Candidate Variants in **MLC1** Gene Causing Megalencephalic Leukodystrophy Using *In silico* Prediction Methods

Mutaz Amin*

Department of Biochemistry, University of Khartoum, Sudan

**Abstract**

Megalencephalic leukodystrophy with subcortical cysts is a type of demyelinating leukodystrophy caused by mutations in **MLC1** gene. Studies in **MLC1** gene are scarce and the few published are limited in either sample size or methods of molecular diagnosis. In *silico* analysis of SNPs in a well-known disease causing gene is a valuable method of analyzing known variants deposited in public databases. This article aimed to analyze all SNPs in **MLC1** gene in order to determine probable disease causing variants that are not yet been evaluated. SNPs in **MLC1** gene were retrieved from NCBI dbSNP. Variants in VCF format were analyzed using Variant Effect Predictor (VEP) of the Ensembl database. The deleterious coding ns SNPs were evaluated by the web programs SIFT, PolyPhen and Mutation Taster in addition to allele frequency and conservation score. Structural analysis was performed in order to evaluate and compare the stability of native and mutant structures. The effect of amino acid changes on the stability of MLC protein was evaluated using I-Mutant 2.0 web server.

**RESULTS**

The **MLC1** gene contains 1471 SNPs, the majority of which were single nucleotide polymorphisms (SNPs) (62%). Database analysis revealed 4 analyzed SNPs resulted nonsense, 18 are indels (frame shift mutations) and 10 potentially disrupt splicing. Six variants in **MLC1** gene were predicted to be pathogenic. The results of our study will facilitate future studies aimed to analyze the genetics of MLC in different families from different populations.

**DISCUSSION**

In *silico* analysis of SNPs in disease causing genes with unknown clinical significance and deposited in public databases is a valuable method screen patients with the respective disease.
using for example SNP microarray. This study aimed to determine potential disease causing variants in \textit{MLC1} gene using in silico prediction methods. The result showed that potential disease causing variants in \textit{MLC1} gene are most commonly indels causing frame shift mutations. This is in accordance with the molecular pathogenesis of the disease which is consequent to a loss of function of MLC protein. Frame shifting mutations are known to disrupt protein synthesis and impair protein’s function [11]. Frame shifting mutations were indeed found in many studies underlying this form of leukodystrophy [12-15]. Variants causing disruption of splicing lead to either retention of an intron or loss of an exon and nonsense variants cause premature termination of protein synthesis. Both nonsense and splicing variants are loss of function mutations [16]. Nonsense variants causing MLC were previously reported [17-19] and so were splicing defects [15,20]. Missense variants causing \textit{MLC1} have been reported before [21]. Judging from the in silico prediction methods which include conservation, allele frequency and tolerance of amino acid changes, the variants reported in our study are very likely to be pathogenic, but they are all novel and their association with MLC is yet to be confirmed. A limit of our study is that neither biological modeling nor segregation in family studies, the only accepted way to evaluate a variant as pathogenic, have been performed. MLC is an autosomal recessive disease and this explains why these variants were not found in a homozygous state in any of the studied populations in the exact database.

Our results will facilitate future epidemiological studies aimed to analyze the genetics of MLC in different families from different populations. Moreover they can have a role in genetic counseling, in particular for an alleged determination of frequency of heterozygote individuals in specific populations and, therefore, of the risk of a known mutation carrier to find a MLC mutation carrying partner. The results of in silico predictive studies, nonetheless the limitation of determine only potential and not validate predictions, will also accelerate variants annotation keeping in advancing genome and exome studies.

**REFERENCES**

Table 2: Missense variants in MLC1 gene predicted to be pathogenic with SIFT (and its score) PolyPhen (and its score), Mutation Taster, allele frequency and conversation.

<table>
<thead>
<tr>
<th>SNP</th>
<th>AA*</th>
<th>SIFT (score)</th>
<th>PolyPhen (score)</th>
<th>Mutation Taster</th>
<th>AF**</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs143061714</td>
<td>V260L</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.987)</td>
<td>Disease causing 0.0006</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>rs568289086</td>
<td>A208V</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.994)</td>
<td>Disease causing 0.0002</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>rs41302601</td>
<td>N218K</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.997)</td>
<td>Disease causing 0.0022</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>rs78644350</td>
<td>V200F</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.997)</td>
<td>Disease causing 0.0002</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>rs555304253</td>
<td>R193W</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.997)</td>
<td>Disease causing 0.0002</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>rs533294413</td>
<td>R20G</td>
<td>Deleterious (0)</td>
<td>Probably damaging (0.994)</td>
<td>Disease causing 0.0004</td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>

*AA = Amino acid change
**AF = Allele frequency

Three of these mutations (A208V, N218K and V260L) lie in the trans-membrane domain of MLC1 protein (not shown).

All potentially pathogenic amino acid changes in our studies (in Table 2) were predicted to significantly decrease the stability of MLC protein using I-Mutant 2.0 web server.