Association of a null allele of SPRN with variant Creutzfeldt–Jakob disease

J A Beck, T A Campbell, G Adamson, M Poulter, J B Uphill, E Molou, J Collinge, S Mead

ABSTRACT

Background: No susceptibility genes have been identified in human prion disease, apart from the prion protein gene (PRNP). The gene SPRN, encodes Shadoo (Sho, shadow of prion protein) which has protein homology and possible functional links with the prion protein.

Methods: A genetic screen was carried out of the open reading frame of SPRN by direct sequencing in 522 patients with prion disease, including 107 with variant Creutzfeldt–Jakob disease (vCJD), and 861 healthy controls.

Results: A common coding variant of SPRN, two further single nucleotide polymorphisms (SNPs) and three rare insertion or deletion variants were found. A single base-pair insertion at codon 46, predicted to cause a frameshift and potentially a novel protein, was found in two patients with vCJD but not in controls (p = 0.01). Two linked SNPs, one in intron 1 and the other a missense variant at codon 7, were associated with risk of sporadic CJD (p = 0.009).

Conclusion: These data justify the functional genetic characterisation of SPRN and support the involvement of Shado in prion pathobiology.

Despite clear evidence from mouse linkage studies of multiple genetic loci affecting incubation periods of prion diseases, no specific human genes have been identified, apart from the prion protein gene (PRNP). A recently characterised highly conserved gene, SPRN, encodes Shadoop (Sho, shadow of prion protein) which has protein homology and possible functional links with the prion protein.

Human prion diseases are a phenotypically diverse range of fatal neurodegenerative disorders that are unique in having sporadic, acquired and inherited aetiologies. Common to the pathogenesis of all prion diseases is a central role for the autocatalytic misfolding of the prion protein (PrP). These uncommon neurodegenerative diseases have come under close scrutiny in recent years because of a public health threat related to the epizootic of bovine spongiform encephalopathy (BSE), its human counterpart variant Creutzfeldt–Jakob disease (vCJD) and sCJD by blood transfusion. Sporadic forms of human prion disease account for approximately 85% of total prion disease cases, most typically appearing as a rapidly progressive multifocal dementia with myoclonus (sCJD). Multiple distinct prion strains in sporadic disease associate with a variety of clinicopathological phenotypes. Conversely, vCJD is associated with a distinct and unique prion strain, and has been described to date in around 200 patients worldwide with most (>160 cases) occurring in the UK. As a group, patients with vCJD differ from those with sCJD in their relatively early age of onset and longer duration, prominent psychiatric and sensory features at presentation, and characteristic neuropathology and prominent lymphoreticular prion infection.

Although there is only one known familial concurrence of vCJD, there is a strong likelihood that genetic susceptibility has played a part in determining why particular individuals succumbed following the widespread population exposure to dietary BSE prions. Epidemiological case control studies have shown no clear-cut evidence for unusual dietary or occupational exposure. PRNP is polymorphic in Europeans at codon 129, between methionine (~60% allele frequency) and valine. All clinical vCJD cases examined have been homozygous for methionine, representing the strongest association of a common genotype with any disease. Support for the existence of novel genetic factors is derived from mouse quantitative locus studies, which have shown multiple regions, unlinked to Prnp, that control the incubation time to BSE prions.

The normal function of PrP, its molecular interactions in the healthy cell, and its pathogenesis and neurotoxic pathways remain unclear. Proteins encoded by genetic loci that modify susceptibility to, or incubation time in, prion disease might do so by involvement with the normal functional pathways of PrP, either as a ligand or homologue, and several putative PrP ligands have been proposed. The recently characterised prion protein family, including Doppel and Shadoo, are pre-eminent candidate functional homologues of PrP, but several lines of evidence contradict involvement of PRND genetic variation and Dpl in human prion disease.

SPRN, the gene encoding Sho, was identified by comparative gene analysis, and this gene and its encoded protein have several characteristics suggestive of a role in prion biology. The SPRN gene comprises two exons, the latter containing the entire open reading frame (ORF). Also in common with PRNP, high levels of conservation of this gene exist between many species from fish to mammals, with highest levels of SPRN conservation present in its hydrophobic domain. In addition, both proteins are predicted to have glycosphosphatidylinositol (GPI) anchor attachment, a single transmembrane domain and a highly conserved N-terminal sequence. Recent work has shown overlapping expression patterns of PrP and Sho in the CNS, the similarity of Sho transgenes to PrP in...
a role for SPRN to the patients with vCJD, 137 UK healthy blood donor consisted of 92 healthy UK controls matched for age and sex clinical criteria and imaging findings. The control cohort biopsy, post-mortem analysis of brain tissue, or established WHO diagnostic criteria and were diagnosed either by tonsillar definite or probable patients with vCJD studied also met Health Organization (WHO) diagnostic criteria. The 107 who were definite or probable cases of sCJD according to World and a control group. The sCJD group comprised 415 patients, participants gave informed consent.

Neurology and National Hospital for Neurology and Ethics approval was obtained from the UCL Institute of 

METHODS
Ethics approval was obtained from the UCL Institute of Neurology and National Hospital for Neurology and Neurosurgery Local Research Ethics Committee, and all participants gave informed consent. The study comprised patients with sCJD, patients with vCJD and a control group. The sCJD group comprised 415 patients, who were definite or probable cases of sCJD according to World Health Organization (WHO) diagnostic criteria. The 107 definite or probable patients with vCJD studied also met WHO diagnostic criteria and were diagnosed either by tonsillar biopsy, post-mortem analysis of brain tissue, or established clinical criteria and imaging findings. The control cohort consisted of 92 healthy UK controls matched for age and sex to the patients with vCJD, 157 UK healthy blood donor controls, 446 UK healthy controls from the European Collection of Cell Cultures (ECACC) collection, and 186 unrelated healthy control individuals from the Centre d’Etude du Polymorphisme Humain (CEPH) collection.

We performed direct gene resequencing on peripheral blood samples from cases and controls to investigate the possibility of a significant association between genetic variation within the SPRN ORF and susceptibility to sCJD or vCJD. Genomic DNA was isolated using a commercial extraction kit (Nucleon BACC2; GE Healthcare UK, Little Chalfont, Buckinghamshire, UK) following the supplied protocol. DNA concentrations were determined using a spectrophotometer (Nanodrop ND-1000; Thermo Fisher Scientific, Waltham, Massachusetts, UK), and diluted in Tris/EDTA buffer in aliquots of 20 ng/μl to prepare 96-well plates for PCR. For DNA sequencing, oligonucleotide primers were designed to amplify two overlapping amplicons to span the entire ORF. Genomic DNA was amplified with sense 5′-CTTCTGCTCTCAGCTTG-3′ (8404-8422) and antisense 5′-AGCCCGCGGCCAGGCC-3′ (599-626) described below). Amplicon 1 was 359 bp (reference to AL161645 are given in parentheses after each primer accession number AL161645 (exact location of primers with

Table 1 Genetic variants found in the study

<table>
<thead>
<tr>
<th>Variation</th>
<th>DNA</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-11A→G</td>
<td>A:G</td>
<td>vCJD</td>
</tr>
<tr>
<td>c.20C→T</td>
<td>T:C</td>
<td>sCJD</td>
</tr>
<tr>
<td>c.136_137insG</td>
<td>wt:ins</td>
<td>HC</td>
</tr>
<tr>
<td>c.183C→T</td>
<td>C:T</td>
<td>*1</td>
</tr>
<tr>
<td>c.216_227delAGC</td>
<td>wt:del</td>
<td>*1A</td>
</tr>
<tr>
<td>c.239_240insGCC</td>
<td>wt:ins</td>
<td>*1B</td>
</tr>
<tr>
<td>c.293_294insGCC</td>
<td>wt:ins</td>
<td>*2</td>
</tr>
<tr>
<td>c.302_303insGCC</td>
<td>wt:ins</td>
<td>*2A</td>
</tr>
<tr>
<td>c.310_311insGCC</td>
<td>wt:ins</td>
<td>*2B</td>
</tr>
<tr>
<td>c.321_322insGCC</td>
<td>wt:ins</td>
<td>*2C</td>
</tr>
</tbody>
</table>

Alleles and allele counts, together with a key to the allele counts, are displayed horizontally. Insertion and deletion alleles are coded as wild type (wt) or ‘ins/del’ depending on the absence or presence of the indel variation compared with contig AL161645, respectively.

Table 2 Results from tests performed using genotypic and allelic (trend) models

<table>
<thead>
<tr>
<th>Variation</th>
<th>MAF</th>
<th>Genotypic</th>
<th>Allelic (trend)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-11A→G/c.20T→C</td>
<td>0.29</td>
<td>vCJD</td>
<td>Counts</td>
</tr>
<tr>
<td>c.136_137insG</td>
<td>0.01</td>
<td>sCJD</td>
<td>30/172/213</td>
</tr>
<tr>
<td>c.183C→T</td>
<td>0.33</td>
<td>Control</td>
<td>81/405/375</td>
</tr>
<tr>
<td>c.216_227delAGC</td>
<td>0.01</td>
<td>vCJD</td>
<td>0/2/105</td>
</tr>
<tr>
<td>c.239_240insGCC</td>
<td>0</td>
<td>sCJD</td>
<td>0/0/415</td>
</tr>
<tr>
<td>c.293_294insGCC</td>
<td>0</td>
<td>Control</td>
<td>0/0/861</td>
</tr>
</tbody>
</table>

Denotes Fisher Exact test applied because cell counts were small. Genotypic counts and allele counts are displayed AA/AB/BB and A/B respectively, where A = 5′-11G (c.20C) or c.136_137insG. Bold type denotes significance.
and amplified with sense 5′-TTTCTCTCTGCACACAGCGCG-3′ (8257–8276) and antisense 5′-GCGATGGAGCGTGGC-TGGG-3′ (7566–7585). PCR reactions containing a final concentration of 1 mol/l betaine were thermally cycled with an initial denaturation step at 95°C for 5 minutes, followed by 35 cycles of 95°C for 30 s, 65.5°C (amplicon 1) or 63.1°C (amplicon 2) for 30 s and 72°C for 45 s, and then a final extension step at 72°C for 5 minutes. Cleaned PCR products were sequenced in reactions including betaine 1 mol/l (final concentration) and separated by electrophoresis on an automated sequencer (3130xl; Applied Biosystems, Foster City, California, USA). Data analysis was performed using Seqscape software V.2.5 (Applied Biosystems) with analysis filter settings adjusted to allow assembly of poor data due to insertions of deletions, maximum mixed bases 95%, maximum uncalled bases 95%, minimum clear length of 1 bp, and minimum sample score of 1). Poor data or failed reactions were removed from projects by visual inspection. To assess the expression level of the frameshift allele (p.Ala46GlyfsX294) in postmortem frontal cortex, total RNA was extracted from 100 mg of frozen tissue using (TRIzol Plus Purification kit; Invitrogen, Paisley, UK). A commercial kit (OmniScript Reverse Transcription Kit; Qiagen, Crawley, West Sussex, UK) and random hexamers (Invitrogen) were used to produce cDNA. Generation of a fluorescently labelled PCR product of both genomic DNA (gDNA) and cDNA was performed in both case and controls (suitable case material) and amplified with sense 5′-AGGTGCGCGGGGCAGT-3′ (8348–8364) and antisense 5′-CTCACGCACCCCTCAGA-3′ (8296–8311) primers. The resultant 69 bp amplicon was separated by electrophoresis on an automated sequencer (3130xl; Applied Biosystems) with LI500 size standard and peak size, height and area from both gDNA and cDNA were calculated using GeneMapper software V.4.0 (Applied Biosystems).

Statistical analysis

Statistical and genetic analyses were performed using Haploview,20 PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml) and SSPS software. For sCJD, the analysis was 80% powered to detect a heterozygous genotype relative risk of 1.25 in a multiplicative model. Haplotype frequencies were estimated with an expectation–maximisation (EM) algorithm.

RESULTS

Six different genetic variations were detected in the 2766 chromosomes sequenced (table 1). Of these, three were of an insertion or deletion (indel) variety, two were non-coding single-nucleotide polymorphisms (SNPs) either inside or outside the SPRN ORF, and one was a coding SNP. This common missense variant was found in the SPRN signal peptide sequence (c.20C→T, p.T7M), and is linked to a G→A transversion 11 bp 5′ of the SPRN translation start methionine. We found 5′-11A linked to p.7M and 5′-11G linked to p.7T. These alleles form the basis of derived haplotypes designated *1 and *2 respectively (table 1). Within this haplotype scheme, the *1 related haplotypes were all associated with c.20T and c.183C of the SPRN ORF, whereas the *2 related haplotypes were associated with c.20C and either position c.183C (haplotypes *2A and *2C) or c.183T (haplotypes *2 and *2C). A 12 bp deletion variation was also detected in four healthy controls (three on the *1 haplotype and one on the *2 haplotype), corresponding to the in-frame deletion of an AAAG amino acid repeat motif located in the highly conserved hydrophobic central domain of Sho, reducing the polypeptide chain to 147aa from the most commonly observed length of 151aa. These variants were designated haplotypes *1A and *2C respectively. A second similar indel variation was detected in a single CEPH control, and involved a 12 bp insertion in the hydrophobic domain resulting in an additional AAAG amino acid repeat motif, increasing the full length protein to 155aa, and was designated haplotype *1B. The third indel variation detected in two patients with vCJD only was an insertion of a single guanine base pair at codon 46. This frameshift mutation creates a potential extended ORF of 823 bp, designated haplotype *2B.

Positive results of statistical analysis between genotypes and alleles of cases versus controls are shown in table 2.

The SNPs 5′-11A→G and c.20T→C were found to be in complete linkage disequilibrium and were therefore grouped. A genotypic model gave p = 0.028, and an allelic model gave p = 0.009, comparing sCJD with controls where an increase of 5′-11A and linked c.20T was observed in cases. A similar increase was observed in vCJD cases although the small numbers meant this did not reach significance. An allelic model was used with the Fisher Exact test when analysing the level of significance of the frameshift allele detected in two vCJD cases (p = 0.01, table 2).

SPRN genotype data was also stratified by PRNP codon 129 status (table 3). No significant association was found between control codon 129 status and SPRN genotype and therefore control data is shown unstratified. Stratification of sCJD SPRN genotype by codon 129 status showed a strongly significant association between c.20T and codon 129 methionine homozygous sCJD (p = 0.007). Further stratification of SPRN

---

### Table 3: Stratification of SPRN genotypes in patients with sporadic Creutzfeldt–Jakob disease

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Control</th>
<th>sCJD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>375</td>
<td>131</td>
</tr>
<tr>
<td>TC</td>
<td>405</td>
<td>98</td>
</tr>
<tr>
<td>CC</td>
<td>81</td>
<td>16</td>
</tr>
<tr>
<td>p Value</td>
<td></td>
<td>0.007</td>
</tr>
</tbody>
</table>

sCJD, sporadic Creutzfeldt–Jakob disease.

Because SPRN SNPs at positions intron 1 5′-11 and c.20 show complete linkage, the genotypes shown are representative of either genotype 5′-11A/5′-11G, or c.20T/c.20C.

---

### Table 4: Stratification of phenotypic data in patients with variant Creutzfeldt–Jakob disease

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Age at disease onset (years)</th>
<th>Duration of disease (months)</th>
<th>Year of onset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Median</td>
</tr>
<tr>
<td>TT</td>
<td>29.8</td>
<td>12.2</td>
<td>28</td>
</tr>
<tr>
<td>TC</td>
<td>29.6</td>
<td>10.3</td>
<td>25</td>
</tr>
<tr>
<td>CC</td>
<td>33.5</td>
<td>9.8</td>
<td>29</td>
</tr>
</tbody>
</table>

SD, standard deviation.
genotype data was performed on vCJD cases by age at onset of disease, duration and year of onset as a measure of susceptibility within the vCJD group (table 4), but no significant associations with these aspects of phenotype were detected. Finally, sCJD SPRN genotype data was stratified by age at onset and duration of disease and PrP<sup>Sc</sup> type (table 5). Again, no significant associations were detected.

Unfortunately, RNA recovered from the single vCJD case possessing the frameshift mutation was of insufficient quality for our assessment of expression level of mutant transcripts. The year of death of the two frameshift mutation patients with vCJD was 1996 and 1995, towards the start of the vCJD epidemic. The clinical history of the patient who died in 1996 and was examined in detail involved depression, behavioural disturbance, altered mood with emotional lability, cognitive decline and progressive chorea of 12 months’ duration. Clinical examination was notable for global rigidity, generalised chorea, distension, altered mood with emotional lability, cognitive disturbance, altered mood with emotional lability, cognitive decline and progressive chorea of 12 months’ duration. Clinical examination was notable for global rigidity, generalised chorea, a broad-based gait and attentional deficits. This patient died aged 28 years. Fewer details are available for the second patient, who resided in different areas and was examined in detail involved depression, behavioural disturbance, altered mood with emotional lability, cognitive decline and progressive chorea of 12 months’ duration.

**DISCUSSION**

We report sequencing of the SPRN ORF in all available samples of UK patients with vCJD, 415 patients with definite or probable sCJD and 861 healthy controls. We found a significant association between a frameshift mutation and vCJD (p = 0.01) and between c.20T and sCJD (p = 0.009). Three major haplotypes exist in the healthy control sample tested and are defined by the three SNPs detected: A/T/C, G/C/T, G/C/C (5′-11A/c.20T/c.183, respectively). Haplotype frequencies were approximately 0.69, 0.21, and 0.10 for these haplotypes, respectively (averaged from cases and controls; table 1). Additional haplotypes exist at low frequency and are associated with indel variations of the hydrophobic core in healthy controls and with a frameshift mutation in two patients with vCJD. Given the low frequency of indel variants discovered in healthy controls, it is not surprising that we did not detect a similar polymorphism in the cases tested.

Our investigations have revealed genetic variants that were discovered in cases only, cases and controls, or controls only. The effect of these variants on Sho biology or prion biology is unknown, but the nature of each and their location within SPRN may provide insight into their likely functionality. The indel variations discovered in five healthy controls all affect the size of the hydrophobic core of the Sho protein from the most common size of 23aa to either 19aa or 27aa depending on alleles containing a 12 bp deletion or insertion respectively. Missense mutations of the hydrophobic core of PrP are associated with disease and have been shown to lead to a transmembrane topological variant of PrP with the C-terminus being luminal and the N-terminus cytoplasmic. The significance of expansion or contraction of the hydrophobic core by insertion or deletion of 4aa in Sho remains unknown, although our finding of such variants in healthy controls suggests toleration of these in Sho.

The allele we report associated with sCJD, for which we also report a trend towards the same allele in vCJD, has two genetic variations that determine haplotype in our scheme. 1<sup>st</sup> haplotypes are associated with 5′-11A/c.20T (p.7M) and we have shown that these are significantly more common in sCJD. It remains possible that either 5′-11A→G or c.20T→C are functional SNPs, or indeed that an as yet undetermined functional polymorphism exists in unscreened regions of SPRN. The SNP 5′-11A→G is located in the short 16 bp untranslated region (UTR) of SPRN exon 1 and has a probability of 0.82 that alternative splicing may occur when guanine is present at bp 5′-11 as determined by Netgene2 (http://www.cbs.dtu.dk/services/NetGene2/) analysis. Conversely, no alternative splicing is predicted for 5′-11A. The biological effect of excision of 5 bp of 5′UTR close to the ORF is unknown, and further complexity exists in deciphering the functional element linked to sCJD due to c.20C→T. In this, an increase of the p.7M allele was found in sCJD, and codon 7 is predicted to be contained within the signal peptide sequence of SPRN. The frequency of these alleles in healthy controls is strongly suggestive that no deleterious transport of Sho occurs due to methionine at residue 7 of the signal sequence, but this does not obviate the effect of subtle influences on protein transport upon establishment and progression of prion disease. Indeed, in a mouse model exploring the role of mutations in the signal peptide sequence of PrP on hydrophobic tract mutants, it has been shown that the double mutant L9R-3AV resulted in abnormal PrP trafficking.

The effects of our final finding of a significant association of a frameshift mutation with vCJD is more easily rationalised. Insertion of a single guanine at codon 46 of SPRN would cause a putative polypeptide chain of 294aa if translation of mutant transcript occurred, and a corresponding reduction in normal Sho protein. Nonsense-mediated mRNA decay (NMD) is a mechanism by which transcripts containing premature translation termination codons are degraded, and as most frameshift mutations result in the presence of a premature termination codon, most mRNA transcripts containing a frameshift mutation are predicted to be degraded by NMD. Despite our efforts to assess the levels of the p.Ala46GlyfsX294 message from frontal cortex tissue of a single patient with vCJD, owing to prolonged post-mortem delay and several sample freeze–thaw cycles we were unable to show whether or not the transcript exists at all, or whether it exists at reduced or equal measure to the wild-type allele. Nevertheless, the presence of a rare frameshift allele in two patients with vCJD remains of interest, particularly as these patients were amongst the most susceptible patients based on time of onset and were among the first British patients. In all other respects, the phenotype of these two cases was unremarkable within the larger vCJD cohort. The possibility that these two patients, who resided in different areas, presented with the earliest diagnoses of vCJD, and the frameshift allele was found in patients with disease onset in 1996 and 1995, is certainly of interest.
parts of the UK, were closely related, was excluded by genome-wide genotype data (500K arrays; Affymetrix, Santa Clara, California, USA; data not shown). More definitive genetic evidence that SPRN null alleles cause susceptibility to vCJD would require an allelic series, which we were unable to show with our necessarily small sample size.

Although confirmation of our findings in further non-UK cases of vCJD and sCJD would be extremely valuable, our data support the hypothesis that SPRN genetic variants are involved in the pathobiology of prion disease. Interactions between PrP and Sho in health and disease are not yet clear, but these genetic data are consistent with an existing model that proposes a neuroprotective role for Sho.5 Our associations of a frameshift mutation, which may lead to a null allele, with vCJD, and *1 SPRN haplotypes with sCJD, are consistent with a direct or indirect interaction between proteins in the disease process. However, the genetic evidence is not straightforward in that different alleles were associated with two categories of human prion disease. Whether this finding relates to our necessarily small sample sizes, as the c.20T allele was more common, but not significantly so, in vCJD, or alternatively reflects the distinct pathogenesis of sporadic and vCJD, is unresolved. Further investigations are required to explore these possibilities, and in particular, creation of null or transgenic cellular or mouse models of SPRN.

Acknowledgements: This work was funded by the MRC and undertaken at the University College London Hospitals and University College London, who received a proportion of funding from the Department of Health’s National Institute for Health Research Biomedical Research Centres funding scheme. We are grateful to D Kaski for providing clinical information. We thank all clinical colleagues who referred patients to the NHS National Prion Clinic, all the patients and their families for kind consent to use their data, and in particular, Onwuazor ON, Beck J, Collinge J, Fisher E. Identification of multiple quantitative trait loci linked to prion disease incubation period in mice. *Proc Natl Acad Sci USA* 2001;98:8279–83.


