Published in final edited form as:
Epilepsy Behav. 2013 July ; 28(0 1): S63–S65. doi:10.1016/j.yebeh.2012.07.006.

Genome Wide Association Studies (GWAS) and Common Forms of Human Epilepsy

Russell J. Buono
Department of Biomedical Science, Cooper Medical School of Rowan University

Abstract

Several GWAS focused on common forms of epilepsy are underway. Currently, only one locus has been published that reached genome wide statistical significance. Two other loci that also reach genome wide statistical significance have been reported as preliminary data and are awaiting publication. Several additional loci identified in these studies fall just short of statistical significance and it is hoped that future large scale meta-analyses will confirm these early findings and identify new loci that influence common forms of human epilepsy. Next generation DNA sequencing (NGS) studies are also underway and in the future will identify rare DNA variations of large effect that also contribute to the final epilepsy phenotypes under study. Finally, these studies have the potential to identify biomarkers of anti-epileptic drug (AED) response as epilepsy patient GWAS and NGS data are stratified based on AED efficacy and tolerability.

Keywords

Epilepsy susceptibility; Genome Wide Association; Anti-epilepsy drug (AED) pharmacogenetics; DNA sequencing

The search for genetic and epigenetic factors that contribute to the cause of common epilepsies or that influence responses to anti-epilepsy drugs (AED) will be greatly accelerated in the next decade as high throughput genomic techniques are utilized. Two such techniques being utilized are GWAS and next generation DNA sequencing (NGS). It is generally agreed that complex human traits including diseases such as epilepsy (or response to AEDs) are caused by interactions between multiple gene variations and environmental factors. GWAS and NGS will identify common DNA variations and rare DNA mutations respectively that interact with each other and the environment to cause common forms of epilepsy and influence AED response.

Although many are abandoning GWAS in favor of NGS, it is reasonable to suggest that utilizing both techniques in a complimentary way is warranted for identification of genetic factors contributing to human epilepsy and AED response. GWAS data have been published since 2007 and in the past five years there have been numerous articles in journals and the lay press, pointing out the limitations of the technique. However, a recent critical review documents that GWAS studies during this time have had remarkable success. GWAS

Corresponding author at: Department of Biomedical Science, Cooper Medical School of Rowan University, 401 South Broadway, Camden, NJ 08103. Phone: 856 361 2863, buonor@rowan.edu.

identified common DNA variation accounting for 10-20% of the genetic liability associated with type two diabetes, Crohn's disease and multiple sclerosis [1]. GWAS data have confirmed loci of importance previously identified by linkage or single gene association and identified new chromosomal loci for many different complex traits. GWAS data also demonstrate that loci associated with disease in one population are often also found in association with the same disease in populations of different ethnic backgrounds, providing evidence that common variation across populations contribute to risk for common disease [1]. GWAS results for some complex traits have been less impressive. This may be related to sample size as when sufficiently large numbers of samples are studied (n=10,000 or more) multiple common variations of small effect size are often discovered. Thus, the future of epilepsy gene discovery will use a combination of GWAS and NGS to identify common and rare DNA variations associated with common forms of human epilepsy and AED response.

Epilepsy Genetics and Phenotype

In the past, genetic linkage analysis demonstrated success in identifying variations in genes that co-segregated with rare seizure disorders inherited in a Mendelian fashion. These variations were then designated as mutations, assuming the rare variation co-segregating with phenotype was causal. NGS is likely to find multiple rare DNA variations that may be causal to any particular phenotype, but proving this causality is very difficult if not impossible. For example, mutations of the sodium channel gene SCN1A were linked to Dravet syndrome as SCN1A mutations were found to co-segregate with ill family members but not found in unaffected relatives or healthy controls. Furthermore, some patients with Dravet syndrome were found to harbor a *de novo* mutations in SCN1A not found in healthy controls [2]. However, recent NGS studies in epilepsy document cases of healthy controls carrying mutations in SCN1A previously linked to Dravet syndrome [3]. There are also Dravet patients with no mutations in SCN1A [2]. These data suggest that while rare DNA variations can act as causal mutations, they do not always cause disease and that loci across the genetic background can modify the effects of any other locus.

Mendelian inherited forms of epilepsy are rare and most clinicians will never see or treat such a patient. In contrast, the vast majority of epilepsy patients (over 95%) suffer from common forms of human epilepsy that are inherited in a non Mendelian fashion. These common forms are divided into two main categories. The first is Genetic Generalized Epilepsy (GGE, formerly called idiopathic generalized epilepsy) and includes entities such as juvenile myoclonic epilepsy (JME) and childhood absence epilepsy (CAE). The second is localization related epilepsy or cryptogenic focal epilepsy (CFE) including temporal lobe epilepsy. These common forms of epilepsy exemplify complex traits that are likely caused by multiple gene variations and interaction with environmental factors. A recent study used meta-analysis on linkage data in patients with GGE and found two genomic regions linked to this phenotype [4]. However, attempts to perform linkage in common forms of epilepsy have not been very successful as they are complicated by non Mendelian patterns of inheritance and linkage data do not resolve to individual genes, but rather large genomic regions. In addition, these studies take extensive resources to collect sufficiently large pedigrees for statistical power to detect linked loci and in epilepsy are confounded by the fact that both focal epilepsy and GGE are often present together in any given pedigree.

There is substantial debate over the appropriate level of phenotyping that is necessary to make genetic discovery. In general, it appears that a broad phenotype has been more successful compared to a restricted phenotype when it comes to identification of important genetic factors related to specific disease via GWAS [1]. Epilepsy phenotyping is quite complex and there are no data sets yet that directly compare genetic results from equal numbers of subjects with a broad versus restricted phenotype. For example, in the GGE

category there are multiple epilepsy subtypes that could be "lumped" together such as JME and CAE, or these could be "split" apart and studied separately. JME or CAE patients could be further subdivided into categories that include seizure frequency, severity, age of onset, presence of myoclonus, or time of day of seizure. Pragmatically, lumping patients with a broad phenotype increases sample size and statistical power where splitting does just the opposite. The extra resources related to recruiting very specific phenotypic categories has not been proven to enhance genetic discovery and is difficult to justify based on presumed advantages of sample homogeneity. Focal epilepsy presents a problem with lumping all patients together for GWAS or other analysis since many patients with focal epilepsy have a known cause of their seizures. For instance, focal epilepsy can be caused by tumors, encephalitis, meningitis or traumatic brain injury. These symptomatic causes of focal epilepsy may be genetically distinct from the causes of cryptogenic focal cases and lumping the two together may be counterproductive. GWAS studies that focus on focal epilepsy patients must decide if symptomatic cases will be included in the analysis.

GWAS in Epilepsy

GWAS typically genotype 500,000 to 1 million DNA markers (single nucleotide polymorphisms, SNPs) in each individual and uses a case control design to search for DNA markers that associate with the phenotype under study. To date only two studies have been published reporting data from GWAS on cohorts of epilepsy patients. The first studied over 3000 patients from Northern Europe with focal epilepsy of mixed symptomatic and cryptogenic types [5]. The results found no genetic markers that reached genome wide significance levels ($p<10^{-8}$) though several markers reached suggestive p values ($p< 10^{-6}$-10^{-7}). The second GWAS studied over 1000 patients of Chinese ancestry with mixed cryptogenic and symptomatic focal epilepsy [6]. This study identified SNP markers in the CAMSAP1L1 gene that reached genome wide significance. This is a good candidate for an epilepsy gene as its protein product plays a role in cytoskeletal function, neuron outgrowth and synaptic plasticity. It is likely that false negative loci are represented in both of these data sets and it is very possible that loci that reached suggestive levels could be relevant.

At least two other GWAS studies are currently in progress, one contains only GGE patients from Europe (n=2000, EpiCure Consortium) and the other has both GGE (n=1000) and non-symptomatic focal epilepsy (n=1000) collected in the USA [7]. Preliminary data have been presented for the USA cohort and two loci reached genome wide significance levels, MYH11 and CNTN4, both good candidates for epilepsy genes based on the function of their encoded proteins [7]. Final results from the USA and EpiCure cohorts should be published in 2012. Several other groups have collected cohorts of epilepsy patients and typed SNP markers across the genome including cohorts in Italy, Australia, and the UK. The International League Against Epilepsy (ILAE) has formed an international consortium and a meta-analysis is in progress that will combine all epilepsy samples for a large GWAS. The consortium will "lump" over 10,000 cases of both focal and GGE patients to be compared with over 30,000 controls. The cohorts will then be stratified into GGE and focal groups for separate GWAS analysis and focal cases stratified further into symptomatic and cryptogenic cases. Thus, a lumping and splitting approach will be performed and results will be compared to determine if a broad or restricted phenotype is more successful.

This approach was taken in recent years by the multiple sclerosis (MS) research community [8]. Prior to the GWAS era, linkage and targeted association studies identified very few genetic factors associated with MS. However, when the MS community formed an international consortium and performed GWAS on ~10,000 cases and ~10,000 controls, many new loci were discovered. The work demonstrated ~50 separate genes that were associated with MS with convincing statistical evidence supporting their relevance to

disease [9]. A large number of these genes encode proteins that play a role in biological pathways controlling T cell function. Thus, these GWAS data confirmed a role for T-cell biology as fundamental to MS. Individual MS patients were stratified based on the subtype of MS they had, however, stratification did not identify any alleles at genome wide significance that were not already detected by the “lumping” method. Therefore, a broad phenotype was able to take advantage of high sample numbers and increased genetic power that lead to a major discovery that T-cell dysfunction is a fundamental problem in MS patients. Individual variation of MS course and progression is likely to be a result of DNA variations outside the T cell pathways identified.

In the future, GWAS and NGS data will identify DNA variations that are associated with common forms of human epilepsy and influence AED response. Identification of these DNA variations would open new avenues of research by discovering possible targets for intervention into the pathophysiology of common forms of epilepsy and for use to predict AED response.

References

[1]. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012; 13(90 (1)):7–24. 2012. [PubMed: 22243964]

[2]. De Jonghe P. Molecular genetics of Dravet syndrome. Dev Med Child Neurol. 2011; 53(Suppl 2): 7–10. [PubMed: 21504425]

[3]. Klassen T, Davis C, Goldman A, Burgess D, et al. Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. Cell. 2011; 24(145(7)):1036–48. [PubMed: 21703448]

[4]. EPICURE Consortium. Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. Epilepsia. 2012; 53(2):308–18. [PubMed: 22242659]

[5]. Kasperaviciūte D, Catarino CB, Heinzen EL, et al. Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. Brain. 2010; 133(Pt 7):2136–47. [PubMed: 20522523]

[6]. Guo Y, Baum LW, Sham PC, et al. Two-stage genome-wide association study identifies variants in CAMSAP1L1 as susceptibility loci for epilepsy in Chinese. Hum Mol Genet. 2011; 21(5): 1184–9. [PubMed: 22116939]

[7]. Buono, RJ.; Helbig, I.; Zhang, H., et al. Identification of epilepsy susceptibility loci by convergence of genome wide association and copy number variation data sets. Society for Neuroscience. 2010. http://www.abstractsonline.com/Plan/ViewAbstract.aspx?sKey=1dfb4638-336e-4544-938e-54c20ae2e4e8&cKey=73347a0d-6950-4b8d-a402-faab516525f7&mKey={E5D5C83F-CE2D-4D71-9DD6-FC7231E090FB}

[8]. Patsopoulos NA, Bayer Pharma MS Genetics Working Group; Steering Committees of Studies Evaluating IFNβ-1b and a CCR1-Antagonist. et al. Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Ann Neurol. Dec. 2011; 70(6):897–912. doi: 10.1002/ana.22609.

[9]. Baranzini SE, Nickles D. Genetics of multiple sclerosis: swimming in an ocean of data. Curr Opin Neurol. Jun; 2012 25(3):239–45. [PubMed: 22516854]

Highlights

- First published epilepsy GWAS data
- GWAS success in past five years
- Broad phenotypes versus restricted phenotypes
- GWAS and NGS as complementary techniques
- Current understanding of “causal” mutations