doi:10.1093/brain/awq192 Brain 2010: 133; 2603–2611 | 2603

BRAIN
A JOURNAL OF NEUROLOGY

Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis

Sergio E. Baranzini,[1] Radhika Srinivasan,[2,*] Pouya Khankhanian,[1,*] Darin T. Okuda,[1] Sarah J. Nelson,[2] Paul M. Matthews,[3] Stephen L. Hauser,[1] Jorge R. Oksenberg[1] and Daniel Pelletier[1]

1 Department of Neurology, University of California, San Francisco, CA 94143, USA
2 Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA 94158, USA
3 GlaxoSmithKline Research and Development, Harlow CM19 5AW, UK

*These authors contributed equally to this work.

Correspondence to: Sergio E. Baranzini,
Department of Neurology,
School of Medicine,
University of California,
San Francisco,
513 Parnassus Ave. Room S-256,
San Francisco, CA 94143-0435, USA
E-mail: sebaran@cgl.ucsf.edu

Glutamate is the main excitatory neurotransmitter in the mammalian brain. Appropriate transmission of nerve impulses through glutamatergic synapses is required throughout the brain and forms the basis of many processes including learning and memory. However, abnormally high levels of extracellular brain glutamate can lead to neuroaxonal cell death. We have previously reported elevated glutamate levels in the brains of patients suffering from multiple sclerosis. Here two complementary analyses to assess the extent of genomic control over glutamate levels were used. First, a genome-wide association analysis in 382 patients with multiple sclerosis using brain glutamate concentration as a quantitative trait was conducted. In a second approach, a protein interaction network was used to find associated genes within the same pathway. The top associated marker was rs794185 ($P < 6.44 \times 10^{-7}$), a non-coding single nucleotide polymorphism within the gene *sulphatase modifying factor 1*. Our pathway approach identified a module composed of 70 genes with high relevance to glutamate biology. Individuals carrying a higher number of associated alleles from genes in this module showed the highest levels of glutamate. These individuals also showed greater decreases in *N*-acetylaspartate and in brain volume over 1 year of follow-up. Patients were then stratified by the amount of annual brain volume loss and the same approach was performed in the 'high' ($n = 250$) and 'low' ($n = 132$) neurodegeneration groups. The association with rs794185 was highly significant in the group with high neurodegeneration. Further, results from the network-based pathway analysis remained largely unchanged even after stratification. Results from these analyses indicated that variance in the activity of neurochemical pathways implicated in neurodegeneration is explained, at least in part, by the inheritance of common genetic polymorphisms. Spectroscopy-based imaging provides a novel quantitative endophenotype for genetic association studies directed towards identifying new factors that contribute to the heterogeneity of clinical expression of multiple sclerosis.

Received January 17, 2010. Revised May 20, 2010. Accepted May 29, 2010

Keywords: glutamate receptor; genetics; multiple sclerosis

Abbreviations: DKS = domain knowledge score; GWAS = genome-wide association study; NAA = *N*-acetylaspartate; PIN = protein interaction network; SNP = single nucleotide polymorphism

Introduction

L-Glutamate is the major excitatory neurotransmitter in the mammalian central nervous system (CNS), acting through both ligand-gated ion channels and G-protein-coupled receptors. Activation of these receptors is responsible for basal excitatory synaptic transmission and many forms of synaptic plasticity, such as long-term potentiation and long-term depression, which underlie the processes of learning and memory. Excess glutamate in the synaptic space triggers a toxic cascade that results in neuroaxonal injury and death in a broad range of disorders (Pitt *et al*., 2000; Ouardouz *et al*., 2009). Glutamate toxicity is also a contributing aetiological factor in multiple sclerosis (OMIM #126200) (Stys, 2005; Dong *et al*., 2009), a common neuroinflammatory disease affecting young adults. Glutamate concentrations are elevated both in acute lesions and in normal-appearing white matter of patients with multiple sclerosis (Srinivasan *et al*., 2005) consistent with a potential role in the neurodegeneration underlying disease progression (De Stefano *et al*., 2007). Using magnetic resonance spectroscopic imaging, it is now possible to measure *in vivo* glutamate levels in the brain with precision (Srinivasan *et al*., 2006). The aim of this study was to test for associations between common variants in the human genome and differences in glutamate concentrations in normal-appearing brain matter of patients with multiple sclerosis.

In the last few years, nearly a dozen non-human leucocyte antigen genes have been associated with this disease including *IL7R* (*CD127*), *IL2RA* (*CD25*), *CD58* and *CLEC16A* (Hafler *et al*., 2007; ANZgene, 2009; Baranzini *et al*., 2009*b*). However, the small effect of most associations demands very large studies to reach adequate statistical power (Sawcer, 2008). An additional challenging aspect in the effort to identify new genetic associations in diseases like multiple sclerosis has been the lack of disease-specific quantitative traits that can be related to DNA variants in a given individual. Such an approach has been successfully used in conditions like hypertension [variation in blood pressure (Cowley, 2006)], erythrocyte phenotypes (Ganesh *et al*., 2009) and circulating fibrinogen levels (Dehghan *et al*., 2009). Recent studies relating *N*-acetylaspartate (NAA) (a marker of neuroaxonal damage) and HLA-DRB1*1501 allele status, and quantitative measures of grey matter density, volume and cortical thickness with genome-wide DNA variants, highlight the potential benefits of integrating genetic and imaging data to illuminate new aspects in the pathogenesis of neurological diseases (Okuda *et al*., 2009; Shen *et al*.). We hypothesized that variation in brain glutamate concentrations measured by *in vivo* magnetic resonance spectroscopy could be used as a quantitative trait and combined with available high-density genotyping data to identify genetic contributions to glutamate-mediated toxicity in multiple sclerosis.

Material and methods

Study participants and genetic data

This study was approved by the institutional review board at the University of California San Francisco (UCSF). After signing appropriate informed consents, all participants were recruited at the UCSF Multiple Sclerosis Centre using established inclusion criteria (Baranzini *et al*., 2009*b*). Baseline clinical characteristics of this patient data set are listed in Table 1. Genotypes were extracted from a recently completed and reported genome-wide association study (GWAS) using the Sentrix® HumanHap550 BeadChip (Baranzini *et al*., 2009*b*).

MRI data acquisition and brain volume measurement

A high-resolution inversion recovery spoiled gradient-echo T_1-weighted isotropic volumetric sequence [3D inversion re-covery spoiled gradient recalled acquisition in steady-state (IRSPGR), $1 \times 1 \times 1\ mm^3$, 180 slices] was acquired for brain volume measurements (echo time/repetition time/interval time = 2/7/400 ms, flip angle = 15°, $256 \times 256 \times 180$ matrix, $240 \times 240 \times 180\ mm^3$ field of view, number of excitations =). Annual per cent whole brain volume change was calculated by structural image evaluation using normalization of atrophy, a fully automated method of longitudinal (temporal) brain change analysis (Smith *et al*., 2002).

Table 1 Cohort characteristics

Variable	Value
Number of cases	382
Gender ratio (M:F)	1:2
Age (years)	44.2 ± 9.9
Disease subtype	
Clinically isolated syndrome	58
Relapse–remitting	279
Secondary progressive	32
Primary progressive	12
Progressive relapsing	1
Clinical	
EDSS	1.96 ± 1.62
MSSS	2.98 ± 2.43
Age of onset (years)	33.8 ± 9.5
Imaging	
T_1 lesion load (cm^3)	4.2 ± 7.4
T_2 lesion load (cm^3)	7.4 ± 13.2

EDSS = expanded disability status scale; MSSS = Multiple sclerosis severity score.

Spectroscopic data acquisition and quantification

Glutamate and NAA concentrations were measured in 382 patients at study entry using two-dimensional echo time-averaged proton spectroscopic imaging. Technical details of this imaging pulse sequence have been reported elsewhere (Srinivasan *et al*., 2006). The spectroscopic data were acquired on a 3T GE Excite scanner (GE Healthcare Technologies, Waukesha, WI, USA) using an 8-channel phased coil immediately following the acquisition of the anatomical images and prior to the administration of the contrast agent. The resulting coil combination data were echo time-averaged and the metabolic concentration of glutamate and NAA were obtained using the LCmodel quantification algorithm and corrected for T_1 (glutamate, NAA) and T_2 (NAA) relaxation times using similar methods described elsewhere (Ratiney *et al*., 2007). Concentration estimations of 'pure' grey and white matter metabolites were derived from segmented high-resolution T_1-weighted images using a hidden Markov random field model with expectation maximization. By modelling the glutamate and NAA concentrations and magnetic resonance relaxation parameters as a linear function of percentage of white matter content, 'pure' grey matter and white matter glutamate and NAA concentrations were extrapolated from the end points of the linear fit.

Association analysis

An association analysis was performed using the linear regression option in PLINK (Purcell *et al*., 2007). Disease duration, age of onset and *HLA-DRB1*1501* status were used as covariates. We only report association *P*-values corresponding to additive effects of allele dosage.

Module (sub-network) analysis

Systematic module searches on a highly curated protein interaction network were conducted as described (Supplementary Fig. 1) (Baranzini *et al*., 2009*a*). Briefly, each gene product in the network was assigned a number corresponding to the *P*-value of the most strongly associated single-nucleotide polymorphism (SNP) for that gene with the trait (only *P*-values <0.05 were considered). Next, the Cytoscape (www.cytoscape.org) plugin *jActive modules* (Ideker *et al.,* 2002) was used to identify groups of interacting gene products that were also associated with *in vivo* glutamate. *jActive modules* converts *P*-values into *z*-scores and uses a greedy algorithm to grow a sub-network (or module) from a random seed node by sequentially incorporating its neighbours in the protein interaction network (PIN). The algorithm then returns the smallest possible module that includes gene products with the most significant associations. A significance (*z*) score is assigned to each reported sub-network after evaluation of 10 000 random networks of similar size. Only modules with a score >3 and of size >5 were considered.

To evaluate whether the significant modules obtained were biologically meaningful, we employed a text-mining strategy by performing automated PubMed searches with each gene product in the module and the terms 'glutamate' OR 'glutamic acid'. An aggregate value, named the domain knowledge score (DKS) was then computed by adding up the number of articles retrieved for all gene products in a given module and dividing it by the number of gene products in the module. To avoid introducing bias due to a single well-characterized molecule, we set a limit of 10 articles per gene.

Computation of genetic scores

The risk allele for each associated SNP was defined as the allele showing the highest frequency among patients with high baseline *in vivo* glutamate levels. For a module of *N* genes, each patient was assigned a genetic score (from 0 to 2**N*) corresponding to the total number of risk alleles carried at the *N* loci. Scores were normalized between 0 and 1 to allow comparison of networks of different sizes. This genetic score was correlated with other magnetic resonance spectroscopic metrics such as NAA concentration change over time and brain atrophy (per cent whole brain volume change values from structural image evaluation using normalization of atrophy).

Computation of expected correlations by simulation

Given three random variables *x*, *y* and *z*, with the following non-zero correlations:

$$r_{xy} = |\mathrm{corr}(x,y)| \quad \text{and} \quad r_{y,z} = |\mathrm{corr}(y,z)|,$$

we expect the correlation between *x* and *z* to be distributed normally with mean

$$|r_{xz}| = r_{xz} * r_{yz}$$

To assign significance to an observed r_{xz} given r_{xy} and r_{yz}, we performed a simulation. One thousand random sets of vectors *x*, *y* and *z* were constructed such that r_{xy} and r_{yz} matched the observed correlations in our data set. r_{xz} was calculated in these random sets and a normal distribution was created. The observed r_{xz} was compared with this normal distribution to assign significance.

Results

Glutamate levels were measured *in vivo* in both the normal-appearing white and grey matter of 382 patients with multiple sclerosis using ^{1}H magnetic resonance spectroscopy imaging. Although at the time of the study most patients had a diagnosis of relapse–remitting multiple sclerosis ($n=279$) or clinically isolated syndrome ($n=58$), all types of the disease were represented in this analysis. Table 1 lists the demographic details of the cohort alongside other radiological and clinical parameters. Approximately 500 000 genotypes were available for each study participant (Baranzini *et al*., 2009*b*) and were used to perform a genome-wide association analysis using brain glutamate concentrations as an endophenotypic continuous trait (Fig. 1). The top associated marker was rs794185 ($P<6.44\times10^{-7}$), a SNP in chromosome 3p26.2 that maps to intron 6 of the gene coding

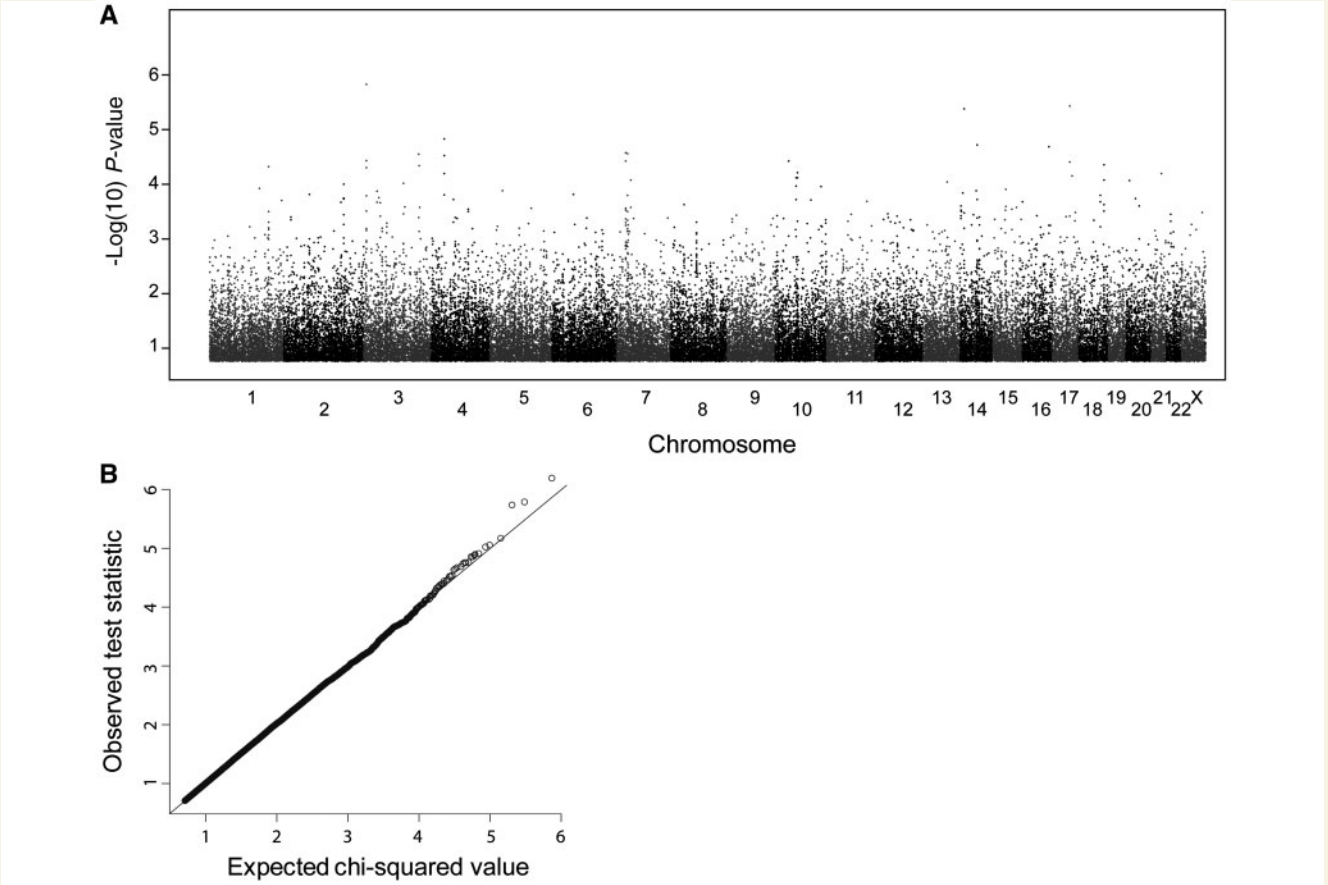

Figure 1 Genome-wide scan for allele frequency differences related to *in vivo* glutamate concentration in multiple sclerosis brains. (**A**) *P*-values from the linear regression with glutamate concentration, controlled by disease duration, age of onset and DRB1 status. (**B**) Quantile–quantile plots of these test statistics.

for sulphatase modifying factor 1 (*SUMF1*). Mutations in *SUMF1* lead to multiple sulphatase deficiency (MIM 272200), a lysosomal storage disorder. DNA variants in this gene may indirectly regulate extracellular glutamate by altering the activity of steroid sulphatases (Shirakawa *et al*., 2005; Gibbs *et al*., 2006; Valenzuela *et al*., 2008). A region spanning ~4 Mb in chromosome 7 and containing 11 SNPs in *HDAC9,* 5 SNPs in *CDCA7L* and 6 SNPs in *DRCTNNB* was found to be modestly associated (*P*-values between 10^{-4} and 10^{-5}). The top 20 associated SNPs are listed in Supplementary Table 1.

To maximize the probability of identifying true associated genes, we used a PIN-based analysis as previously described (Baranzini *et al*., 2009*a*). In this method, evidence of genetic association is combined with evidence of physical interaction of the respective gene products. Thirty-four modules were found to be significantly associated with *in vivo* glutamate concentration. Due to the nature of the searches, significant overlap in the composition of modules is expected. In order to assess the relative importance of these modules, several criteria were considered. First, a literature search was performed to determine the relevance of each of the component genes to glutamate biology. Next, their association with related phenotypes such as NAA decline and brain atrophy change were measured. Module 14 was the top scoring module across all these criteria. This module was composed of 70 genes and included three ionotropic glutamate receptors (*GRID2*, *GRIK2* and *GRIK5*), 17 anchoring proteins required for glutamate receptor and transporter organization (*AKAP5, DLG2, DLG4, SHANK2, PRKCA, LRRC7, PKP4, CTNND2, CDH2, CDH5, DSC3, ARVCF, NLGN4X, DLGAP1, CASK, CASKIN, ACTN2*), two axon guidance molecules (*DAB1*, *DAB2*) and three key regulators of glutamatergic synaptic activity (*ERBB4, PTK2B and PARK2*) (Fig. 2). In addition, seven members of the TGF-β signalling pathway (*SMAD1*, *SMAD2 SMAD3, SMAD6, SMURF1, ERBB2IP and ACVR1*) were also members of this network.

To assess the biological relevance of reported modules a text-mining strategy was implemented by performing automated PubMed searches with each of the component genes of all identified modules and the term 'glutamate' OR 'glutamic acid'. The aggregate number of articles found for all genes in a given module was recorded as its DKS. The top associated genes (by *P*-value) found in the PIN (whether they interact or not) had higher DKS than the top associated genes not found in the PIN, perhaps indicating that genes in the PIN are overall better annotated. However, the DKS of Module 14 (and that of all other modules, data not shown) was significantly higher than that of both the top genes in the PIN and the overall top associated genes (Fig. 3).

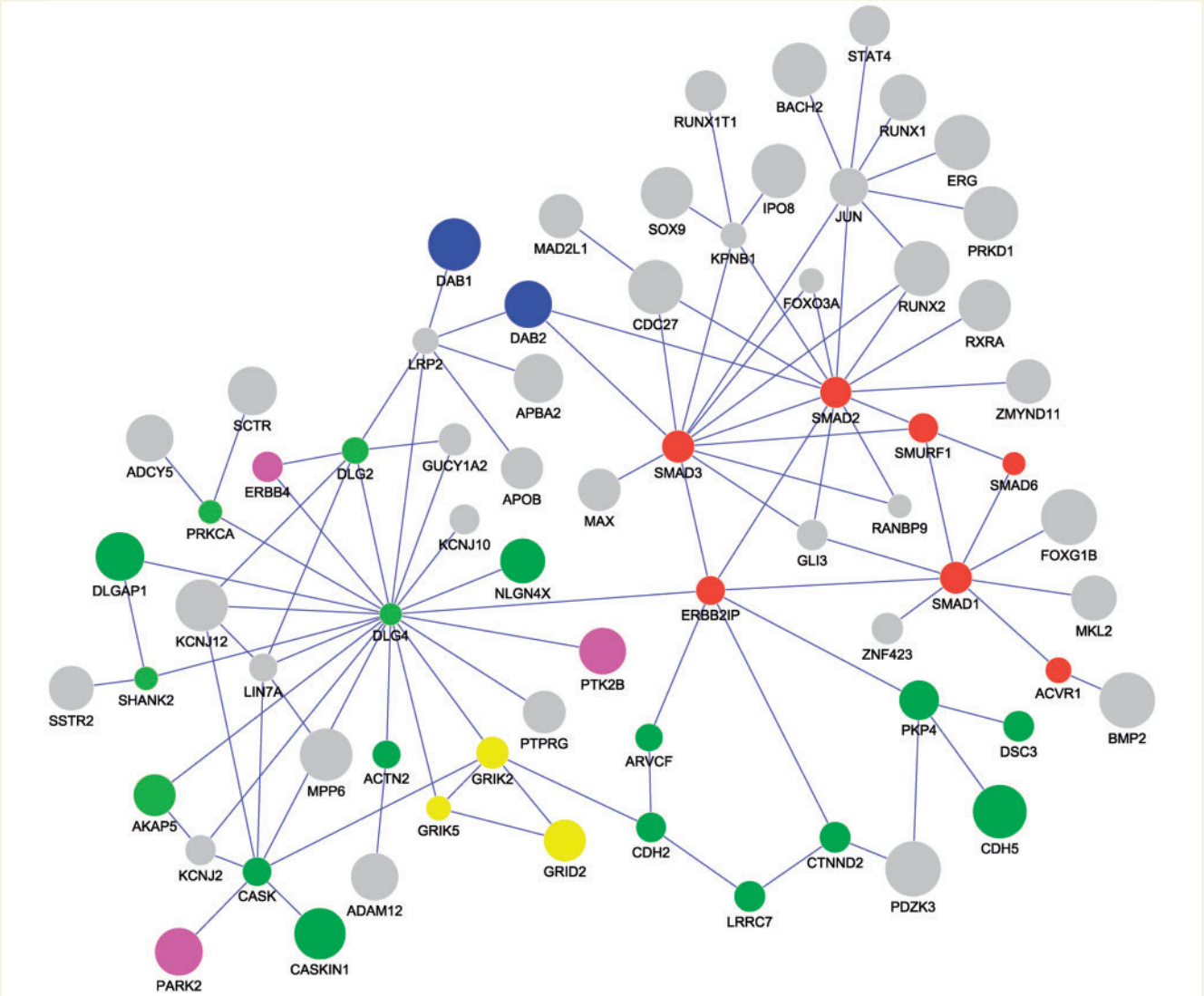

Figure 2 Module 14. A graphical representation of the overall highest scoring module from the protein interaction network. Circles represent proteins and lines represent interactions among them. Proteins are coloured according to their relationship to glutamate. Green = glutamate receptor and transporter organization; red = TGF-β signalling; pink = regulators of glutamatergic synaptic activity; yellow = glutamate receptors; blue = axon guidance; grey = unclassified.

These results support our hypothesis that the network-based approach identifies biologically related genes.

We then tested the association of each module with selected relevant MRI-based metrics (Fig. 4). To this end, a module-specific genetic score was computed for each patient. The score was derived from the number of risk alleles carried at each gene represented in the module. We reasoned that if glutamate concentration was affected by genomic variants, then individuals with the highest glutamate levels would show the highest number of associated alleles in the module. As predicted, patients with the highest glutamate levels in grey matter were more likely to display the highest genetic scores (for Module 14: $R^2 = 42\%$) (Fig. 4A). While this correlation was expected because genetic scores were derived from the regression with the trait, the highly significant *P*-value (2.58×10^{-29}) indicates that most, if not all, of the 70 genes in the module contribute to the phenotype. Interestingly, we observed a significant correlation between the rate of NAA decline in grey matter over the first year after glutamate measurements and the module-specific genetic score ($R^2 = 6.3\%$, $P < 10^{-4}$) (Fig. 4B). Although a relationship between these two variables was anticipated (expected $R^2 = 4.4\%$) due to the existing correlation between glutamate and NAA levels ($R^2 = 10.4\%$, $P < 10^{-7}$), the correlation between genetic scores and NAA decline is higher than expected. Simulations with artificial data sets and conditional regression analysis were conducted to assess the statistical significance of these additional correlations but were not conclusive, possibly reflecting lack of power due to a moderate sample size and the relatively short follow-up time. Similar results were obtained when NAA decline in white matter was considered ($R^2 = 3.2\%$, $P < 0.007$). Finally, we also observed a significant

correlation between brain atrophy over the first year and the module-specific genetic score (observed: $R^2 = 1.2\%$ $P < 0.05$; expected $R^2 = 0.14\%$) (Fig. 4C). Correlations between glutamate-based genetic score and the multiple sclerosis severity score (computed over 2 and 3 years) did not reach significance.

As described in Table 1, the data set included patients of all disease subtypes and presumably with variable degrees of neurodegeneration. It is conceivable that the DNA variants associated with glutamate concentration will have a stronger effect in patients with more neurodegeneration, as evidenced by MRI. To evaluate this hypothesis, we stratified patients based on brain atrophy (i.e. a surrogate of neurodegeneration) and repeated the entire analysis in each group. Patients showing at least 0.2% decline (−0.2% percent whole brain volume change) in structural image evaluation using normalization of atrophy at two or more times during a 3-year follow-up period were considered as the group with 'high' neurodegeneration ($n = 250$), while the remaining 132 individuals were defined as the group with 'low' neurodegeneration. A decline of 0.2% or more was observed in normal ageing from 48 healthy controls scanned annually using the same 3T scanner (D. Pelletier, unpublished data). The top 20 associated variants in the high neurodegeneration group are shown in Supplementary Table 2. It is noteworthy that the top hit from the original GWAS performed in all patients (rs794185 in *SUMF1*, $P < 6.44 \times 10^{-7}$) was the second-most significant marker in the new analysis ($P < 9.92 \times 10^{-6}$). In comparison, the same SNP ranked 329 987 in the GWAS with the low neurodegeneration group (the top 20 variants are shown in Supplementary Table 3), suggesting that most of the statistical significance of the *SUMF1* association in the full data set derives from the group of patients with high neurodegeneration.

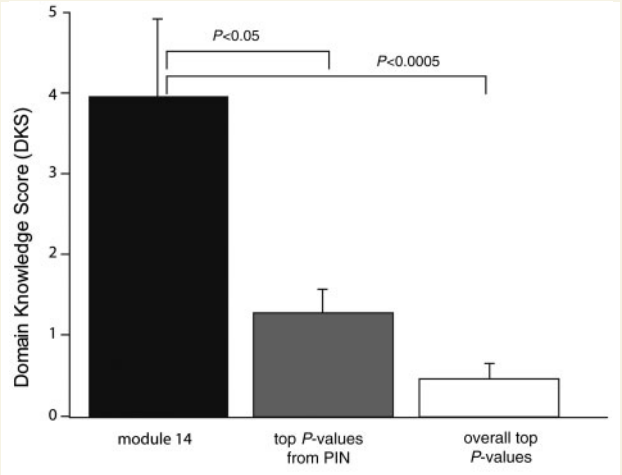

Figure 3 Domain knowledge scores. Mean DKSs were calculated for genes in Module 14 (black bar), and for the top associated genes from the same protein network (whether they interact or not) (grey bar). Also, the mean DKS of the top associated genes from the original GWAS is shown. The mean DKS of genes in Module 14 is significantly higher than those of the other two lists of associated genes (Welch's *t*-test, bars represent SEM).

We then performed the network-based analysis to search for modules enriched in modestly associated variants within functionally related genes. Like in the original GWAS, several networks (i.e. 23) with different degrees of overlapping were significantly enriched in genes associated with glutamate concentrations. Of these, eight networks were also functionally related to glutamate in the group with high neurodegeneration as determined by their DKS. Whereas a similar number of significant modules were found for the group with low neurodegeneration (i.e. 30), only one of them was functionally related to glutamate.

The top associated module in the group with high neurodegeneration was composed of 55 genes (Supplementary Fig. 2) significantly enriched in glutamate biology (Supplementary Fig. 3). Although computing the overlap between any two networks is currently a subject of much debate in graphic theory (nets could be compared on the basis of shared nodes, edges or half a dozen network parameters such as connectivity distribution, clustering

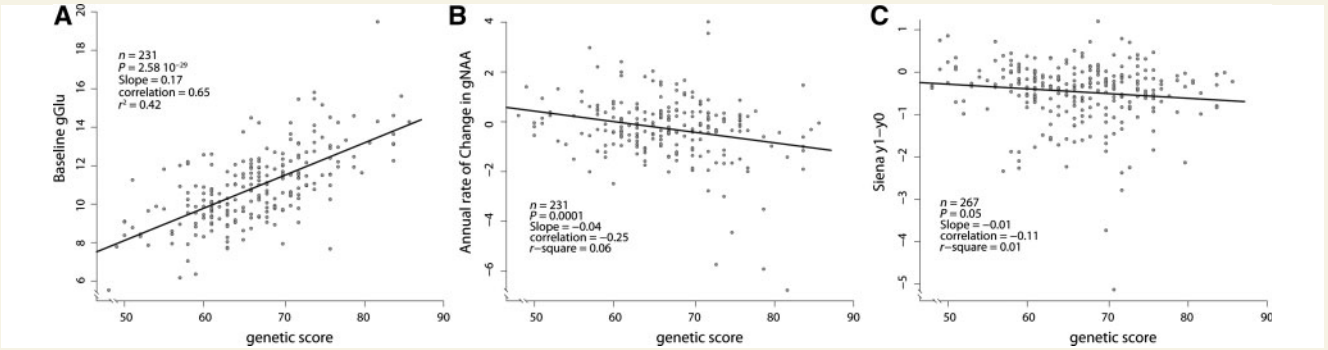

Figure 4 Correlation between glutamate genetic score and relevant variables. (**A**) Correlation of glutamate genetic scores with grey matter glutamate concentration. (**B**) Correlation of genetic scores with NAA change over 1 year. Genetic scores explain more variance in NAA decline than expected given the *a priori* correlation between glutamate level and NAA decline. (**C**) Similarly, correlation between genetic scores and brain atrophy was significant and higher than that expected from *a priori* correlation between glutamate level and brain atrophy.

coefficient, etc.), a rough measure is to count the number of nodes in common. When compared to the network obtained with data from the original GWAS, we identified 13 genes in common, a remarkable overlap considering that the expected number of shared genes between two random networks of comparable sizes is virtually 0.

Discussion

In this study, we performed a genome-wide association analysis using *in vivo* CNS glutamate concentration in patients with multiple sclerosis as a quantitative trait. Despite the relatively small data set, several SNPs appeared associated at significance levels comparable to those reported in genome-wide analyses. For example, the SNP rs794185 in *SUMF1* reached $P < 6.44 \times 10^{-7}$ and a peak of modestly associated markers in three other genes was observed in chromosome 7. *SUMF1* is a key activator of sulphatases; it catalyses the hydrolysis of sulphate esters by oxidizing a cysteine residue in the substrate sulphatase to an active site 3-oxoalanine residue, which is also known as C-α-formylglycine. Dysregulation of *SUMF1* might lead to reduced activity of other sulphatases (Fraldi *et al*., 2007). Interestingly, deficiency of the lysosomal arylsulphatase A causes metachromatic leucodystrophy, a severe neurological disorder that affects myelin integrity. Sulphated steroids are powerful endogenous modulators of the balance between excitatory and inhibitory neurotransmission (Gibbs *et al*., 2006). In particular, they have been shown to regulate synaptic transmission by altering the function of post-synaptic neurotransmitter receptors (Valenzuela *et al*., 2008) and to attenuate α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate cytotoxicity on cortical neurons (Shirakawa *et al*., 2005). Pregnenolone sulphate has been shown to inhibit α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate/kainate as well as a subset of *N*-methyl-D-aspartate receptors (GRIN2C and GRIN2D) highly expressed in oligodendrocytes (Salter and Fern, 2005) and myelin (Micu *et al*., 2006). Reduced activity of pregnenolone sulphate might lead to an increased expression of these receptors, thus accelerating degeneration of white matter elements. Intriguingly, pregnenolone sulphate is highly expressed during pregnancy, a period where relapse rate is significantly reduced (Vukusic *et al*., 2004).

The use of a protein interaction network to find groups of functionally related genes has been shown to be an effective method to identify genetic modules significantly associated with a trait (Wang *et al*., 2007; Torkamani *et al*., 2008; Baranzini *et al*., 2009*a*). Using data from a GWAS performed on glutamate in 382 patients, we identified and described a sub-network of the global protein interaction ensemble (Module 14) composed primarily of molecules related to glutamate biology. GRID2 is a component of the *N*-methyl-D-aspartate receptor complex and GRIK2 and GRIK5 belong to the kainate-responsive family of glutamate receptors. Variation in the sequence of these genes could lead to an altered glutamate homoeostasis, contributing then to the primary susceptibility and progression underlying multiple sclerosis. Glutamate receptors also associate with a number of accessory proteins whose function is to recruit other members and to stabilize the receptor complex. Five of these genes are also part of Module 14 (*AKAP5, DLG2, DLG4, SHANK2, PRKCA*) along with two others (DAB1 and DAB2) whose main function is to interact with protein kinase pathways to regulate neuronal positioning in the developing brain. Neuregulin-1 signalling participates in numerous neurodevelopmental processes and variants in this gene have been associated with schizophrenia (Stefansson *et al*., 2002). More recently, the neuregulin-1 receptor *ERBB4* (which is a member of Module 14) has been shown to control synapse maturation and plasticity, thus contributing to the glutamatergic hypofunction characteristic in that disease (Li *et al*., 2007). *DLG4* encodes a member of the membrane-associated guanylate kinase family. As shown in Fig. 2, it interacts with several proteins including membrane-associated guanylate kinase members DLG2 and CASK, and it is then recruited into *N*-methyl-D-aspartate receptor and potassium channel clusters. DLG2 and CASK interact at post-synaptic sites to form a multimeric scaffold for the clustering of receptors, ion channels and associated signalling proteins. Interestingly, changes in *DLG4* expression have also been associated with schizophrenia (Clinton and Meador-Woodruff, 2004; Funk *et al*., 2009). Finally, the protein tyrosine kinase 2 β (*PTK2B)* was also a component of Module 14. This gene encodes a cytoplasmic protein tyrosine kinase that is involved in calcium-induced regulation of ion channels and activation of the map kinase signalling pathway. PTK2B represents an important signalling intermediate between neurotransmitter receptors that increase calcium flux and the downstream signals that regulate neuronal activity. Of relevance, PTK2B has been found to regulate glutamate release through the modification of actin dynamics (Ohnishi *et al*., 2001). Seven members of the TGF-β signalling pathway were also part of Module 14, strongly suggesting the involvement of this pleiotropic cytokine in glutamate metabolism. TGF-β signalling has been previously associated with neuroprotection via regulation of Ca^{2+} homoeostasis by *N*-methyl-D-aspartate receptors (Meucci and Miller, 1996), although a recent report related this pathway to neuronal activation and glutamate-mediated CNS injury (Luo *et al*., 2006). These observations underscore the complex regulatory relationship between TGF-β and glutamate. Altogether, Module 14 is largely composed of genes involved in glutamate biology, thus suggesting that the network-based search for significant associations does select biologically relevant genes.

We also performed two additional genome-wide association studies, stratifying the study participants by high or low brain atrophy (neurodegeneration). These data revealed that *SUMF1* is highly associated with glutamate concentrations only in the group with high neurodegeneration. This is noteworthy because it suggests that DNA variants in this sulphatase-modifying enzyme are more likely to influence the glutamate concentration in a subset of patients while other mechanisms may be involved in the regulation and maintenance of glutamate concentrations in patients with less neurodegeneration. Indeed, while more evidence for gene networks associated with glutamate concentrations was found in the group with high neurodegeneration, at least one significant network of genes related to glutamate was also found in the group of patients with low neurodegeneration.

In order to quantify the genetic load of a given individual, we developed a genetic score that takes into account the number of alleles associated with the trait (glutamate concentration). Thus, individuals with higher genetic scores are more likely to show higher glutamate levels and other related traits. The significant correlation between genetic scores and NAA decline over the following year is slightly above what would be expected given the relationship between glutamate and NAA levels. This suggests that any effect these genes have on NAA is mediated by glutamate levels. Similarly, we observed that variants in these genes also explain a non-negligible amount of the variance in brain atrophy measured over 1 year. As deduced from the slope of the graph in Fig. 4C, 0.01% of brain volume loss can be attributed to carrying a single additional risk allele from Module 14. Since patients at the extremes of the distribution of genetic scores can differ in as many as 30 units, we estimate that as much as 0.3% (one-third) of the yearly main brain volume loss observed in patients with multiple sclerosis (annual per cent whole brain volume change $\leq -1\%$) can be attributed to the common variation found in these genes. Correlations between the multiple sclerosis severity score (computed over 2 and 3 years) and the genetic score did not reach significance, possibly due to a lack of power as data from fewer patients were available for these later time points. Another possibility is that in order to observe a causal relationship between glutamate and the multiple sclerosis severity score, a longer follow-up period may be needed. While larger and independent cohorts will be desirable to validate these findings, the exact number of individuals to test for an approach like this is unclear as traditional power calculations to estimate the sample size may not be directly applicable. While increased glutamate can originate from either a defect in transporters or by increased production by activated inflammatory cells (Piani *et al.*, 1991), the exact source could not be determined since the method used here only quantifies total glutamate. Finally, the text-mining approach used to establish biological significance is subject to some limitations including the introduction of false positives and enrichment in well-annotated genes.

In summary, we have identified genetic variation in genes associated with *in vivo* glutamate measured using ^{1}H magnetic resonance spectroscopic imaging in the grey matter of patients with multiple sclerosis. We have also shown that common variations in a limited group of functionally related genes contribute significantly to NAA decay and brain volume in multiple sclerosis. These results highlight the importance of glutamate in multiple sclerosis biology and generate new hypotheses that link genetic variation with disease progression. Studies of this nature represent the logical next step following the discovery of susceptibility genes and are ideally suited to begin explaining the heterogeneity observed in multiple sclerosis and other neurodegenerative disorders.

Acknowledgements

We would like to acknowledge all the patients that participated in this study. We also thank the anonymous reviewers for helpful suggestions. S.E.B. and D.P. are Harry Weaver Neuroscience scholars of the US National Multiple Sclerosis Society

Funding

U.S. National Multiple Sclerosis Society (RG2901D9/1, CA1035-A-7, RG3517/3); U.S. National Institutes of Health (RO1 NS26799); U.S. National Multiple Sclerosis Society (PP1182 to R.S.). DP and SEB are Harry Weaver Fellows from the US NAtional multiple sclerosis Society.

Supplementary material

Supplementary material is available at *Brain* online.

References

ANZgene. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 2009; 41: 824–8.

Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. Hum Mol Genet 2009a; 18: 2078–90.

Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009b; 18: 767–78.

Clinton SM, Meador-Woodruff JH. Abnormalities of the NMDA receptor and associated intracellular molecules in the thalamus in schizophrenia and bipolar disorder. Neuropsychopharmacology 2004; 29: 1353–62.

Cowley AW Jr. The genetic dissection of essential hypertension. Nat Rev Genet 2006; 7: 829–40.

De Stefano N, Filippi M, Miller D, Pouwels PJ, Rovira A, Gass A, et al. Guidelines for using proton MR spectroscopy in multicenter clinical MS studies. Neurology 2007; 69: 1942–52.

Dehghan A, Yang Q, Peters A, Basu S, Bis JC, Rudnicka AR, et al. Association of novel genetic Loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. Circ Cardiovasc Genet 2009; 2: 125–33.

Dong XX, Wang Y, Qin ZH. Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. Acta Pharmacol Sin 2009; 30: 379–87.

Fraldi A, Biffi A, Lombardi A, Visigalli I, Pepe S, Settembre C, et al. SUMF1 enhances sulfatase activities in vivo in five sulfatase deficiencies. Biochem J 2007; 403: 305–12.

Funk AJ, Rumbaugh G, Harotunian V, McCullumsmith RE, Meador-Woodruff JH. Decreased expression of NMDA receptor-associated proteins in frontal cortex of elderly patients with schizophrenia. Neuroreport 2009; 20: 1019–22.

Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet 2009; 41: 1191–8.

Gibbs TT, Russek SJ, Farb DH. Sulfated steroids as endogenous neuromodulators. Pharmacol Biochem Behav 2006; 84: 555–67.

Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007; 357: 851–62.

Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 2002; 18 (Suppl 1): S233–40.

Li B, Woo RS, Mei L, Malinow R. The neuregulin-1 receptor erbB4 controls glutamatergic synapse maturation and plasticity. Neuron 2007; 54: 583–97.

Luo J, Lin AH, Masliah E, Wyss-Coray T. Bioluminescence imaging of Smad signaling in living mice shows correlation with excitotoxic neurodegeneration. Proc Natl Acad Sci USA 2006; 103: 18326–31.

Meucci O, Miller RJ. gp120-induced neurotoxicity in hippocampal pyramidal neuron cultures: protective action of TGF-beta1. J Neurosci 1996; 16: 4080–8.

Micu I, Jiang Q, Coderre E, Ridsdale A, Zhang L, Woulfe J, et al. NMDA receptors mediate calcium accumulation in myelin during chemical ischaemia. Nature 2006; 439: 988–92.

Ohnishi H, Yamamori S, Ono K, Aoyagi K, Kondo S, Takahashi M. A src family tyrosine kinase inhibits neurotransmitter release from neuronal cells. Proc Natl Acad Sci USA 2001; 98: 10930–5.

Okuda DT, Srinivasan R, Oksenberg JR, Goodin DS, Baranzini SE, Beheshtian A, et al. Genotype-Phenotype correlations in multiple sclerosis: HLA genes influence disease severity inferred by 1HMR spectroscopy and MRI measures. Brain 2009; 132: 250–9.

Ouardouz M, Coderre E, Zamponi GW, Hameed S, Yin X, Trapp BD, et al. Glutamate receptors on myelinated spinal cord axons: II. AMPA and GluR5 receptors. Ann Neurol 2009; 65: 160–6.

Piani D, Frei K, Do KQ, Cuenod M, Fontana A. Murine brain macrophages induced NMDA receptor mediated neurotoxicity in vitro by secreting glutamate. Neurosci Lett 1991; 133: 159–62.

Pitt D, Werner P, Raine CS. Glutamate excitotoxicity in a model of multiple sclerosis. Nat Med 2000; 6: 67–70.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559–75.

Ratiney H, Noworolski SM, Sdika M, Srinivasan R, Henry RG, Nelson SJ, et al. Estimation of metabolite T1 relaxation times using tissue specific analysis, signal averaging and bootstrapping from magnetic resonance spectroscopic imaging data. MAGMA 2007; 20: 143–55.

Salter MG, Fern R. NMDA receptors are expressed in developing oligodendrocyte processes and mediate injury. Nature 2005; 438: 1167–71.

Sawcer S. The complex genetics of multiple sclerosis: pitfalls and prospects. Brain 2008; 131: 3118–31.

Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. Neuroimage 2010; 51: 542–54.

Shirakawa H, Katsuki H, Kume T, Kaneko S, Akaike A. Pregnenolone sulphate attenuates AMPA cytotoxicity on rat cortical neurons. Eur J Neurosci 2005; 21: 2329–35.

Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 2002; 17: 479–89.

Srinivasan R, Cunningham C, Chen A, Vigneron D, Hurd R, Nelson S, et al. TE-averaged two-dimensional proton spectroscopic imaging of glutamate at 3 T. Neuroimage 2006; 30: 1171–8.

Srinivasan R, Sailasuta N, Hurd R, Nelson S, Pelletier D. Evidence of elevated glutamate in multiple sclerosis using magnetic resonance spectroscopy at 3 T. Brain 2005; 128: 1016–25.

Stefansson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigmundsson T, Ghosh S, et al. Neuregulin 1 and susceptibility to schizophrenia. Am J Hum Genet 2002; 71: 877–92.

Stys PK. General mechanisms of axonal damage and its prevention. J Neurol Sci 2005; 233: 3–13.

Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. Genomics 2008; 92: 265–72.

Valenzuela CF, Partridge LD, Mameli M, Meyer DA. Modulation of glutamatergic transmission by sulfated steroids: role in fetal alcohol spectrum disorder. Brain Res Rev 2008; 57: 506–19.

Vukusic S, Hutchinson M, Hours M, Moreau T, Cortinovis-Tourniaire P, Adeleine P, et al. Pregnancy and multiple sclerosis (the PRIMS study): clinical predictors of post-partum relapse. Brain 2004; 127: 1353–60.

Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet 2007; 81: 1278–83.