Published in final edited form as:

Lancet Neurol. 2008 November ; 7(11): 1067–1072. doi:10.1016/S1474-4422(08)70241-2.

Genome-wide association studies in neurological disorders

Javier Simón-Sánchez, MS and

Molecular Genetics Section, Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, Unidad de Genética Molecular, Departamento de Genómica y Proteómica, Instituto de Biomedicina de Valencia-CSIC, Valencia, Spain

Andrew Singleton, PhD

Molecular Genetics Section, Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, and Public Health Sciences and Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Abstract

Background—During the past decade, the genetic causes of monogenic forms of disease have been successfully defined; this work has helped the progression of basic scientific investigation into many disorders, and has helped to characterise several molecular biological processes. An important goal of genetic research is to extend this work and define genetic risk factor loci for complex disorders. The aim is for these data not only to offer further basic understanding of the disease process, but also to provide the opportunity to obtain genetic risk assessments that could be generalised to the public.

Recent developments—The development of resources such as the Human Genome Project and the International Human Haplotype Map Project, coupled with technological advances in ultra-high-throughput genotyping, have provided the basis for genome-wide association studies (GWAS). This approach has been successful for several complex disorders in a short time. Although GWAS are still a new method, these studies have been used for a small number of neurological disorders and, despite varied results for these conditions, GWAS can usefully show the power and limitations of this approach.

Where next?—GWAS have the potential to show and emphasise common genetic variability associated with disease. However, a challenge of this approach is that large sample series and considerable resources are required. One important consideration will be the interpretation of the results of GWAS in a clinically meaningful way and to discern the implications for all therapy areas, including neurological disorders; this challenge will require specialised skills and resources from both the medical and the scientific communities.

Genome-wide association studies—fad or fab?

The perceived objective of science is commonly thought to be susceptible to fashion, and novel techniques and biological events attract the attention of the scientific community. The most recent approach to receive such attention is genome-wide association studies (GWAS). The application of these studies to define risk variability was given the title “Breakthrough of the Year”,¹ and GWAS have been well documented not only in scientific journals but also in

Correspondence to: A Singleton, Molecular Genetics Section, Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, singleta@mail.nih.gov.

Contributors: Both authors equally contributed to the preparation of this Review.

Conflicts of interest: We have no conflicts of interest.

Lancet Neurol 2008; 7:1067-72

newspapers and on television. Therefore, is this excitement about GWAS just hyperbole or is there substance to this approach; what does this technology really deliver; and will findings from this approach ever matter in the context of clinical practice?

Before GWAS, the search for genetic variants that alter risk for disease was dominated by candidate gene-association studies, which are usually focused on genetic variants that alter the coding sequence of a gene. In general, positive associations would be extensively reported, whereas negative associations would normally not be published, unless the data refuted previous results. Genuine success required sufficient understanding of the disease process to enable selection of the correct gene, the right variants to type within the gene, and—most importantly—the presence of variability within the gene that could alter function or expression. Out of an assumed 30 000 genes and a couple of million common variants, the likelihood of success for any investigator who undertakes such an endeavour is low, and the net failure of more than a decade of such work is well known.

GWAS have been used to prevent the inherent bias and low throughput of single-gene experiments by offering association across the entire genome, including the intragenic and intergenic regions. This approach has been facilitated by two recent developments that build on previous efforts by scientific research, as well as our increased understanding of the genetic complexity that underlies diseases. First is the implementation of the International Haplotype Map Project—a study designed to catalogue and correlate, by means of calculating linkage disequilibrium, genetic variability that is specific to the study population.² This project uses single nucleotide polymorphism (SNP) genotypes to show genotype information about proximal genetic variants; thus, a large proportion of the information content in several million SNP variants can be shown with a smaller set of several hundred thousand informative tagging SNPs. Second is the development of reliable high-throughput genotyping approaches, which enable cost-effective typing of millions of SNP loci simultaneously. Both of these developments enable the application of SNP GWAS to help understand the genetic basis of complex diseases.

For more on the **International Haplotype Map Project** see <http://www.hapmap.org>

In general, new technologies are seldom encouraged by the whole scientific community, which commonly polarises into opposing groups. Therefore, that there are those who believe that GWAS are the answer to many scientific needs, and those who feel that these studies are fatally flawed endeavours that are draining research funding is not a surprise. An understanding of the arguments from both sides is important for the interpretation of GWAS; so too is an awareness of what a well-designed GWAS should find and—perhaps more importantly—what this type of study is not designed to find.

GWAS design, application, and problems

GWAS are designed for a single, specific purpose: to identify common genetic variability associated with a trait, usually a disease. The use of these studies to discover and type copy number polymorphisms has been reviewed elsewhere;^{3–5} indeed, GWAS can detect many types of variability, including copy number variants, repeat variants, and other SNP variants.

There are several potential problems with GWAS, mainly population stratification and population admixture, which are apparent when the case and control groups are not well-matched genetically or where there are several distinct, but unrecognised, sub-populations in a cohort. In addition, factors such as non-reported relatedness in the population might also be a problem.

Although these problems do have the potential to mask an association and produce false-positive or false-negative results, there are now sensitive analytical techniques to distinguish and account for such difficulties with the SNP data obtained from GWAS. The objectives of these strategies range from identifying population outliers or distantly related individuals and removing them from further analysis, to adjustments to association statistics on the basis of admixture measurements. The empirical evidence generated so far supports the idea that population effects are unlikely to mask all genetic associations; furthermore, these data suggest that genetic heterogeneity (different risk genes in different populations) and allelic heterogeneity (different risk alleles in the same gene in different populations) are not sufficient to mask associations when pooling white northern European and North American populations. However, more distantly related populations would need to be analysed separately.^{6–8}

The success of GWAS is dependent on many factors, particularly the frequency of risk alleles, sample size, and individual effect sizes. The size of each genetic effect within the sample being analysed and the number of samples analysed are primary limitations on any study; the smaller the risk associated with a common variant, the greater the number of samples required to identify this variant. Risk allele frequency is also important: the more common the variant, the more power available to detect association at that variant.⁹ Risk allele frequency is, of course, related to effect size: a risk variant that has a high odds ratio but is rare in the population is much more difficult to identify with GWAS than a common variant linked to moderate or mild risk within the population. The low frequency of a risk variant in the population causes little attributable risk—the smaller incidence that would be observed if the population was not exposed to the risk variant.

For more on the **Coriell repository** see <http://ccr.coriell.org/>

When these considerations are taken into account, the description of a GWAS as underpowered is frequently misused, and it is perhaps more appropriate to describe the size effect that a study is adequately powered to detect. Even studies with small sample sizes are adequately powered to see common genetic modifiers of disease that exert a large effect. The classic example of such a study was one of the first GWAS, where variability at the gene that encodes complement factor H (*CFH*) was identified as a risk factor for age-related macular degeneration (AMD) in only 96 cases and 50 controls.¹⁰ In the studied population, the risk allele was common and also exerted a substantial effect, which increased disorder risk by about four to seven times. However, this magnitude effect is an exception, and common genetic risk variants for most complex traits will confer a small risk (odds ratio of 1.1 to 2.0, with only a few (if any) loci conferring larger risk effect detectable in only a few hundred individuals. The members of the Wellcome Trust Case Control Consortium¹¹ suggest that GWAS in complex traits should include at least 2000 cases and 2000 controls.¹¹

One important caveat to the usefulness of GWAS is that the genetic coverage has a high statistical price: each GWAS relies on the application of millions of statistical tests, with most studies assessing models of genetic risk. This means that low p values are recorded as a result of chance, rather than biological association. Thus, although significance values of $p=1\times10^{-6}$ would ordinarily be statistically significant, an uncorrected p value of this order of magnitude is most likely to be a false-positive association in the context of a GWAS. Although there are several methods that apply test correction or generate more realistic test statistics, independent replication is the gold standard for detection of genuine association. Coverage of the genome is not absolute, and different platforms provide varying efficiencies for tagging untyped genomic variation in the genome; this coverage varies by population. Thus, although the platforms with the highest density to date provide 90% coverage of the SNPs from the International Haplotype Map Project in white populations, the coverage is lower in Asian and African populations.

Another important requirement for the success of GWAS is accurate phenotyping, which is related to the aetiological heterogeneity of disease. GWAS are not sensitive to low levels of misdiagnosis or unidentified subtypes of disease; therefore, patients who are mistakenly diagnosed with the disorder can be included in the GWAS. However, large numbers of patients who are misdiagnosed and have a non-specific genotype have the potential to distort associations. Solutions to this problem include deep endophenotyping as close to the molecular level as possible¹² and the collection of DNA from patients who have had a standardised diagnosis. A good example of such a repository is the Coriell Institute for Medical Research database (Camden, NJ, USA)—funded by the US National Institute of Neurological Disorders and Stroke—which banks DNA from patients with various diseases and from controls.

Neurological disorders

Although GWAS have started to be used only recently, several have been done in patients with neurological disorders, including periodic limb movements in sleep (PLMS),¹³ restless legs syndrome (RLS),¹⁴ amyotrophic lateral sclerosis (ALS),^{15–19} multiple sclerosis (MS),^{20,21} Parkinson's disease (PD),^{22,23} stroke,²⁴ and Alzheimer's disease²⁵ (table and panel). These studies have produced a range of results, including definitive association, a scarcity of detectable association, and claims of association that are subsequently refuted by other groups. Although the examination of each of these individually is beyond the scope of this Review, it is useful to comment on some aspects of these studies.

For a list of **GWAS** see <http://www.genome.gov/26525384>

In terms of identifying genes, the most successful of the current GWAS in neurological disorders were the two studies of patients with RLS and PLMS.^{13,14} Two groups independently identified a common genetic risk factor: variability at the locus that contains *BTBD9*, which encodes a protein of unknown function. Variability at *BTBD9* seems to increase the risk of RLS by 50%¹⁴ and also doubles the risk of PLMS. Independent replication of results in different populations is an important support for genuine association, and the report of the same disease locus by two independent groups is certainly convincing. These studies also benefited from large sample sizes, and at least one moderate-effect risk loci was apparent for this disorder. As well as confirming *BTBD9* as a genetic risk factor for RLS, the second group¹⁴ identified *MEIS1* (a gene involved in myeloid leukemia) and a region that contains the transcription factor *LBXCOR1* and the *MAP2K5* mitogen-activated protein kinase as risk factors for this syndrome, with effect sizes of 2.76 and 1.5, respectively. Winkelmann and co-workers¹⁴ estimated that the joint attributable risk fraction for these three loci (*BTBD9*, *MEIS1*, and *MAP2K5* and *LBXCOR1*) was about 70%. Statistical analysis by Vilarino-Guell and co-workers²⁶ confirmed the association with risk of PLMS with *BTBD9* and *MEIS1*, but the association with variants in the *MAP2K5* and *LBXCOR1* loci was insignificant, with odds ratios that ranged from 1.8 to 2.9 for *BTBD9*, and from 2.6 to 3.7 for *MEIS1*.²⁶ These studies also showed the effect of phenotyping on GWAS: the strength of association and size of effect were both different if the symptoms of PLMS and RLS were regarded separately or jointly.

For more on the **Database of Genotype and Phenotype** see <http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gap>

For more on **public data sharing** see Lancet Neurol 2006; 5: 895

Panel: Publicly shared results of GWAS

ALS^{15,16,19}

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000127.v1.p1

<http://www.alscentrum.nl/index.php?id=GWA>

Parkinson's disease^{22,23}

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000089.v1.p1

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000048.v1.p1

Stroke²⁴

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000102.v1.p1

Alzheimer's disease²⁵

http://www.tgen.org/research/neuro_gab2.cfm

Substantial associations have also been found for MS.²¹ In a multistage GWAS by The International Multiple Sclerosis Genetics Consortium, the data showed evidence that certain alleles of the genes that encode the α chain of the interleukin 2 and interleukin 7 receptors (*IL2RA* and *IL7RA*) were strongly associated with MS ($p=2.96 \times 10^{-8}$ and $p=2.94 \times 10^{-7}$, respectively) with odds ratios that ranged from 1.11 to 1.36 for *IL2RA*, and 1.11 to 1.26 for *IL7RA*.²¹ Because these risk ratios are modest, each of the alleles of *IL2RA* and *IL7RA* would cause less than 0.2% of the heritable influence on MS susceptibility. An extension of this study²⁷ identified three markers within these two genes to be highly associated with MS, with a modest risk for MS (odds ratios of about 1.2). These three markers were typed in a further 20708 individuals from different populations, of which 17 showed that the risk allele of these variants was over-represented.

The authors of studies with GWAS in patients with ALS^{15–19} have described two genetic loci that are associated with this disorder. Variability at *ITPR2*, which encodes the inositol 1,4,5-triphosphate type 2 receptor, was suggested as a risk factor for ALS by the authors of one study.¹⁸ Although this study included replication, the results derived from the Belgian population were not wholly consistent, and further investigation is required to establish that this as a genuine risk locus. Additionally, the authors of two studies described an association with ALS at *DPP6*, the gene that encodes dipeptidyl peptidase. Although these results might show replication, both studies used overlapping publicly available data and thus should not be thought of as independent studies.

Most GWAS in neurological disorders have been small in scope and only designed to detect moderate to large effect sizes. This was the case for the studies on PD, ALS, and stroke that were done by our group,^{15,16,22,24} which did not show any convincing disease-associated loci. Although we were aiming to identify risk loci, another objective of this work was to provide sets of publicly available genome-wide SNP data for others to work with and augment. These data have been downloaded several hundred times and used in several publications that have originated from outside our group. The release of data from GWAS to the scientific community has been much discussed, and some funding agencies now require data release. The US National Center for Biotechnology Information has developed the Database of Genotype and Phenotype, where this information can be deposited and accessed. This facilitates the construction of well-characterised phenotype and genotype data throughout populations; such a data repository enables the collection and augmentation of information by other researchers, and ensures maximum benefit from the substantial resources invested in GWAS.

Therefore, the results of GWAS can vary. The responsibility for the thorough but realistically conservative interpretation of these data lies with authors, journal editors, and reviewers; however, the pressure to publish positive associations is substantial, and the reader should be aware of this in articles that describe the results of GWAS. Independent and consistent

replication is a gold standard for genuine association and, although there is no minimum number of samples required for a convincing GWAS, the successful experiments in complex diseases have used thousands rather than hundreds of samples.^{6–8,11} This number is likely to increase substantially as investigators start to examine gene–gene interactions and gene–environment interactions within disease groups. Finally, failure to detect true association in a GWAS does not necessarily show true failure of the study; in addition to the potential reasons discussed above, no such common genetic risk factors might be apparent in the disease, and this is a possibility that should be considered. Such a negative finding does not preclude that other genetic or epigenetic influences in the disease are apparent, nor should such a result be used to support environmental influences. Although the absence of association is a compelling argument for other plausible aetiological hypotheses, such an argument is based on circumstantial evidence.

Too much information: what to do with it?

The “drinking from the fire hose” analogy by Hunter and Kraft²⁸ is one that suitably describes the excess of information associated with GWAS that medical professionals increasingly have to manage. The challenge is no longer to obtain data, but to determine what—if anything—to do with this information, and what the effect of these data might be for treatment and diagnosis. Loci that increase risk by 20–60% are not informative for disease prediction—particularly for rare disorders—and, in most cases, there is no preventative therapy that could be applied to a population deemed to be at risk for that disorder. The most immediate benefits of data from GWAS will be the understanding of the molecular aetiology of diseases, and the potential to identify at-risk individuals for clinical trials and epidemiological studies. In the long term, the use of genetic risk profiles could be implemented for individuals and, particularly, for predictions on different responses to therapy on the basis of genetic make up. Although there is little immediate effect of the results of GWAS on treatment, clinicians should be increasingly aware of this technology. The launch of several companies that specialise in direct-to-consumer, personalised genomics that offer genetic risk profiling means that there will be an increasing number of patients who present to their physician with concerns and questions that are specific to the results of such tests. Understanding the basis for these tests and the limitations of the information from such studies is already important, but understanding these test data will be more necessary as the genetic prediction of disease risk, prognosis, and treatment becomes more accurate.

Search strategy and selection criteria

References for this Rapid Review were identified through searches of PubMed by use of the search terms “genome-wide”, “whole-genome”, “association analysis”, “restless syndrome”, “periodic limb movement”, “amyotrophic lateral sclerosis”, “Parkinson's disease”, “stroke”, “Alzheimer's disease”, “multiple sclerosis”, and “age-related macular degeneration”. Only papers published in English were reviewed. The final reference list was generated on the basis of relevance to the topics covered in this Review.

Conclusions

Although this is still a new specialty, we predict that GWAS will identify common genetic risk variants for complex neurological diseases, where there are such variants. We also propose that future genomic technologies—such as whole genome resequencing and genome-wide measures of epigenetic variability and somatic variation—are likely to change medicine and alter our perception of genetic determinism.

In the long term, the data from GWAS and other genomic approaches will be clinically meaningful with respect to risk profile, disease prognosis, and response to treatment. This information will be extensive and complex, even in its most basic form. Although some interpretation of the results could be done in silico (eg, the automatic generation of risk profiles), clinicians will need to have a firm understanding of basic and complex genetic principles, to translate this complex information to their patients in a meaningful way.

Acknowledgments

This work was supported by the Intramural Research Program of the National Institute on Aging, the National Institutes of Health, and the Department of Health and Human Services, and relates to concepts developed under Intramural project Z01 AG000949-02.

References

1. Pennisi E. Breakthrough of the year: human genetic variation. *Science* 2007;318:1842–43. [PubMed: 18096770]
2. A second generation human haplotype map of over 3.1 million SNPs. *Nature*; The International HapMap Consortium; 2007. p. 851–61.
3. Blauw HM, Veldink JH, van Es MA, et al. Copy-number variation in sporadic amyotrophic lateral sclerosis: a genome-wide screen. *Lancet Neurol* 2008;7:319–26. [PubMed: 18313986]
4. Matarin M, Simon-Sanchez J, Fung HC, et al. Structural genomic variation in ischemic stroke. *Neurogenetics* 2008;9:101–08. [PubMed: 18288507]
5. Simon-Sanchez J, Scholz S, Matarin Mdel M, et al. Genomewide SNP assay reveals mutations underlying Parkinson disease. *Hum Mutat* 2008;29:315–22. [PubMed: 17994548]
6. Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007;316:1331–36. [PubMed: 17463246]
7. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007;316:1341–45. [PubMed: 17463248]
8. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007;316:1336–41. [PubMed: 17463249]
9. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005;6:109–18. [PubMed: 15716907]
10. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;308:385–89. [PubMed: 15761122]
11. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*; Wellcome Trust Case Control Consortium; 2007. p. 661–78.
12. Tan NC, Mulley JC, Scheffer IE. Genetic dissection of the common epilepsies. *Curr Opin Neurol* 2006;19:157–63. [PubMed: 16538090]
13. Stefansson H, Rye DB, Hicks A, et al. A genetic risk factor for periodic limb movements in sleep. *N Engl J Med* 2007;357:639–47. [PubMed: 17634447]
14. Winkelmann J, Schormair B, Lichtner P, et al. Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nat Genet* 2007;39:1000–06. [PubMed: 17637780]
15. Schymick JC, Scholz SW, Fung HC, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol* 2007;6:322–28. [PubMed: 17362836]
16. Cronin S, Berger S, Ding J, et al. A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum Mol Genet* 2008;17:768–74. [PubMed: 18057069]
17. Duncldey T, Huentelman MJ, Craig DW, et al. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N Engl J Med* 2007;357:775–88. [PubMed: 17671248]
18. van Es MA, Van Vught PW, Blauw HM, et al. *ITPR2* as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol* 2007;6:869–77. [PubMed: 17827064]

19. van Es MA, van Vught PW, Blauw HM, et al. Genetic variation in *DPP6* is associated with susceptibility to amyotrophic lateral sclerosis. *Nat Genet* 2008;40:29–31. [PubMed: 18084291]
20. Goris A, Williams-Gray CH, Foltynie T, Compston DA, Barker RA, Sawcer SJ. No evidence for association with Parkinson disease for 13 single-nucleotide polymorphisms identified by whole-genome association screening. *Am J Hum Genet* 2006;78:1088–90. [PubMed: 16685662]
21. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genome-wide study. *N Engl J Med* 2007;357:851–62. [PubMed: 17660530]
22. Fung HC, Scholz S, Matarin M, et al. Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol* 2006;5:911–16. [PubMed: 17052657]
23. Maraganore DM, de Andrade M, Lesnick TG, et al. High-resolution whole-genome association study of Parkinson disease. *Am J Hum Genet* 2005;77:685–93. [PubMed: 16252231]
24. Matarin M, Brown WM, Scholz S, et al. A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. *Lancet Neurol* 2007;6:414–20. [PubMed: 17434096]
25. Coon KD, Myers AJ, Craig DW, et al. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry* 2007;68:613–18. [PubMed: 17474819]
26. Vilarino-Guell C, Farrer MJ, Lin SC. A genetic risk factor for periodic limb movements in sleep. *N Engl J Med* 2008;358:425–27. [PubMed: 18216367]
27. Refining genetic associations in multiple sclerosis. *Lancet Neurol*; International Multiple Sclerosis Genetics Consortium (IMSGC); 2008. p. 567-69.
28. Hunter DJ, Kraft P. Drinking from the fire hose—statistical issues in genome-wide association studies. *N Engl J Med* 2007;357:436–39. [PubMed: 17634446]

Table 1
Summary of the study design and results obtained for published GWAS on neurological disorders and age-related macular degeneration

	Patients (controls) [*]	Ethnic origin (exploratory sample)	Method	Patients (controls) [†]	Ethnic origin (replication sample)	Gene	Chromosome	Odds ratio	p value	Attributable risk fraction
Restless legs syndrome and periodic limb movements										
Stefansson and co-workers ¹³	306 (15633)	Icelandic	Multistage	123 (1233)	Icelandic	<i>BTBD9</i>	6p21.2	1.8	4×10 ⁻⁴	50%
Restless legs syndrome				188 (662)	American	<i>BTBD9</i>	6p21.2	1.5	4×10 ⁻³	..
	401 (1644)	German	Multistage	903 (891)	German	<i>MEIS1</i>	2p14-p13	275	5.87×10 ⁻²⁰	68-6%
						<i>BTBD9</i>	6p21.2	1.66	1.6×10 ⁻⁶	
						<i>MAP2K5</i> and <i>LBXCOR1</i>	15q23	1.50	4.67×10 ⁻⁷	
ALS				255 (287)	French-Canadian	<i>MEIS1</i>	2p14-p13	2.36	8.51×10 ⁻⁷	74.2%
						<i>BTBD9</i>	6p21.2	1.66	4.14×10 ⁻³	
						<i>MAP2K5</i> and <i>LBXCOR1</i>	15q23	1.50	2.75×10 ⁻⁵	
				..	..	..	..	..	..	..
[‡] Schymick and co-workers ¹⁵	276 (271)	North American white	Preliminary	..	..	<i>DPP6</i>	7q36.2	1.37	2.53×10 ⁻⁶	37%
[‡] Cronin and co-workers ¹⁶	222 (217)	Irish	Joint analysis	276 (271)	North American white (Schymick and co-workers ¹⁵)					
				461 (450)	Dutch (A van Es and co-workers ¹⁸)		..	..	..	..
Dunckley and co-workers ¹⁷	386(542)	White	Multistage	766 (750)	White, non-white [§] , and unknown	<i>FGGY</i>	1p32.1	1.35	3×10 ⁻⁴	..
van Es and co-workers ¹⁸				135 (275)	North American white		..	..	..	..
	461 (450)	Dutch	Multistage	272 (336)	Dutch	<i>ITPR2</i>	12p11	1.85	0.0017	..
				291 (267)	Belgian		..	0.85	0.413	..
				313 (303)	Swedish		..	2.12	0.0007	..
[‡] van ES and co-workers ¹⁹	461 (450)	Dutch	Joint analysis	276 (271)	North American white (Schymick and co-workers ¹⁵)	<i>DPP6</i>	7q36.2	1.3	4.3×10 ⁻⁵	..
				272 (336)	Dutch		..	1.26	0.04	..
				467 (437)	Swedish		..	1.31	0.006	..
				291 (320)	Belgian		..	1.21	0.11	..

	Patients (controls)*	Ethnic origin (exploratory sample)	Method	Patients (controls) [†]	Ethnic origin (replication sample)	Gene	Chromosome	Odds ratio	p value	Attributable risk fraction
Parkinson's disease										
‡ Fung and co-workers ²²	267 (270)	North American white	Preliminary	..	..	..	..	..	..	..
‡ Maraganore and co-workers ²³	443 (300)	North American white	Multi-stage	332 (332)	North American white	<i>SEMA5A</i>	5p15.2	1.7	7.62×10 ⁻⁶	..
Stroke										
‡ Matarin and co-workers ²⁴	278 (275)	North American white	Preliminary	..	..	..	..	..	..	..
Alzheimer's disease										
‡ Coon and co-workers ²⁵	644 (422)	North American white	Single stage	..	..	<i>APOE</i>	19q13.2	4.01	5.3×10 ⁻³⁴	..
Multiple sclerosis										
International Multiple Sclerosis Genetics Consortium ²¹	931 trios [¶]	..	Multi-stage and joint analysis	2322 cases, 1540 trios (5418)	..	<i>IL2RA</i>	10p15-p14	1.19–1.25	2.16×10 ⁻⁷	0.2%
						<i>IL2RA</i>	10p15-p14	1.19–1.25	2.96×10 ⁻⁸	0.2%
						<i>IL7RA</i>	5p13	1.18	2.94×10 ⁻⁷	0.2%
						<i>HLA-DRA</i>	6p21.3	1.99	8.94×10 ⁻⁸¹	..
Age-related macular degeneration										
Klein and co-workers ¹⁰	96 (50)	White	..	..	..	<i>CFH</i>	1q31	4.6–7.4	4.1×10 ⁻⁸	45–80%
										1.4×10 ⁻⁶

* Exploratory sample size,
† Replication sample size.
‡ For further information, please refer to the publicly shared links listed in the panel.
§ Hispanic, Asian, American Indian, Pacific Islander, and black ethnic origins.
¶ An affected person and both of his or her parents. ALS=amyotrophic lateral sclerosis.