European Journal of Human Genetics (2000) 8, 946–954

www.nature.com/ejhg

ARTICLE

Mutational spectrum of the *EPM2A* gene in progressive myoclonus epilepsy of Lafora: high degree of allelic heterogeneity and prevalence of deletions

Pilar Gómez-Garre[1], Yolanda Sanz[1], Santiago Rodríguez de Córdoba[2,3] and José M Serratosa[1]

[1]*Servicio de Neurología, Fundación Jiménez Díaz;* [2]*Unidad de Patalogía Molecular, Fundación Jiménez Díaz;* [3]*Departmento de Inmunología, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain*

Progressive myoclonus epilepsy of the Lafora type (Lafora disease) is an autosomal recessive disease characterised by epilepsy, myoclonus, progressive neurological deterioration and the presence of glycogen-like intracellular inclusion bodies (Lafora bodies). We recently cloned the major gene for Lafora disease (*EPM2A*) and characterised the corresponding product, a putative protein tyrosine phosphatase (LAFPTPase). Here we report the complete coding sequence of the *EPM2A* gene and the analysis of this gene in 68 Lafora disease chromosomes. We describe 11 novel mutations: three missense (F84L, G240S and P301L), one nonsense (Y86stop), three < 40 bp microdeletions (K90fs, Ex1-32bpdel, Ex1-33bpdel), and two deletions affecting the entire exon 1 (Ex1-del1 and Ex1-del2). In addition, we have identified three patients with a null allele in non-exonic microsatellites EPM2A-3 or EPM2A-4, suggesting the presence of two distinct > 3 kb deletions affecting exon 2 (Ex2-del1 and Ex2-del2). Considering these mutations, a total of 25 mutations, 60% of them generating truncations, have been described thus far in the *EPM2A* gene. In spite of this remarkable allelic heterogeneity, the R241stop *EPM2A* mutation was found in approximately 40% of the Lafora disease patients. We also report the characterisation of five new microsatellite markers and one SNP in the *EPM2A* gene and describe the haplotypic associations of alleles at these sites in normal and EPM2A chromosomes. This analysis suggests that both founder effect and recurrence have contributed to the relatively high prevalence of R241stop mutation in Spain. The data reported here represent the first systematic analysis of the mutational events in the *EPM2A* gene in Lafora disease patients and provide insight into the origin and evolution of the different EPM2A alleles. *European Journal of Human Genetics* (2000) **8**, 946–954.

Keywords: mutation; polymorphism; *EPM2A*; Lafora; progressive; epilepsy; PTPase; genetics; myoclonus

Introduction

Among the epilepsies, the progressive myoclonus epilepsies (PMEs) constitute a rare, heterogeneous subgroup characterised by the presence of progressive neurologic deterioration, myoclonus, and epilepsy.[1] PME of the Lafora type or Lafora disease (EPM2 [MIM 254780]) is a well known form of progressive myoclonus epilepsy characterised by the presence of typical periodic acid Schiff-positive intracellular inclusion bodies (Lafora bodies).[2,3] Lafora bodies consist of an abnormal glucose polymer that accumulates in the central and peripheral nervous system, among other tissues.[4,5] Lafora disease initially manifests during adolescence, the most common age of onset being between 10 and 17 years. Generalised tonic-clonic seizures, absences, drop attacks, or partial visual seizures are usually the first manifestation, followed soon after by asymmetric as well as massive myoclonic jerks. As the disease progresses, the myoclonus

Correspondence: José M Serratosa, MD, PhD, Servicio de Neurología, Fundación Jiménez Díaz, Avda Reyes Católicos, 2, 28040 Madrid, Spain. Tel: +34 91 550 4854; Fax: +34 91 549 7381; E-mail: serratosa@jet.es
Received 13 March 2000; revised 7 August 2000; accepted 23 August 2000

increases in frequency and becomes constant. A rapidly progressive dementia with apraxia, aphasia, and visual loss ensues, leading patients to a vegetative stage and death, usually within less than a decade of first symptoms.[6] Lafora disease occurs worldwide, but is particularly common in the Mediterranean countries of Southern Europe and Northern Africa, in Southern India and in the Middle East. As in other autosomal recessive diseases, consanguinity is common in Lafora disease. Linkage analysis and homozygosity mapping first localised a major gene for Lafora disease to chromosome 6q24.[7] However, other gene or genes than *EPM2A* may cause Lafora disease when mutated, since in approximately 10–20% of the families linkage to chromosome 6q24 can be excluded.[8,9] Recently, we and others have cloned the *EPM2A* gene and have shown that affected individuals were homozygous or compound heterozygous for loss-of-function mutations.[10,11]

The human *EPM2A* gene spans approximately 70 kb and codes for a transcript of about 3 kb which is split into 4 exons ranging from 112 to 2328 bp.[10,11] Exon 4 includes two alternative splicing transcripts (4 and 4b) (Figures 1 and 2).

Here we have investigated the *EPM2A* gene in 68 Lafora disease chromosomes from 34 unrelated patients and in 126 normal chromosomes of Spanish origin. We report the detection of 11 novel *EPM2A* mutations and the characterisation of several *EPM2A* polymorphisms. The analysis of the associations between the *EPM2A* mutations and the *EPM2A* polymorphisms provides useful data for population genetic studies and for investigating the history of the *EPM2A* mutations.

Materials and methods

Patients

The clinical diagnosis of Lafora disease was based on the presentation of epilepsy, myoclonus, rapidly progressive neurological deterioration, and a slow background with polyspike-wave complexes in the electroencephalogram.[6,12] In addition, we required a biopsy of skin, muscle, liver, or brain showing the characteristic periodic acid Schiff-positive Lafora bodies.[2,3] Information on age of onset was available from 29 patients belonging to 19 families. Age of onset ranged from 4 to 16 years (mean = 12 years). The first symptom of the disease was either generalised tonic-clonic seizures, absences, or simple partial seizures with visual symptoms. The severity of the myoclonus was mild at onset and increased in severity as the disease progressed. Progressive dementia followed in all patients. All families included in this study were genotyped with the chromosome 6q24 microsatellites markers contained in the region of the Lafora disease's gene; 30 families showed segregation of markers flanking *EPM2A* with the Lafora disease phenotype and four were not informative because of the low number of available family members. Five families in which linkage to chromosome 6q24 markers could be excluded were not included in this study. These families represent 13% of our Lafora disease family database.

Samples

Thirty-four affected probands were included in this study. The patients originated from Spain, Italy, Australia, Holland, the United States of America, North Africa, Turkey, and France. Forty-eight healthy Spanish individuals were screened to verify that the mutations found were absent in the normal population. To calculate allele frequencies at the EPM2A polymorphic sites and to estimate haplotype frequencies we used 126 control chromosomes. DNA samples were obtained from peripheral blood lymphocytes, using standard methods. The study was approved by the Ethics Committee of the Fundación Jiménez Díaz. Blood was collected from patients and their relatives after informed consent.

PCR amplification and sequencing of exons

Exons 1, 2, 3 and 4 of the *EPM2A* gene[11] were amplified from genomic DNA using specific primers derived from 5' and 3' intronic sequences (Table 1) using standard methods. The initiating ATG of exon 1 was obtained by PCR amplification from a human placenta cDNA (Clontech, Palo Alto, CA, USA). The annealing temperature for all primer pairs was 60°C. The corresponding PCR products were purified by agarose gel electrophoresis and extracted with the Qiaquick Gel Extraction Kit (Qiagen, Hilden, Germany). Direct sequencing of PCR products was performed with a dye-terminator cycle-sequencing kit (Perkin-Elmer, Warrington, UK) using *Taq* FS DNA polymerase. Sequences were resolved on an ABI PRISM 377 automatic sequencer, and the results analysed with the ABI Analysis software (version 3.1).

Microsatellite analysis

Analysis of polymorphisms at the EPM2A-1, EPM2A-2, EPM2A-3, EPM2A-4, EPM2A-5, and D6S1703 microsatellites was performed by PCR, using total human genomic DNA. Amplification was performed in a total volume of 10 μl containing 40 ng of genomic DNA, 3 pmol of each primer (Table 2), 0.3 U *Taq* polymerase (Promega, Madison, WI, USA), 200 μM each dATP, dGTP and dTTP, 2.5 μM dCTP, 0.7 μCi [^{32}P] α-dCTP at 300 Ci mmol^{-1}, and 1.5 mM $MgCl_2$. PCR conditions were one cycle at 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 55.5°C for 1 min, and 74°C for 15 s, and one last cycle of 74°C for 6 min. Samples were resolved on a 6.5% polyacrylamide sequencing gels and exposed on Kodak XAR film for 1–14 h.

SSCP analysis

SSCP analysis[13] was performed by PCR, using total genomic DNA, using the GenePhor DNA Electrophoresis System (Amersham Pharmacia Biotech, Uppsala, Sweden). Amplification was performed in a total volume of 10 μl containing 60 ng of genomic DNA, 12.5 pmol of each primer (Table 3), 1 U of *Taq* polymerase (Promega), 200 μM each dATP, dCTP,

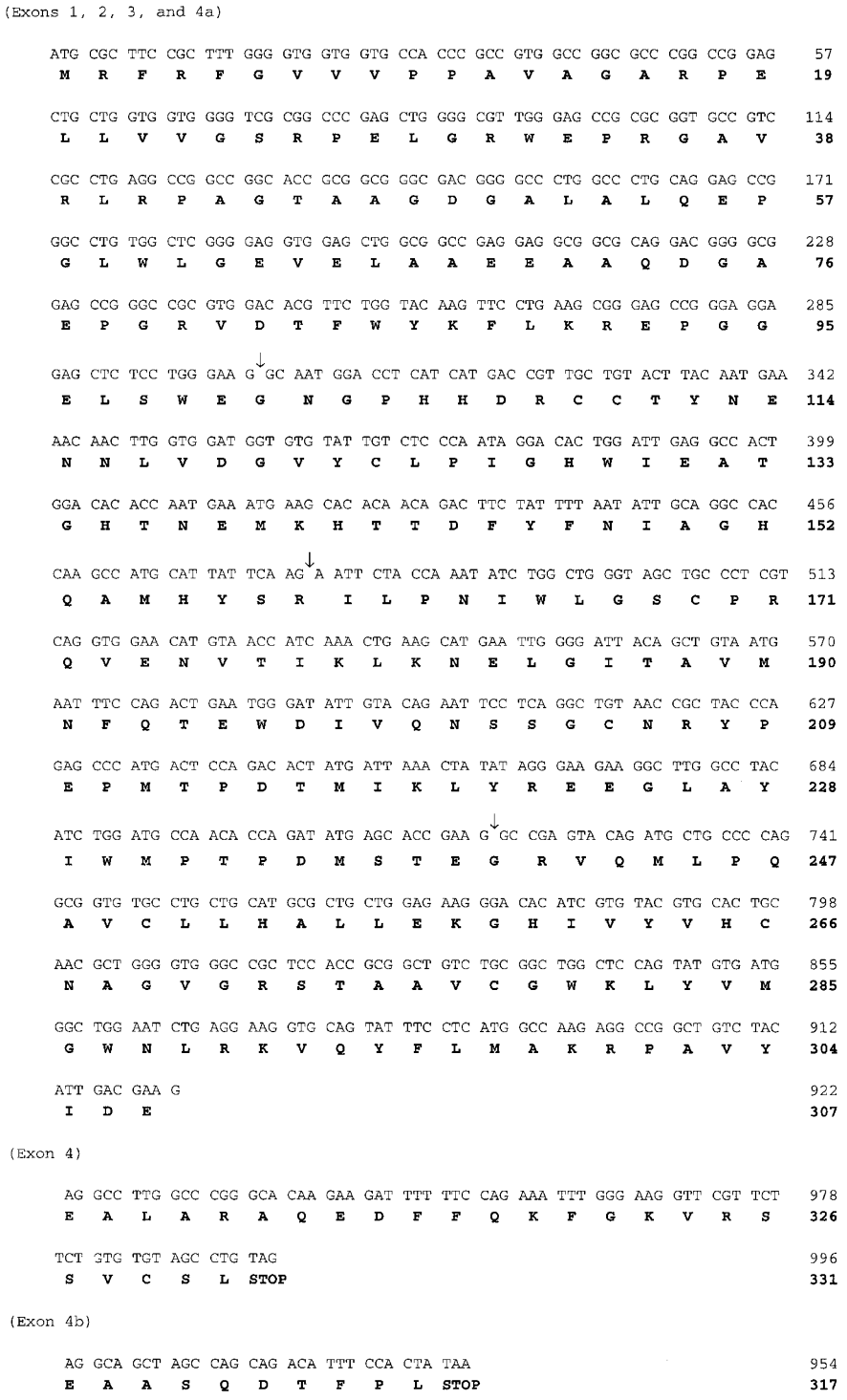

Figure 1 cDNA sequence and derived amino acid sequence of *EPM2A*. Nucleotide sequence of the cDNA corresponding to EPM2A exons 1, 2, 3 and 4. Exon junctions are indicated with small vertical arrows. The conceptual translation of the EPM2A ORF is shown below the cDNA sequence in one-letter code. The sequence begins with the translation ATG start codon.

dGTP and dTTP, and 1.5 mM $MgCl_2$. PCR conditions were one cycle at 94°C for 2 min, followed by 30 cycles of 94°C for min, 60°C for 1 min, and 74°C 1 min, one cycle of 74°C for 3 min and one last cycle of 25°C for 2 min. Samples were resolved on 12.5% non-denaturing polyacrilamide gels with the GeneGel Excel 12.5/24 kit (Amersham Pharmacia Biotech)

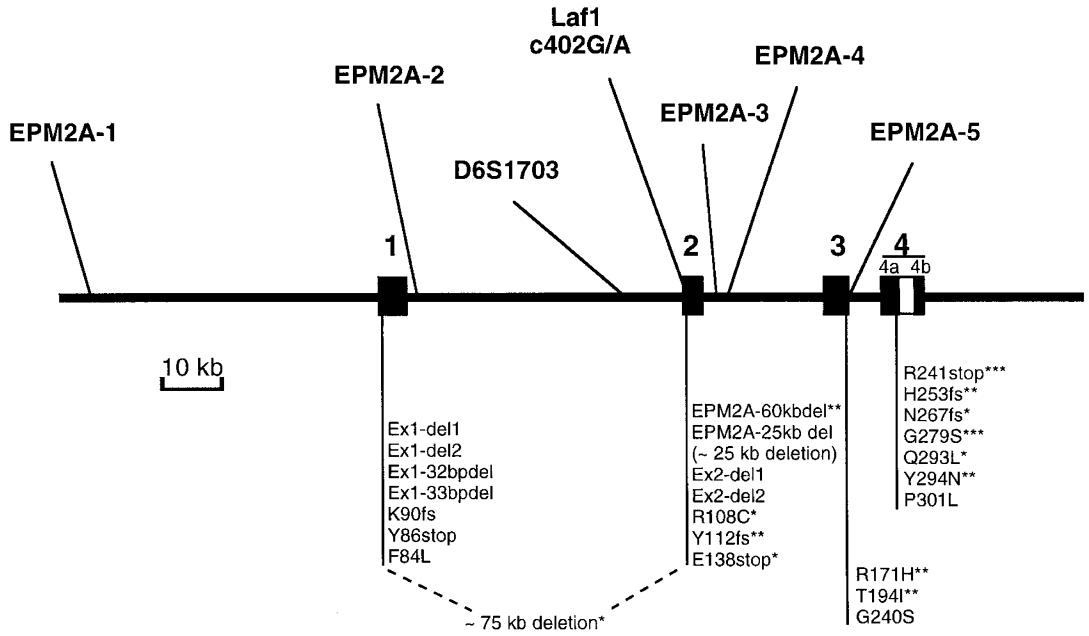

Figure 2 Localisation of *EPM2A* mutations and associated polymorphic markers at the human EPM2A locus. The human *EPM2A* gene is shown as a horizontal line, with the exons indicated by short, thick vertical lines and labelled 1–4. Exon 4 includes two alternative splicing transcripts: one includes the complete exon 4 and presents a stop codon 75 bp after the end of segment 4a; the other one is formed by segments 4a and 4b and lacks the central part of exon 4. Below the structure of the *EPM2A* gene, the thin lines localise the position of each of the 25 *EPM2A* mutations characterised so far. Above the structure of the EPM2A gene, the thin lines show the relative positions of seven polymorphic sites: EPM2A-1, EPM2A-2, EPM2A-3, EPM2A-4, EPM2A-5, and D6S1703 are SSRs; Laf-1 is a SNP. EPM2A-25kbdel and ~25 kb deletion[16] may be the same mutation; *from reference 10; **from reference 11; ***from both references 10 and 11.

and silver staining using the PlusOne™ DNA Silver Staining kit (Amersham Pharmacia Biotech).

Mutation hot spot analysis

Identification of potential mutation hot spots was performed with the computer program MUTPRED, which is based on empirical data and designed to predict the location of point mutations within gene coding regions causing human genetic disease.[14]

References for links using databases

The major gene for Lafora disease codes the LAFPTPase, a putative protein tyrosine phosphatase (PTPase).[11] The nucleotide sequences for the human *EPM2A* gene have been deposited under EMBL/GenBank accession numbers AJ130763, AJ130764, and AF084535.

Results

Initiating ATG of exon 1

The human *EPM2A* gene codes for a transcript of about 3 kb for which the 5' end sequence was only partially characterised in humans.[10,11] Here we report the complete coding sequence of the *EPM2A* gene including the ATG initiation codon region (Figure 1). We obtained a 450 bp PCR fragment from a human placenta cDNA covering the complete exon 1 and part of exon 2. This fragment contained an ATG codon 241 bp upstream from the beginning of the consensus cDNA sequence we proposed in our initial cloning report and extends the previously communicated sequence of exon 1.[10,11] The presence of an in-frame stop codon 66 nucleotides before this ATG codon and the high sequence identity with the murine ATG translation initiation codon

Table 1 PCR primers for genomic amplification of the *EPM2A* exons (5'→3') for sequencing

Exon	*Primer forward*	*Primer reverse*	*Product size (bp)*
1	TGCGCTTCCGCTTTGGG	AGGGACGCGGGCAAAAAGC	400
2	GTATCAGCTGCTTGAGGATA	CTTGTCCTACTTCTATGCCTA	291
3	CTACATGTTTTATGCAGCTCC	ATTTATTCCATTTCTACCATTCAT	431
4	GAGAGAGCCTCTGGCCTC	CAGAAGAACGAACCTTCCCA	483

Table 2 Novel polymorphic markers within the *EPM2A* gene

Marker	*ID*[a]	*Type*	*Het*[b]	*Primers (5'→3')*	*Allele*	*FQ1*[c]	*FQ2*[d]
EPM2A-1	9992791	(CA)n	0.746	F: CTCCTTGCTTCTCAAGCTT	175	0.008	0.020
				R: TGGACTTACAGAATGCTTATA	177	0.016	0.040
					179	0.016	0.040
					181	0.008	0.020
					183	0.057	0.020
					185	0.074	0.040
					187	0.377	0.180
					189	0.303	0.500
					191	0.082	0.100
					193	0.041	0.040
					195	0.016	0.000
EPM2A-2	9992793	(CA)n	0.725	F: CTCATAATCCTATTATGCAGGA	152	0.008	0.020
				R: AAAGTCTCAGGTTTACAGCTAA	154	0.083	0.059
					156	0.008	0.020
					158	0.275	0.196
					160	0.042	0.078
					162	0.075	0.196
					164	0.042	0.059
					166	0.417	0.451
					168	0.050	0.020
					170	0.000	0.020
EPM2A-3	9992834	(CA)n	0.851	F: GTGTATGTATACCTGGAGTATAA	134	0.057	0.061
				R: GATAAATCATAACATGGAATGG	136	0.156	0.041
					138	0.229	0.306
					140	0.016	0.020
					142	0.033	0.020
					144	0.066	0.082
					146	0.180	0.102
					148	0.123	0.143
					150	0.049	0.000
					152	0.049	0.020
					154	0.033	0.000
					156	0.008	0.020
					158	0.000	0.061
					160	0.000	0.122
EPM2A-4	9992864	(AAAT)n	0.518	F: ATGTAACTAACCTGCACATTGTG	143	0.051	0.000
				R: GATAAATCATAACATGGAATGG	151	0.393	0.275
					157	0.000	0.039
					159	0.530	0.686
					167	0.026	0.000
EPM2A-5	9992874	(TGTC)n	0.102	F: CCATGACTCCAGACACTATG	199	0.986	0.906
				R: TTTATTCCATTTCTACCATTCAT	195	0.014	0.094
Laf-1	–	C402G→A (G134G)	0.219	F: GTATCAGCTGCTTGAGGATA	G	0.783	0.594
				R: CTTGTCCTACTTCTATGCCTA	A	0.217	0.406

[a]ID: accession number from Généthon; [b]Het: calculated heterozygosity; [c]FQ1: allele frequency in normal population; [d]FQ2: allele frequency in affected population.

region[15] suggests that this is the ATG translation initiation codon in *EPM2A*. In addition, the length of the mRNA present in northern blots is in agreement with this ATG being the initiation codon.

Mutations in EPM2A

In our previous study reporting the identification of the *EPM2A* gene we described the mutational analysis of the *EPM2A* gene in 20 Lafora disease pedigrees that resulted in the identification of 13 Lafora disease chromosomes carrying a total of eight different EPM2A mutations.[11] Probably because those mutational studies were incomplete and focused only on *EPM2A* exons 2, 3 and 4a, we failed to identify *EPM2A* mutations in 70% of the Lafora disease chromosomes. In this report 68 chromosomes from 34 unrelated Lafora disease patients, including those partially characterised before, were screened for mutations in the *EPM2A* gene. This time, all four exons of the *EPM2A* gene, including their flanking intronic sequences, were amplified from genomic DNA of each Lafora disease patient, using the PCR primers specified in Table 1. SSCP analysis was performed for all exons and those presenting band shifts were subsequently sequenced in both strands.

Table 3 PCR primers for SSCP analysis of the *EPM2A* gene

DNA fragment	*Primer forward*	*Primer reverse*	*Product size (bp)*
Exon 1	F: TCGGGGAGGTGGAGCT	R: CAGGCGTCTGCTGGCAATA	140
	F: TGCGCTTCCGCTTTGGG	R: GTCCTGCGCCGCCTCCTC	220
Exon 2	F: GTATCAGCTGCTTGAGGATA	R: CTTGTCCTACTTCTATGCCTA	291
	F: CTACATGTTTTATGCAGCTCC	R: AGGAATTCTGTACAATATCCC	208
Exon 3	F: ACCAAATATCTGGCTGGGTAG	R: TGCTCATATCTGGTGTTGGC	229
	F: CCATGACTCCAGACACTATG	R: ATTTATTCCATTTCTACCATTCAT	199
	F: CGGTATCTGGTGGTTTAGTTA	R: TGCACCTTCCTCAGATTCCA	214
Exon 4	F: GAGAAGGGACACATCGTGTA	R: CCTTAGGGAAATCAGGAG	258
	F: CCTAATTAAACTATGATGCGTAT	R: GGACTTCACTTTACTTAATGCT	271

The analysis of the 68 chromosomes from 34 Lafora disease probands resulted in the identification of *EPM2A* mutations in 49 of them (72%). Twenty-three probands presented mutations in both chromosomes (affected individuals were homozygous or double heterozygous). In four probands mutations were identified in only one of the two copies of the *EPM2A* gene. Therefore, the total number of families presenting *EPM2A* mutations was 27 (79%). It is thus possible that 7 of the 34 families from this study harbour mutations in a non-*EPM2A* gene since no mutated alleles were identified in these families. However, the existence of four families where mutations have been identified in only one chromosome supports the idea that there are indeed unidentified mutations.

A total of 20 *EPM2A* mutations, 11 of them novel, were characterised (Figure 2 and Table 4). Mendelian inheritance of the mutations was confirmed by SSCP and sequence analysis of the relevant amplified DNA fragment in family

Table 4 *EPM2A* mutations identified in Lafora disease patients

Mutation	*Exon*	*Type*	*Nucleotide change*	*Amino acid change/ predicted consequence*	*No. of chromo-somes*	*No. of probands*	*Country of origin*	*Reference*
Ex1-del1	1	deletion	deletion exon 1	Truncation at exon 2	1	1	Spain	Present report
Ex1-del2	1	deletion	deletion exon 1	Truncation at exon 2	1	1	Spain	Present report
Ex1-33bpdel	1	microdeletion	c.91-123del	Truncation after gly 30	3	2	Spain	Present report
Ex1-32bpdel	1	microdeletion	c.102-133del	Truncation after gly 36	2	2	SpaIn	Present report
F84L	1	missense	c.252C→G	phe 84 leu	2	1	France	Present report
Y86stop	1	nonsense	c.258C→G	tyr 86 Stop	2	1	Spain	Present report
K90fs	1	frameshift	c.269-275del	Truncation after lys 90	2	1	Holland	Present report
R108C	2	missense	c.322C→T	arg 108 cys	2	1	Spain	Present report
					2	1	–	Minassian *et al*[10]
Y112fs	2	frameshift	c.335-336insA	Truncation after thr 112	2	1	Italy	Serratosa *et al*[11]
E138stop	2	nonsense	c.412G→T	glu 138 Stop	2	1	–	Minassian *et al*[10]
EPM2A-60kbdel	2	deletion	deletion exon 2	Truncation at exon 3	2	1	Italy	Serratosa *et al*[11]
EPM2A-25kbdel	2	deletion	deletion exon 2	Truncation at exon 3	2	1	USA	Present report
					2	1	Iran	Minassian *et al*[16]
Ex2-del1	2	deletion	deletion exon 2	Truncation at exon 3	1	1	USA	Present report
Ex2-del2	2	deletion	deletion exon 2	Truncation at exon 3	1	1	Italy	Present report
R171H	3	missense	c.512G→A	arg 171 his	1	1	USA	Serratosa *et al*[11]
T194I	3	missense	c.581C→T	thr 194 ile	1	1	Spain	Serratosa *et al*[11]
G240S	3	missense	c.718G→A	gly 240 lys	1	1	Australia	Present report
R241stop	4	nonsense	c.721C→T	arg 241 Stop	13	10	Spain	Serratosa *et al*[11]
					1	1	Italy	
					2	1	Turkey	
					9	5	–	Minassian *et al*[10]
N267fs	4	frameshift	c.799-800insA	Truncation after asn 267	2	1	–	Minassian *et al*[10]
H253fs	4	frameshift	c.759T→CATGCA	Truncation after his 253	1	1	Australia	Serratosa *et al*[11]
G279S	4	missense	c.835G→A	gly 279 ser	2	2	Spain	Serratosa *et al*[11]
					2	2	Italy	
					1	1	–	Minassian *et al*[10]
Q293L	4	missense	c.878A→T	gln 293 leu	2	1	–	Minassian *et al*[10]
Y294N	4	missense	c.880T→A	tyr 294 asn	1	1	Spain	Serratosa *et al*[11]
P301L	4	missense	c.902C→T	pro 301 leu	1	1	France	Present report
about 75kbdel		deletion	deletion exons 1 and 2	Truncation in exon 3	2	1	–	Minassian *et al*[10] Minassian *et al*[16]

members. One of the novel mutations is a 32 bp deletion (c.102–133del) affecting the ATG region of exon 1 and resulting in a stop codon in exon 2. This mutation confirms that this region is part of the *EPM2A* transcript.

Three of the novel mutations reported here are missense mutations, resulting in the following amino acid substitutions: F84L, G240S and P301L. We do not know yet whether these mutations are loss-of-function mutations. However, our control population of unaffected unrelated individuals, which included 96 chromosomes, showed none of these mutations, suggesting that these changes do not represent polymorphisms in our population. Another novel mutation, Y86stop, is a nonsense mutation; three are deletions of less than 40 bp (K90fs, Ex1-32bpdel, Ex1-33bpdel); two are deletions of yet undetermined size affecting exon 1 (Ex1-del1 and Ex1-del2); and two are >3 kb deletions that include exon 2 (Ex2-del1, Ex2-del2). In a recent report[16] a new mutation has been communicated: a 25 kb deletion affecting exon 2 which may be identical to a 25 kb deletion we have identified and denote *EPM2A*-25kbdel.

Adding these data to those reported earlier,[10,11,16] a total of 25 EPM2A mutations have been identified in 44 unrelated Lafora disease patients. Nine of the *EPM2A* mutations are missense mutations, three are nonsense mutations, three are frameshift mutations, and ten are deletions of the different sizes (Table 4). The R241stop mutation was encountered in almost 40% of the Lafora disease probands.

Polymorphic sites in the EPM2A region

Here we describe the identification and analysis of six novel polymorphic sites at the EPM2A locus. These novel polymorphic sites correspond to three variable dinucleotide repeats (EPM2A-1, EPM2A-2 and EPM2A-3), two variable tetranucleotide repeats (EPM2A-4 and EPM2A-5), and one SNP (Laf-1). The localisation of each simple sequence repeat (SSR) and SNP within the *EPM2A* region is shown in Figure 2. All five SSRs were encountered by inspection of a 100-kb-long sequence of the 150 kb 466P17 clone (The Sanger Centre, Cambridge, UK) encompassing exons 1–2 and by sequencing part of intron 3. EPM2A-1, EPM2A-2, and EPM2A-3 are perfect $(CA)_n$ repeats. EPM2A-4 is a perfect $(AAAT)_n$ repeat and EPM2A-5 is a $(TGTC)_{2\text{-}3}$ repeat. EPM2A-1 is located 50.5 kb upstream the 5' end of exon 1, EPM2A-2 is located 2866 bp downstream of the 3' end of exon 1, EPM2A-3 is located 2767 bp downstream of the 3' end of exon 2, and EPM2A-4 is located 2763 bp downstream of EPM2A-3 in intron 2. EPM2A-5 is a rare polymorphism located in the 5' end of intron 3. During SSCP mutation analysis experiments, we found one SNP (c.402G/A in exon 2) which we excluded from being an *EPM2A* mutation because it was present in the normal population (in heterozygosis and homozygosis) and represents no change in the aminoacid sequence. The allele frequencies of these six novel polymorphisms in Lafora disease patients and in the control population are shown in Table 2.

Haplotype analysis of polymorphic sites within the *EPM2A* gene

Haplotype associations between alleles at seven polymorphic sites within or near the *EPM2A* gene were established for a total of 126 control chromosomes. This analysis included the D6S1703 site, a microsatellite from the Généthon collection.[17] The haplotypes generated from this small sample of chromosomes are almost all unique.

To establish the haplotypic associations between EPM2A polymorphisms and *EPM2A* mutations, all seven polymorphic markers (EPM2A-1, EPM2A-2, D6S1703, Laf-1, EPM2A-3, EPM2A-4, and EPM2A-5) were characterised in the Lafora disease probands and their families. We observed that there are five different haplotypes associated with the R241stop mutation (Table 5). One of them is present in 10 of 16 Lafora disease chromosomes carrying this mutation. The second most common mutation, G279S, appears associated with three different haplotypes. These data suggest that both R241stop and G279stop are recurrent mutations. In addition, these data indicate that the high prevalence of the R241stop mutation is a consequence of both founder effect and recurrence. All other mutations are present in three or less Lafora disease chromosomes and are associated each to only one haplotype. A complete description of the EPM2A haplotypes for 47 EPM2A chromosomes where a mutation has been identified, including 13 Lafora disease chromosomes reported in our initial communication,[11] is presented in Table 5.

Discussion

We and others reported, elsewhere, that a gene coding a protein tyrosine phosphatase (PTPase) is the gene for 6q-linked Lafora disease.[10,11] In the two initial reports the partially characterised 5' end of the gene and 13 different mutations were communicated. We described eight mutations in our Lafora disease family set, but failed to identify *EPM2A* mutations in 70% of the affected Lafora disease chromosomes because of the incompleteness of our preliminary study. In this report we describe the sequence of exon 1 containing the ATG translation initiation codon and the systematic mutational analysis of the *EPM2A* gene in 68 chromosomes from 34 unrelated Lafora disease probands. In this sample we have identified a total of 20 EPM2A mutations, eleven of them novel, and have characterised the EPM2A haplotypes associated with each EPM2A mutation.

Five of the 11 novel mutations described here are deletions affecting a few nucleotides or even entire exons. Ex1-32bpdel is a deletion (c.102–133del) resulting in a frameshift after Gly36 that results in the posterior truncation of the protein. Ex1-del1, Ex1-del2, Ex2-del1, and Ex2 del2 are deletions affecting entire exons (1 or 2). All these deletions result in truncation of the *EPM2A* gene product. Y86stop is a nonsense mutation (c. 258C $\rightarrow$ G) that results in a stop codon and, thus, in a truncated protein as well. Ex1-33bpdel is a deletion resulting in the substitution of one amino acid and the

Table 5 EPM2A haplotypes associated with *EPM2A* mutations

					Associated polymorphisms				
Mutation	*Freq.*	*EPM2A-1*	*EPM2A-2*	*D6S1703*	*Laf-1*	*EPM2A-3*	*EPM2A-4*	*EPM2A-5*	*Origin*
P301L	1	189	162	151	G	146	159	199	France (50a)
Y294N	1	189	160	153	G	138	159	199	Spain (102a)
G279S	4	189	166	149	A	156	151/159	199	Spain (13a)
		187	160	157	G	138	159	199	Italy (57a)
		185	160	157	G	138	151	195	Italy (9557a)
		185	160	157	G	136	151	195	Spain (109a)
H253fs	1	187	168	149	A	156	159	199	Australia (119a)
R241stop	16	187	152	151	G	138	151	199	Turkey (113a)
		187	152	151	G	138	151	199	Turkey (113b)
		189	156	157	G	138	159	199	Italy (57b)
		189	160	151	G	152	159	199	Spain (14a)
		189	166	149	A	160	159	199	Spain (16a)
		189	166	149	A	160	159	199	Spain (16b)
		189	166	149	A	160	159	199	Spain (109b)
		189	166	149	A	160	159	199	Spain (3a)
		189	166	149	A	160	159	199	Spain (3b)
		189	166	149	A	160	159	199	Spain (12a)
		189	166	149	A	160	159	199	Spain (15a)
		189	166	149	A	160	159	199	Spain (15b)
		189	166	149	A/G	160	159	199	Spain (8a)
		189	166	149	A	160	159/151	199	Spain (13b)
		191	166	155	G	144	159	199	Spain (116a)
		191	166	155	G	144	159	199	Spain (116b)
G240S	1	189	166	155	A	146	159	199	Australia (119b)
T194I	1	189	162	153	G	148	159	199	Spain (12b)
R171H	1	191	166	153	A	148	159	199	USA (2a)
Y112fs	2	187	166	157	G	144	159	199	Italy (114a)
		187	166	157	G	144	159	199	Italy (114b)
R108C	2	193	164	151	A	144	159	199	Spain (5a)
		193	164	151	A	144	159	199	Spain (5b)
EPM2A-	2	189	?	155	0	0	0	199	USA (1a)
25kbel		189	?	155	0	0	0	199	USA (1b)
Ex2-del1	1	183	158	153	0	0	159	199	USA (2b)
Ex2-del2	1	187	154	151	0	0	151	199	Italy (9557b)
K90fs	2	189	166	149	A	146	159	199	Holland (110a)
		189	166	149	A	146	159	199	Holland (110b)
Y86Stop	2	181	160	147	G	140	157	199	Spain (4a)
		181	160	147	G	140	157	199	Spain (4b)
F84L	2	179	158	161	G	138	151	199	North Africa (65a)
		179	158	161	G	138	151	199	North Africa (65b)
Ex1-33bpdel	3	189	166	153	A	134	159	199	Spain (101a)
		189	166	153	A	134	159	199	Spain (101b)
		189	166	153	A	134	159	199	Spain (107b)
Ex1-32bpdel	2	189	164	151	A	148	159	199	Spain (102b)
		187?	164	151	A	148	159	199	Spain (108b)
Ex1-del1	1	177	158	163	G	138	151	195	Spain (107a)
Ex1-del2	1	175	158	155	G	138	151	199	Spain (108a)

elimination of the subsequent ten amino acids without producing a posterior frameshift. K90fs is a mutation found in homozygosis in a patient born of a consanguineous mating. This mutation is the consequence of a seven-nucleotide deletion in exon 1 (c.269–275del).

Therefore, eight of the novel EPM2A mutations result in truncated polypeptides that are most likely non-functional. The consequences of the other three novel mutations are more difficult to foresee. F84L, G240S, and P301L are missense resulting in changes of amino acids in different exons. F84L is a mutation carried in homozygosis by one patient from a consanguineous family that results from the substitution c.252C→G in exon 1. G240S (c.718G→A) was found in one patient who carries the H253fs insertional substitution (c.759T→CATGCA) in the other Lafora disease chromosome. P301L is the consequence of the c.902C→T substitution and was also identified in heterozygosis. All novel mutations described here were not found in a sample of 96 unaffected chromosomes, indicating that they are not polymorphisms.

As a whole, 25 different mutations in *EPM2A* have been identified in 46 unrelated Lafora disease patients (including those described elsewhere,[10,11,16] and in this report). These mutations are distributed throughout the whole length of the *EPM2A* gene sequence (Figure 2, Table 4). It is noticeable, however, that the occurrence of missense mutations in

EPM2A (36%) is somehow reduced compared with that observed in other genes causing autosomal recessive traits (60–70%) and that most missense mutations in *EPM2A* (9 out of 11) are clustered within exons 3 and 4. These two exons of *EPM2A* encode an amino acid sequence with significant homology to the catalytic domain of PTPases.[10,11] Although no functional data is yet available for this putative enzymatic activity of the *EPM2A* gene product, the clustering of missense mutations suggests that the peptide region coded by exons 3 and 4 is crucial for the function of the *EPM2A* product, providing further support to the concept that *EPM2A* indeed codes for a PTPase.

Among all *EPM2A* mutations, the R241stop and G279S are present in Lafora disease chromosomes from different Mediterranean regions. R241stop is the most prevalent *EPM2A* mutation. It is present in 23.5% (16/68) of EPM2A chromosomes and has been found in 32.5% of Spanish EPM2A chromosomes (13/40). A total of 40% of the Lafora disease patients carry this mutation. The analysis of the haplotypic associations of alleles at the EPM2A-1, EPM2A-2, D6S1703, Laf-1, EPM2A-3, EPM2A-4, and EPM2A-5 polymorphic sites demonstrated the existence of different EPM2A haplotypes associated with this mutation. One of them (189–166–149-A–160–159–199) is the most common haplotype associated with the R241Stop mutation (Table 5). Our finding of distinct EPM2A haplotypes associated with the R241Stop mutation suggest that R241stop has different phylogenetic origins and, therefore could be a recurrent mutation. Interestingly, analysis of the nucleotide sequence at this mutation site identifies a CpG dinucleotide. Moreover, analysis of the complete *EPM2A* sequence using the MUTPRED program[14] denotes this site as a potential mutation hot spot, where c. 322C mutates to c. 322T, strongly supporting the concept that R241Stop is a recurrent mutation. The second most common mutation found in our family set, G279S, was associated with three different haplotypes and also coincides with a CpG dinucleotide, suggesting that it is a recurrent mutation too.

The Ex1-33bpdel (c.91-123del) is a mutation that has been encountered in only two unrelated families from one geographical Spanish region (Table 5) and in all instances associated with identical EPM2A haplotypes. This argues in favour of a common origin for the EPM2A chromosomes carrying the Ex1-33bpdel mutation. The fact that no variation is detected among these EPM2A chromosomes suggests that Ex1-33bpdel originated recently.

The data presented in this article represent the first extensive account of *EPM2A* mutations and associated polymorphisms and provide a general understanding of the variability at the EPM2A locus in the Lafora disease patient population.

Acknowledgements

The authors wish to thank all the patients and family members who participated in this study. We also thank Dr D Lindhout, Dr PB Augustijn, Dr CA Tassinari, Dr R Michelucci, Dr A Malafosse, Dr M Topcu, Dr D Grid, Dr C Dravet and Dr SF Berkovic for identifying Lafora disease patients. This work was supported by the Asociación Lafora España, the Fundación Jose Antonio de Castro, the Spanish Comisión Interministerial de Ciencia y Tecnología (SAF96/0318, SAF96/0055 and SAF99-0013-C02-02), the Fondo de Investigaciones Sanitarias (FIS98/0687), the Junta de Comunidades de Castilla-La Mancha and the Comunidad Autónoma de Madrid (08.6/0015/1997 and 08.5/0008/1998). P G-G was supported by a fellowship from the Association France Lafora.

References

1 Berkovic SF, Andermann F, Carpenter S *et al*: Progressive myoclonus epilepsies: specific causes and diagnosis. *N Engl J Med* 1986; **315**: 296–305.
2 Lafora GR: Über das Vorkommen amyloider Körperchen im Innern der Ganglienzellen; zugleich ein zum Studium der amyloiden Substanz im Nervensystem. *Virchows Arch*. [Pathol. Anat.] 1911; **205**: 295–303.
3 Lafora GR, Glueck B: Beitrag zur Histopathologie der myoklonischen Epilepsie. *Z. Gesamte Neurol Psychiatr* 1911; **6**: 1–14.
4 Yokoi S, Austin J, Witmer F *et al*: Studies in myoclonus epilepsy (Lafora body forms). I. Isolation and preliminary characterization of Lafora bodies in two cases. *Arch Neurol* 1968; **19**: 15–33.
5 Sakai M, Austin J, Witmer F *et al*: Studies in myoclonus epilepsy (Lafora body form). II. Polyglucosans in the systemic deposits of myoclonus epilepsy and in corpora amylacea. *Neurology* 1970; **20**: 160–176.
6 Van Heycop Ten Ham MV: Lafora disease, a form of progressive myoclonus epilepsy. In: Vinken PJ and Bruin GW (eds). *Handbook of Clinical Neurology*, vol 15. North Holland: Amsterdam, 1974, pp 382–422.
7 Serratosa JM, Delgado-Escueta AV, Posada I *et al*: The gene for progressive myoclonus epilepsy of the Lafora type maps to chromosome 6q. *Hum Mol Genet* 1995; **4**: 1657–1663.
8 Gómez-Garre P, Anta B, Castro-Gago *et al*: Reduction of the Lafora disease candidate gene region to a 2 cM interval in chromosome 6q24 and evidence for genetic heterogeneity. *Eur J Hum Genet* 1998; **6**: 152 (abstract).
9 Minassian BA, Sainz J, Serratosa JM *et al*: Genetic locus heterogeneity in Lafora's progressive myoclonus epilepsy. *Ann Neurol* 1999; **45**: 262–265.
10 Minassian BA, Lee JR, Herbrick JA *et al*: Mutations in a gene coding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. *Nat Genet* 1998; **20**: 171–174.
11 Serratosa JM, Gómez-Garre P, Gallardo ME *et al*: A novel protein tyrosine phosphatase gene is mutated in progressive myoclonus epilepsy of the Lafora type (EPM2). *Hum Mol Genet* 1999; **8**: 345–352.
12 Tassinari CA, Bureau-Paillas M, Dalla Bernardina B *et al*: La maladie de Lafora. *Rev EEG Neurophysiol* 1978; **8**: 107–122.
13 Orita M, Iwahana H, Kanazawa H *et al*: Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphism. *Proc Natl Acad Sci USA* 1989; **86**: 2766–2770.
14 Cooper DN, Krawczak M: The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions. *Hum Genet* 1990; **85**: 55–74.
15 Ganesh S, Amano K, Delgado-Escueta AV *et al*: Isolation and characterization of mouse homologue for the human epilepsy gene, EPM2A. *Biochem Biophys Res Commun* 1999; **257**: 24–28.
16 Minassian BA, Ianzano L, Delgado-Escueta AV, Scherer SW: Identification of new and common mutations in the EPM2A gene in Lafora disease. *Neurology* 2000; **54**: 488–490.
17 Dib C, Fauré S, Fizames C *et al*: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 1996; **380**: 152–154.