Biochemical and Biophysical Research Communications **257,** 24–28 (1999)
Article ID bbrc.1999.0402, available online at http://www.idealibrary.com on IDEAL®

Isolation and Characterization of Mouse Homologue for the Human Epilepsy Gene, *EPM2A*

Subramaniam Ganesh,* Kenji Amano,* Antonio V. Delgado-Escueta,† and Kazuhiro Yamakawa*,[1]
**Laboratory for Neurogenetics, Brain Science Institute, Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan; and †Comprehensive Epilepsy Program, UCLA School of Medicine and West Los Angeles DVA Medical Center, 11301 Wilshire Boulevard, Los Angeles, California 90073*

Received February 15, 1999

Mutations in the novel gene, *EPM2A*, have been shown recently to cause the progressive myoclonus epilepsy of Lafora type. *EPM2A* is predicted to encode a putative protein-tyrosine phosphatase but its specific role in normal brain function and in the Lafora disease is not known. As a first step towards understanding the cellular function of *EPM2A* in an animal model, we have isolated cDNA clones for mouse *EPM2A* and analyzed its expression. Sequence analyses of the mouse cDNA clones revealed a complete ORF that supports the 5′ coding sequence predicted for human *EPM2A* from the genomic sequence. When compared to *EPM2A*, the mouse homologue, named *Epm2a*, shows 86% identity at the nucleotide level and 88% identity and 93% similarity at the amino acid level. Similar to the human counterpart, *Epm2a* showed ubiquitous expression in Northern with a major transcript size of 3.5 kb. We have mapped the *Epm2a* to the proximal region of mouse chromosome 10 which is the syntenic region for human chromosome band, 6q24. Our results suggest that *EPM2A* is highly conserved in mammals and might have a conserved function.

The progressive myoclonus epilepsy of the Lafora type is a rare but lethal autosomal recessive disorder, characterized by the presence of periodic acid stain positive intracellular inclusion bodies (1, 2). Lafora disease (LD) usually manifests as epileptic seizures during adolescence with a mean age at onset about 15 years and a rapid and progressive dementia follows often with psychotic features and the patient die within a decade after the onset (3). Earlier, Serratosa *et al.* (4) and Sainz *et al.* (5) have mapped the LD locus to chromosome 6q24 and recently, by positional cloning, Minassian *et al.* (6) identified at 6q24 a novel gene, named *EPM2A*, that encodes a putative protein for protein-tyrosine phosphatase. *EPM2A* transcripts are alternatively spliced and are detected in all the tissues analyzed including brain. Mutations in *EPM2A* have been found in a number of LD families and they cosegregate with the disease phenotype (6, 7).

Malfunctions of protein-tyrosine phosphatases (PTP) are linked to a number of oncogenic disorders (8, 9), but the discovery of *EPM2A* has been the second example for their involvement in an non-neoplastic disease (6). While PTPs are known to play critical roles in a number of cellular signal transduction pathways, the specific function of *EPM2A* in the normal brain and in LD is not yet known. In recent years, cellular role of various genes associated with neuronal disorders have been understood using model systems and therefore, such studies may reveal the function of *EPM2A* as well. To investigate the *EPM2A* function in an animal model, we have cloned the *EPM2A* homologue of mouse and report here its characterization, expression pattern and chromosomal localization.

MATERIALS AND METHODS

Screening of cDNA library. Based on the partial exonic sequences reported for the human *EPM2A* (6), primers were designed to amplify exons 2 and 4 from the human genomic DNA. The nucleotide sequences for the forward and reverse primers respectively for exon 2 and 4 are: 5-GCTCTTTACAAATTCCTACCGTTCTGT-3, 5-AGTGAGGCACTGCAGTTTCGA-3, 5-TGTGATGGGCTGGAATCTGAG-3, and 5-GAGCAGGAA CTGCACGCATTA-3. The partial coding sequence of human *EPM2A* was amplified by PCR and the purified PCR products were used to screen mouse brain cDNA library of ICR out-bred strain (Stratagene). The library was constructed from cDNAs primed with oligo(dT) and random primers and cloned into Lambda ZAP-II vector. The established method of Sambrook *et al.* (10) was used for the library screening. Phage lifts from 20 plates containing approximately 2×10^6 plaques were hybridized overnight and positive plaques were identified using AlkPhos direct labeling and detection kit (Amersham Lifesciences). The filters were washed at 55°C in the recommended wash buffer and processed for chemiluminescent detection as per the instruc-

[1] To whom correspondence should be addressed at Laboratory for Neurogenetics, Brain Science Institute, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan. Fax: 81-48-462-4796. E-mail: yamakawa@brain.riken.go.jp.

0006-291X/99 $30.00

tions. Selected phages were transformed into plasmids by *in vivo* excision as recommended (Stratagene) and the nucleotide sequence was determined by dye-terminator method using ABI autosequencer type 377.

Northern blot analysis. Mouse multiple tissue Northern blot was purchased from Clontech laboratories and hybridization was carried out as specified. An 1.4 kb insert of the mouse *EPM2A* cDNA clone LDM3 was labeled with [^{32}P]dCTP and hybridized overnight with ExpressHyb hybridization solution (Clontech). The blot was washed finally in 0.1× SSC/0.1% SDS at 55°C and exposed to X-ray film at −80°C.

Screening of genomic library. Mouse C57BL/6N strain genomic library constructed in EMBL3SP6/T7 vector (Clontech) was screened using the 1.4 kb mouse *Epm2a* cDNA clone, LDM3, as a probe. Approximately, 1×10^6 plaques were screened using the AlkPhos labeling and detection method as described above. The clones were partially sequenced using primers derived from the cDNA sequence of mouse. For sizing the insert, 500 ng of the DNA was digested with 3 U of Sal I enzyme for 4 hr to release the vector arms and the samples were run in 1% agarose gel (SeaKem GTG, FMC) on a CHEF mapper (Bio-Rad).

Fluorescence in situ hybridization (FISH). The direct R-banding FISH method of Matsuda *et al.* (11) was followed for the chromosomal assignment of mouse *EPM2A* gene. Mitogen-stimulated spleen lymphocyte culture was synchronized by thymidine block and 5-bromodeoxyuridine was incorporated during the late replication stage. R banding was performed by exposure to UV light after staining with Hoechst 33258. For hybridization, the chromosome slides were denatured at 70°C in 70% formamide with 2× SSC. The DNA of mouse 15 kb genomic clone B6LDM8 was labeled with biotin 16-dUTP by nick translation (Boehringer-Mannheim) and hybridized to the R-banded slides at 37°C in the presence of Cot-I DNA and salmon sperm DNA. The hybridization solution constitutes 50% formamide, 2× SSC, 10% dextran sulfate and 2 mg/ml BSA. The slides were washed in 50% formamide in 2× SSC at 37°C followed by 2× and 1× SSC at room temperature. Subsequently the slides were stained with fluorescein-streptavidin and counterstained with propidium iodide. The image was captured by Photometrics Cooled-CCD camera using the Leica QFISH system (Leica).

RESULTS

Isolation of mouse EPM2A homologue (Epm2a). Screening of the mouse cDNA library with the human *EPM2A* PCR products yielded three different clones of insert sizes 0.6, 0.8 and 1.4 kb and nucleotide sequence analysis confirmed that two smaller ones were in fact part of the larger clone. An ORF prediction of the sequence from the 1.4 kb cDNA clone LDM3 (GenBank accession number: AF124044) revealed an extended coding frame (Figure 1). The previously reported cDNA clones for human *EPM2A* is partial since they lacked the 5′ coding region (6, 7). However, a potential first exon was suspected in the genomic sequence as it extended the ORF of the most 5′ sequence of cDNA clone and also showed all features expected for an eukaryotic gene (6). Interestingly, the 5′ coding sequence of the mouse cDNA clone LDM3 showed very high homology to the predicted first exon of the human gene in nucleotide as well as to its deduced amino acid sequence (Figure 2). The start site of the coding region for the mouse clone has been defined on aligning the deduced amino acid sequence with the sequence predicted for human *EPM2A*. Although the 66 bases preceding the predicted start site did not show an in-frame stop codon, the sequence flanking the presumed start site

```
                                                  gcgccgccccggact   -51
cgcgccgccgccgctgccgctcgccaggctcccaggactcgggccaccgcc     0

ATGCTCTTCCGCTTCGGCGTGGTGGTGCCGCCCGCGGTGGCCGGCGCTCGG    51
 M  L  F  R  F  G  V  V  V  P  P  A  V  A  G  A  R     17

CAGGAGCTGCTGCTCGCCGGCTCGCGGCCCGAGTTAGGGCGCTGGGAGCCG    102
 Q  E  L  L  L  A  G  S  R  P  E  L  G  R  W  E  P     34

CACGGCGCCGTGCGTCTCAGGCCCGCGGGTACCGCGGCGGGCGCCGCTGCG    153
 H  G  A  V  R  L  R  P  A  G  T  A  A  G  A  A  A     51

CTGGCGCTGCAGGAGCCCGGCCTGTGGCTCGCCGAGGTGGAGCTGGAGGCG    204
 L  A  L  Q  E  P  G  L  W  L  A  E  V  E  L  E  A     68

TACGAGGAGGCCGGCGGGGCGGAGCCGGGCCGCGTTGACACGTTCTGGTAC    255
 Y  E  E  A  G  G  A  E  P  G  R  V  D  T  F  W  Y     85

AAGTTCCTGCAGCGCGAGCCTGGAGGCGAGCTGCACTGGGAAGGAAATGGA    306
 K  F  L  Q  R  E  P  G  G  E  L  H  W  E  G  N  G    102

CCTCACCATGACCGTTGCTGCACATATAATGAGGACAACTTGGTGGATGGT    357
 P  H  H  D  R  C  C  T  Y  N  E  D  N  L  V  D  G    119

GTTTATTGTCTCCCAGTAGGACACTGGATTGAGGCCACTGGGCACACCAAT    408
 V  Y  C  L  P  V  G  H  W  I  E  A  T  G  H  T  N    136

GAAATGAAGCGCACAACAGACTTCTATTTTAATATTGCTGGCCACCAAGCC    459
 E  M  K  R  T  T  D  F  Y  F  N  I  A  G  H  Q  A    153

ATGCACTATTCAAGAATTCTACCAAATATCTGGCTGGGTAGCTGCCCTCGC    510
 M  H  Y  S  R  I  L  P  N  I  W  L  G  S  C  P  R    170

CAACTGGAACATGTGACCATCAAACTGAAGCATGAACTGGGAGTTACAGCT    561
 Q  L  E  H  V  T  I  K  L  K  H  E  L  G  V  T  A    187

GTCATGAATTTCCAGACTGAATGGGATATCATCCAGAATTCTTCAGGCTGC    612
 V  M  N  F  Q  T  E  W  D  I  I  Q  N  S  S  G  C    204

AACCGCTACCCTGAACCCATGACTCCAGACACCATGATGAAGCTGTATAAG    663
 N  R  Y  P  E  P  M  T  P  D  T  M  M  K  L  Y  K    221

GAAGAAGGCTTGTCCTACATCTGGATGCCCACTCCAGACATGAGCACTGAG    714
 E  E  G  L  S  Y  I  W  M  P  T  P  D  M  S  T  E    238
                                                  ↓
GGCCGAGTGCAGATGCTGCCACAGGCTGTGTGTCTCCTGCACGCGCTTCTG    765
 G  R  V  Q  M  L  P  Q  A  V  C  L  L  H  A  L  L    255

GAGAATGGACACACGGTGTATGTCCACTGCAACGCTGGCGTGGGTCGCTCC    816
 E  N  G  H  T  V  Y  V  H  C  N  A  G  V  G  R  S    272

ACAGCTGCAGTGTGCGGCTGGCTCCACTATGTGATTGGCTGGAATCTGCGC    867
 T  A  A  V  C  G  W  L  H  Y  V  I  G  W  N  L  R    289

AAGGTGCAGTACTTCATCATGGCCAAAAGGCCTGCGGTCTACATTGACGAG    918
 K  V  Q  Y  F  I  M  A  K  R  P  A  V  Y  I  D  E    309

GACGCTTTGGCTCAAGCACAACAAGACTTTTCTCAGAAGTTCGGGAAGGTT    969
 D  A  L  A  Q  A  Q  Q  D  F  S  Q  K  F  G  K  V    323

CACTCTTCCATATGCGCTTTGTAGgtgatcagccctccattccgtctagct   1020
 H  S  S  I  C  A  L  *                               330

gccttagtaaggaacctgggggtgtggttagtggaagacctggagacaaag   1071
gggactgtggctggtgtcacctggagattgccctctattgtggcttggttg   1122
ggcttgtgtttgaaatgctttacaaaggaaaactgcatgccacatgagaaa   1173
acttcaaagtacacttgcaatgcaagtgtatgcctgtagttgctgctgcca   1224
ccggcaggggggcctttgcttacagggtctgcataggttgttatggatttgt   1275
gactttggcatgcagaagaccagagtcctctcttgggggttgctgggagga   1326
cttggtgtgatgataaagatgaaagct                           1353
```

FIG. 1. The nucleotide sequences of the 1.4 kb cDNA clone, LDM3, and predicted amino acid sequence of mouse *Epm2a* gene (GenBank accession number: AF124044). Upper and lowercase letters in the nucleotide sequence designate the coding and noncoding regions of the transcript respectively. The boldface letters of the amino acid sequence define the putative phosphatase domain. The start site of the coding region has been defined on aligning the deduced amino acid sequence expected for the human first exon (see Figure 2) but the 66 bases preceding the proposed start site did not show any in-frame stop codon. Arrow defines the last exon as identified by the partial sequencing of the genomic clone, B6LDM8.

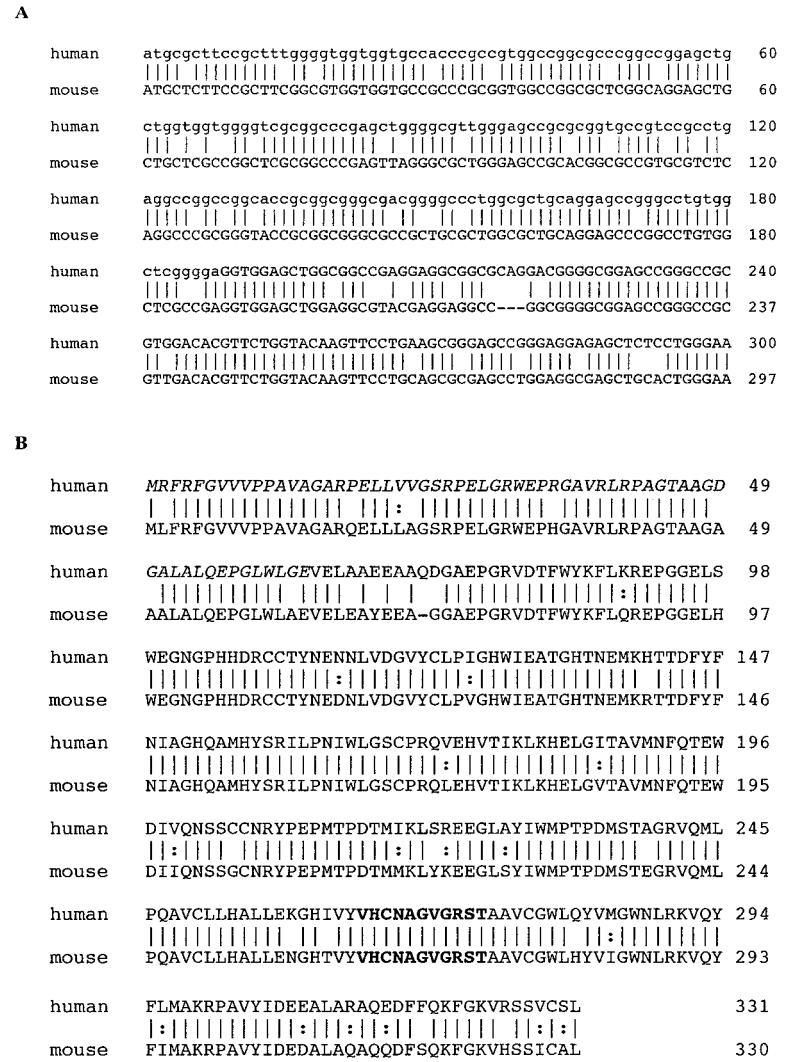

FIG. 2. Nucleotide and amino acid sequence comparison between the human and mouse *EPM2A* homologues. (A) The nucleotide sequence of coding region for the predicted first exon of human *EPM2A* (6) was aligned with the corresponding coding region of mouse *Epm2a*. For the human, the 5′ sequence of partial cDNA (upper case letters; Acc. No. AF084535) and the preceding genomic sequence of the expected first exon (lower case letters; Acc. No. AL023806) were combined together to generate a consensus sequence. Vertical bars indicate the identities. (B) Amino acid sequence alignment for the *EPM2A* gene of human and mouse. The amino acids shown in italics for human are deduced from the predicted coding sequence of the first exon (shown in A). Vertical lines represent identities and dotted lines denote the conservative substitutions. The boldface defines the putative protein-tyrosine phosphatase domain. A gap has been inserted at position 73 of the mouse sequence to compensate for an additional amino acid predicted for the human. This alignment shows 93% similarity and 88% identity between the two sequences.

(ACCGCCATGC) has features of the Kozak's consensus sequence suggested for an eukaryotic translation start site (12). Moreover, in the genomic sequence, the predicted start codon for human *EPM2A* was preceded by an in-frame TAG at 70 bases 5′ to it (6). The deduced *EPM2A* protein of mouse harbors the putative protein-tyrosine phosphatase (PTP) domain and encodes 331 amino acids, one amino acid less to the predicted protein of human (Figure 2B). When compared to human *EPM2A*, the mouse homologue, named *Epm2a*, showed 86% identity at the nucleotide level in the coding region and 88% identity and 93% similarity at the amino acid level (Figure 2).

Isolation of genomic DNA clone for mouse Epm2a. Screening of the mouse genomic library with *Epm2a* cDNA probe yielded two positive clones. End sequencing and restriction patterns revealed that the two clones are apparently identical. Pulsed-field gels estimated the insert size to be 15 kb and hybridization with the PCR products for exon 4 of human *EPM2A* gene suggested that both the clones harbor an exon coding for the putative PTP domain. Therefore, primers were designed for the coding region and one of the genomic clones (B6LDM8) was sequenced for the flanking regions. This partial sequencing revealed the 5′ intron/exon junction (tttctctagGCCGA)

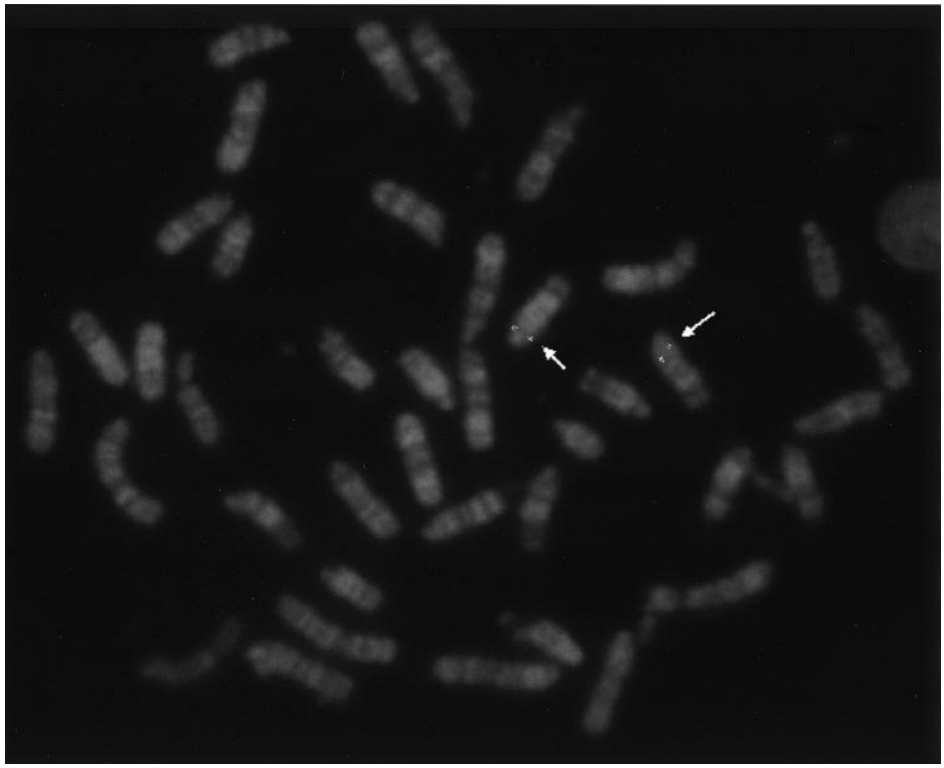

FIG. 3. Chromosomal localization of mouse *Epm2a* gene by FISH. The 15 kb genomic clone, B6LDM8, for mouse *Epm2a* was used as probe on the R banded metaphase spreads made from lymphocytes of an adult mouse. Symmetrical fluorescent signals were observed on the proximal region of chromosome 10 (arrows).

and the 3′UTR sequence for the last exon that codes for the putative PTP domain (see Figure 1). The genomic organization for the last exon of mouse *Epm2a* is identical to the human gene for having the minor consensus sequence AG/GC of mRNA splicing and in both the species, this exon encodes 91 amino acids (6, 13).

Chromosomal localization of Epm2a. The chromosomal locus of *Epm2a* was determined by FISH using the 15 kb genomic clone (B6LDM8) harboring the last exon and the 3′ UTR. On the R-banded metaphase spreads, the genomic probe revealed distinct signals for *Epm2a* on mouse chromosome 10A, a syntenic region for the human chromosome band, 6q24 (Figure 3).

Epm2a expression. To determine the tissue distribution of *Epm2a* transcripts, the 1.4 kb cDNA clone, LDM3, was hybridized to mouse multiple tissue Northern blot (Figure 4). As a control, the same blot was hybridized to actin probe. *Epm2a* showed relatively higher level of expression in the tissues of heart, brain, liver, skeletal muscle and kidney when compared to spleen, lung and testis. While all the tissues analyzed showed a single discrete band of 3.5 kb size, an additional transcript of 1.5 kb was detected only in testis.

DISCUSSION

In the present report, we have described the cloning and characterization of the mouse homologue for the human epilepsy gene, *EPM2A*. We demonstrate that the mouse *Epm2a*, which encodes a putative protein-tyrosine phosphatase, is located on the proximal region of chromosome 10 that is syntenic to 6q24 of human and also the locus for *EPM2A* (6, 14). Unlike the human counterpart (6, 7), the cDNA sequence obtained for mouse *Epm2a* is predicted to contain a complete ORF that supports the 5′ coding sequence derived from the genomic sequence for human *EPM2A* (6). Since no full-length cDNA clone has been identified for the human gene and the complete coding sequence is yet to be identified (6, 7), from this sequence comparison it is reasonable to assume that the human *EPM2A* encodes 332 amino acids and the 5′ genomic sequence represents the first exon.

The predicted protein for *Epm2a* showed 88% identity and 93% similarity to the human homologue and shared 100% identity at the consensus sequence for catalytic site, called the PTP domain (6). Interestingly, *Epm2a* showed ubiquitous expression even in mouse with a major transcript size of 3.5 kb which roughly equals the size of human transcript (6). When com-

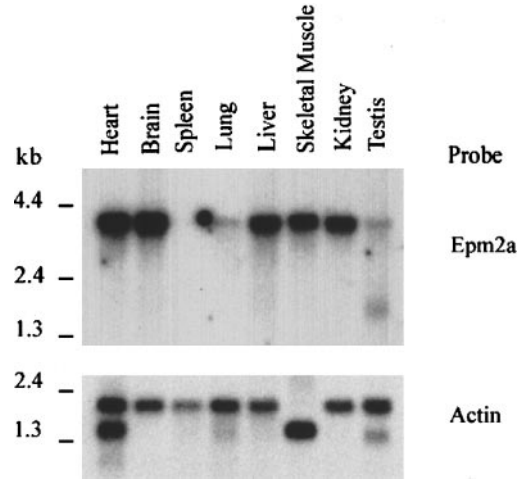

FIG. 4. Northern blot analysis for the mouse *Epm2a* gene. The 1.4 kb insert of the cDNA clone, LDM3, was labeled and used as the probe. The size of the *Epm2a* transcript is around 3.5 kb, but an additional transcript of 1.5 kb is detected in testis. Control hybridization with the actin probe is shown in the lower panel.

pared to other tissues analyzed, testis showed an additional transcript for *Epm2a* although its overall expression was lower in that tissue. Since differentially spliced transcripts were identified for the human *EPM2A* (6, 7), it is likely that the minor transcript of testis represents an isoform of *Epm2a* in mouse. Nevertheless, their very high sequence homology and identical expression in both the species suggest that *EPM2A* is functionally conserved in mammals.

The Lafora bodies found in the LD patients have the properties of an acid mucopolysaccharide and therefore it has been suggested that it could be polyglucosan in nature and an enzyme defect might lead to these depositions and the disorder (15, 16). Since glycogen metabolism is known to be regulated by protein phosphatases (8, 9), a putative role for the *EPM2A* in the formation of Lafora body is speculated (6, 7). Present study on the characterization of *Epm2a* enables the mouse as an attractive animal model for testing the above hypothesis. As has been shown recently for cystatin B gene that causes EPM1 (17), generation of mutant mouse for *Epm2a* should help to understand the cellular functions of *EPM2A* and its role in the pathogenesis of LD.

ACKNOWLEDGMENTS

We thank Dr. S. Itohara for providing the mouse genomic library and extending his animal facility and Ms. E. Miyazaki for her technical help.

REFERENCES

1. Lafora, G. R. (1955) Proceedings of the 2nd International Congress of Neuropathology. Excerpta Medica, Part I, pp. 5–9, Amsterdam.
2. Schwarz, G. A., and Yanoff, M. (1965) *Arch. Neurol.* **12,** 172–188.
3. Van Heycop Ten Ham, M. W. (1974) *in* Handbook of Clinical Neurology (Vinken, P. J., and Bruyn, G. W., Eds.), pp. 171–89. North-Holland, New York.
4. Serratosa, J. M., Delgado-Escueta, A. V., Posada, I., Shih, S., Drury, I.; Berciano, J., Zabala, J. A., Antunez, M. C., and Sparkes, R. S. (1995) *Hum. Mol. Genet.* **4,** 1657–1663.
5. Sainz, J., Minassian, B. A., Serratosa, J. M., Gee, M. N., Sakamoto, L. M., Iranmanesh, R., Bohlega, S., Baumann, R. J., Ryan, S., Sparkes, R. S., and Delgado-Escueta, A. V. (1997) *Am. J. Hum. Genet.* **61,** 1205–1209.
6. Minassian, B. A., Lee, J. R., Herbrick, J. A., Huizenga, J., Soder, S., Mungall, A. J., Dunham, I., Gardner, R., Fong, C. Y., Carpenter, S., Jardim, L., Satishchandra, P., Andermann, E., Snead, O. C., Lopes-Cendes, I., Tsui, L. C., Delgado-Escueta, A. V., Rouleau, G. A., and Scherer, S. W. (1998) *Nature Genet.* **20,** 171–4.
7. Serratosa, J. M., Gomez-Garre, P., Gallardo, M. E., Anta, B., Beltran-Valero de Bernabe, D., Lindhout, D., Augustijn, P. B., Tassinari, C. A., Malafosse, R. M. A., Topcu, M., Grid, D., Dravet, C., Berkovic, S. F., and Rodriguez de Cordoba, S. (1999) *Hum. Mol. Genet.* **8,** 345–352.
8. Tonks, N. K., and Neel, B. G. (1996) *Cell* **87,** 365–368.
9. Zhang, Z. Y. (1998) *Crit. Rev. Mol. Biol.* **33,** 1–52.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Laboratory Press, Cold Spring Harbor, NY.
11. Matsuda, Y,, Harada. Y. N., Natsuume-Sakai, S., Lee, K., Shiomi, T., and Chapman, V. M. (1992) *Cytogenet. Cell Genet.* **61,** 282–5.
12. Kozak, M. (1996) *Mamm. Genome* **7,** 563–574.
13. Jackson, I. J. (1991) *Nucleic Acids Res.* **19,** 3795–3798.
14. Debry, R. W., and Seldin, M. F. (1996) *Genomics* **33,** 337–351.
15. Yokoi, S., Austin, J., Witmer, F., and Sakai, M. (1968) *Arch. Neurol.* **19,** 15–33.
16. Sakai, M., Austin, J., Witmer, F., and Trueb, L. (1970) *Neurol.* **20,** 160–176.
17. Pennacchio, L. A., Bouley, D. M., Higgins, K. M., Scott, M. P., Noebels, J. L., and Myers, R. M. (1998) *Nature Genet.* **20,** 251–258.