NEWS AND COMMENTARY

Amyotrophic Lateral Sclerosis

Genome-wide association studies in amyotrophic lateral sclerosis

Nicolas Dupré and Paul Valdmanis

European Journal of Human Genetics (2009) 17, 137–138;
doi:10.1038/ejhg.2008.205; published online 5 November 2008

On page 213 of this issue, Cronin *et al*¹ report their findings of a replication study of the top 27 associated single nucleotide polymorphisms (SNPs) identified when combining results from several whole genome association studies (WGASs) in amyotrophic lateral sclerosis (ALS). Although the results of the study are largely negative, the insight that is gained from their interpretation is timely as researchers are addressing the strengths and weaknesses of WGASs in ALS and which genes should be followed up from a functional biology perspective. Although the cause of this deadly disease is not yet fully understood, genetic studies have lead the way so far in identifying pathways that may increase its understanding, and the hope is that genes identified in WGASs will eventually help in this regard.

ALS, also known as Lou Gehrig's disease, is a motor neuron disease characterized by rapidly progressive paralysis leading to death due to respiratory failure, typically within 3–5 years of symptom onset. It is the most common adult onset motor neuron disease with an incidence of 2.1 per 100 000 person-years.² Familial ALS is responsible for 1–5% of ALS cases as a whole depending on the population studied, and it is caused by highly penetrant genes that are most often inherited in an autosomal dominant manner.³ The greatest contribution toward an understanding of ALS thus far has come from the discovery of mutations in the superoxide dismutase 1 (*SOD1*) gene on chromosome 21q22.11, which accounts for 10–20% of autosomal dominant familial ALS cases. Other than linkage analysis, which has allowed the mapping of *SOD1* in familial

ALS, other approaches have been used to identify genes that are involved in non-familial or 'sporadic' ALS. These include mainly candidate gene association studies and WGASs.

Candidate gene association studies have been performed to identify genes based on *a priori* hypotheses.³ The genes identified pertain to broad categories such as hypoxia and oxidative stress (vascular endothelial growth factor, paraoxonase); cytoskeletal structure (neurofilament heavy chain subunit, dynactin); motor neuron survival (survival motor neuron, ciliary neurotrophic factor, leukemia-inhibitory factor) and neurodegenerative disorders (hemochromatosis, apolipoprotein E4 allele, *TARDBP*).

As genotyping on a large scale has become affordable, WGASs have enabled the hunt for genes involved in complex diseases without *a priori* hypotheses about the function of the gene.⁴ WGASs have been successful in identifying genetic factors that underlie common diseases such as diabetes⁵ and breast cancer.⁶ The variants identified in these studies, however, have odds ratios not much above 1, which means that their discovery necessitates the compilation of very large sample collections numbering in the thousands or even tens of thousands.

The first reported WGAS in ALS was performed in a small cohort of 276 American sporadic cases and 275 neurologically normal American control samples.⁷ It identified 34 SNPs that were significantly associated with increased risk of developing ALS, though none of these SNPs exceeded the Bonferroni threshold for multiple testing. Another group⁸ shortly after performed a WGAS using

766 955 SNPs in 386 white patients with sporadic ALS and 542 neurologically normal white controls, while replicating findings in two stages: one of 744 cases and 750 controls, and a follow-up of 135 cases and 275 controls. Their most significant association with disease was found for an SNP near an uncharacterized gene known as *FLJ10986*, which was expressed in the spinal cord and cerebrospinal fluid of patients and controls. Another group, from the Netherlands, performed a WGAS in sporadic ALS using an initial sample of 461 cases and 450 controls, while replicating their findings in an independent sample of 876 cases and 906 controls from the Netherlands, Belgium and Sweden; they found *ITPR2* to be significantly associated with ALS.⁹ This same group then extended their initial analysis to include additional samples and found that variants and haplotype blocks surrounding the *DPP6* gene were strongly associated with ALS susceptibility in 1767 cases and 1916 controls of European ancestry.¹⁰

In their initial WGAS of sporadic ALS, Cronin *et al*¹¹ used a more homogeneous population by conducting their study in 221 cases and 211 controls from Ireland, identifying 35 potentially associated loci. They then used for replication a joint analysis of genome-wide data from the publicly available Dutch and US datasets⁷ totaling 958 ALS cases and 932 controls. Their strongest association was also a variant in the gene encoding *DPP6*, a component of type A neuronal transmembrane potassium channels. Cronin *et al*¹ present, in the current issue of the *European Journal of Human Genetics*, a replication study of their initial WGA scan. They examine a total of 27 SNPs that had the same allelic directionality and approached nominal significance in the previous US, Dutch and Irish WGA sample sets. Their replication population was a slightly expanded cohort of 91 patients and 48 controls from Ireland and an independent set of 218 patients and 356 controls from Poland. Their combined data analysis of 1267 cases and 1336 control subjects did not identify any SNPs reaching Bonferroni-corrected significance for association.

The most salient finding from the present study of Cronin *et al*¹ is that the *DPP6*

allele that was over-represented in the initial WGAS is now under-represented in the Polish cohort. This swap of the disease-associated allele is often a strong warning sign when interpreting the WGAS results. We can thus conclude with some certainty that the SNP itself that was examined (rs10260404) does not cause ALS. However, as was observed in the Dutch *DPP6* study,¹⁰ a full linkage disequilibrium block incorporating several adjacent SNPs was associated, such that an SNP nearby and perhaps still within the *DPP6* genomic interval may have an effect. This is critical as the identification of this causal variant is the ultimate goal of this genetic study. Alternatively, as the authors point out, lack of replication can suggest that the initial result was a false positive, or that population structure may be at play. However, the main argument against the latter is that the initial studies were performed in quite diverse populations already (US, Ireland, the Netherlands, Sweden and Belgium). How specific can the Polish population be?

At the present stage, one has to be very careful in interpreting results originating from WGASs, since these studies have limited resolution owing to small sample sizes. The reporting of negative replication results is important in analyzing and interpreting the significance of candidate genes that emerge as associated WGASs. Public availability of WGAS genotype datasets is invaluable in increasing the sample size

and thus the power to detect SNPs, conferring a relatively small risk. However, this has resulted in tremendous overlap of datasets from separate supposedly independent publications in which a substantial portion of the signal comes from data already present in other studies. What is lacking therefore is truly independent validation from additional populations. While we await these results, replication of the top associated SNPs in a separate population is nonetheless relevant ■

N Dupré is at the Department of Neurological Sciences, Faculty of Medicine, Laval University, CHA (Enfant-Jésus), 1401, 18th Street, Quebec City, QC, Canada G1J 1Z4.

Tel: +1 418 649 0252 ext. 3177;

Fax: +1 418 649 5915;

E-mail: nicolas.dupre.cha@ssss.gouv.qc.ca

P Valdmánis is at the CHUM Research

Center and Department of Medicine,

Center of Excellence in Neuromics of the

University of Montréal, University of

Montréal, QC, Canada

References

- 1 Cronin S, Tomik B, Bradley DG *et al*: Screening for replication of genome-wide SNP associations in sporadic ALS. *Eur J Hum Genet* 2008; **17**: 213–218.
- 2 Beghi E, Logroscino G, Chiò A *et al*: The epidemiology of ALS and the role of population-based registries. *Biochim Biophys Acta* 2006; **1762**: 1150–1157.
- 3 Valdmánis PN, Rouleau GA: Genetics of familial amyotrophic lateral sclerosis. *Neurology* 2008; **70**: 144–152.
- 4 Schymick JC, Talbot K, Traynor BJ: Genetics of sporadic amyotrophic lateral sclerosis. *Hum Mol Genet* 2007; **16** (Spec no. 2): R233–R242.
- 5 Sladek R, Rocheleau G, Rung J *et al*: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881–885.
- 6 Easton DF, Pooley KA, Dunning AM *et al*: Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007; **447**: 1087–1093.
- 7 Schymick JC, Scholz SW, Fung H-C *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**: 322–328.
- 8 Dunckley T, Huentelman MJ, Craig DW *et al*: Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N Engl J Med* 2007; **357**: 775–788.
- 9 van Es MA, van Vught PWJ, Blauw HM *et al*: ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol* 2007; **6**: 869–877.
- 10 van Es MA, van Vught PWJ, Blauw HM *et al*: Genetic variation in *DPP6* is associated with susceptibility to amyotrophic lateral sclerosis. *Nat Genet* 2008; **40**: 29–31.
- 11 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**: 768–774.