] present a novel approach to reconcile the discrepancy between detailed phenotyping and large-scale studies. Using the Epilepsy Phenome/Genome Project (EPGP) dataset, the authors demonstrate the surprising aspects of epilepsy genetics when phenotypic features are taken into account. The EPGP dataset is remarkable as one of the few epilepsy datasets where detailed and uniform phenotyping has been performed in every participant. The authors ask the question whether certain phenotypic features in focal epilepsies are seen more often in affected relatives than expected. There is some precedence to suggest that such an approach is worthwhile: the *LGI1* gene was initially discovered in familial lateral temporal lobe epilepsy, where phenotypic details such as the presence of auditory features determined whether a family was positive for segregation with this gene.[5]

To approach the question of familial aggregation of focal seizures symptoms, 149 families with nonacquired focal epilepsy in the EPGP cohort were investigated.[4] The authors found that motor, autonomic, psychic, and aphasic seizure semiologies clustered in families. Even subtle phenotypic features such as whole body posturing, diaphoresis, dyspnea, fear, and déjà vu/jamais vu showed familial aggregation. Also, the type of focal seizure, including simple partial seizures, complex partial seizures, and secondarily generalized seizures, clustered in families. The long list of phenotypic features that demonstrate familial aggregation in the EPGP dataset is remarkable. It conflicts with our current understanding of disease genetics and suggests a component in human epilepsy that is yet to be identified. Many of the features identified as present in up to 40% of patients have an odds ratio of 2 or higher—no genetic factor known to date in human epilepsy is compatible with such a model. Accordingly, the study by Tobochnik and collaborators raises 2 main questions.

First, can the identified phenotypic features help in discovering novel genes? Particularly in milder nonlesional epilepsies, few genetic risk factors are known and novel strategies to find causative genes are needed. The current study builds upon the authors' previous work on familial clustering in seizure types and epilepsy syndromes.[6] Various genetic models for familial clustering have been discussed; for example, allelic heterogeneity of a major gene, multiple interactions between genes, imprinting, interaction with environmental factors, and interaction between a major gene and modifier genes. So far, the discussion had been theoretical given the lack of available genetic data. However, the EPGP cohort has been exome-sequenced within the Epi4K project.[7] Combining the findings by Tobochnik et al. with the existing genetic data is therefore feasible in the near future.

Second, what are the future possibilities of phenotyping in epilepsy genetic studies? It is tempting to speculate that detailed phenotypes in large-scale studies may help prioritize novel genes. However, the traditional phenotyping approach is labor- and cost-intensive. The cost for phenotyping a single patient

See page 22

From the Division of Neurology (I.H.), The Children's Hospital of Philadelphia, PA; Department of Genetics (D.L.), University Medical Center Utrecht; and Stichting Epilepsie Instellingen Nederland (SEIN) (D.L.), Heemstede, the Netherlands.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the editorial.

in the EPGP cohort approached $1,000—equaling the cost of a human genome sequence in 2017. But there are interesting advances in the phenotyping field. Novel technologies using data extraction from electronic medical health records or standardized phenotypic languages such as Human Phenotype Ontology may facilitate the inclusion of phenotypic datasets.[8] Furthermore, integration of imaging and EEG datasets as in the ENIGMA or iEEG project may result in new insights through cross-modal analyses.[9,10]

The study by Tobochnik and collaborators demonstrates that systematic phenotyping in large epilepsy datasets can reveal phenotypic aggregations that are beyond our current understanding of disease genetics. It is a good reminder that even in the era of genome-first study, there is a primacy of the phenome that needs to direct future studies.

STUDY FUNDING

No targeted funding reported.

DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

REFERENCES

1. Epi4K-Consortium, Epilepsy-Phenome/Genome-Project. De novo mutations in epileptic encephalopathies. Nature 2013;501:217–221.
2. Escayg A, MacDonald BT, Meisler MH, et al. Mutations of *SCN1A*, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet 2000;24:343–345.
3. Steinlein OK, Mulley JC, Propping P, et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet 1995;11:201–203.
4. Tobochnik S, Fahlstrom R, Shain C, Winawer MR, for the EPGP Investigators. Familial aggregation of focal seizure semiology in the Epilepsy Phenome/Genome Project. Neurology 2017;89:22–28.
5. Kalachikov S, Evgrafov O, Ross B, et al. Mutations in *LGI1* cause autosomal-dominant partial epilepsy with auditory features. Nat Genet 2002;30:335–341.
6. Winawer MR, Marini C, Grinton BE, et al. Familial clustering of seizure types within the idiopathic generalized epilepsies. Neurology 2005;65:523–528.
7. Epi4K Consortium. Epi4K: gene discovery in 4,000 genomes. Epilepsia 2012;53:1457–1467.
8. Kohler S, Doelken SC, Mungall CJ, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res 2014;42:D966–D974.
9. Medland SE, Jahanshad N, Neale BM, Thompson PM. Whole-genome analyses of whole-brain data: working within an expanded search space. Nat Neurosci 2014;17:791–800.
10. Wagenaar JB, Worrell GA, Ives Z, Dümpelmann M, Litt B, Schulze-Bonhage A. Collaborating and sharing data in epilepsy research. J Clin Neurophysiol 2015;32:235–239.