Lafora Disease: Molecular Etiology
Lafora Hastalığı: Moleküler Etiyoloji
S. Hande ÇAĞLAYAN¹,²

¹Department of Molecular Biology and Genetics, Boğaziçi University, Istanbul, Turkey
²International Biomedicine and Genome Center, Dokuz Eylül University, İzmir, Turkey

Summary
Lafora Disease (LD) is a fatal neurodegenerative condition characterized by the accumulation of abnormal glycogen inclusions known as Lafora bodies (LBs). Patients with LD manifest myoclonus and tonic-clonic seizures, visual hallucinations, and progressive neurological deterioration beginning at the age of 8-18 years. Mutations in either EPM2A gene encoding protein phosphatase laforin or NHLRC1 gene encoding ubiquitin-ligase malin cause LD. Approximately, 200 distinct mutations accounting for the disease are listed in the Lafora progressive myoclonus epilepsy mutation and polymorphism database. In this review, the genotype-phenotype correlations, the genetic diagnosis of LD, the downregulation of glycogen metabolism as the main cause of LD pathogenesis and the regulation of glycogen synthesis as a key target for the treatment of LD are discussed.

Key words: EPM2A and NHLRC1 gene mutations; genotype-phenotype relationship; Lafora progressive myoclonus epilepsy; LD pathogenesis.

Clinical Characteristics
LD is an autosomal recessive progressive myoclonus epilepsy (PME) characterized by onset of progressive neurodegeneration between the age of 8 and 18 years. The initial features can include headache, difficulties in school work, myoclonic jerks, generalized seizures, depressed mood, cognitive deficits, and frequent visual hallucinations. EEG shows normal or slow background. Photosensitivity is common.¹ Many affected individuals experience isolated febrile or nonfebrile convulsions in infancy or early childhood.² Intra- and interfamilial variability in age at onset is frequently seen.³⁻⁵ Myoclonus, seizures, and hallucinations gradually worsen and become intractable. Dysarthria, ataxia, and spasticity, accompanied by progressive cognitive decline and dementia, are found on neurological examination. EEG shows slow background, paroxysms of generalized irregular spike-wave discharges with occipital dominance, and focal abnormalities. About 10 years after onset, affected individuals are in near-continuous myoclonus with absence seizures, frequent generalized seizures, and profound dementia or a vegetative state.⁶⁻⁷

Diagnosis is usually based on clinical and EEG findings. The presence of periodic acid-Schiff-positive (PAS+) polyglucosan inclusion bodies called Lafora bodies (LB) is a hallmark...
of the disease. Besides axons and dendrites of the central nervous system, LBs are also found in the retina, heart, liver, muscle, and skin.[6] LD occurs worldwide, but is relatively common in the Mediterranean Basin and in other parts of the world with a high rate of consanguinity.[6] At present, no preventive or curative treatment for LD is available.

**Molecular Genetics**

At least three loci are implicated in LD. The first one, EPM2A at 6q24.3 was discovered in 1998/1999,[9,10] and the second gene, NHLRC1 (EPM2B) at 6p22.3 was discovered in 2003.[11,12] In about 80% of cases defects have been found in EPM2A and NHLRC. In one family with three biopsy-confirmed LD patients and no identifiable pathogenic variants in both genes, linkage and haplotype analysis excluded both loci from disease association and provided indirect evidence for a third locus for LD.[13] The findings were supported by the lack of EPM2A mutations in the Japanese population.[14] However, no studies have been conducted to identify the gene of interest.

Early-onset LD typically presenting at around 5 years of age caused by pathogenic variants in PRDM8 is a newly recognized condition characterized by progressive myclonus epilepsy and LBs.[15]

### EPM2A

EPM2A encodes a dual-specificity protein phosphatase, called laforin that localizes in the rough endoplasmic reticulum. EPM2A has four exons spanning 130kb and is known to encode two distinct proteins by differential splicing: phosphatase active cytoplasmic isoform-a, and phosphatase inactive nuclear isoform-b.[9,10] Both isoforms of the laforin protein have the unique C termini. The carboxy terminal of isoform-b targets laforin to the nucleus. The pathogenesis of LD is caused by the physiologic functions of laforin isoform-a.[16,17] Laforin contains an N-terminal carbohydrate-binding domain (CBD), encoded mainly by exon 1, and a dual-specificity phosphatase domain (DSPD) spanning exons 3 and 4, common to both isoforms (Fig. 1). Three additional splice variants with potential to code for five distinct proteins in alternate reading frames exist. Two novel isoforms, when ectopically expressed in cell lines, show distinct subcellular localization, and interact with and serve as substrates of malin. Alternative splicing could possibly be one of the mechanisms through which EPM2A regulates the cellular functions of the protein it codes for.[17]

To date, 102 different pathogenic variants in EPM2A have been reported in more than 100 families. Nucleotide substitutions resulting in missense, nonsense, and frameshift variants and indels form 74% and 26% of the total, respectively. One splice site variant has been reported for EPM2A. An overview of the different pathogenic variants can be found in the Lafora PME mutation and polymorphism database (http://projects.tcag.ca/lafora; updated Sep 5, 2017). Of all the types of pathogenic variants in EPM2A described to date, 45% represent missense variants. All the known missense variants target either the CBD or DSPD of laforin.[18] Except for the larger deletions, all the pathogenic variants are evenly distributed across EPM2A. The only exception is the high prevalence of the nonsense c.721C<T variant (p.Arg241X), the so called Spanish pathogenic variant, in over 23 families which is the result of both a founder effect and recurrent events.[9,10] The variant is also found in two Turkish patients; one as an unpublished observation and the other as reported in Salar et al. 2012.[19] Exon 2 deletion is the second frequently observed mutation occurring in eight families. Other variants recurring 5-6 times are exon 1-2 deletion; c.258C<G (p.Tyr86X); c.322C>T (p.Arg108Cys); and c.512G>A (p.Arg171His).

Several polymorphisms in EPM2A have also been described. Among these, c.136G>C (p.Ala46Pro) is specific to the Japanese and Chinese populations.[20] The c.163C>A (p.Gln55Lys) polymorphism in EPM2A was found in two affected persons who were also heterozygous for a deletion in NHLRC1, in seven of the 500 individuals without LD, and in a person with adult onset disease, also in the heterozygous state. It is unclear whether this change constitutes a rare benign SNP of no consequence, whether it may cause LD when homozygous, or whether it could predispose to NHLRC1 deletion in certain situations.[4]

The pathogenic variants (nonsense, missense, and insertions and deletions) located in the DSPD and CBD of EPM2A have the functional consequence of a “null effect” with the loss of phosphatase activity.[21] Although, all aspects of the
protein function have not been tested for each missense variant, transfection experiments overexpressing missense mutants that resulted in ubiquitin-positive cytoplasmic aggregates, suggest that they were folding mutants destined for degradation. It is also observed that missense mutants affect the subcellular localization of laforin and disrupt the interaction of laforin with protein targeting glycogen (PTG) and malin proteins that interact with laforin in vivo.

**NHLRC1 (EPM2B)**

NHL-repeat-containing 1 gene (NHLRC1, EPM2B) encodes E3 ubiquitin-ligase malin, a 395-amino acid protein. Malin contains a zinc finger of the RING type and six NHL-repeat protein-protein interaction domains (Fig. 2) and is involved in the ubiquitin-mediated proteolysis cascade. Malin colocalizes with laforin in the endoplasmic reticulum. Laforin and malin interact with misfolded proteins and promote their degradation in the ubiquitin-proteasome system. Malin also ubiquitinates laforin, leading to its degradation. Thus, one of the critical functions of malin is to regulate the cellular concentration of laforin by ubiquitin-mediated degradation.

**Fig. 2.** A schematic representation of NHLRC1 gene and its product, malin.

Nearly all pathogenic variants in NHLRC1 are predicted to result in the loss of function of malin. Six benign variants have also been reported in the database.

**PRDM8**

A new PME associated with LB, early-onset Lafora body disease, was reported to map to locus 4q21.21 by linkage analysis. The candidate PRDM8 gene sequencing identified a c.718T>C, p.Phe262Leu mutation that co-segregated with the disorder in the family. PRDM8 belongs to a family of PR domain-containing histone methyltransferases with strong expressions in brain and testis. It was shown that the PRDM8 protein interacted with laforin and malin and relocalized them to the nucleus. A gain-of-function effect in the mutant protein resulted in over-sequestration of laforin and malin in the nucleus leading to an effective deficiency of cytoplasmic laforin and malin.

**Genotype-Phenotype Correlations**

Establishing genotype-phenotype correlations is an ongoing effort that requires comprehensive description of the clinical phenotype and the prediction of the functional effect of the genetic variant in patients. There are difficulties as the presence of compound heterozygotes in different combinations is common, and within an ethnic group of individuals sharing the same pathogenic variant, the phenotype can be highly variable or similar. Additionally, several modifiers of the phenotype possibly exist since laforin and malin interact with a diverse set of other proteins, and one Turkish family. The second most common pathogenic variant in NHLRC1 is c.468-469delAG. It has been identified in 14 individuals belonging to the same genetic isolate of tribal Oman. All shared a common haplotype, suggesting a founder effect. Both c.205C>G and c.468-469delAG pathogenic variants have been identified in different ethnic groups, suggesting a recurrent mutation event. Thus, these two sites represent hot spots for NHLRC1 pathogenic variants. Another missense pathogenic variant c.76T>A is prevalent in French Canadian ethnic isolates and there are 10 entries in the database.

To date, 84 pathogenic variants have been reported in more than 125 families (http://projects.tcag.ca/lafora). The majority of variants are missense; however, insertions, deletions, and nonsense variants, making up approximately 14% of the total mutations, have also been found. A heterozygous deletion of the entire NHLRC1 gene has been reported in an Italian and a Serbian family.

The missense variant p.Pro69Ala (c.205C>G) in the RING finger domain is the most common missense variant observed in 32 families (http://projects.tcag.ca/lafora). It is repeatedly reported in affected persons of Italian, French, and Spanish descent. The variant is also observed in one Turkish family. The second most common pathogenic variant in NHLRC1 is c.468-469delAG. It has been identified in 14 individuals belonging to the same genetic isolate of tribal Oman. All shared a common haplotype, suggesting a founder effect. Both c.205C>G and c.468-469delAG pathogenic variants have been identified in different ethnic groups, suggesting a recurrent mutation event. Thus, these two sites represent hot spots for NHLRC1 pathogenic variants. Another missense pathogenic variant c.76T>A is prevalent in French Canadian ethnic isolates and there are 10 entries in the database.

Nearby, 84 pathogenic variants have been reported in more than 125 families (http://projects.tcag.ca/lafora). The majority of variants are missense; however, insertions, deletions, and nonsense variants, making up approximately 14% of the total mutations, have also been found. A heterozygous deletion of the entire NHLRC1 gene has been reported in an Italian and a Serbian family.

To date, 84 pathogenic variants have been reported in more than 125 families (http://projects.tcag.ca/lafora). The majority of variants are missense; however, insertions, deletions, and nonsense variants, making up approximately 14% of the total mutations, have also been found. A heterozygous deletion of the entire NHLRC1 gene has been reported in an Italian and a Serbian family.

The missense variant p.Pro69Ala (c.205C>G) in the RING finger domain is the most common missense variant observed in 32 families (http://projects.tcag.ca/lafora). It is repeatedly reported in affected persons of Italian, French, and Spanish descent. The variant is also observed in one Turkish family. The second most common pathogenic variant in NHLRC1 is c.468-469delAG. It has been identified in 14 individuals belonging to the same genetic isolate of tribal Oman. All shared a common haplotype, suggesting a founder effect. Both c.205C>G and c.468-469delAG pathogenic variants have been identified in different ethnic groups, suggesting a recurrent mutation event. Thus, these two sites represent hot spots for NHLRC1 pathogenic variants. Another missense pathogenic variant c.76T>A is prevalent in French Canadian ethnic isolates and there are 10 entries in the database.

To date, 84 pathogenic variants have been reported in more than 125 families (http://projects.tcag.ca/lafora). The majority of variants are missense; however, insertions, deletions, and nonsense variants, making up approximately 14% of the total mutations, have also been found. A heterozygous deletion of the entire NHLRC1 gene has been reported in an Italian and a Serbian family.

The missense variant p.Pro69Ala (c.205C>G) in the RING finger domain is the most common missense variant observed in 32 families (http://projects.tcag.ca/lafora). It is repeatedly reported in affected persons of Italian, French, and Spanish descent. The variant is also observed in one Turkish family. The second most common pathogenic variant in NHLRC1 is c.468-469delAG. It has been identified in 14 individuals belonging to the same genetic isolate of tribal Oman. All shared a common haplotype, suggesting a founder effect. Both c.205C>G and c.468-469delAG pathogenic variants have been identified in different ethnic groups, suggesting a recurrent mutation event. Thus, these two sites represent hot spots for NHLRC1 pathogenic variants. Another missense pathogenic variant c.76T>A is prevalent in French Canadian ethnic isolates and there are 10 entries in the database.

To date, 84 pathogenic variants have been reported in more than 125 families (http://projects.tcag.ca/lafora). The majority of variants are missense; however, insertions, deletions, and nonsense variants, making up approximately 14% of the total mutations, have also been found. A heterozygous deletion of the entire NHLRC1 gene has been reported in an Italian and a Serbian family.
variations in genes that code for these interacting proteins could contribute to the phenotype. For instance, a sequence variant in PPP1R3C, which codes for the protein PTG contributes to a milder course of LD. To date, no correlations between the phenotype and variant type (missense or truncating) or location of the pathogenic variant in the gene have been established.

A genotype-phenotype correlation may exist between NHLRC1 mutations and adult-onset LD. Individuals with pathogenic variants, particularly p.Asp146Asn, in NHLRC1 appear to live longer than those with pathogenic variants in EPM2A. However, some patients with pathogenic variants in NHLRC1 have extremely severe phenotypes.

Within the Italian and Japanese populations, pathogenic variants in NHLRC1 are more common than pathogenic variants in EPM2A. Conversely, EPM2A pathogenic variants are more common in the Spanish and French populations. Within the Indian and Arab populations, the distribution of pathogenic variants in the two genes is more or less even.

**Molecular Pathogenesis**

The mechanism through which pathogenic variants in either EPM2A or NHLRC1 result in LD and the exact role of LBs in the pathogenesis of LD have been the subject of intensive research efforts over the past years.

Glycogen is formed through coordinated actions of glycogen synthase (GS) and glycogen branching enzyme as a tightly packed glycogen sphere, containing up to 55,000 glucose units. The spherical structure provides glycogen solubility, preventing it from precipitating. Glycogen is digested by glycogen phosphorylase and glycogen debranching enzyme. Polyglucosans are malformed glycogen molecules with excessively long strands, and inadequate branching and lack spherical form. Mainly in dendrites, they precipitate and aggregate into concretized masses called LBs. The extent of LB deposition correlates with neuronal cell death and seizure frequency, suggesting that LD pathogenesis may be due to a defect or multiple defects in glycogen metabolism, e.g., misregulated glycogen synthesis or degradation.

Both laforin and malin may be involved in defective glycogen metabolism. Laforin and malin form a complex that interacts with PTG (PPP1R3C) and targets the pleiotropic phosphatase protein phosphatase 1 (PP1) to dephosphorylate and activate GS. PTG is a molecular scaffold that directly binds laforin, GS, phosphorylase, and phosphorylase kinase and assembles the glycogen machinery. Laforin is ubiquitinated in a malin-dependent manner, leading to its degradation. Thus, malin regulates the protein concentration of laforin through polyubiquitin-dependent degradation. The absence of laforin or malin is associated with increased PTG, most likely through the loss of malin-mediated ubiquitination and the degradation of GS and PTG, and resultant increased GS activity (Fig. 3).

Polyglucosans are then formed due to an imbalance in glycogen elongation and branching, since branching enzyme activity does not increase with increased GS activity. Conversely, GS adds phosphogluco to glycogen instead of glucose, and laforin corrects this by removing the phosphates. The loss of laforin results in glycogen hyperphosphorylation and prevents the formation of its spherical structure which is essential for solubility. The malformed glycogen accumulates into LB.

The malin-laforin complex have additional functions to that of the regulation of glycogen synthesis, such as the participation in the control of autophagy, a process by which damaged components of the cell are sequestered and degraded in the lysosome. The loss of either laforin or malin presents autophagy impairment. Thus, this defect in autophagy resulting in the accumulation of glycogen could be due to the neurodegeneration seen in LD. Laforin has been shown to be a positive regulator of autophagy, via the mammalian target of rapamycin kinase, and autophagy is impaired in laforin knock-out mice. However, the exact mechanism for the laforin effect on autophagy is still elusive. Malin has also been suggested to regulate autophagy and the misfolded protein response.
In malin knockout mice with impaired glycogen synthesis, autophagy impairment and neurodegeneration are rescued.[44] Treating LD through the downregulation of GS has been shown effective in different mouse models. Genetically removing brain GS from laforin-lacking LD mice resulted in correction of the LD phenotype, including elimination of LB, neurodegeneration, and seizure predisposition. The same result was obtained through partial reduction of glycogen synthesis by genetically removing PTG, a protein that activates GS. This was shown effective in laforin-deficient LD mice as well as malin-deficient LD mice.[45]

All of the above experimental evidence point out that the regulation of glycogen synthesis could be a key target for the treatment of LD.

**Molecular Genetic Testing**

Pathogenic variants can be detected by sequencing of coding and flanking intronic regions of genomic DNA and may include small intragenic deletions/insertions and missense, nonsense, and splice site variants in *EPM2A* and *NHLRC1*. Sequence analysis alone detects 88%–97% of pathogenic variants in these two genes. Exon or whole-gene deletion/duplications not detectable by sequence analysis can be detected by quantitative PCR, long-range PCR, multiplex ligation-dependent probe amplification, and array CGH.

The parents of an affected child are obligate heterozygotes, i.e., carriers of one *EPM2A* or *NHLRC1* pathogenic variant. Heterozygotes are asymptomatic and are not at risk of developing the disorder. Compound heterozygosity arises when parents carry two different mutations in a heterozygous condition. If a pathogenic variant is identified in homozygous condition in either NPLRC1 or *EPM2A*, and if one parent lacks the mutation, deletion/duplication analysis of that gene should be considered.

A multi-gene panel that includes *EPM2A*, *NHLRC1*, and other genes of interest or whole-exome sequencing (WES) can also be conducted for molecular diagnosis. For WES, some drawbacks exist because it may not be as accurate as targeted single-gene or multigene panel testing, and positive results need to be confirmed by Sanger sequencing method. Nucleotide repeat expansions, epigenetic alterations, and deletions/duplications larger than 8–10 nucleotides cannot be detected by WES.

Genetic counseling in LD is possible once the *EPM2A* and *NHLRC1* pathogenic variants in an affected family member are identified. Afterwards, carrier testing for at-risk relatives, prenatal testing, and preimplantation genetic diagnosis also become available. Each sibling of the proband’s parents is at a 50% risk of being a carrier of an *EPM2A* or *NHLRC1* pathogenic variant, and first cousins of the proband are at a 25% risk of being carriers. Because of early onset and rapid deterioration, patients with LD typically do not reproduce.

**Conclusion**

LD is an autosomal recessive PME caused by mutations in the *EPM2A* and *NHLRC1* genes that code for laforin and malin, respectively. Laforin and malin interacts with PTG that regulates glycogen synthesis. Defects in either laforin or malin causes degradation of PTG and downregulates glycogen synthesis resulting in the accumulation of LBs. In the majority of cases molecular genetic diagnosis is possible by detecting pathogenic variants through sequencing of the coding and flanking intronic regions of both genes and consequently, enables carrier testing and prenatal diagnosis in affected families. Although, 102 and 84 pathogenic variants in EPM2A and NHLRC1 genes, respectively, are listed in the recent update of Lafora Mutations database (http://projects.tcag.ca/lafora), a clear association between the phenotype and the variant type have not been observed. However, a variable frequency of pathogenic variants for the two genes exist among different populations and a possible association between *NHLRC1* mutations and milder forms and adult onset LD have been recognized. Gain-of-function mutations in a recently identified PRDM8 gene causes early-onset LD because the mutant product interacts and over-sequesters laforin and malin in the nucleus; thereby, regulates the cytoplasmic quantities of LD enzymes. Although autophagy impairment has been suggested as a possible pathomechanism for the neurodegeneration in LD, stronger experimental evidence exist that LB accumulation results from the downregulation of glycogen metabolism. Consequently, regulation of glycogen synthesis is regarded as the key target for the treatment of LD.

**Conflict of interest**

The author declares that there is no conflict of interest.

**References**

1. Jansen AC, Andermann E. Progressive Myoclonus Epilepsy,


23. Gentry MS, Worby CA, Dixon JE. Insights into Lafora disease: malin is an E3 ubiquitin ligase that ubiquitinates and promotes the degradation of laforin. Proc Natl Acad Sci U S A 2005;102(24):8501–6. [CrossRef]


43. Mittal S, Dubey D, Yamakawa K, Ganesh S. Lafora disease proteins malin and laforin are recruited to aggresomes in response to proteasomal impairment. Hum Mol Genet 2007;16(7):753–62. [CrossRef]