International Journal of Applied Research in Bioinformatics
Volume 12 • Issue 1

Investigating Variations/SNPs in AUH Gene Causing 3-Methylglutaconic Aciduria, Type I

Malik Muhammad Sajjad, NFC Institute of Engineering and Technology, Multan, Pakistan
Sarah Bukhari, NFC Institute of Engineering and Technology, Multan, Pakistan
Omer Aziz, NFC Institute of Engineering and Technology, Multan, Pakistan

ABSTRACT

Single nucleotide polymorphisms (SNPs) are a source variation in a genome. The AUH gene gives guidance about how to generate an enzyme named 3-methylglutaconyl-CoA hydratase. Mutations in AUH gene leads to 3-Methylglutaconic aciduria type I disease. The authors used multiple bioinformatics tools SIFT, Provean, PolyPhen, PHD-SNP, I-Mutant, ConSurf server, and Project HOPE to isolate missense SNPs that should be deleterious to the structure and function of the AUH protein. This research aims to analyze the impact of missense SNPs on the structure and function of AUH protein. There have been a total of 259 Missense SNPs obtained, of which 13 mutations were identified as deleterious to the structure and function of the AUH protein. This is the first study in relation to AUH gene missense SNPs where most damaging SNPs associated with the AUH gene were examined using computational analysis. This research could be useful in designing specific medicines for treatment of genomic variation diseases.

KEYWORDS

Computational Analysis, Consurf, I-Mutant, Polyphen 2, Project Hope, Sift, Single Nucleotide Polymorphism

INTRODUCTION

A Single nucleotide polymorphism (SNP) is a change of the base of a genome. SNPs are closely associated to human genetic polymorphism. The coding region of human genome contains SNPs that are much important. Among these the nsSNPs (missense SNPs) are highly significant due to their huge effect on the amino acid changing's and a great association with many diseases. Functional differences could have a damaging or neutral impact on the role and structure of proteins(Arshad et al., 2018).

The AUH gene also known as MG-CoA hydratase offers directions for a 3-methylglutaconyl-CoA hydratase enzyme/protein to be produced. This enzyme is discovered in cell structures called mitochondria, which converts food energy into a form that can be used by cells. This enzyme plays significant role within mitochondria in breaking down proteins into smaller molecules that can be used by cells to generate energy (Ijlst et al., 2002).

3-methylglutaconyl-CoA hydratase deficiency is caused by mutations in the AUH gene(Ijlst et al., 2002).Deficiency inside the methylglutaconyl-CoA hydratase causes 3-Methylglutaconic aciduria type

DOI: 10.4018/IJARB.2022010104

I disease(Nga Ly et al., 2003).This is a part of the leucine metabolism pathway's inborn errors. The 3Methylglutaconic aciduria type I (MGA1) causes neurological issues such as movement disorders and problems with ability to think(cognition) that occur during childhood but may delay the diagnosis until adulthood (Tavasoli et al., 2017).

Numerous previous studies have made known that missense SNPs are legally responsible for about fifty percent of mutations that are related to many genetics problems. A study examined polymorphisms in the ABCA1gene and calculated their harmful impacts like producing tanger disorder. A closely related analysis revealed that polymorphisms in STEAP2 which trigger prostate cancer to develop upregulation. Due to their damaging effects on protein structural features, SNPs in the NKX2-5 gene have been reported to be linked with congenital heart defects. A recent study indicated that nsSNPs may cause malignant melanoma in the MITF gene. (Arshad et al., 2018).

Basing on the significance of AUH gene in inflammatory disease like 3-Methylglutaconic aciduria type I (MGA1) we performed computational analysis by the help of missense SNP effect predictors like Imutant2, PolyPhen2,SIFT, Provean and PhD-SNP. Consurf conservation method was used to further evaluate largely harmful nsSNPs. To examine protein 3D structures Project HOPE tool was used.

First of all, this study has defined and evaluated the effects of the SNPs that exist in the AUH coding area using computational tools. There have been a total of 259 Missense SNPs obtained, of which only 15 mutations were identified as deleterious by the SIFT, PhD-SNP, PolyPhen, Provean and I-mutant2 and among these 15 mutations, only 13 were highly conserved by ConSurf server. rs1425044760, rs931179337, rs775913288, rs1165561302, rs1409932485, rs148157596, rs752044125, rs1294255949, rs755299132, rs773299021, rs748318386, rs1398261487, rs964265074 are the reference IDs of damaging missense SNPs and the amino acid changes are L143F, D195N, L207R, G220R, G220E, R226C, R226H, E235K, A248T, A271P, A275V, F278S, and G324R.

Several papers explaining the AUH gene association with disease are written. The computational analysis of missense SNPs in the AUH gene have not yet been analyzed. In this study, a computational analysis of missense SNPs in AUH gene was done by different bioinformatics tools and we analyzed the missense SNPs of AUH gene, that could perform an important part in causing disease using several computational Bioinformatics software.

MATERIAL AND METHODS DATA RETRIEVAL

The human AUH gene SNP information was collected by visiting NCBI webiste (https://www.ncbi.nlm.nih.gov/snp/). The corresponding AUH gene missense SNPs were extracted according to their rsIDs(Sherry, 2001).

Figure 1 illustrates the workflow, methods and repositories used to identify potential functional SNPs within the human AUH gene.

IDENTIFICATION OF DAMAGING SNPS

We were used different in silico tools in order to separate the deleterious missense SNPs and for checking the impact of these SNPs on protein function/structure. The study analyzes deleterious missense SNPs only instead of neutral ones.

SIFT (Sorting Intolerant From Tolerant)

SIFT (http://sift.jcvi.org) a web tool was used to detect the deleterious nsSNPs from the tolerated SNPs on the base of the sequence homology. A range of value is used for the prediction (*Accounting for Human Polymorphisms Predicted to Affect Protein Function*, n.d.)(Ng & Henikoff, 2001).