ORIGINAL ARTICLE

Genetic mapping of a new Lafora progressive myoclonus epilepsy locus (*EPM2B*) on 6p22

E M Chan, D E Bulman, A D Paterson, J Turnbull, E Andermann, F Andermann, G A Rouleau, A V Delgado-Escueta, S W Scherer, B A Minassian

See end of article for authors' affiliations

Correspondence to: Dr B Minassian, Division of Neurology, Department of Pediatrics, Program in Genetics and Genomic Biology, The Hospital for Sick Children and The University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada; bminass@sickkids.ca

J Med Genet 2003;**40**:671–675

Revised version received 9 May 2003
Accepted for publication 12 May 2003

Background: Lafora disease is a progressive myoclonus epilepsy with polyglucosan accumulations and a peculiar neurodegeneration with generalised organellar disintegration. It causes severe seizures, leading to dementia and eventually death in early adulthood.
Methods: One Lafora disease gene, *EPM2A*, has been identified on chromosome 6q24. Locus heterogeneity led us to search for a second gene using a genome wide linkage scan in French-Canadian families.
Results: We mapped a second Lafora disease locus, *EPM2B*, to a 2.2 Mb region at 6p22, a region known to code for several proteins, including kinesins. Kinesins are microtubule dependent motor proteins that are involved in transporting cellular components. In neurones, they play a major role in axonal and dendritic transport.
Conclusion: Analysis of the present locus in other non-*EPM2A* families will reveal whether there is further locus heterogeneity. Identification of the disease gene will be of major importance towards our understanding of the pathogenesis of Lafora disease.

Lafora disease (LD) is the principal progressive myoclonus epilepsy with onset in adolescence. Onset is characterised by the appearance of sporadic myoclonus, occipital lobe seizures (visual hallucinations, photoconvulsive seizures), and/or generalised seizures, with intervening insiduous behavioral changes and cognitive decline. Within 3 years, myoclonus is almost constant, seizures frequent, and dementia evident. Symptoms then progress slowly in a protracted fashion often over more than 10 years. Most patients die in status epilepticus.[1 2]

In a number of organs, pathological analysis reveals cytoplasmic inclusions known as Lafora bodies (LB). These consist of dense aggregates of polyglucosans, which are malformed glycogen molecules lacking normal regular branching.[3] Polyglucosan formation appears to initiate in association with the rough endoplasmic reticulum (RER);[4] in brain, LB are found in the same neuronal compartments as RER,[5] namely perikarya and dendrites, but not axons.[6] Neuronal degeneration in LD occurs through an unusual non-apoptotic generalised organellar disintegration as documented in a murine model.[7]

In ~70% of patients, the disease is caused by mutations in the *EPM2A* gene.[8] Laforin, the protein product of *EPM2A*, localises to the RER,[9 10] and contains a carbohydrate binding module[11] with which it appears to bind glycogen.[12] Laforin also contains a dual specificity phosphatase domain,[9 10] the phosphoprotein substrate of which is still unknown.

The remaining LD families do not genetically link to the *EPM2A* locus.[13] The number of loci containing genes that cause LD is not known, but the mode of inheritance is consistent with autosomal recessive in all families studied to date. Here, we report the identification of a second LD locus.

PATIENTS AND METHODS

The genome scan was based on four families of French-Canadian (F-C) origin with eight affected members. The parents in three families (LD7, LD28, LD27) were first

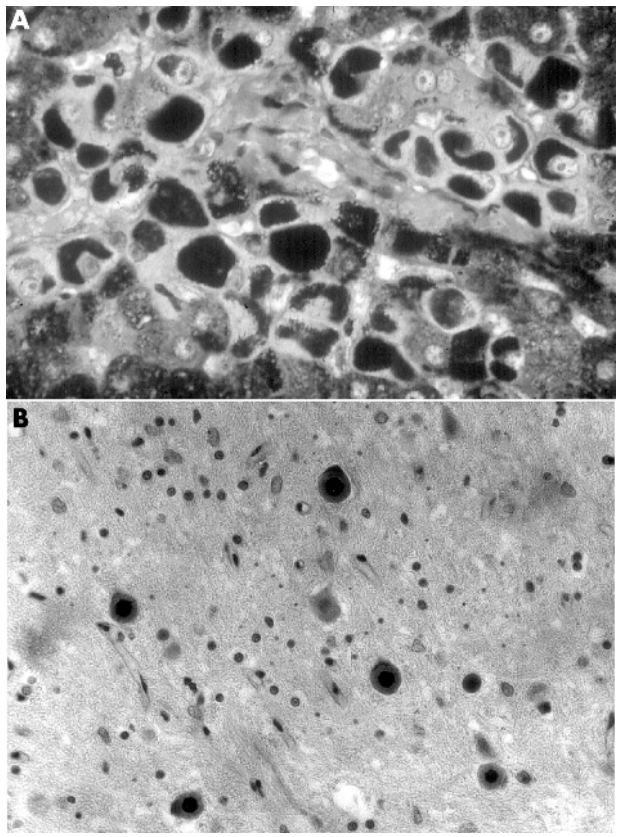

Figure 1 (A) Intense staining of large cytoplasmic inclusions (darkest structures in image), with peripheral displacement of nuclei, in hepatocytes from a periportal region (patient A, liver biopsy). (B) Numerous Lafora bodies with a dark core and paler peripheral zone in dentate nucleus (autopsy, another study patient). Periodic acid Schiff, original magnification × 500.

Abbreviations: LB, Lafora bodies; LD, Lafora disease; RER, rough endoplasmic reticulum

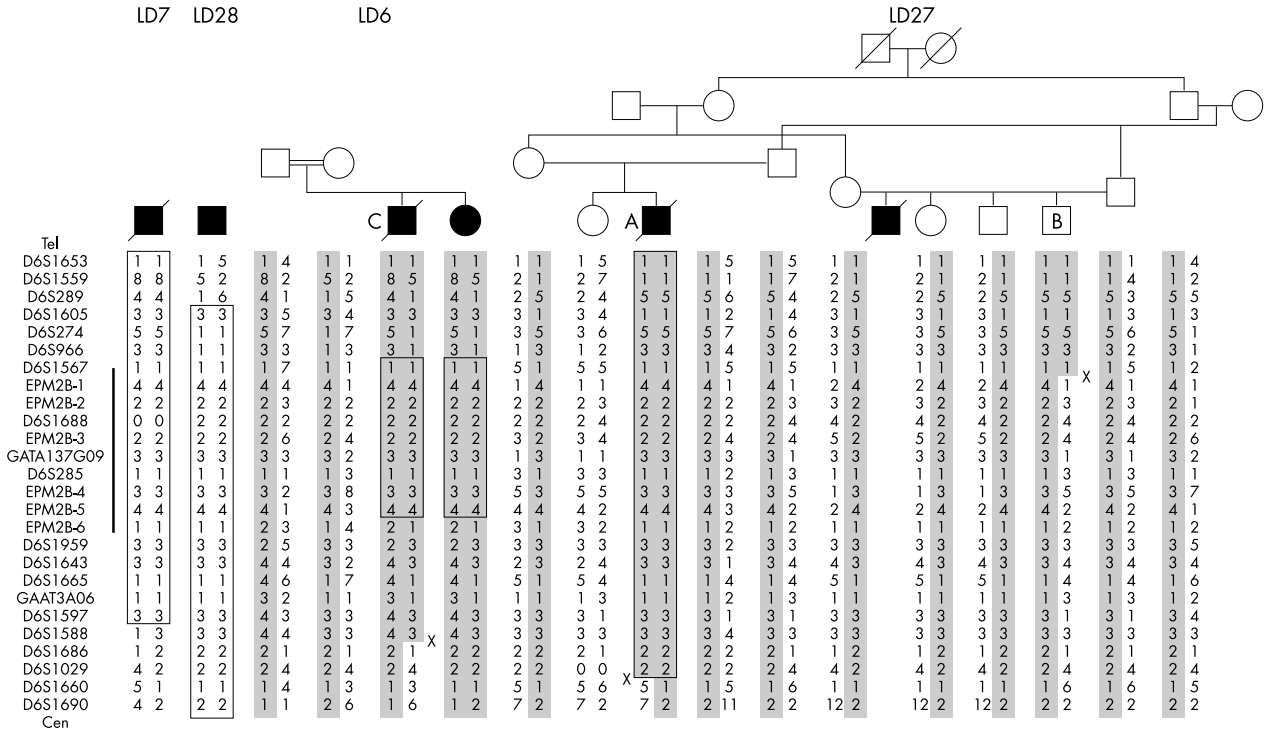

Figure 2 Haplotypes of French Canadian pedigrees LD6 and LD27, and one affected individual from each of families LD7 and LD28. Boxes indicate regions of homozygosity in affected individuals. X marks recombination intervals. Shading in LD6 and LD27 indicates chromosomes carrying mutations, and their contributions to recombinants. No recombinants were present in LD7 and LD28 across the region. Recombination analysis places the telomeric boundary of the critical region at D6S1567 (individual B) and the centromeric boundary at D6S1686 (individual C). Homozygosity analysis further refines the centromeric border to EPM2B-6: in LD6, this marker flanks the nine-marker homozygous haplotype shared among all patients in all families. The vertical bar indicates the final critical interval: D6S1567 to EPM2B-6.

degree cousins. Although the exact pedigree structure is unknown in the fourth family, LD6, consanguinity was confirmed through a review of church records. Linkage to the *EPM2A* locus on 6q24 was excluded in these families,[13] and sequencing of the *EPM2A* gene revealed no mutations.

All patients followed a typical clinical course with disease onset within a year around the 12th birthday. All had LB on

Table 1 Two-point lod scores for microsatellite markers on 6p22

Marker	Sex averaged distance (KosambicM)*	Physical distance (Mb)†	Lod score at θ values of								
			0	0.01	0.05	0.1	0.2	0.3	0.4	Z Max	θ Max
D6S1653	26.71	14.51	0.43	1.07	1.41	1.35	0.97	0.55	0.20	1.41	0.060
D6S1559	29.93	15.21	−∞	0.68	1.58	1.65	1.24	0.71	0.28	1.67	0.081
D6S289	29.93	15.34	−1.41	0.53	1.43	1.52	1.16	0.68	0.28	1.53	0.085
D6S1605	31.01	16.37	−∞	1.82	2.16	2.03	1.47	0.87	0.35	2.16	0.049
D6S274	32.62	16.80	−∞	3.08	3.29	2.97	2.08	1.19	0.46	3.32	0.035
D6S966	33.43	17.46	2.31	2.25	2.01	1.72	1.15	0.64	0.24	2.31	0.001
D6S1567	33.43	17.50	1.36	1.31	1.11	0.86	0.45	0.16	0.025	1.36	0.001
EPM2B-1	—	17.86	2.33	2.25	1.91	1.51	0.80	0.32	0.081	2.33	0.001
EPM2B-2	—	18.17	1.89	1.83	1.62	1.36	0.88	0.48	0.19	1.89	0.001
D6S1688	34.23	18.34	0.89	0.87	0.77	0.66	0.46	0.28	0.13	0.89	0.001
EPM2B-3	–	18.40	4.45	4.34	3.86	3.28	2.16	1.17	0.42	4.45	0.001
GATA137G09	34.23	18.55	2.36	2.30	2.05	1.74	1.15	0.64	0.25	2.36	0.001
D6S285	34.23	18.63	2.43	2.37	2.12	1.82	1.22	0.71	0.30	2.43	0.001
EPM2B-4	—	19.39	3.61	3.50	3.08	2.55	1.58	0.78	0.24	3.61	0.001
EPM2B-5	—	19.61	3.27	3.18	2.83	2.40	1.57	0.86	0.31	3.27	0.001
EPM2B-6	—	19.74	3.08	2.99	2.64	2.21	1.40	0.73	0.26	3.08	0.001
D6S1959	34.23	19.97	2.84	2.76	2.45	2.06	1.34	0.71	0.24	2.84	0.001
D6S1643	35.66	20.44	4.14	4.02	3.57	3.00	1.92	0.99	0.33	4.14	0.001
D6S1665	36.37	21.05	5.22	5.10	4.61	4.00	2.78	1.63	0.66	5.22	0.001
GAAT3A06	36.37	21.27	2.91	2.84	2.56	2.20	1.50	0.85	0.32	2.91	0.001
D6S1597	37.79	21.78	2.55	2.48	2.20	1.86	1.22	0.67	0.25	2.55	0.001
D6S1588	38.24	22.10	−1.04	−0.35	0.14	0.25	0.21	0.11	0.029	2.57	0.122
D6S1686	39.20	22.13	−∞	−0.54	0.45	0.63	0.47	0.22	0.54	6.35	0.108
D6S1029	39.20	22.14	−∞	−1.43	−0.41	−0.20	−0.24	−0.30	−0.22	—	—
D6S1660	40.14	23.37	−∞	−1.87	−0.12	0.42	0.61	0.45	0.21	6.11	0.182
D6S1691	42.27	23.98	−∞	−2.40	−0.65	−0.091	0.17	0.13	0.049	1.69	0.219

*Map locations from the Center for Medical Genetics (http://research.marshfieldclinic.org/genetics/); the sequence of all primers for all markers are in Genbank.
†Physical distance from University of California, Santa Cruz (www.genome.ucsc.edu/) human genome assembly browser (November 2002).

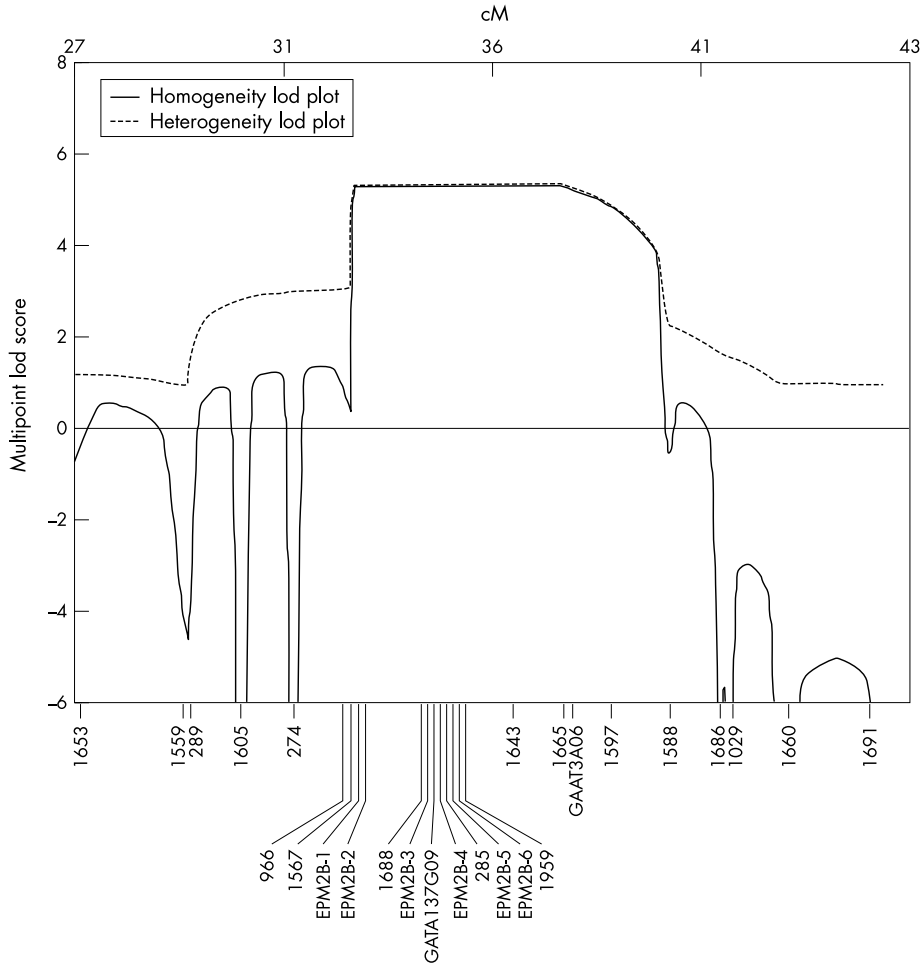

Figure 3 Multipoint lod scores with microsatellite markers in 6p22.3. Homogeneity and heterogeneity lod scores are shown. The *x* axis indicates the relative microsatellite marker positions in cM from the 6p telomere, and the *y* axis, the lod scores. All marker numbers are D6S except for those indicated in figure. A maximum lod score of 5.37 was generated across 13 markers flanked by D6S1567 and D6S1597.

biopsy or at autopsy (fig 1). As an illustrative example, patient A (fig 2) developed normally until the age of 12 years, when he had his first generalised tonic-clonic seizure, followed soon after by the appearance of myoclonic jerks. EEG showed multiple spike and polyspike-slow wave generalised discharges enhanced by photic stimulation. With seizure intractability and mental decline, a liver biopsy was performed (fig 1A), and the diagnosis of LD was made. By the age of 17 years, the patient was bedridden, tube fed and demented. He died at 20 years of age. Brain autopsy was declined, but histology of the spinal cord revealed numerous pathognomonic neuronal LB (shown in reference 2). Patient A had a first cousin who died from LD at 19 years of age. Their fathers are brothers, their mothers sisters, and the fathers are first degree cousins of the mothers (fig 2).

The genome wide linkage scan was performed using Weber panel version 8A (www.marshfieldclinic.org/research/genetics). A total of 387 fluorescently labelled, highly polymorphic microsatellite markers were included in the screening set, with an average heterozygosity of 0.77 calculated from CEPH families. Markers had an average intermarker spacing distance of approximately 10 centimorgans (cM). Chromosomal regions that generated a maximum two point lod score of >1.0 with the disease were genotyped using additional flanking markers. Linkage analysis was performed, specifying an autosomal recessive mode of inheritance with no phenocopies allowed, and a disease allele frequency of 0.001. Marker allele frequencies for the genome wide linkage scan were calculated using the pedigrees being analysed (n = 9-18), and a panel of ninety F-C control individuals were genotyped for five markers at the 6p22 region. Pairwise linkage analysis was calculated using the programs MLINK and ILINK.[14–16] In the analysis, inbreeding loops were broken according to Terwilliger and Ott.[17] The lod scores were calculated for recombination fractions ranging from 0 to 0.5 and under the assumption that there was no sex difference in recombination rates. Multipoint linkage analysis was performed using GeneHunter (version 2.1_r3 beta)[18 19] with inbreeding loops kept intact. Intermarker, sex averaged distances were obtained from published maps.[20]

RESULTS

The genome wide scan of 387 markers was initiated in the F-C families, and after 303 markers were genotyped, linkage to the D6S1969 microsatellite on chromosome 6p22 was detected with a two-point lod score of 2.84 at a recombination fraction of $\theta = 0.00$. To fine map the putative LD locus, 25 additional microsatellites spanning a region of 16 cM were genotyped in the four families. This resulted in maximum two-point lod scores ranging from 0.89–6.35 for 25 microsatellites at $\theta = 0.001$–0.219 (table 1). The EPM2B markers

are new microsatellites identified in this study (Genbank accession numbers: EPM2B-1, BVO12563; −2, BVO12564; −3, BVO12565; −3, BVO12566; −4, BVO12567; BVO12568). Multipoint linkage analysis was analysed by GeneHunter using the 26 markers spanning a genetic distance of 16 cM. This generated the same maximum homogeneity and heterogeneity lod scores of 5.37 across 13 markers in the critical interval (fig 3) and produced a heterogeneity parameter, α, of 1.00 across this region.

Haplotype analysis revealed a critical recombinant between D6S1567 and EPM2B-1 in the paternal chromosome of individual B, establishing the telomeric boundary of the disease gene region at D6S1567 (fig. 2). This person is unaffected and is 15 years older than the age of disease onset in these families. The centromeric boundary of the critical interval is defined by a recombination event observed in the maternal chromosome of individual C between D6S1588 and D6S1686. Regions of homozygosity shared by affected individuals in all four families allowed further fine mapping of the centromeric border. A nine marker haplotype (D6S1567, EPM2B-1, EPM2B-2, D6S1688, EPM2B-3, GATA137G09, D6S285, EPM2B-4, EPM2B-5) is identical and homozygous in all patients. Loss of homozygosity in the affected individuals in family LD6 beyond EPM2B-5 places the centromeric boundary at the next marker, EPM2B-6. The final critical interval is therefore between D6S1567 (telomeric) and EPM2B-6 (centromeric), spanning a distance of 2 235 308 bases. A five marker haplotype present in affected individuals of families LD7, LD27 and LD28 at markers D6S1567, D6S285, D6S1959, D6S1665, and GAAT3A06 spanning our critical interval was found to be absent in 180 F-C control chromosomes; no significant pairwise linkage disequilibrium was observed between these markers ($P>0.17$, alleles<20% frequency were pooled).

DISCUSSION

We have identified a second LD gene locus, named *EPM2B*, at 6p22 based on a study of LD families from a F-C isolate. Affected individuals from these families are probably segregating a single mutation. All four families are of exclusively F-C descent, and all originate from the same region of east Quebec (lower Saint-Lawrence, southern shore; P. Rioux *et al*., in preparation), their linkage to the new locus is highly significant (two-point maximum lod score = 5.2), and all share the same nine marker disease haplotype. All but two of the patients in this study have died since they were recruited.

The current genetic mapping data localises the *EPM2B* locus to a 2.2 Mb critical region. Known proteins encoded in this region include the acute myeloid leukaemia associated protein, DEK,[21] thiopurine s-methyl transferase (TPMT), responsible for the TPMT deficiency trait,[22] and KIF13A, a kinesin heavy chain protein similar to *Drosophila* kinesin-73.[23] Kinesins are microtubule dependent motor proteins that are involved in transporting cellular components. In neurones, they play a major role in axonal and dendritic transport.[24] Hypothetical proteins in the region include a zinc finger, RING domain containing protein (AL832329) and an amino oxidase domain containing protein (FLJ34109). Mutation analysis of all known and putative genes in the region is in progress.

EPM2A disclosed a novel RER associated glycogen binding tyrosine phosphatase and recently its first interacting partner.[25] The *EPM2B* protein should further unravel this biological pathway and elucidate the cellular mechanisms underlying the unique RER associated disturbance of glycogen metabolism, the non-apoptotic neurodegeneration, and the florid neuronal hyperexcitability and intractable epilepsy of LD.

Analysis of the present locus in other non-*EPM2A* families will reveal whether there is further locus heterogeneity. Identification of the *EPM2B* gene will improve patient diagnosis, prenatal diagnosis and genetic counselling for this devastating epilepsy.

ACKNOWLEDGEMENTS

We would like to thank Dr P Rioux and C Truchon for sharing the genealogical data prior to publication, the Quebec Lafora Disease Association families, patients and president (Mrs O Malenfant) for their unwavering support despite very trying times, and the Centre for Applied Genomics (http://tcag.bioinfo.sickkids.on.ca/) at The Hospital for Sick Children for technical support. This work was funded by the Canadian Institute of Health Research (CIHR) and Genome Canada. ADP holds a Canada Research Chair in Genetics of Complex Diseases. SWS is a Scholar of CIHR and International Scholar of the Howard Hughes Medical Institute.

Authors' affiliations

E M Chan, Program in Genetics and Genomic Biology, The Hospital for Sick Children and Department of Molecular and Medical Genetics, The University of Toronto, 555 University Ave, Toronto, Ontario M5G 1X8, Canada
D E Bulman, Division of Neurology, Ottawa Health Research Institute, Ottawa Hospital - General Campus, 501 Smyth Rd, Ottawa, Ontario K1H 8L6, Canada
A D Paterson, Program in Genetics and Genomic Biology, The Hospital for Sick Children, 555 University Ave, Toronto, Ontario M5G 1X8, Canada
J Turnbull, Program in Genetics and Genomic Biology, The Hospital for Sick Children, 555 University Ave, Toronto, Ontario M5G 1X8, Canada
E Andermann, Neurogenetics Unit, Montreal Neurological Hospital and Institute, Departments of Neurology and Neurosurgery, and Human Genetics, McGill University, 3801 University St, Montreal, Quebec H3A 2B4, Canada
F Andermann, Epilepsy Service, Montreal Neurological Hospital and Institute and Departments of Neurology and Neurosurgery, and Paediatrics, McGill University, 3801 University St, Montreal, Quebec H3A 2B4, Canada
G A Rouleau, Centre for Research in Neuroscience, McGill University Health Centre Research Institute, Montreal General Hospital, 1650 Cedar Ave, Montreal, Quebec H3G 1A4, Canada
A V Delgado-Escueta, Comprehensive Epilepsy Program, Department of Neurology and Brain Research Institute, University of California, Los Angeles School of Medicine, West Los Angeles DVA Medical Center, 11301 Wilshire Blvd, Los Angeles, California 90073, USA
S W Scherer, Program in Genetics and Genomic Biology, The Hospital for Sick Children and The University of Toronto, 555 University Ave, Toronto, Ontario M5G 1X8, Canada
B A Minassian, Division of Neurology, Department of Pediatrics, Program in Genetics and Genomic Biology, The Hospital for Sick Children and Department of Molecular and Medical Genetic, The University of Toronto, 555 University Ave, Toronto, Ontario M5G 1X8, Canada

REFERENCES

1 **Lafora GR**, Gluck B. Beitrag zur histopathologie der myoklonischen epilepsie. *Z Ges Neurol Psychiat* 1911;**6**:1–14.
2 **Minassian BA**. Progressive myoclonus epilepsy with polyglucosan bodies: Lafora disease. *Adv Neurol* 2002;**89**:199–210.
3 **Sakai M**, Austin J, Witmer F, Trueb L. Studies in myoclonus epilepsy (Lafora body form). II.Polyglucosans in the systemic deposits of myoclonus epilepsy and in corpora amylacea. *Neurology* 1970;**20**:160–76.
4 **Collins GH**, Cowden RR, Nevis AH. Myoclonus epilepsy with Lafora bodies. An ultrastructural and cytochemical study. *Arch Pathol* 1968;**86**:239–54.
5 **Cavanagh JB**. Corpora-amylacea and the family of polyglucosan diseases. *Brain Res Brain Res Rev* 1999;**29**:265–95.
6 **Steward O**, Banker GA. Getting the message from the gene to the synapse: sorting and intracellular transport of RNA in neurons. *Trends Neurosci* 1992;**15**:180–6.
7 **Ganesh S**, Delgado-Escueta AV, Sakamoto T, Avila MR, Machado-Salas J, Hoshii Y, Akagi T, Gomi H, Suzuki T, Amano K, Agarwala KL, Hasegawa Y, Bai DS, Ishihara T, Hashikawa T, Itohara S, Cornford EM, Niki H, Yamakawa K. Targeted disruption of the *epm2a* gene causes formation of Lafora inclusion bodies, neurodegeneration, ataxia, myoclonus epilepsy and impaired behavioral response in mice. *Hum Mol Genet* 2002;**11**:1251–62.

8 **Minassian BA**, Lee JR, Herbrick JA, Huizenga J, Soder S, Mungall AJ, Dunham I, Gardner R, Fong CY, Carpenter S, Jardim L, Satishchandra P, Andermann E, Snead OC 3rd, Lopes-Cendes I, Tsui LC, Delgado-Escueta AV, Rouleau GA, Scherer SW. Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. *Nat Genet* 1998;**20**:171–4.
9 **Ganesh S**, Agarwala KL, Ueda K, Akagi T, Shoda K, Usui T, Hashikawa T, Osada H, Delgado-Escueta AV, Yamakawa K. Laforin, defective in the progressive myoclonus epilepsy of Lafora type, is a dual-specificity phosphatase associated with polyribosomes. *Hum Mol Genet* 2000;**9**:2251–61.
10 **Minassian BA**, Andrade DM, Ianzano L, Young EJ, Chan E, Ackerley CA, Scherer SW. Laforin is a cell membrane and endoplasmic reticulum-associated protein tyrosine phosphatase. *Ann Neurol* 2001;**49**:271–5.
11 **Minassian BA**, Ianzano L, Meloche M, Andermann E, Rouleau GA, Delgado-Escueta AV, Scherer SW. Mutation spectrum and predicted function of laforin in Lafora's progressive myoclonus epilepsy. *Neurology* 2000;**55**:341–6.
12 **Wang J**, Stuckey JA, Wishart MJ, Dixon JE. A unique carbohydrate binding domain targets the Lafora disease phosphatase to glycogen. *J Biol Chem* 2002;**277**:2377–80.
13 **Minassian BA**, Sainz J, Serratosa JM, Gee M, Sakamoto LM, Bohlega S, Geoffroy G, Barr C, Scherer SW, Tomiyasu U, Carpenter S, Wigg K, Sanghvi AV, Delgado-Escueta AV. Genetic locus heterogeneity in Lafora's progressive myoclonus epilepsy. *Ann Neurol* 1999;**45**:262–5.
14 **Lathrop GM**, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA* 1984;**81**:3443–6.
15 **Schaffer AA**, Gupta SK, Shriram K, Cottingham RW Jr. Avoiding recomputation in linkage analysis. *Hum Hered* 1994;**44**:225–37.
16 **Cottingham Jr RW**, Idury RM, Schäffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet* 1993;**53**:252–63.
17 **Terwilliger JD**, Ott J. *Handbook of human genetic linkage*. Baltimore: Johns Hopkins University Press, 1994.
18 **Kruglyak L**, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified mulitpoint approach. *Am J Hum Genet* 1996;**58**:1347–63.
19 **Kruglyak L**, Lander ES. Faster multipoint linkage analysis using Fourier transforms. *J Comput Biol* 1998;**5**:1–7.
20 **Broman KW**, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998;**63**:861–9.
21 **Fu GK**, Grosveld G, Markovitz DM. DEK, an autoantigen involved in a chromsomal translocation in acute myelogenous leukemia, binds to the HIV-2 enhancer. *Proc Nat Acad Sci* 1997;**94**:1811–15.
22 **Krynetski EY**, Schuetz JD, Galpin AJ, Pui CH, Relling MV, Evans WE. A single point mutation leading to loss of catalytic activity in humans. A single point mutation leading to loss of catalytic activity in human thiopurine S-methyltransferase. *Proc Natl Acad Sci USA* 1995;**92**:949–53.
23 **Jamain S**, Quach H, Fellous M, Bourgeron T. Identification of the human KIF13A gene homologous to Drosophila kinesin-73 and candidate for schizophrenia. *Genomics* 2001;**74**:36–44.
24 **Goldstein LS**, Yang Z. Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. *Annu Rev Neurosci* 2000;**23**:39–71.
25 **Ianzano L**, Zhao XC, Minassian BA, *et al*. Identification of a novel protein interacting with laforin, the EPM2A progressive myoclonus epilepsy gene-product. *Genomics* 2003;**81**:579–87.