

ARTICLE

Screening for replication of genome-wide SNP associations in sporadic ALS

Simon Cronin^{*,1,2}, Barbara Tomik³, Daniel G Bradley⁴, Agnieszka Slowik³ and Orla Hardiman^{2,5}

¹Department of Clinical Neurological Sciences, Royal College of Surgeons in Ireland, Dublin 2, Ireland; ²Department of Neurology, Beaumont Hospital, Dublin 9, Ireland; ³Department of Neurology, Jagiellonian University, Krakow, Poland; ⁴Smurfit Institute of Genetics, Trinity College, Dublin 2, Ireland; ⁵Trinity College Institute of Neuroscience, Trinity College, Dublin 2, Ireland

We recently reported a joint analysis of genome-wide association (GWA) data on 958 sporadic amyotrophic lateral sclerosis (ALS) cases and 932 controls from Ireland and the publicly available data sets from the United States and the Netherlands. The strongest pooled association was rs10260404 in the *dipeptidyl-peptidase 6 (DPP6)* gene. Here, we sought confirmation of joint analysis signals in both an expanded Irish and a Polish ALS cohort. Among 287 522 autosomal single-nucleotide polymorphisms (SNPs), 27 were commonly associated on joint analysis of the Irish, US and Dutch GWAs. These 27 SNPs were genotyped in an expanded Irish cohort (312 patients with SALS; 259 controls) and an additional Polish cohort (218 patients; 356 controls). Eleven SNPs, including rs10260404, reached a final *P*-value below 0.05 in the Irish cohort. In the Polish cohort, only one SNP, rs6299711, showed nominal association with ALS. Pooling of data for 1267 patients with ALS and 1336 control subjects did not identify any association reaching Bonferroni significance ($P < 1.74 \times 10^{-7}$). The present strategy did not reveal any consistently associated SNP across four populations. The result for *DPP6* is surprising, as it has been replicated elsewhere. We discuss the possible interpretations and implications of these findings for future ALS GWA studies both within and between populations.

European Journal of Human Genetics (2009) 17, 213–218; doi:10.1038/ejhg.2008.194; published online 5 November 2008

Keywords: amyotrophic lateral sclerosis; genome-wide study; DPP6; genetics

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative condition that leads to paralysis and death often within 5 years of symptom onset.¹ Following the discovery in 1993 that mutations of the *SOD1* gene account for 10% of familial ALS,² there has been an intensive effort to under-

stand the genetics of both familial and sporadic ALS (SALS). Significant progress has been made in familial ALS, with mutations identified in several genes through linkage studies.³ However, 90% of ALS cases occur as apparently sporadic illness, and here progress using candidate gene paradigms has been limited by lack of replication between populations.⁴

The recent advent of robust high-density genotyping arrays has promised a new era for the identification of genetic risk factors for SALS. The genome-wide association (GWA) approach has the advantage over candidate gene studies of being unbiased by *a priori* hypotheses about

*Correspondence: Dr S Cronin, The Irish ALS Research Group, Department of Neurology, Beaumont Hospital, Dublin 9, Ireland.
Tel: +353 1 8092174; Fax: +353 1 8092302;

E-mail: scronin@rcsi.ie

Received 21 May 2008; revised 18 August 2008; accepted 12 September 2008; published online 5 November 2008

disease biology. GWA has been successful in identifying a single gene of major effect in macular degeneration, using a sample size of less than 100 affected individuals. However, in conditions in which the phenotype may be variable, and where there may be multiple susceptibility genes of small effect, sample size is of critical importance.⁵ The incidence of ALS, which is ~2 per 100 000 person-years, limits the availability of large numbers of patients for study.⁶ However, the past year has seen the publication of four GWA studies for SALS, each with a different list of most associated single-nucleotide polymorphisms (SNPs).^{7–10} These studies were designed such that each has the power only to identify variants of large effect size. As it is now clear that no single locus drives genetic risk for SALS, the contribution of common genetic variability to the phenotype must rest upon the contribution of many genes, each conferring a relatively modest increase in risk. To identify such risk factors, investigators have turned to mining the individual GWA data sets for commonly associated SNPs.^{10,11}

Using such an approach, van Es *et al*¹¹ in the Netherlands identified 15 SNPs commonly associated with SALS in the US and Dutch GWA data sets, and investigated replication of these SNPs in three additional SALS control populations from Sweden, Belgium and the Netherlands.¹¹ They were the first to identify a common association of rs10260404, an intronic variant in the gene encoding *dipeptidyl-peptidase 6 (DPP6)*, among each of their five study populations.¹¹ Using a similar methodology, we recently reported a joint analysis of the Irish, US and Dutch GWAs.¹⁰ Interestingly, the same variant was the top hit in our analysis, confirming and extending the supportive data for *DPP6* to the Irish.

Here, we present a follow-up on our joint analysis signals in an expanded Irish population and an independent Polish population. Using a list of SNPs that are commonly associated in three GWAs, we test for confirmation of association with ALS by sample addition.

Materials and methods

Study design and aims

In phase I of the study, we sought to identify SNPs commonly associated with SALS in the three publicly available genome-wide data sets from Ireland, the United States and the Netherlands.^{8–10} In phase II, we genotyped this list of SNPs in 913 additional samples from the Irish and Polish populations. We hypothesized that a truly associated SNP would attain a genome-wide level of significance in the expanded data set (phase I and II).

Participants

Table 1 shows the demography of the three publicly available GWA series used in phase I. The Irish series included 432 unrelated participants of self-declared Irish

ethnicity for at least three generations (221 ALS cases and 211 neurologically normal controls).¹⁰ Genotyping was performed using Illumina HumanHap 550K SNP chips (Illumina Inc., San Diego, CA, USA) and raw sample-level genotyping data have been made available at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000127.v1.p1. The US series comprised 276 unrelated, white, non-Hispanic individuals diagnosed with ALS and 271 neurologically normal individuals of similar ethnicity drawn from across the United States.⁸ Genotyping was undertaken using Illumina HumanHap 550K SNP chips and full phenotype and raw genotyping data are available at <https://queue.coriell.org/Q/index.asp>. The Netherlands series included 911 unrelated individuals with all four grandparents born in the Netherlands (461 ALS cases and 450 neurologically normal controls). This cohort was genotyped using the Illumina HumanHap 300K platform.⁹ Minor allele frequencies (MAFs) and allelic *P*-values for each genotyped SNP have been made available at <http://www.alscentrum.nl/index.php?id=GWA>. Stringent quality control criteria have been applied to these data sets as previously described.^{8–10} In particular, samples with genotyping call rates below 95% and cryptically related individuals have been removed.

The demography of the phase II study participants is also given in Table 1. All Irish DNA samples were collected at Beaumont Hospital in Dublin, Ireland, and the Polish DNA samples were collected at the Department of Neurology, MND Clinic, Jagiellonian University, Krakow, Poland. All patients fulfilled the 1994 El Escorial criteria¹² for probable or definite SALS and were phenotyped by physicians with expertise in ALS. Patients with a family history of ALS have been excluded from the study. Control DNA samples were collected from healthy unrelated neurologically normal individuals, either spouses of ALS patients or those accompanying non-ALS patients. Controls were matched for age, gender and ethnicity. All participants gave written informed consent and local ethics review boards approved all procedures.

Phase I: selection of commonly associated SNPs from the GWAs

The power of the individual Irish, US and Dutch genome-wide ALS association studies is low to detect loci with moderate effect sizes. Nonetheless, the signal from truly associated SNPs might be present, although weak. The size of the Dutch cohort is approximately double that of each of the Irish and US cohorts. Thus, to select markers for replication, we identified those SNPs commonly associated for the same allele at an allelic *P*-value below 0.05 in the Dutch GWA study and at an allelic *P*-value below 0.1 in each of the Irish and US GWA studies. SNPs with MAFs below 0.01 or deviations from Hardy–Weinberg equilibrium (HWE) in controls below 0.01 were excluded. The selected SNPs were then genotyped in phase II.

Table 1 Characteristics of the study populations

	Total	Male (%)	Female (%)	Age at onset (years, mean)	Spinal onset (%)	Bulbar onset (%)
<i>Phase I – SNP selection populations</i>						
Ireland 550K GWA						
Patients with sporadic ALS	221	54	46	61	72	28
Controls	211	53	47	58	—	—
US 550K GWA						
Patients with sporadic ALS	276	63	37	55	76	24
Controls	271	48	52	68	—	—
The Netherlands 300K GWA						
Patients with sporadic ALS	461	59	41	59	69	31
Controls	450	59	41	60	—	—
<i>Phase II – replication populations</i>						
Ireland (additional)						
Patients with sporadic ALS	91	46	54	59	66	34
Controls	48	51	49	56	—	—
Poland						
Patients with sporadic ALS	218	52	48	56	70	30
Controls	356	49	51	58	—	—

GWAS, genome-wide association study.

Phase II genotyping

Genotyping of the selected SNPs was performed in 139 additional Irish participants (91 patients with SALS and 48 control subjects) and 574 Polish participants (218 patients with SALS and 356 controls). DNA was extracted from fresh venous blood using standard procedures. Genotyping was performed by KBiosciences (Herts, UK) using KASPar chemistries. Quality control criteria included mixing of case and control samples on individual plates, inclusion of water controls as negatives, sample duplicates between plates and blinding of data. The mean SNP call rate for the Irish study population was 98.6% and the lowest call rate was 96.1%. The mean SNP call rate for the Polish study population was 97.9% and the lowest call rate was 94.1%. To facilitate confirmation of strand and direction between the Illumina and KASPar genotyping assays, we also genotyped 27 SNPs in 10 replicate individuals from the Irish genome scan using the KASPar platform. The genotyping concordance rate between the Illumina and KASPar assays was 99.3%, with the non-concordant calls being accounted for by two failed assays on the KASPar platform.

Statistical analysis

Each SNP was tested for allelic association with SALS by the χ^2 test of independence. Estimations of departures from HWE were calculated by the χ^2 test in controls. Analyses were computed using PLINK v 1.01 software.¹³ Power statistics were calculated assuming a MAF of 0.4 for SNPs conferring an odds ratio of 1.37.¹⁴ For phase I, the power of the Irish and US GWA data sets to detect such SNPs is 74 and 83%, assuming an α of 0.1, whereas the Dutch GWA

has 90% power to detect the same SNPs assuming an α of 0.05.

In phase II, we report uncorrected allelic *P*-values for the Irish ($n = 139$) and Polish ($n = 574$) replication samples. We have not applied Bonferroni correction for the 27 SNPs tested in the replication populations as only nominal trends are likely to be detected within individual populations at the sample sizes we have used. Therefore, meaningful conclusions can only be drawn from strengthening of association across the study by addition of the replication samples to the GWA data. Pooled genotype data for the SNPs common to all three GWAs (phase I) have only 34% power to reach genome-wide Bonferroni threshold (*P*-value below 1.73×10^{-7} accounting for 287 522 tests); by contrast, the final study data (phase I and II) has 66% power to reach this level of significance. Importantly, *a priori*, we set genome-wide Bonferroni significance in the combined phase I and II data as the threshold to declare significant association of any SNP with ALS.

Results

In phase I, we compared the Irish (221 SALS cases; 211 controls), US (276 SALS cases; 271 controls) and Dutch (461 SALS cases; 450 controls) genome-wide data sets.^{8–10} There were 287 522 autosomal SNPs in common between these studies that were successfully genotyped in all the three cohorts. Among these, there were 27 SNPs associated for the same allele at a *P*-value below 0.1 in the Irish and US and below 0.05 in the Dutch cohort (Table 2 and Supplementary Table 1). As previously reported,¹⁰ the

Table 2 Minor allele frequencies and *P*-values for the 27 selected SNPs in the pooled Irish, US and Dutch GWA studies (phase I) and following addition of the replication Irish and Polish samples to the GWA data (phase I and II)

Chromosome	SNP	Gene ^a	Risk allele	Phase I (GWAs)				Phase I and II (combined GWAs and additional)			
				ALS MAF (n 958)	Control MAF (n 932)	<i>P</i> -value	OR (95% CI)	ALS MAF (n 1267)	Control MAF (n 1336)	<i>P</i> -value	OR (95% CI)
18	rs2159942	ZNF519 (42 kb)	G	0.27	0.22	3.96×10^{-5}	1.37 (1.18–1.59)	0.26	0.21	1.35×10^{-5b}	1.33 (1.17–1.51)
4	rs10024717	PPM1K (14 kb)	G	0.34	0.41	4.86×10^{-5}	1.31 (1.15–1.5)	0.34	0.40	2.06×10^{-5b}	1.28 (1.14–1.43)
6	rs9351470	Intergenic	G	0.23	0.18	9.02×10^{-5}	1.37 (1.17–1.61)	0.22	0.17	1.27×10^{-4}	1.31 (1.14–1.5)
18	rs1942239	GALNT1 (92 kb)	A	0.38	0.34	0.002	1.23 (1.08–1.41)	0.38	0.33	1.48×10^{-4b}	1.25 (1.11–1.4)
6	rs6922711	ARHGAP18	C	0.47	0.42	0.001	1.23 (1.08–1.4)	0.47	0.42	1.83×10^{-4b}	1.23 (1.1–1.38)
11	rs2405657	Intergenic	G	0.29	0.35	4.5×10^{-5}	1.33 (1.16–1.52)	0.30	0.35	2.07×10^{-4}	1.25 (1.11–1.4)
7	rs10260404	DPP6	C	0.43	0.35	2.53×10^{-6}	1.37 (1.2–1.56)	0.41	0.36	2.62×10^{-4}	1.23 (1.1–1.38)
3	rs9289100	Intergenic	G	0.08	0.05	2.81×10^{-4}	1.65 (1.26–2.16)	0.07	0.05	9.66×10^{-4}	1.47 (1.17–1.86)
4	rs12233843	MIST	G	0.42	0.48	2.54×10^{-4}	1.27 (1.12–1.44)	0.43	0.48	0.0010	1.20 (1.08–1.34)
3	rs6779441	Intergenic	G	0.08	0.05	1.34×10^{-4}	1.69 (1.29–2.22)	0.07	0.05	0.0010	1.46 (1.17–1.84)
21	rs2834908	Intergenic	A	0.47	0.41	8.08×10^{-4}	1.25 (1.1–1.42)	0.46	0.42	0.0010	1.20 (1.08–1.34)
13	rs2478046	LOC80183	T	0.17	0.13	4.38×10^{-4}	1.38 (1.15–1.66)	0.16	0.13	0.0012	1.29 (1.11–1.5)
16	rs17823157	LOC92017 (55 kb)	T	0.29	0.22	1.42×10^{-5}	1.38 (1.2–1.6)	0.28	0.24	0.0017	1.22 (1.08–1.38)
3	rs2279812	Intergenic	A	0.15	0.19	9.17×10^{-5}	1.41 (1.18–1.67)	0.15	0.18	0.0018	1.27 (1.09–1.47)
8	rs558889	ANK1	A	0.37	0.32	1.71×10^{-4}	1.29 (1.13–1.48)	0.37	0.33	0.0023	1.19 (1.07–1.34)
10	rs7899260	UPF2	A	0.08	0.12	4.04×10^{-5}	1.58 (1.27–1.97)	0.08	0.10	0.0027	1.34 (1.11–1.62)
20	rs458032	LOC140733	G	0.35	0.30	2.37×10^{-4}	1.29 (1.13–1.48)	0.34	0.31	0.0028	1.19 (1.06–1.34)
11	rs1981410	NELL1	C	0.38	0.43	6.94×10^{-4}	1.25 (1.1–1.43)	0.38	0.42	0.003	1.18 (1.06–1.32)
4	rs6851312	FBXO8	A	0.34	0.40	7.34×10^{-5}	1.31 (1.14–1.49)	0.34	0.38	0.0032	1.19 (1.06–1.33)
3	rs16833234	NLGN1	C	0.15	0.19	6.89×10^{-4}	1.34 (1.13–1.59)	0.15	0.18	0.0061	1.23 (1.06–1.42)
1	rs1289830	MAN1A2 (65 kb)	C	0.25	0.20	1.15×10^{-4}	1.35 (1.16–1.58)	0.24	0.21	0.012	1.18 (1.04–1.35)
5	rs7705483	Intergenic	A	0.52	0.46	6.99×10^{-4}	1.25 (1.1–1.42)	0.51	0.47	0.017	1.14 (1.02–1.27)
7	rs10954418	EXOC4	C	0.35	0.30	5.07×10^{-4}	1.27 (1.11–1.46)	0.33	0.30	0.018	1.15 (1.02–1.29)
4	rs2251316	HPGD (15 kb)	G	0.28	0.34	1.02×10^{-4}	1.31 (1.14–1.51)	0.29	0.32	0.023	1.15 (1.02–1.29)
8	rs10503355	Intergenic	T	0.04	0.02	4.36×10^{-4}	2.11 (1.38–3.23)	0.03	0.02	0.031	1.45 (1.03–2.04)
10	rs1259590	LOC439985 (70 kb)	T	0.38	0.43	0.001	1.24 (1.09–1.41)	0.38	0.41	0.053	1.12 (1–1.25)
14	rs7156012	Intergenic	G	0.46	0.52	1.44×10^{-4}	1.28 (1.13–1.46)	0.47	0.50	0.08	1.10 (0.99–1.23)

ANK1, ankyrin 1; ARHGAP18, Rho GTPase-activating protein 18; DPP6, dipeptidyl-peptidase 6; EXOC4, SEC8 protein; FBXO8, F-box only protein 8; GALNT1, polypeptide *N*-acetylgalactosaminyltransferase 1; HPGD, hydroxyprostaglandin dehydrogenase 15-NAD; HWI, Hardy–Weinberg equilibrium *P*-values in controls; LOC, hypothetical protein locus; MAF, minor allele frequency; MAN1A2, mannosidase; MIST, mast cell immunoreceptor signal transducer; NELL1, nell-like precursor 1; NLGN1, neuroligin 1; PPM1K, protein phosphatase 1K; SNP, single-nucleotide polymorphism; UPF2, UPF regulator of nonsense transcripts; ZNF519, zinc-finger protein 519.

^aThe nearest gene with distance in parentheses.

^bStrengthening of allelic association in the combined phase I and phase II data vs phase I alone.

strongest pooled association was for an intronic variant, rs10260404, in the *DPP6* gene.

In phase II, the 27 selected SNPs were genotyped in the additional Irish (91 cases and 48 controls) and Polish (218 cases and 356 controls) populations. Eleven of these SNPs, including rs10260404, reached a final uncorrected allelic *P*-value below 0.05 in the expanded Irish data set (Supplementary Table 2). None of these also showed evidence of association in the Polish ALS population. Indeed, only rs6922711, an intronic marker in the gene encoding *Rho GTPase-activating protein 18 (ARHGAP18)*, reached a final uncorrected allelic *P*-value below 0.05 in the Polish sample (Supplementary Table 2).

The most important results are the effects of combining phase I and phase II. On analysis of combined genotyping data, for a total of 1267 patients with ALS and 1336 controls, only 4 of the 27 SNPs (rs6922711, rs1942239, rs10024717 and rs2159942) gained in strength of allelic association compared with the phase I GWA data alone (Table 2). No SNP reached a final level of genome-wide significance in the combined analysis.

Discussion

Several lines of evidence support the hypothesis that genetic susceptibility factors contribute to ALS pathogenesis. Epidemiological studies have established that between 2 and 5% of ALS cases are familial, most often with an autosomal dominant pattern of inheritance.³ Linkage studies have identified causal variants in a portion of these families, establishing that the phenotype may develop as a monogenic process.^{2,3} The expression of mutant SOD1 in mouse models produces an ALS phenotype.¹⁵ Finally, twin concordance studies indicate that the heritability of SALS is between 0.38 and 0.85.¹⁶ These observations have led to the suggestion that SALS is a complex oligogenic disease, requiring the interaction of several genes in a single patient, each contributing a portion of the underlying risk.¹⁷

We hypothesized that common SNPs contributing such a modest effect would be detected in the recent GWA studies of SALS. Hoping to achieve genome-wide significance for 27 SNPs, we genotyped additional Irish and Polish individuals. One SNP, rs6922711 in *ARHGAP18*, showed nominal association in the Polish cohort and gained in overall strength of association. However, no SNP showed a convincing pattern of association across all individual study populations, nor did any SNP achieve Bonferroni significance overall.

Perhaps the most unanticipated finding from our study is the lack of replication of the rs10260404 variant in the *DPP6* gene. A compelling replication at this locus has been demonstrated in five previous populations of Northern European ancestry¹¹ and by us in the Irish GWA.¹⁰ Importantly, and for the first time in SALS genetics, the

same risk allele of the same SNP was seen in each of the populations. However, in this paper, the non-risk allele was markedly over-represented among ALS cases in a Polish population. Although we concede that sampling error and chance may account for these findings, the degree of difference in allele frequencies in the Polish groups, despite a relatively uniform pattern in six other European case-control data sets, leads us to hypothesize a mechanism for population-specific difference that could account for the *DPP6* findings.

If motor neuron degeneration is regarded as one pole of a continuum of normal variation, then it may be that a constellation of common SNPs, when inherited together, contributes to the onset of ALS. In this instance, it seems likely that the risk polymorphisms will be common among populations of European ancestry.⁵ By contrast, SALS could equally result from rare mutations, whether single point, copy number or epigenetic, with variable penetrance or *de novo* change accounting for the non-familial pattern of incidence.^{3,4} Unlike SNPs, which have relatively uniform frequency between Caucasian populations,¹⁸ mutations may segregate within major ethnic pockets. For example, the frequency of *LRRK2* gene mutations in several series of patients with sporadic Parkinson disease is 0.6–2% in the United States, 4–8% in Spain and Portugal, 10–18% in the Ashkenazi Jewish population and over 40% in Arab populations.^{19–23} In a sporadic mutation model of ALS, common SNPs may still modulate the penetrance or phenotype of a particular mutation through epistasis.²⁴ Such SNP-mutation epistasis has been demonstrated in a recent GWA of unselected breast cancer cases (over 4300 case-control pairs) that detected that SNPs in the genes *FGFR2* and *MAP3K1* increase breast cancer susceptibility.²⁵ A follow-up study segregating phenotypes by *BRCA* mutation indicated that these SNPs acted exclusively in *BRCA2* mutation carriers and were not risk factors for *BRCA1*-related breast cancer.²⁵ SNP-mutation interactions mean that the findings of a case-control association study of SNPs in a given population will vary depending on the underlying prevalent mutations in that population.^{25,26} This may account for the failure of replication in some geographical distributions and poses challenges for the GWA approach to identify risk variants for SALS.

An alternative view is that SNPs should associate equally with disease risk across populations and that the *DPP6* association in Northern Europeans is false positive. Indeed, the obvious limitation of our study is the relatively small sample size of the GWA data sets used to identify markers for phase II, although at least for the *DPP6* variant, supportive data are available from additional populations. Even with the combined phase I and II data, power remains low to detect the oligogenic influence of SNPs. It is also important to bear in mind that the primary aim of our analysis was to add additional sample size to the GWA data. Consequently, nominal *P*-values reported for individual

populations must be interpreted with caution. Given the modest effects now expected for SNPs in SALS, the next step should be the expansion of sample size.

In summary, we tested whether commonly associated SNPs from three recent SALS genome-wide studies would attain greater significance when tested in an expanded cohort. Our findings caution against drawing conclusions, positive or negative, from the presently available GWA data sets. Further whole-genome genotyping both within and between populations will be necessary to fully define the contribution of SNPs to SALS.

Acknowledgements

We are sincerely grateful to the ALS patients, their families and referring neurologists. The original genotyping for the NIH study used data from the SNP Resource at the NINDS Human Genetics Resource Center DNA and Cell Line Repository (<http://ccr.coriell.org/ninds/>). The original genotyping was performed in the laboratory of Drs Traynor, Singleton and Hardy (NIA, LNG), Bethesda, MD, USA. Access to the samples and to these data is available from the Coriell website. We thank LNG also for assistance in genotyping the Irish GWA samples at NIH. The original genotyping for the Dutch ALS GWA was conducted at the Laboratory of Dr van den Berg, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, the Netherlands. This study was funded by the Muscular Dystrophy Association (US), by the Irish Institute of Clinical Neurosciences Travel Award and by the Irish Motor Neurone Disease Research Foundation.

Conflict of interest

None declared.

References

- Shaw PJ: Molecular and cellular pathways of neurodegeneration in motor neurone disease. *J Neurol Neurosurg Psychiatry* 2005; **76**: 1046–1057.
- Rosen DR, Siddique T, Patterson D *et al*: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 1993; **362**: 59–62.
- Valdmanis PN, Rouleau GA: Genetics of familial amyotrophic lateral sclerosis. *Neurology* 2008; **70**: 144–152.
- Schymick JC, Talbot K, Traynor BJ: Genetics of sporadic amyotrophic lateral sclerosis. *Hum Mol Genet* 2007; **16**: R233–R242.
- de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D: Efficiency and power in genetic association studies. *Nat Genet* 2005; **37**: 1217–1223.
- Beghi E, Logroscino G, Chiò A *et al*, EURALS Consortium: The epidemiology of ALS and the role of population-based registries. *Biochim Biophys Acta* 2006; **1762**: 1150–1157.
- Dunckley T, Huentelman MJ, Craig DW *et al*: Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N Engl J Med* 2007; **357**: 775–788.
- 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**: 322–328.
- 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**: 869–877.
- 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.
- 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.
- Brooks BR: El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases. *J Neurol Sci* 1994; **124**: S96–S107.
- Purcell S, Neale B, Todd-Brown K *et al*: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- Skol AD, Scott LJ, Abecasis GR, Boehnke M: Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; **38**: 209–213. Genetic power calculator available at <http://www.sph.umich.edu/csg/abecasis/CaTS/index.html>.
- Gurney ME, Pu H, Chiu AY *et al*: Motor neuron degeneration in mice that express a human CuZn superoxide dismutase mutation. *Science* 1994; **264**: 1772–1775.
- Graham AJ, Macdonald AM, Hawkes CH: British motor neuron disease twin study. *J Neurol Neurosurg Psychiatry* 1997; **62**: 562–569.
- Seng KC, Seng CK: The success of the genome-wide association approach: a brief story of a long struggle. *Eur J Hum Genet* 2008; **16**: 554–564.
- The International HapMap Consortium: The International HapMap Project. *Nature* 2003; **426**: 789–796.
- Gilks WP, Abou-Sleiman PM, Gandhi S *et al*: A common LRRK2 mutation in idiopathic Parkinson's disease. *Lancet* 2005; **365**: 415–416.
- Kay DM, Zabetian CP, Factor SA *et al*: Parkinson's disease and LRRK2: frequency of a common mutation in US movement disorder clinics. *Mov Disord* 2006; **21**: 519–523.
- González-Fernández MC, Lezcano E, Ross OA *et al*: Lrrk2-associated parkinsonism is a major cause of disease in Northern Spain. *Parkinsonism Relat Disord* 2007; **13**: 509–515.
- Ozelius LJ, Senthil G, Saunders-Pullman R *et al*: LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. *N Engl J Med* 2006; **354**: 424–425.
- Lesage S, Dürr A, Tazir M *et al*: LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. *N Engl J Med* 2006; **354**: 422–423.
- Valdmanis PN, Kabashi E, Dion PA, Rouleau GA: ALS predisposition modifiers: knock NOX, who's there? SOD1 mice still are. *Eur J Hum Genet* 2008; **16**: 140–142.
- Antoniou AC, Spurdle AB, Sinilnikova OM *et al*: Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet* 2008; **82**: 937–948.
- Lyon HN, Emilsson V, Hinney A *et al*: The association of a SNP upstream of INSIG2 with body mass index is reproduced in several but not all cohorts. *PLoS Genet* 2007; **3**: e61.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)