**ABSTRACT**

Mutation in the three genes, namely CACNA1A, ATP1A2, and SCN1A causes FHM1, FHM2, and FHM3 respectively. It is an inherited autosomal heterogeneous disease with an aura that causes temporary weakness or hemiplegia. There is no known cure for this disease, but the chances are reduced by maintaining a healthy lifestyle, and balanced diet, by reducing stress or changes in sleep patterns.

We used eight different bioinformatics methods SIFT, SNAP2, GVGD, Mutation Assessor, Panther, PolyPhen-2, Pon-P2, and Snp&Go to estimate the functional impact of missense variants sourced from the NCBI database. As per the study's findings. The SNPs that were estimated to be deleterious by all eight in silico approaches were declared high-risk variants and underwent additional research.

This report contains highlights of the review of the literature to be carried out during the study. Various computational tools have also been listed that can be helpful in the study.

**Chapter1:Introduction**

**1.1 Research Background**

More than 10% of the general population is affected by migraine, which is a common type of episodic (headache that occurs in discrete attacks rather than continuous ) headache disorder.

Migraine is a type of disorder that can occur with or without aura. Migraine with aura (MA) is a type of disorder that mainly involves the symptoms of visual disturbances, and sensory changes before the headache phase of the migraine.MA symptoms last between 5 to 60 minutes and are also reversible.  Migraine without aura (MO) is also a common subtype of migraine that doesn't involve the symptoms before the headache. People having MO experience severe headaches followed by other symptoms such as nausea, and vomiting. The duration of this condition varies from a few hours to several days.

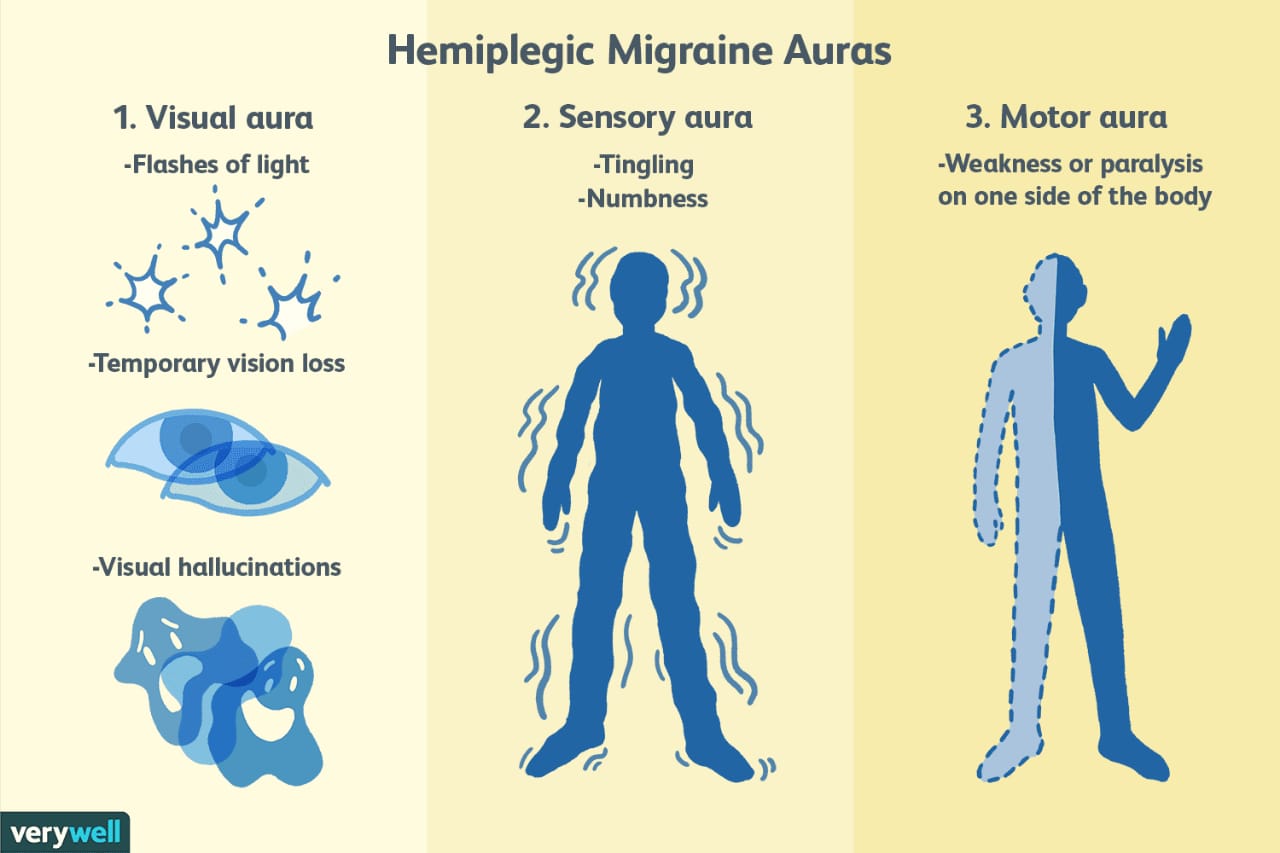


Fig1: Symptoms of aura in Familial Hemiplegic migraine

[https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.verywellhealth.com%2Fhemiplegic-migraines]

Familial Hemiplegic Migraine (FHM) is a very uncommon type of migraine that is distinguished by the presence of aura, temporary weakness, or paralysis on one of the sides of the body. Mainly four aura symptoms are found in FHM - motor(contraction of muscle), aphasic(difficulties with language), visual(vision changes), and sensory(numbness). It is a genetically heterogeneous diverse kind of migraine, meaning that multiple genetic variants can cause the condition, and it is also inherited autosomally dominantly, which means that a single copy of the mutant gene from one parent is enough to induce the illness. In FHM, hemiplegic refers that this is usually temporary and reversible, mainly for an hour to a few days.

FHM, a genetically heterogeneous disorder, is caused mainly by mutations in 3 genes, namely CACNA1A, ATP1A2, and SCN1A.

**1.2 Objective**

**Types of disorder-**

* Familial Hemiplegic Migraine (FHM1) is a genetic condition that occurs infrequently and is linked to mutations in the CACNA1A gene present at chromosome 19p13. The CACNA1A gene provides instruction for creating a protein known as CaV2.1, which is a type of voltage-gated calcium channel located on the presynaptic nerve terminals.
* Familial hemiplegic migraine type 2 (FHM2) is an uncommon genetic condition that arises from mutations in the ATP1A2 gene present at 1q23. This gene provides instructions for creating the alpha-2 subunit of the sodium-potassium ATPase pump (known as Na+/K+-ATPase), which is located in the cell membranes of astrocytes, a type of glial cell in the brain.
* Familial hemiplegic migraine type 3 (FHM3) is a hereditary disorder that arises from mutations in the SCN1A gene present at 2q24, which codes for the voltage-gated sodium channel Nav1.1of alpha-1subunit. The Nav1.1 channel is primarily located in the axon terminal of the presynaptic cell at the surface of the neurons.

This study attempts to analyze nonsynonymous SNPs (Single nucleotide polymorphisms) of three genes and their structural and functional impacts of it on the channel protein via docking analysis with various ligand molecules. We used multiple bioinformatics techniques for this goal to identify the most deleterious variants of the Ca ion channel protein by using the three-dimensional structure of CaV2.1 and its mutated variations.

These genes are involved in many different diseases, and they are also involved in additional migraine-related neurological ailments. The table below indicates the disorders these genes are involved with, as well as the mechanism they use.

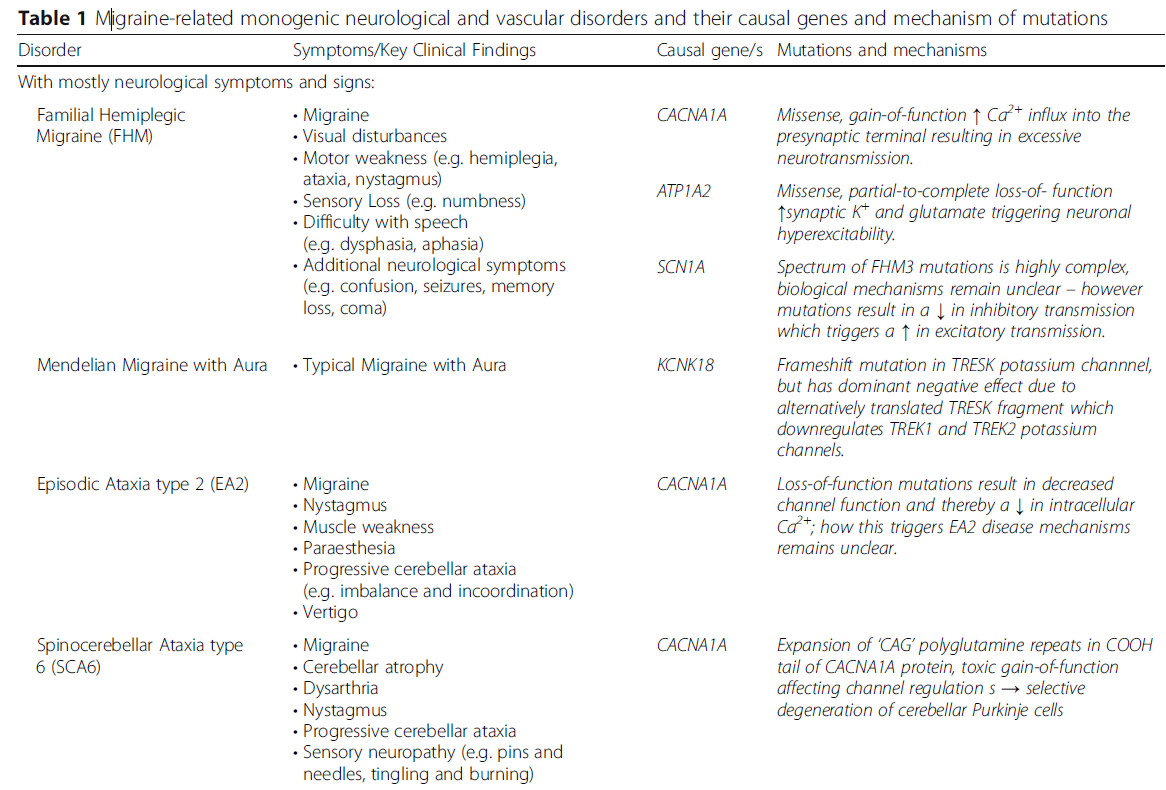


Table1: Migraine-related monogenic neurological disorder causes genes and mechanism of mutation

[Sutherland, H. G., Albury, C. L., & Griffiths, L. R. (2019). Advances in the genetics of migraine. *The Journal of headache and Pain*, *20*(1), 1-20.]

**CHAPTER 2: LITERATURE REVIEW**

At rest, the inside of the neuron cells is negatively charged relative to the outside, as the K+ is present in low concentration and Na+ in high concentration in the extracellular membrane, and Na+ and K+ in the low and high concentration respectively in the axon. When a sensory signal is received by the cell it triggers the opening of voltage-gated ion channels on the cell membrane, leading to a flow of ions into or out of the cell. When there is an inflow of sodium ions (positively charged) into the cell, it depolarizes the membrane potential and triggers the opening of more sodium channels, leading to a rapid depolarization of the membrane potential. If the depolarization exceeds a particular threshold, it opens the voltage-gated potassium channels, which allows the positively charged potassium ions to exit the cell leading to the repolarization of the membrane potential. The rapid influx and efflux of ions result in a transient electrical signal, and this phenomenon is known as ‘Action Potential’.

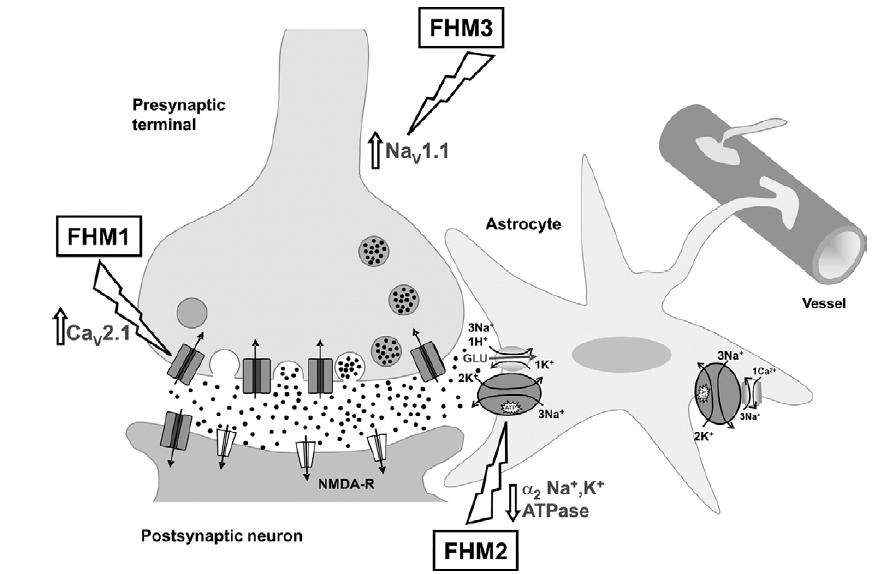


Fig2 . Gain-of-function mutations in Cav2.1 channels(FHM1) or Nav1.1 channels(FHM3) and loss-of-function mutations in Na+, K+ -ATPases(FHM2) may all render the brain by causing either excessive synaptic glutamate release(FHM1) or excessive extracellular K+ (FHM3) or decreased removal of K+ and glutamate from the synaptic cleft (FHM2) of released glutamate and in the clearance of K+ from the extracellular space during neuronal activity. Nav1.1 channels are located mainly at the soma and are critical for action potential firing.

[Pietrobon, D. (2007). Familial hemiplegic migraine. *Neurotherapeutics*, *4*, 274-284]

**2.1 Mechanism of all three channels-**

Presynaptic nerve terminals are specialized structures that release neurotransmitters, such as glutamate, into the synaptic cleft. The release of glutamate by presynaptic neurons is a critical step in the transmission of nerve impulses in the brain. Mutations in the CACNA1A gene can result in the gain of function and production of a faulty CaV2.1 protein, which can alter the function of presynaptic nerve terminals. The mutation can cause increased calcium influx into the presynaptic neuron, leading to an increased release of glutamate into the synaptic cleft. Increased release of glutamate activates the postsynaptic neurons leading to FHM1 symptoms. It is believed that glutamate release may cause changes in blood flow, inflammation, and neuronal excitability in the brain, ultimately leading to migraine symptoms.

The release of glutamate at the synapse can depolarize the postsynaptic neuron and trigger the release of potassium ions. Glutamate and potassium ions are transported into nearby astrocytes through specific channels, which are regulated by the alpha-2 subunit of the Na+/K+-ATPase pump encoded by the ATP1A2 gene. Mutations in this gene can lead to a loss of function in the pump, resulting in reduced removal of excess glutamate and potassium ions from the synapse. However, excessive glutamate release can be toxic to neurons and can lead to a condition known as excitotoxicity. When the excess potassium is not removed by astrocytes this lead to an increase in extracellular potassium concentration, leading to migraine attack.

Astrocytes are connected to blood vessels and are involved in regulating cerebral blood flow, so dysfunction in astrocytes due to mutations in the ATP1A2 gene may contribute to the development of migraine pain.

Mutations in the SCN1A gene can result in Nav1.1 channel function gain, which is primarily located at the soma of neurons and plays a critical role in the generation of action potentials that leads to an increased influx of sodium ions into the cell. This gain of function leads to excessive activity of the Nav1.1 channel, resulting in an increase in extracellular sodium ions, which in turn decreases the concentration of sodium outside the axon. In addition, mutations in the ATP1A2 gene (FHM2) can result in high levels of extracellular potassium ions. Together, these changes can reverse the membrane potential, leading to a state of hyperexcitability and increased risk for the development of Familial Hemiplegic Migraine type 3 (FHM3).

**Chapter 3: Materials and Methods**

**3.1** **Data Retrieval**

Information regarding the human CACNA1A, ATP1A2, SCN1A gene and its protein sequence in FASTA format was gathered from NCBI and UniProtKB, respectively. Additionally, Single Nucleotide Polymorphism(SNP) present in the three genes is obtained from the dbSNP database. The selection criteria for the SNPs from the NCBI database were limited to missense nsSNPs because these variants can alter the amino acid sequence of the encoded protein, potentially disrupting its structural conformation and functionality.

**3.2** **Prediction of deleterious nsSNP by using various tools**

Several online-based tools were used to find out the deleterious missense nsSNPs and also analyze the functional and structural effect of nsSNPs on CACNA1A, ATP1A2, and SCN1A genes.

**SIFT**

Sorting Intolerant From Tolerant(SNP) determines the score based on homologous sequence, sequence alignment, and physical characteristics of amino acids and works on Sorting Intolerant from Tolerant algorithm, which can determine whether a substitution will impact how a protein functions. Deleterious have a probability score less than 0.05 and tolerated SNPs have a probability score greater than 0.05. It accepts the rsID as an input.

**PolyPhen-2**

Polymorphism Phenotyping version 2(Polyphen-2) takes the input in the form of rsID/accession number and also takes the details of the substitution as input. Its score lies between 0 to 1 where 0 represents tolerated and 1 shows deleterious effect. It generates PSSM (Position specific substitution matrix) based on blast query and provides a score by estimating the potential effects using a comparison of phylogenetics.

**PANTHER**

Evolutionary conservation is the principle of the PANTHER (Position specific evolutionary

preservation) method. Sequences Alignment and phylogenetic trees can be found in the

PANTHER database. To predict the impact of variants, BLASTP is used to compare the query

protein sequence to the known homologous protein sequences. Ancestor sequences are used to calculate conservation scores.The impact of non-synonymous SNP on a protein's function is estimated using this tool based on information about the protein's evolutionary details as it is possible to take evolutionary aspects into account while deciding whether or from the neutral variants of the SNP will be damaging or benign.

**Pon-P2**

PON-P2 is exceptionally rapid and accurate at identifying potentially harmful mutations in

proteins. This tool is supported by this machine learning classifier which is trained by various

information like evolution conservation, gene ontology annotations, and characteristics of

amino acids. Pathogenic, Neutral, and undefined are the three results given by this tool. This

tool determines how a specific SNP is deleterious or if its exact function is unknown. The

random forest probability score serves as the foundation for this classification.

**Align GVGD**

Align GVGD is used to predict the missense substitution. It forms the range as a scoring value from C0, C15 to C65 where C65 is considered deleterious and C15 is considered less likely to affect. It takes the input in the form of the FASTA sequence of the protein and also takes the amino acid substitution. It considers the Multiple sequence alignment for the prediction.

**SNAP2**

Screening for Not Acceptable Polymorphisms takes the input file in the form of FASTA format. It classifies all the nsSNPs based on their sequence-based information as either damaged or neutral. It gives the score ranges from, -100 to +100 where -100 shows the neutral effect and +100 shows the strong effect of this substitution on protein function and its structure.

**SNP&Go**

This tool predicts whether or not the substitution is likely to be associated with a disease. For this, it analyzes information such as protein sequence information, phylogenetic relationships, and functional annotations encoded by the protein. A higher score ensures that the variant is disease-associated whereas a lower score indicates less chance of disease.

**3.3 Prediction of the effect of Missense mutants on protein stability by I-Mutant 3.0**

It is important to see the substitution impact on protein stability. I-Muatnt is used to predict the stability of mutated protein due to a single mutation using the DDG value. DDG value decides whether the protein stability increases or decreases. Delta Delta G value is the change in Gibbs free energy and it is the difference in free energies of the mutant structure and the wild-type. I-Muatant takes the input of deleterious SNPs and predicts the value of DDG which infers the stability of the protein. If DDG>0.5Kcal/mol then protein stability increases, DDG <-0.5Kcal/mol shows protein stability decreases, or neutral if -0.5< DDG<0.5Kcal/mol.

**Chapter 4: Results**

The nsSNPs of three genes were examined in this study and were retrieved from the SNP database of NCBI. 8 different tools SIFT, GVGD, SNAP2, PANTHER, PolyPhen-2, PANTHAR, PON-P2, and SNPs&GO were used for computational analysis of variants.

For CACNA1A – The database consists of a total of 121605 SNPs, out of which 2230 SNPs were missense, the ATP1A2 – Database consists of a total of 12160 SNPs, out of which 678 SNPs were missense and SCN1A– Database consists of a total of 60645 SNPs, out of which 2042 SNPs were missense. For our investigation, only missense SNPs were selected since they have a deleterious impact on protein structure and its function.

**4.1 Screening of deleterious nsSNPs**

**CACNA1A**

Firstly for the prediction of the nsSNP, SIFT was used that screened the 139 nsSNPS out of 2230 SNP as tolerated or deleterious both, the rest are not found in the SIFT. Among them, SHIFT found 50 as deleterious, and 89 were tolerated.SNAP2 found 50 nsSNPS having an effect and remaining as neutral.PANTHER shows 24 nsSNPS as possibly damaging and 115 as probably damaging.SNP&Go found 42 nsSNPs as diseased and 97 nsSNPs as the neutral impact on the structure. Whereas Align GVGD predicted 47 nsSNPs as pathogenic and remaining as neutral or less pathogenic.

By using all 7 prediction tools, 15 nsSNPs were selected as the most deleterious of all the prediction tools.

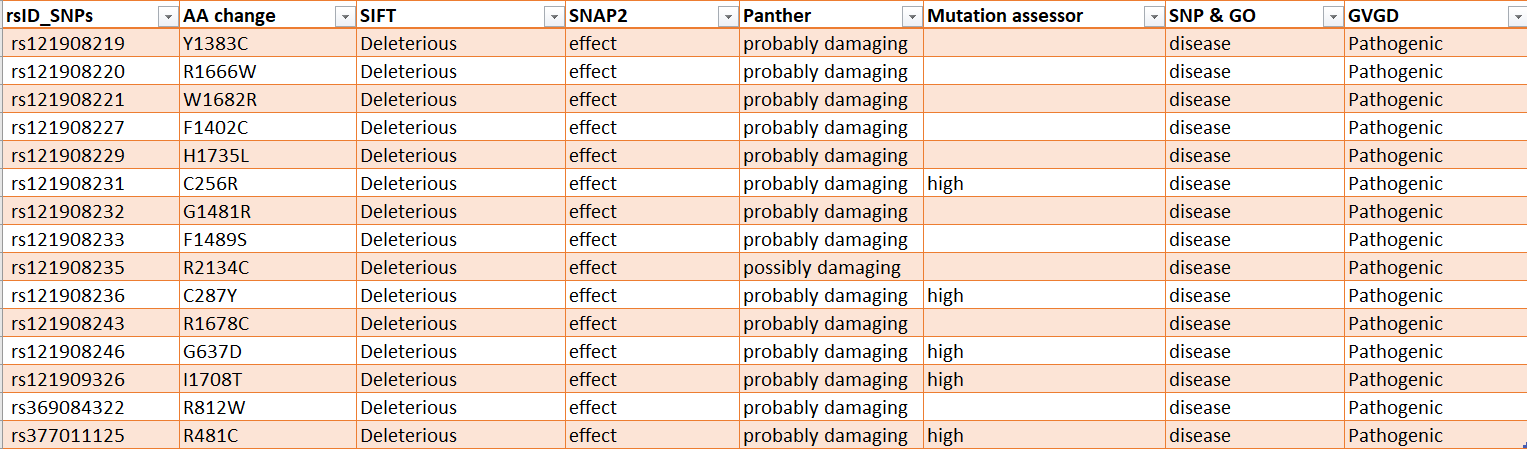


Table1: List of CACNA1A gene SNPs expected to be harmful

**ATP1A2**

SIFT was used that screened the 68 nsSNPS out of 678 SNP as tolerated or deleterious both, the rest are not found in the SIFT. Among them, SHIFT found 25 as deleterious and 43 were tolerated. SNAP2 found 25 nsSNPS having an effect and remaining as neutral.PANTHER shows 10 nsSNPS as possibly damaging and 58 as probably damaging. SNP&Go found 22 nsSNPs as diseased and 46 nsSNPs as the neutral impact on the structure. Whereas Align GVGD predicted 28 nsSNPs as pathogenic and remaining as neutral or less pathogenic. The mutation assessor screened 15 nsSNPs has a high effect and the remaining 53 has medium and less impact.Polyphen2 predicted 28 nsSNPs as probably damaging.Pon-P2 found 35 nsSNPs as pathogenic and the rest were neutral.

By using all 8 prediction tools, 8 nsSNPs were selected as the most deleterious of all the prediction tools.

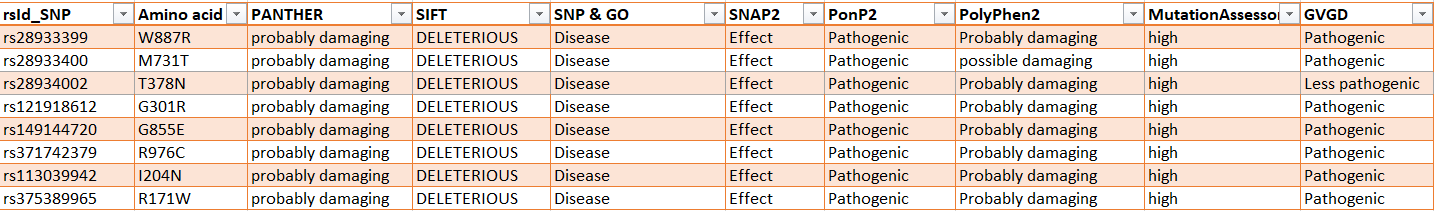
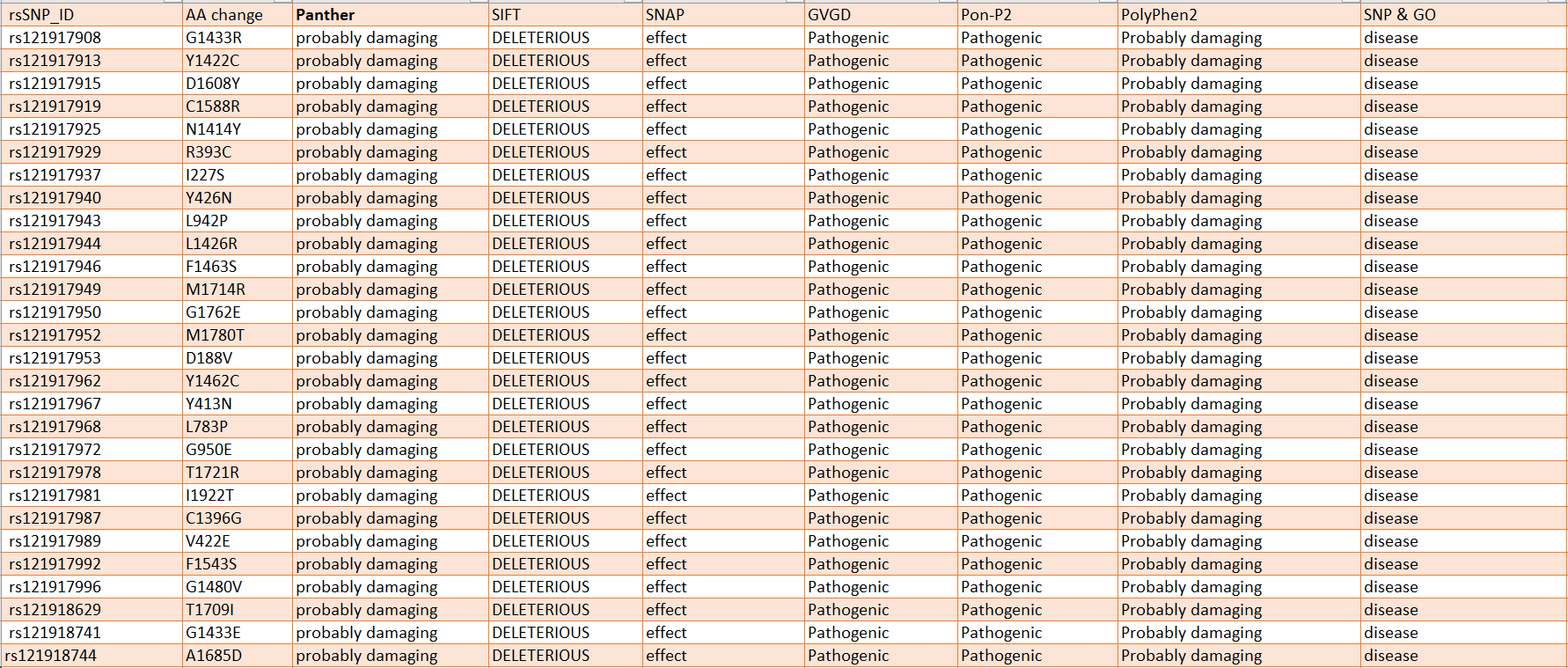


Table2: List of ATP1A2 gene SNPs expected to be harmful

**SCN1A**

SIFT was used that screened the 311 nsSNPS out of 2042 SNP as tolerated or deleterious both, the rest are not found in the SIFT. Among them, SHIFT found that 190 as deleterious, and 121 were tolerated. SNAP2 found 25 nsSNPS having an effect and remaining as neutral.PANTHER shows 35 nsSNPS as possibly damaging and 276 as probably damaging. SNP&Go found 224 nsSNPs as diseased and 87 nsSNPs as the neutral impact on the structure. Whereas Align GVGD predicted 28 nsSNPs as pathogenic and remaining as neutral or less pathogenic. The mutation assessor screened 15 nsSNPs has a high effect and the remaining 53 had a medium and less impact.Polyphen2 predicted 28 nsSNPs as probably damaging.Pon-P2 found 35 nsSNPs as pathogenic and the rest were neutral.

By using all 8 prediction tools, 56 nsSNPs were selected as the most deleterious of all the prediction tools.



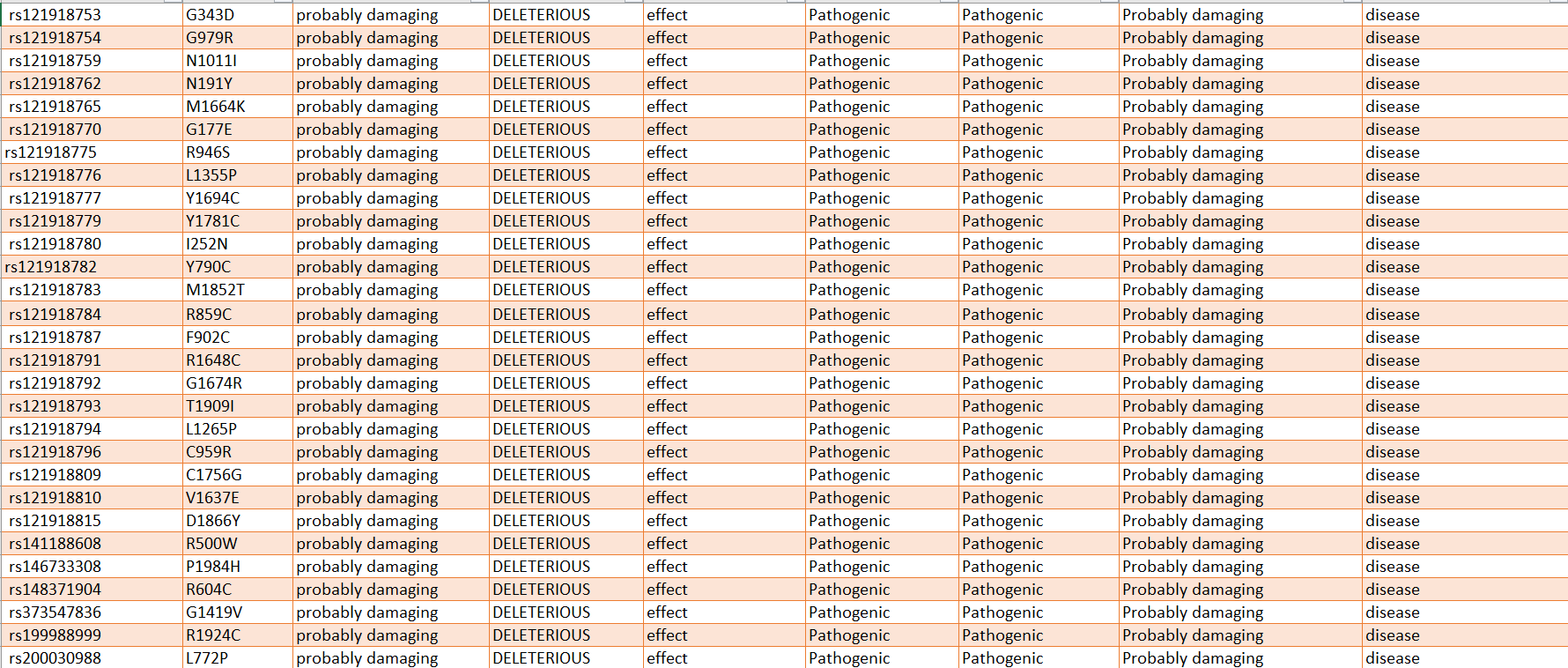


Table3: List of SCN1A gene SNPs expected to be harmful

**4.2 Prediction of effect of nsSNPs on protein structure stability by I- Mutant 3.0**

I – Mutant helps to analyze the effect of non-synonymous substitution on protein stability either increasing or decreasing. If DDG>0.5Kcal/mol then protein stability increases, DDG <-0.5Kcal/mol shows protein stability decreases, or neutral if -0.5< DDG<0.5Kcal/mol.

**CACNA1A**

The SNP prediction tools find out 15nsSNPs as the most deleterious in the CACNA1A gene. Out of which the I-Mutant predicts the 5 variants (Y1383C, R1666W, H1735L, C287Y, G637D) as they improve protein structure stability as their DDG > -0.5 and the rest 10 reduces protein structure stability.

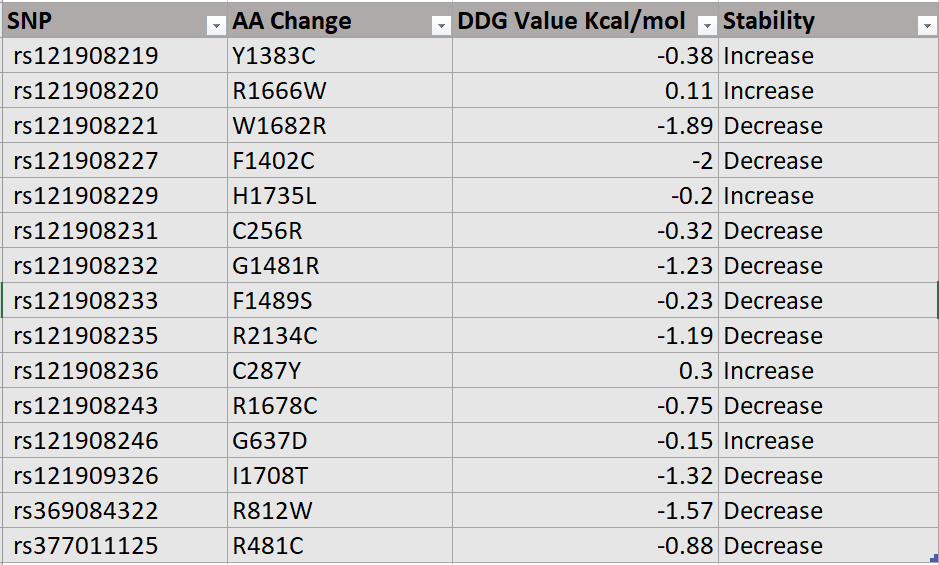


Table 4: Analysis of protein stability of deleterious nsSNPs of CACNA1A gene using I- Mutant3.0

**ATP1A2**

The SNP prediction tools find out 8 nsSNPs as the most deleterious in the ATP1A2 gene. Out of which the I-Mutant predicts that all the variants (W887R, M731T, T378N, G301R, G855E, R976C, I204N, R171W) which reduce protein structure stability as their DDG < -0.5.

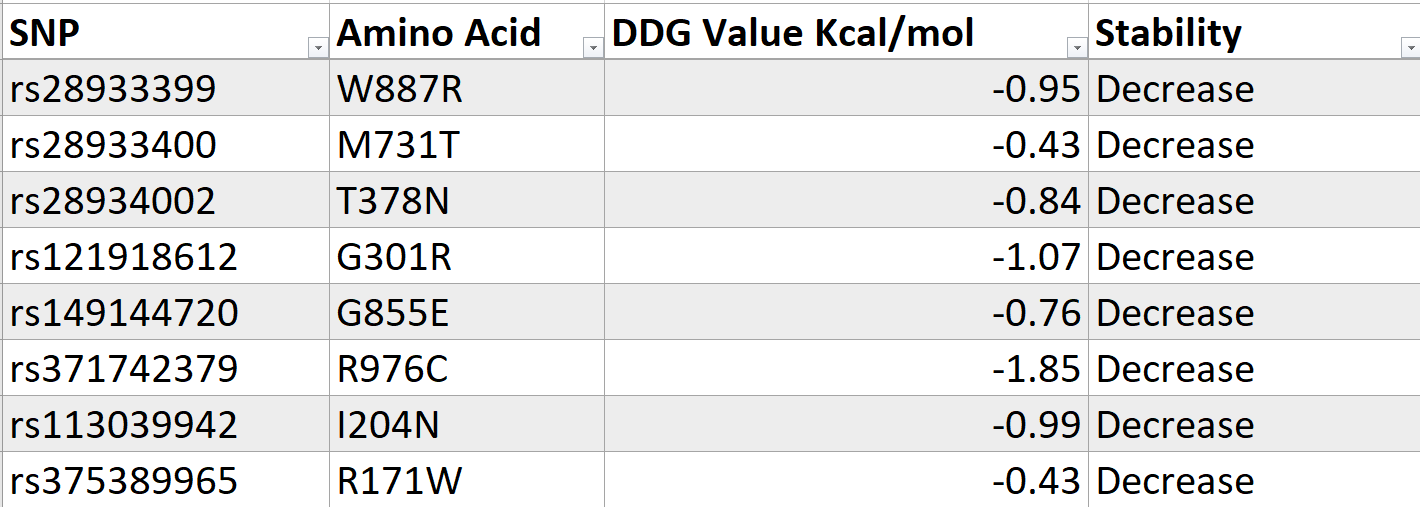
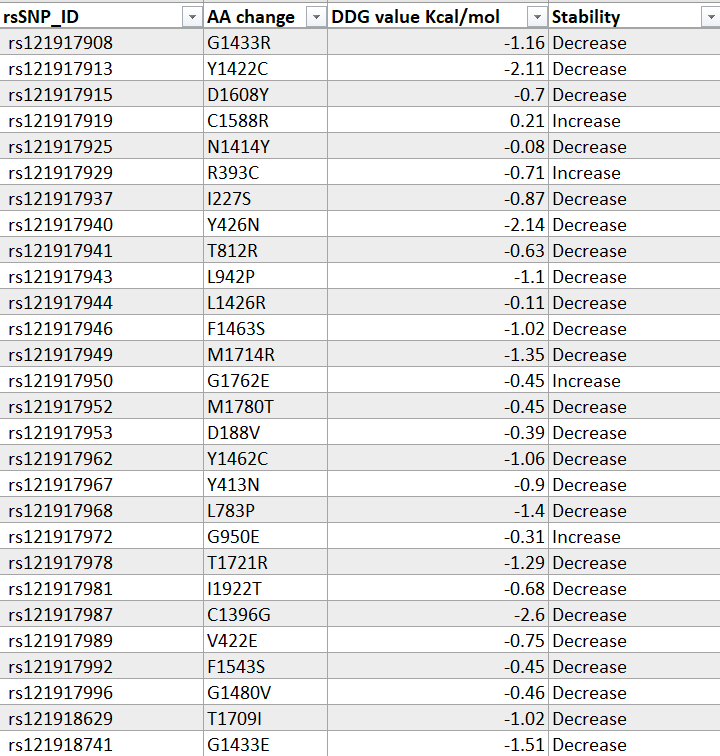


Table 5: Analysis of protein stability of deleterious nsSNPs of ATP1A2 gene using I- Mutant3.0

**SCN1A**

The SNP prediction tools find out 56 nsSNPs as the most deleterious in the SCN1A gene. Out of which the I-Mutant predicts that 10nsSNPs increase protein structure stability as their DDG > -0.5.



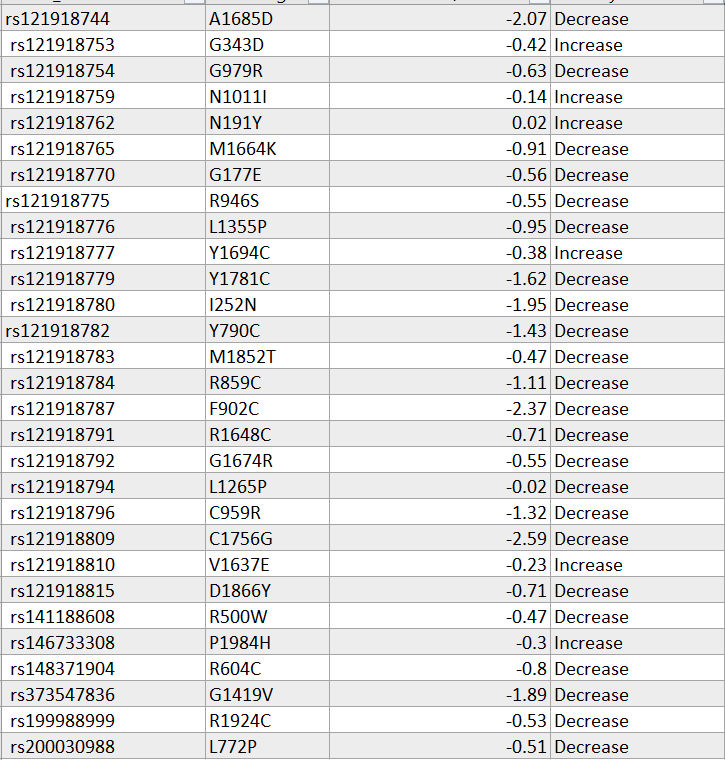


Table 6: Analysis of protein stability of deleterious nsSNPs of SCN1A gene using I- Mutant3.0

**Chapter 5: Conclusion**

As a Familial Hemiplegic Migraine, the CACNA1A, ATP1A2, and SCN1A gene plays a crucial role in various migraine problems. As a result, mutations in this gene have been linked to a variety of illnesses.

To discover the deleterious nsSNPs in the CACNA1A, ATP1A2, and SCN1A, many web-based algorithm techniques based on sequence and structural conservation were applied. Using various bioinformatics tools, 2230 SNPs of CACNA1A were found missense, out of which 15 nsSNPs were predicted as the most harmful in all the tools. Same as in the ATP1A2 gene out of 678 SNPs missense, 8 nsSNPs were predicted deleterious in all the tools, and in the SCN1A gene out of 2042 missense SNPs, 56 nsSNPs were predicted as deleterious. Also, we find that the mutations in the highly conserved region cause potentially damaging and also affect the protein structure stability. The use of several bioinformatics methods to predict pathogenic nsSNPs might save money and time, but confirmation of these nsSNPs' roles requires experimental validation.

Using the I-Mutant prediction tool protein structure stability was found. It shows that out of 15nsSNPs that are found most deleterious of the CACNA1A gene, 5 of them increase protein stability, and the rest 10 decrease protein stability. In the ATP1A2 gene, all decrease the protein stability and in the SCN1A gene, 10 nsSNPs were found which increases the protein stability.

**CHAPTER 6: FUTURE WORK**

Molecular docking –

This project mainly focused on harmful SNPs that are predicted using various tools. Future research carrying out molecular docking of the deleterious SNPs with the help of Autodock or Pyrx.Molecular docking helps in understanding the interactions between the small molecule (ligands) and the variant(mutated protein).

Ligands finding-

To do the molecular docking the first step was to find the ligand so to perform the virtual screening of the ligand we will use the SWISSADME tool to see the Bioavailability Radar and the Boiled egg to see where the ligand can cross the Blood Brain Barrier or not. And try to find the phytochemicals which were already used in migraine disorders.

Development of Scoring function –

One potential work for future direction is the development of an enhanced scoring function because knowledge-based approaches have limitations in accurately capturing interactions of ligands and target proteins. So exploring machine learning techniques, such as deep learning could to the creation of a more accurate scoring function for molecular docking.

**References**

[1] Pietrobon, D. (2007). Familial hemiplegic migraine. *Neurotherapeutics*, *4*, 274-284.

[2] Jackman, S. L., & Regehr, W. G. (2017). The mechanisms and functions of synaptic facilitation. *Neuron*, *94*(3), 447-464.

[3] Rozario, L. T., Sharker, T., & Nila, T. A. (2021). In silico analysis of deleterious SNPs of human MTUS1 gene and their impacts on subsequent protein structure and function. *Plos One*, *16*(6), e0252932.

[4] Venkata Subbiah, H., Ramesh Babu, P., & Subbiah, U. (2020). In silico analysis of non-synonymous single nucleotide polymorphisms of human DEFB1 gene. *Egyptian Journal of Medical Human Genetics*, *21*(1), 1-9.

[5] Elkhattabi, L., Morjane, I., Charoute, H., Amghar, S., Bouafi, H., Elkarhat, Z., ... & Barakat, A. (2019). In silico analysis of coding/noncoding SNPs of human RETN gene and characterization of their impact on resistin stability and structure. *Journal of Diabetes Research*, *2019*.

[6] Sutherland, H. G., Albury, C. L., & Griffiths, L. R. (2019). Advances in genetics of migraine. *The journal of headache and pain*, *20*(1), 1-20.

[7] Nanou, E., & Catterall, W. A. (2018). Calcium channels, synaptic plasticity, and neuropsychiatric disease. *Neuron*, *98*(3), 466-481.

[8] Benarroch EE (2010) Glutamate transporters: diversity, function, and

involvement in neurologic disease. Neurology 74(3):259–264

[9] Bassi MT, Bresolin N, Tonelli A, Nazos K, Crippa F, Baschirotto C et al (2004)A novel mutation in the ATP1A2 gene cause alternating hemiplegia of

childhood. J Med Genet 41(8):621–628

[10] Deprez L, Weckhuysen S, Peeters K, Deconinck T, Claeys KG, Claes LR et al(2008) Epilepsy as part of the phenotype associated with ATP1A2 mutations.

[11] Mantuano E, Romano S, Veneziano L, Gellera C, Castellotti B, Caimi S et

al (2010) Identification of novel and recurrent CACNA1A gene mutations in fifteen patients with episodic ataxia type 2. J Neurol Sci 291(1–2):30–36

[12] De Fusco M, Marconi R, Silvestri L, Atorino L, Rampoldi L, Morgante L et al (2003)Haploinsufficiency of ATP1A2 encoding the Na+/K+ pump alpha2 subunit associated with familial hemiplegic migraine type 2. Nat Genet 33(2):192–196

[13] Vanmolkot KR, Stroink H, Koenderink JB, Kors EE, van den Heuvel JJ, van den Boogerd EH et al (2006) Severe episodic neurological deficits and

permanent mental retardation in a child with a novel FHM2 ATP1A2

mutation. Ann Neurol 59(2):310–314

[14] Unekawa M, Ikeda K, Tomita Y, Kawakami K, Suzuki N (2018) Enhanced

susceptibility to cortical spreading depression in two types of Na(+),K(+

)-ATPase alpha2 subunit-deficient mice as a model of familial hemiplegic

migraine 2. Cephalalgia 38(9):1515–1524

[15] Isaksen TJ, Lykke-Hartmann K (2016) Insights into the pathology of the

alpha2-Na(+)/K(+)-ATPase in neurological disorders; lessons from animal

models. Front Physiol 7:161