ARTICLE

A refined physical and transcriptional map of the *SPG9* locus on 10q23.3–q24.2

Cristiana Lo Nigro[1], Roberto Cusano[1], Monica Scaranari[1], Roberta Cinti[1], Paola Forabosco[1], Vincenzo Brescia Morra[2], Giuseppe De Michele[2], Lucio Santoro[2], Sally Davies[3], Jane Hurst[4], Marcella Devoto[5], Roberto Ravazzolo[1,5] and Marco Seri[1]

[1]*Laboratory of Molecular Genetics, Gaslini Institute, Genoa, Italy;* [2]*Department of Neurological Science, Federico II University, Naples, Italy;* [3]*Institute of Medical Genetics, University Hospital of Wales, Cardiff, UK;* [4]*Department of Clinical Genetics, Oxford Radcliffe Hospital Trust, The Churchill, Oxford, UK;* [5]*Department of Oncology, Biology and Genetics, University of Genoa, Italy*

Hereditary spastic paraplegia (HSP) is a genetically heterogeneous disorder characterised by progressive spasticity of the lower limbs. Beside 'pure' forms of HSP, 'complicated' forms are reported, where spasticity occurs associated with additional symptoms. We recently described an Italian family with a complicated dominant form of HSP (SPG9) and we mapped the gene responsible to 10q23.3–q24.2, in a 12 cM interval between markers D10S564 and D10S603. The phenotypic manifestations in our family are reminiscent of those already described in a smaller British pedigree. We typed individuals from this British family using markers located in the *SPG9* critical interval and haplotype reconstruction showed the disorder co-segregating with SPG9. To characterise the *SPG9* region better, we constructed a contig of 22 YACs, assigned it to 18 polymorphic markers and positioned 54 ESTs. Furthermore, we searched for ESTs containing a trinucleotide repeat sequence, since anticipation of symptoms was reported in both families. Finally, analysis of a muscle biopsy specimen from one patient was normal, suggesting that, contrary to SPG7, mitochondrial disturbance could not be a primary feature of SPG9. *European Journal of Human Genetics* (2000) 8, 777–782.

Introduction

Hereditary spastic paraplegia (HSP) is a genetically heterogeneous group of inherited neuro-degenerative disorders, with an estimated prevalence ranging from 3 to 10×10^{-5}.[1] Autosomal dominant, autosomal recessive and X-linked forms of HSP have been described.[2] The 'pure' forms of HSP are clinically characterised by lower limb slow progressive weakness and spasticity, hyper-reflexia and Babinski signs. The age of onset and severity of the disorder vary even among individuals from the same family.[2,3] Many patients exhibit 'complicated' forms of HSP, which are characterised by the presence of additional neurological and non-neurological symptoms.[2]

Recently, we described a form of dominant spastic paraplegia in a large Italian pedigree, complicated by bilateral cataracts, gastro-oesophageal reflux with persistent vomiting and amyotrophy.[4] The considerable variability in the age of onset within a family and in the severity of symptoms observed was suggestive of genetic anticipation. Our previous genome-wide screening located this disorder on chromosomal band 10q23.3–q24.2, in a 12 cM interval between markers D10S564 and D10S603. This locus was named SPG9 (MIM601162) and represents the first autosomal dominant form of complicated HSP to be mapped in the human genome.

In 1996, Slavotinek *et al*[5] reported a pedigree with recurrence of congenital cataracts, motor system disorder, short stature, learning difficulties and skeletal abnormalities

(MIM 601162); the similarity of phenotypic manifestations suggested that both families could be affected by the same genetic disorder.

We report here the linkage analysis in this second pedigree and the construction of a combined physical and transcriptional map, towards identification of the disease gene.

Materials and methods

Patients
Clinical features of the Italian pedigree and of the smaller British pedigree have been described previously.[4,5]

Genotyping
Genotyping was performed in the British family using polymorphic markers (http://www.genethon.fr/) at the *SPG9* locus for the available DNA samples. PCR was carried out under standard conditions, labelling the forward primer with fluorescent amidites; amplification products were loaded into an automated sequencer ABI mod. 377 in order to define allele size using GeneScan software.

Linkage analysis
Classic two-point lod score analysis was carried out on the British pedigree, under the assumption of autosomal dominant inheritance with 90% penetrance of the segregating trait and an estimated disease allele frequency of 0.001. Linkage analysis was performed using the MLINK program included in the LINKAGE package.[6]

Physical map
Eighteen polymorphic markers were used to construct a contig of 22 YAC clones. DNA extraction from YACs and PCR amplifications were carried out using standard protocols.[7]

Transcript map
Fifty-four ESTs among those mapped in GeneMap'99[8] between D10S564 and D10S603 were tested with all YAC clones within the SPG9 contig, following standard conditions of PCR amplification. Primer sequences were retrieved from the Genome Database (http://gbwww.dkfz-heidelberg.de/) or the Stanford Human Genome Center (http://www-shgc.stanford.edu/).

Analysis of muscle biopsy
Muscle biopsy was obtained from left biceps brachii muscle. A muscle specimen was frozen in isopentane, cooled in liquid nitrogen. Serial cryostatic sections (10 μm thick) were stained with hematoxylin and eosin, modified Gomori trichrome, and the histochemical reactions for routine ATPase, ATPase with pre-incubation at pH 4.6 and 4.3, NADH-TR, PAS, Sudan black, non-specific esterase, phosphorylase, cytochrome *c* oxidase, and succinate dehydrogenase were performed.

Results

Genotyping and linkage analysis in the British pedigree
Linkage analysis using markers located within the SPG9 critical interval and haplotype reconstruction in individuals from the British pedigree showed that all affected individuals shared the same 'at risk' haplotype at 10q23.3–q24.2 (Figure 1). Unfortunately, no recombination event occurred in the members of this second family, thus permitting no further restriction of the critical region. Given the limited number of meiotic events observed (six meioses, one phase unknown), the maximum lod score that can be obtained in the linkage analysis is only 1.50, and this expected maximum is reached in the 2-point linkage analysis at informative markers and in the multipoint linkage analysis on the whole region (data not shown).

Construction of a YAC contig and genetic refinement of the *SPG9* locus
To characterise the *SPG9* region in more detail, we selected 16 overlapping YAC clones from the STS-formatted YAC contig WC10.7 available at the Whitehead Institute/MIT (http://www.genome.wi.mit.edu/) and previously described contigs.[9,10] In addition, six YACs (967 G 4, 885 H 11, 764 E 1, 709 A 1, 809 A 9, 890 B 11) were selected from the CEPH human Mega-YAC library to improve clone coverage in some portions of the contig. We assigned a total of 18 CA repeat markers to our YAC contig by PCR amplification and we were able to put 9 of them in clear order (Figure 2).

On the basis of the physical map produced in this study, the relative order of markers at the centromeric boundary is D10S564, D10S536, D10S583 and D10S1755. Therefore the recombination event in the Italian pedigree actually occurred between markers D10S536 and D10S583. The SPG9 critical region thus lies between the recombinant markers D10S536 and D10S603, and its physical length is estimated to be about 6.5–7.0 Mb.

Transcript map
The *SPG9* critical region appears to be gene-rich and several gene loci and expressed sequence tags (ESTs) have been determined within 10q24.[9–11] Different neurological disorders and other diseases have been mapped to this region, the critical intervals of some of them overlapping a large part of the *SPG9* locus (Figure 2).

Moreover, sequencing of a significant proportion of this region is in progress at the Sanger Centre (http://www.sanger.ac.uk/), as part of the Human Genome Project. The entire SPG9 locus is included in Chr 10ctg15, where more than 80 genomic clones are located between markers D10S536 (located in clone bA56M3) and D10S603 (in clone bA34D15).

In particular, 215 ESTs were found in GeneMap'99[8] between markers D10S564 and D10S603. To identify transcribed sequences potentially implicated in *SPG9*, we selected ESTs expressed in brain or with a possible function in brain
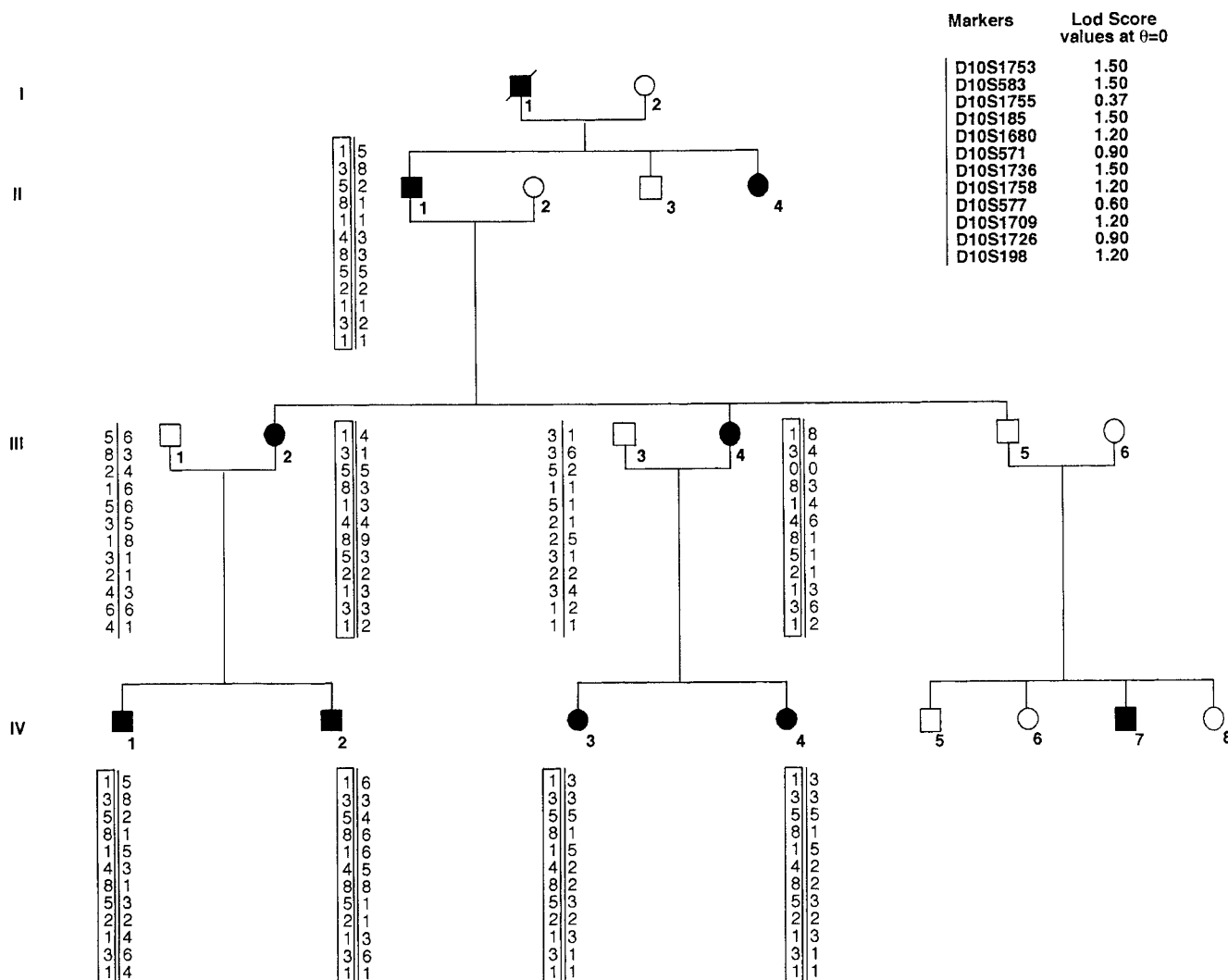
Markers	Lod Score values at θ=0
D10S1753	1.50
D10S583	1.50
D10S1755	0.37
D10S185	1.50
D10S1680	1.20
D10S571	0.90
D10S1736	1.50
D10S1758	1.20
D10S577	0.60
D10S1709	1.20
D10S1726	0.90
D10S198	1.20

Figure 1 Haplotype reconstruction in the British pedigree. Individual IV-3, previously considered healthy, was found on clinical evaluation to be affected. Haplotype reconstruction was performed using 12 informative markers at the SPG9 locus for the available DNA samples. The 'at-risk' haplotype is boxed. Lod score values at θ = 0 are reported. Markers D10S583, D10S185 and D10S1736 gave a maximum combined lod score of 7.15 in the two pedigrees.

and EST clones apparently containing trinucleotide repeats. In total, we positioned 54 ESTs within the YAC contig, 22 of which represent known genes (HTR7, PPP1R5, TAF-172, IDE, HHEX, KNSL1, RPB4, LGI1, PDE6C, CYP2C, CLIM1, CD39, DNTT, MEGF4, FRAT1, HPS, CPN1, CMOAT, COX15, GOT1, CYB561, SCD) and 32 are unidentified transcripts (Figure 2). Further details on EST mapping and expression are available on request from genseq@tin.it.

Search for ESTs containing a trinucleotide repeat

As we observed genetic anticipation of symptoms in both families, we collected ESTs H06060 and R42638 (IMAGE clones 43585 and 30865), already reported to map to the critical region and to contain a CAG and a CTG repeat,

respectively.[12] Both clones were completely sequenced and failed to reveal any trinucleotide repeat sequences.

Experiments using the Repeat Expansion Detection (RED) procedures, following standard protocols,[13] gave the same negative results (data not shown).

Muscle biopsy analysis

Since analysis of muscle biopsies from patients carrying mutations in paraplegin showed typical signs of mitochondrial OXPHOS defects,[14] we searched the *SPG9* region for transcripts encoding proteins with mitochondrial activity. Interestingly, CYP2C and COX15 have been physically mapped to our YAC contig. Nonetheless, morphological and histochemical examination of a muscle biopsy specimen from the left biceps brachii of one patient from the Italian
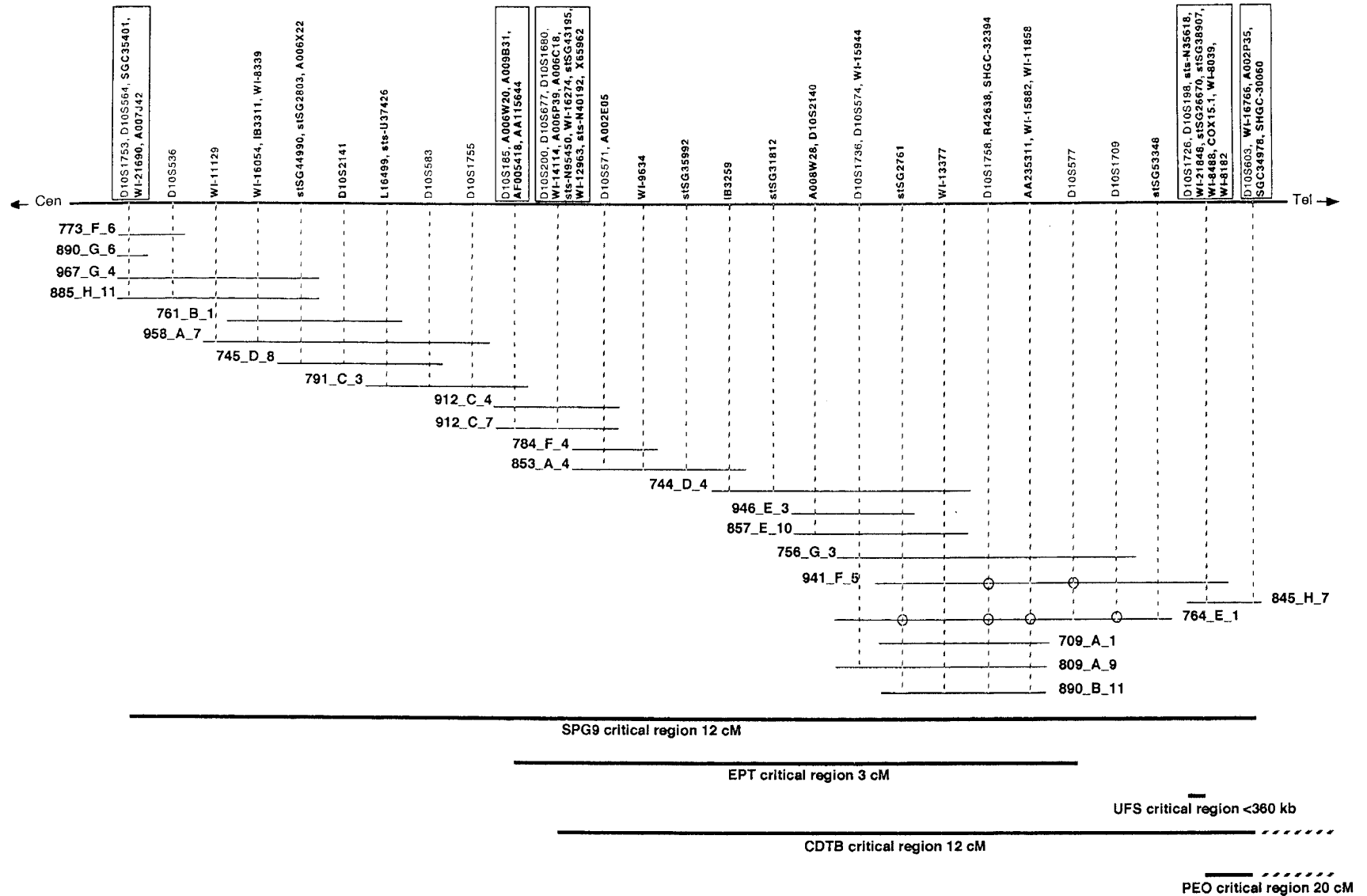
Figure 2 Combined physical and transcriptional map of the SPG9 locus. A library reference number identifies each YAC. Tel = telomere; Cen = centromere. YACs 941 F 5 and 764 E 1 denote the presence of rearrangements during clone propagation or chimerism, since some inconsistencies are present in EST assignments (open circles). ESTs assigned to YACs are printed in bold; relative order between markers and ESTs inside boxes is not defined. Position of 18 Généthon markers are also reported. At the bottom of the figure, bars indicate the critical region for *SPG9*, for partial epilepsy (EPT, MIM600512), urofacial syndrome (UFS, MIM236730), Thiel-Behnke corneal dystrophy (CDTB, MIM602082) and progressive external ophthalmoplegia (PEO, MIM157640).

pedigree revealed no 'ragged-red' fibres and normal staining for NADH, SDH and COX.

Discussion

Our previous study on the Italian pedigree localised the *SPG9* gene to a 12 cM region defined by markers D10S564 and D10S603.[4] In the present study, the genotyping of individuals from the family described by Slavotinek *et al*[5] and haplotype reconstruction showed this second family to co-segregate with the SPG9 locus, suggesting that the *SPG9*-linked HSP could represent a genetically homogeneous syndrome.

Here, we constructed a YAC contig of the 6.5–7 Mb critical region and assigned 18 polymorphic markers to it, putting nine of them in clear order. On the basis of this physical map, we determined the SPG9 critical region to lie between the recombinant markers D10S536 and D10S603.

In order to undertake a 'positional candidate gene approach' for identification of the disease gene, we assigned a total of 54 ESTs on the YAC contig. Some of them were selected because of their expression in brain, because they apparently contain a trinucleotide repeat sequence, or they might belong to nuclear genes encoding proteins with mitochondrial function.

Genetic anticipation of the age of onset had been observed in the two families, suggesting that this form of the disease may be caused by expansion of a trinucleotide repeat. No evidence of the presence of a trinucleotide repeat sequence within the SPG9 critical region came out of the present study. Similarly, genetic anticipation, though much debated, had been reported in several families with autosomal dominant HSP linked to the SPG4 locus.[15] Hazan *et al*[16] exhaustively searched for CAG repeat sequences in the entire SPG4 genomic interval and failed to detect any. In fact, the identification of *SPAST*,[17] which is the gene responsible for SPG4 HSP, definitively proves that expansion of repeat sequences is not involved in the pathogenesis of the disorder.

Casari *et al*[14] identified a novel nuclear-encoded mitochondrial protein, paraplegin, which is mutated in an autosomal recessive form of HSP. This finding, revealing the presence of an impairment of mitochondrial function, raised the question of whether this could be a common mechanism in HPSs and suggested muscle biopsy as a useful diagnostic tool for HSP. Following this hypothesis, we obtained, with informed consent, a muscle biopsy specimen from the left biceps brachii of one patient of the Italian pedigree. Morphological and histochemical examination revealed no 'ragged-red' fibres and normal staining for NADH, SDH and COX. These findings seem to suggest that mitochondrial disturbance is not a primary feature of *SPG9* and that defective oxidation–phosphorylation is not a common pathogenic mechanism for all genetic types of HSP as for SPG4[17] and as recently

reported for the autosomal dominant HSP linked to 8q, SPG8.[18]

In summary, we have constructed a physical map of the *SPG9* locus and achieved a high-resolution transcript map. These data, integrated with the sequences available in public databases, provide important information for the identification of the *SPG9* gene as well as for other disease genes mapped to the same region.

References

1 Filla A, De Michele G, Marconi R *et al*: Prevalence of hereditary ataxias and spastic paraplegias in Molise, a region of Italy. *J Neurol* 1992; **239**: 351–353.
2 Reid E: Pure hereditary spastic paraplegia. *J Med Genet* 1997; **34**: 499–503.
3 Durr A, Brice A, Serdaru M *et al*: The phenotype of 'pure' autosomal dominant spastic paraplegia. *Neurology* 1994; **44**: 1274–1277.
4 Seri M, Cusano R, Forabosco P *et al*: Genetic mapping to 10q23.3–q24.2, in a large Italian pedigree, of a new syndrome showing bilateral cataracts, gastro-esophageal reflux and spastic paraparesis with amyotrophy. *Am J Hum Genet* 1999; **64**(2): 586–593.
5 Slavotinek AM, Pike M, Mills K, Hurst JA: Cataracts, motor system disorder, short stature, learning difficulties, and skeletal abnormalities: a new syndrome? *Am J Med Genet* 1996; **62**: 42–47.
6 Lathrop GM, Lalouel JM, Julier C, Ott J: Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA* 1984; **81**(11): 3443–3446.
7 Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
8 Deloukas P, Schuler GD, Gyapay G *et al*: A physical map of 30,000 human genes. *Science* 1998; **282**: 744–746.
9 Gray IC, Fallowfield J, Ford S, Nobile C, Volpi EV, Spurr NK: An integrated physical and genetic map spanning chromosome band 10q24. *Genomics* 1997; **43**: 85–88.
10 Nobile C, Manca A, Pisano M, Angius A, Gray IC, Spurr NK: A refined physical and EST map spanning 7.4 Mb of 10q24, a region involved in neurological disorders. *Mamm Genome* 1998; **9**: 835–837.
11 Nobile C, Pitzalis S: Expression analysis of 21 transcripts physically anchored within the chromosomal region 10q24. *Genomics* 1999; **62**: 86–89.
12 Bulle F, Chiannilkulchai N, Pawlak A, Weissenbach J, Gyapay G, Guellaen G: Identification and chromosomal localization of human genes containing CAG/CTG repeats expressed in testis and brain. *Genome Res* 1997; **7**: 705–715.
13 Shalling M, Hudson TJ, Beutow KH, Housman DE: Direct detection of novel expanded trinucleotide repeats in the human genome. *Nat Genet* 1993; **4**: 135–139.
14 Casari G, De Fusco M, Ciarmatori S *et al*: Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell* 1998; **93**: 973–983.
15 Durr A, Davoine CS, Paternotte C *et al*: Phenotype of autosomal dominant spastic paraplegia linked to chromosome 2. *Brain* 1996; **119**: 1487–1496.

16 Hazan J, Davoine C-S, Mavel D *et al*: A fine integrated map of the SPG4 locus excludes an expanded CAG repeat in chromosome 2p-linked autosomal dominant spastic paraplegia. *Genomics* 1999; **60**: 309–319.

17 Hazan J, Fonknechten N, Mavel D *et al*: Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. *Nat Genet* 1999; **23**: 296–303.

18 Hedera P, DiMauro S, Bonilla E, Wald J, Eldevik OP, Fink JK: Phenotypic analysis of autosomal dominant hereditary spastic paraplegia linked to chromosome 8q. *Neurology* 1999; **53**(1): 44–50.