EUROPEAN ACADEMIC RESEARCH
Vol. II, Issue 12/ March 2015

Impact Factor: 3.1 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

# Functional SNPs within Exons and 3-untranslated regions of *PAH* gene associated with Phenylketonuria: using bioinformatics methods

SAFA A. A. ABDALGALEEL
Faculty of Medical Laboratory Sciences
Al Zaiem Al Azhari University, Khartoum, Sudan
NANSI M. ELSHAFIE
Faculty of medical Laboratory Sciences
Al Zaiem Al Azhari University, Khartoum, Sudan
KHITMA A. SIRELKHATIM
Departments of Biochemistry and Food Science
University of Khartoum, Khartoum, Sudan
HIND G.A. IDRIS
Faculty of Medical Laboratory Sciences
Sudan University of Science and Technology, Khartoum, Sudan
LEENA A. E. OSMAN
Faculty of Chemistry
United Arad Emirates University, Al-Ain, United Arab Emirates
SOFIA B. MOHAMED
Department of Vector and Biomedical Studies
Tropical Medicine Research Institute, Khartoum, Sudan
MOHAMED M.HASSAN[1]
Faculty of Medical Laboratory Sciences
Al Zaiem Al Azhari University, Khartoum, Sudan

**Abstract:**

*Phenylketonuria (PKU) is an inherited autosomal recessive human genetic disorder, many previous studies proved the association of PKU disease and PAH gene. Mutations within PAH gene could resulting to deficiency in activity of the hepatic phenylalanine hydroxylase enzyme (PAH), which also could leads to irreversible severe physical, neurological and cognitive abnormalities. Most human genetic variation is represented by single nucleotide polymorphisms*

[1] Corresponding author: m_rbx@hotmail.com

*(SNPs), and many human SNPs are believed to cause phenotypic differences between individuals. These variations are considered as being the cause of diseases, differences in response to treatment, susceptibility to diseases or being neutral having no impact at all. In this study for the nsSNPs of PAH gene, comprehensive analysis were done to filtrate deleterious from benign nsSNPs at both functional and structural level, used SIFT and PolyPhen servers. Furthermore, analyzed the SNPs in 3'UTR (microRNAs binding and there target sites) used PolymiRTS database. A major findings in this study were 30 damaging SNPs detection by SIFT and Polyphen-2 servers, and six functional SNPs disturbed or create a new miRNA binding sits. Based on this study investigation, shown the potential predicted of nsSNPs and mirSNPs used insilico methods.*

**Key words:** Nonsynonymous polymorphisms (nsSNPs), miRNA-binding site polymorphism, *PAH* gene, Bioinformatics method, Phenylketonuria disease (PKU).

## Introduction

Phenylketonuria (PKU) is an autosomal recessive disorder caused by mutations in the phenylalanine hydroxylase (PAH) gene (Pietro and Daniela, 2014). PAH, one of the highly heterogeneous hepatic`s enzyme, mainly function to convert phenylalanine to tyrosine in the existence of co-factor (Scriver et al., 1995). Deficiency of PAH enzyme or its cofactor, leads to defect in breakdown of excess accumulates of phenylalanine (necessary part of the human diet) and converted into phenyl pyruvate (also known as phenyl ketone), which can be detected in urine (Williams et al., 2008). PAH gene is located on the long arm of chromosome 12 of 79, 277 bases long. Its cytogenetic address is 12q22-q24.2 .By the same token PAH protein composed of 452 amino acid and primarily expressed in liver, kidney and spleen (Charles R and Scriver, 2007).

With rapid advances in high-throughput genotyping and next generation sequencing technologies, a vast amount of genetic variation has been discovered and deposited in databases (Eric et al., 2011). One of the major challenges in human genetic variation analysis is to distinguish functional from non-functional variants. The simplest form of genetic variation is the substitution of a single nucleotide coined as Single Nucleotide Polymorphism (SNPs). SNPs occur at a frequency of approximately to every 100 to 300 base pairs throughout the genome (Lee et al., 2005). There is an effectively and efficiently need to identify functional non-synonymous SNPs (SNPs within exon regions) which may be deleterious or disease causing and to identify their molecular effects. Phenotype prediction of nsSNPs using computational analysis may provide a good way to explore the function of nsSNPs and its susceptibility relationship to disease. For this purpose, number of bioinformatics tools, based on recent findings from evolutionary biology (amino acid sequence), protein structure research and computational biology may provide useful information in assessing the functional importance of SNPs (Rajith and George, 2011). Currently, most molecular studies are focusing on SNPs located in coding and regulatory regions, yet many of these studies have been unable to detect significant associations between SNPs and disease susceptibility. The non-protein coding parts of the genome have been recognized as key players in the regulation of gene expression. Among various classes of non-coding RNAs, the best known are micro RNAs (miRNA), small (~22 nucleotide) RNA molecules. MiRNAs bind to complementary target sequences usually located in the 3' untranslated region (3'UTR) of messenger RNAs and act predominantly by negatively regulating gene expression (Barte, 2009). According to the miRBase database, a catalog of miRNA sequences in animals, plants and viruses, so far over 2500 mature human miRNAs have been characterized. Since

miRNAs are predicted to regulate over 60% of human protein-coding genes (Agnieszka, 2014). It is not surprising that they have been shown to regulate a plethora of biological processes, including cell proliferation, apoptosis, differentiation and metabolism. ). The main objective of this study is to screen and identifies the nsSNPs and miSNPs (micro RNA SNPS) in *PAH* gene which could be responsible for Phenylketonuria disease using bioinformatics methods.

## Material and Methods

### Dataset

In this study for *PAH* gene, a universal stock for single-nucleotide alternatives database (dbSNP) was used to obtained SNPs data (http://www.ncbi .nlm.nih.gov/SNP/), and universal protein UniProt database used to get proteins sequences (http://www.uniprot.org). Homo sapiens SNPs within PAH gene were separated from hole species SNPs, and then two SNPs batches of - non-synonymous and 3'UTR SNPs- were prepared to submit to subsequent insilico tools.

### Bioinformatics processing and data analysis

Total nsSNPs submitted respectively to SIFT and Polyphen servers, then double positive (damaging by two servers) SNPs were submitted to I-mutant tool, after that protein sequences of double positive results were homology modeling to predict the effect of these SNPs in the protein structural level.

### Analysis of Functional Effect on Protein used SIFT

There are many web-based resources available that allows predicting whether non-synonymous coding SNPs that have functional effects on proteins function level. SIFT is a conservation analysis and predict tool, used to predict the phenotypic effect of amino acid substitutions. The underlying

principle of this program generates alignments with a large number of homologous sequences and assigns scores to each residue, ranging from zero to one. SIFT scores were classified as intolerant when score is below or equal to 0.05, and tolerated if the score is greater than 0.05 (Ng and Henikoff, 2003). SIFT version 5.1 is available at: http://sift.bii.a-star.edu.sg/index.html.

**Evaluation of functional change in nsSNPs used PolyPhen**

A computational tool used to identify functional nsSNPs on proteins. Predictions are based on a combination of structural and sequence annotation information characterizing a substitution and its position in the protein. PolyPhen predictions (score between 0-1) were classified as probably damaging (> 0.8), possibly damaging (0.6 - 0.8), or benign (< 0.6) (Jones and Mohrenweiser, 2004). PolyPhen version 2 is available at: http://genetics.bwh.harvard.edu/pph2/index.shtml.

**Protein Stability Analysis**

I-Mutant 3.0 software was used to predict nsSNP causing protein stability change. I-Mutant 3.0 is a support vector machine- (SVM) based tool for the automatic prediction of protein stability change upon single amino acid substitution. The software computed the predicted free energy change value or sign (DDG) which is calculated from the unfolding Gibbs free energy value of the mutated protein minus unfolding Gibbs free energy value of the native protein (kcal/mol). A positive DDG value indicates that the mutated protein possesses high stability and vice versa (Capriott et al., 2005). http://gpcr.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I Mutant3.0.cgi

**Protein 3D structural Modeling**

Modeling SNPs on the 3D structure of the proteins is a very helpful action in order to predict the effect that SNPs may cause on the structural level. Therefore we used CPH models 3.2 server to predict the 3D structure for those proteins with an unknown 3D structure model. It's a protein homology modeling server, where the template recognition is based on profile-to-profile alignment, guided by secondary structure and exposure predictions (Nielsen et al., 2010).

**Modeling amino acid substitution**

UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data. Chimera (version 1.8) software was used to scan the 3D (three-dimensional) structure of specific protein, and hence modifies the original amino acid with the mutated one to see the impact that can be produced. The outcome is then a graphic model depicting the mutation (Hassan et al., 2014). Chimera (version 1.8) currently available within the Chimera package and available from the Chimera web site http://www.cgl.ucsf.edu/chimera/

**PolymiRTS data base 3.0 to Polymorphism in microRNA Target Site**

PolymiRTS database was designed specifically to analyze SNPs of 3'UTR region, it aims to identify single-nucleotide polymorphisms (SNPs) that affect miRNA targeting in human and mouse. We used this server at this stage to determine SNPs .That may alter miRNA target sites (Hassan et al., 2014).

**Result and Dissection**

The *PAH* gene investigated in this work contained a total of 4157 Homo sapiens SNPs, of which 474 were nsSNPs, and 29 SNPs in the 3'UTR. Figuer.1 shows the distribution of SNPs in

*PAH* gene within dbSNP, where green column shows 25.5% total Homo sapiens SNPs of all species SNPs, in addition red column (11.4%) non-synonymous and blue column (0.7%) 3UTR SNPs of total Homo sapiens. Total SNPs contain Homo sapiens SNPs of 3′UTR/5′UTR near to gene, non-coding 3′UTR/5′UTR, intron, coding synonymous and coding non-synonymous (frame shift, missense, nonsense and stop gained) regions.

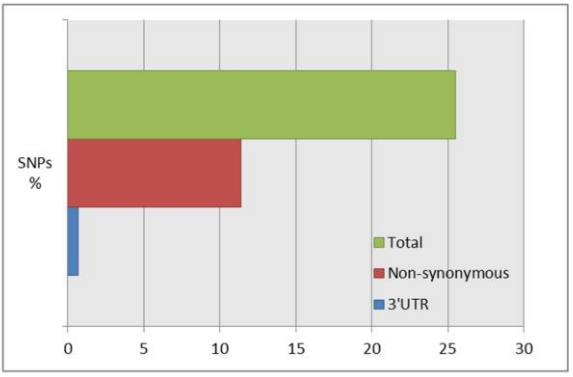



**Fig.1. Shows SNPs distribution of *PAH* gene**

*Prediction of Deleterious or Damaged nsSNPs by SIFT and PolyPhen*

474 nsSNPs were submitted independently to the SIFT and PolyPhen serves, result showed that, 30 nsSNPs were predicted as deleterious having the tolerance index score of <0.05 and damaging with PSIC score of >0.700 by SFIT and PolyPhen respectively. (Table1).

**Table 1. SIFT and PolyPhen prediction results**

| Gene name | SNP ID | Amino Acid change | Polyphen-2 Result | PolyPhen score | SIFT Result | SIFT score |
|---|---|---|---|---|---|---|
| **PAH** | Rs5030847 | R252W | Probably damaging | 1.000 | damaging | 0.02 |
| | rs5030849 | R261P | Probably damaging | 0.999 | damaging | 0.02 |

| | | | | | |
|---|---|---|---|---|---|
| rs62507265 | Y343C | Probably damaging | 1.000 | damaging | 0.02 |
| rs62507326 | E78K | Probably damaging | 0.981 | damaging | 0.01 |
| rs62508595 | C357G | Probably damaging | 0.990 | damaging | 0.04 |
| rs62508695 | R71H | Possibly damaging | 0.938 | damaging | 0.02 |
| rs62514931 | P211T | Probably damaging | 0.970 | damaging | 0.03 |
| rs62514933 | G218V | Probably damaging | 1.000 | damaging | 0.02 |
| rs62514934 | E221G | Possibly damaging | 0.860 | damaging | 0.01 |
| rs62514957 | A322T | Probably damaging | 1.000 | damaging | 0.02 |
| rs62514958 | A322G | Possibly damaging | 0.895 | damaging | 0.01 |
| rs62516098 | P366H | Probably damaging | 1.000 | damaging | 0.01 |
| rs62516102 | D394A | Possibly damaging | 0.870 | damaging | 0.01 |
| rs62516149 | M276V | Possibly damaging | 0.517 | damaging | 0.03 |
| rs62517198 | L212P | Probably damaging | 1.000 | damaging | 0.01 |
| rs62642912 | A313T | Probably damaging | 1.000 | damaging | 0.03 |
| rs62642928 | L41F | Probably damaging | 0.992 | damaging | 0.02 |
| rs62642934 | I306V | Possibly damaging | 0.901 | damaging | 0.03 |
| rs62642935 | A309V | Probably damaging | 1.000 | damaging | 0.02 |
| rs62642939 | R297H | Possibly damaging | 0.836 | damaging | 0.02 |
| Rs62642940 | P314H | Probably damaging | 1.000 | damaging | 0.01 |
| rs62644467 | R413C | Probably damaging | 0.991 | damaging | 0.04 |
| rs62644471 | Y417H | Probably | 1.000 | damaging | 0.02 |
| rs63048261 | S70P | Probably damaging | 1.000 | damaging | 0.03 |
| rs76394784 | R68S | Probably | 1.000 | damaging | 0.04 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | damaging | | | |
| | rs76687508 | R241C | Probably damaging | 1.000 | damaging | 0.05 |
| | rs78655458 | Y277D | Probably damaging | 1.000 | damaging | 0.02 |
| | rs79931499 | R413P | Possibly damaging | 0.805 | damaging | 0.01 |
| | rs180819807 | F392I | Probably damaging | 1.000 | damaging | 0.02 |
| | rs192592111 | Q172H | Possibly damaging | 0.705 | damaging | 0.04 |

*Identification protein stability change due to nsSNP*

All damaging nsSNPs (by SIFT and PolyPhen) were submitted to I-Mutant server. The results showed, 28 nsSNPs were found to be decreasing the protein stability and two nsSNPS were predicted to be increase protein stability (Table 2). Above results demonstrated that all proteins stability was changed due to SNPs alteration.

**Table 2. Prediction of I-Mutant 3.0**

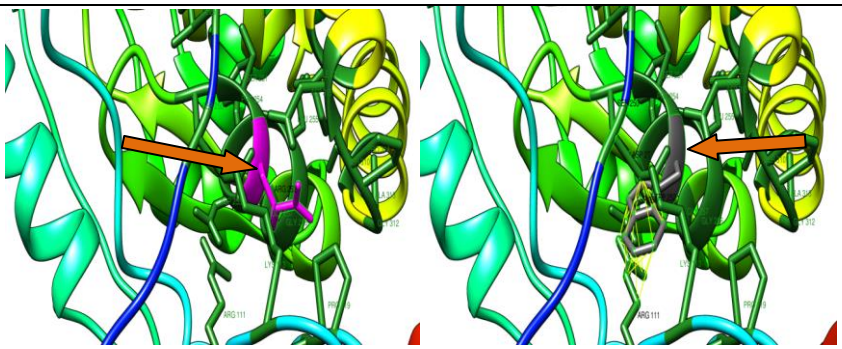
| Gene name | SNP ID | Amino acid Position | WT | MT | SVMS2 prediction Effect | DDG value prediction Kcal/mol | RI |
|---|---|---|---|---|---|---|---|
| **PAH** | rs5030847 | 252 | R | W | Decrease | -0.37 | 3 |
| | rs5030849 | 261 | R | P | increase | 0.09 | 1 |
| | rs62507265 | 343 | Y | C | Decrease | -0.54 | 5 |
| | rs62507326 | 78 | E | K | Decrease | -0.31 | 5 |
| | rs62508595 | 357 | C | G | Decrease | -1.13 | 8 |
| | rs62508695 | 71 | R | H | Decrease | -1.47 | 9 |
| | rs62514931 | 211 | P | T | Decrease | -1.18 | 8 |
| | rs62514933 | 218 | G | V | Decrease | -0.41 | 6 |
| | rs62514934 | 221 | E | G | Decrease | -1.19 | 7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| rs62514957 | 322 | A | T | Decrease | -0.60 | 5 |
| rs62514958 | 322 | A | G | Decrease | -1.21 | 7 |
| rs62516098 | 366 | P | H | Decrease | -1.04 | 6 |
| rs62516102 | 394 | D | A | Decrease | -0.88 | 8 |
| rs62516149 | 276 | M | V | Decrease | -0.81 | 7 |
| rs62517198 | 212 | L | P | Decrease | -1.48 | 5 |
| rs62642912 | 313 | A | T | Decrease | -0.60 | 6 |
| rs62642928 | 41 | L | F | Decrease | -1.20 | 8 |
| rs62642934 | 306 | I | V | Decrease | -0.68 | 8 |
| rs62642935 | 309 | A | V | Decrease | 0.05 | 2 |
| rs62642939 | 297 | R | H | Decrease | -1.43 | 9 |
| rs62642940 | 314 | P | H | Decrease | -1.25 | 8 |
| rs62644467 | 413 | R | C | Decrease | -0.98 | 4 |
| rs62644471 | 417 | Y | H | Decrease | -1.76 | 8 |
| rs63048261 | 70 | S | P | Increase | -0.41 | 2 |
| rs76394784 | 68 | R | S | Decrease | -1.58 | 9 |
| rs76687508 | 241 | R | C | Decrease | -0.78 | 3 |
| rs78655458 | 277 | Y | D | Decrease | -1.14 | 4 |
| rs79931499 | 413 | R | P | Decrease | -0.69 | 3 |
| rs180819807 | 392 | F | I | Decrease | -1.06 | 9 |
| rs192592111 | 172 | Q | H | Decrease | -0.73 | 7 |

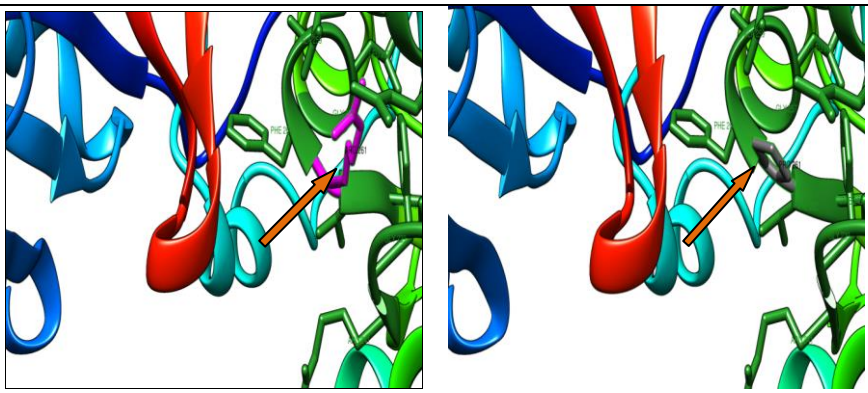
For all the predictions, pH and temperature were selected as 7.0 pH and 25 °C, respectively. WT: Wild type amino acid, MT Mutant type amino acid, DDG: DG (New Protein)-DG (Wild Type) in Kcal/mol (DDG<0: Decrease stability, DDG>0: Increase stability), RI: Reliability index

*Homology modeling new and wide amino acids of deleterious nsSNPs*

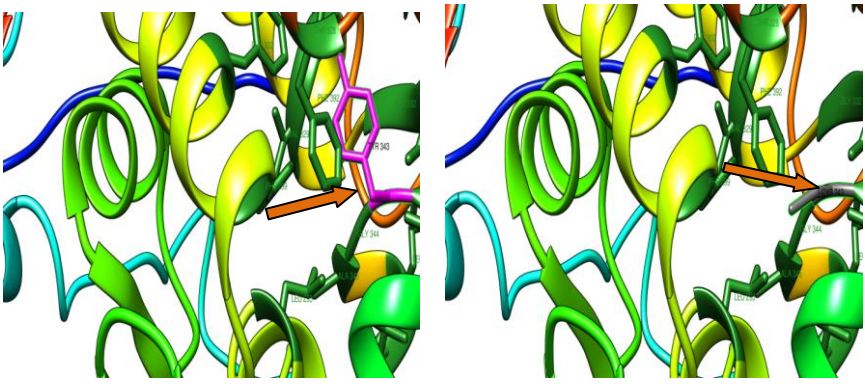
Protein 3D structural information is an important feature for predicting the effects of deleterious nsSNPs. Analysis of the protein structure provides information about the environment of the mutation. The structure of native and newly predicted structure of *PAH* gene were shown in (Figure 2).



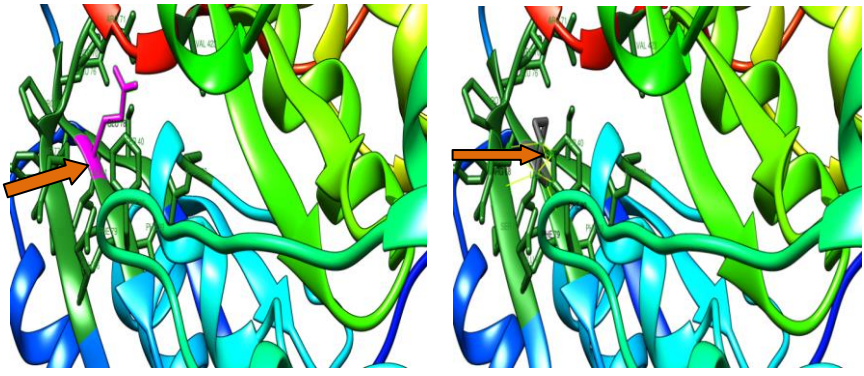
**rs5030847, P: 252, WT:R, NT:W**



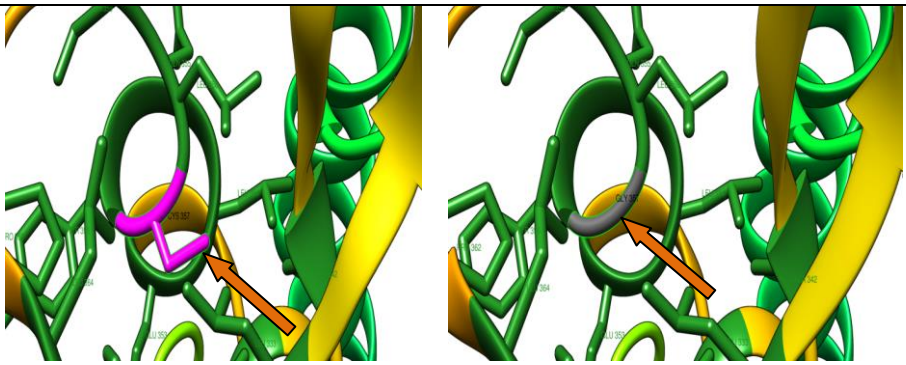
**rs5030849, P:261, WT:R, NT:P**



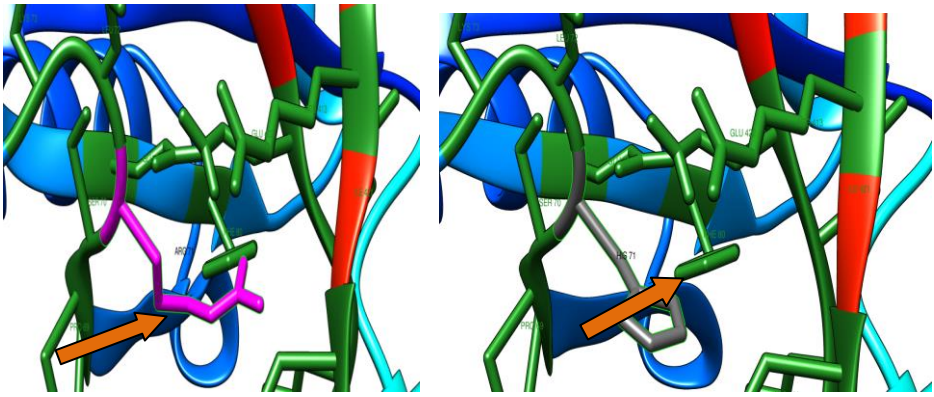
**rs62507265, P:343, WT:Y, NT:C**



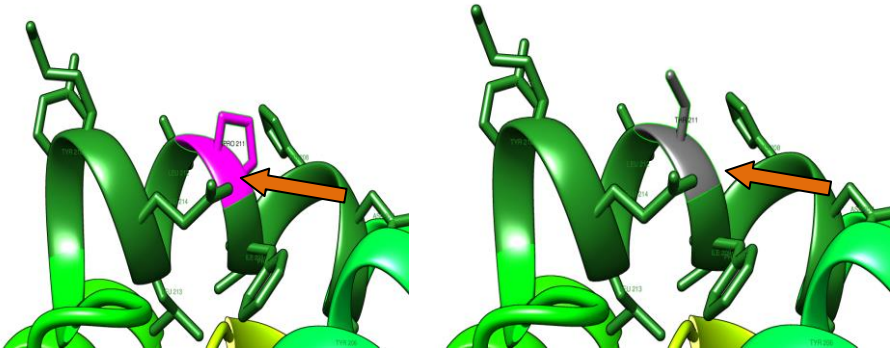
**rs62507326, P:78, WT:E, NT:K**



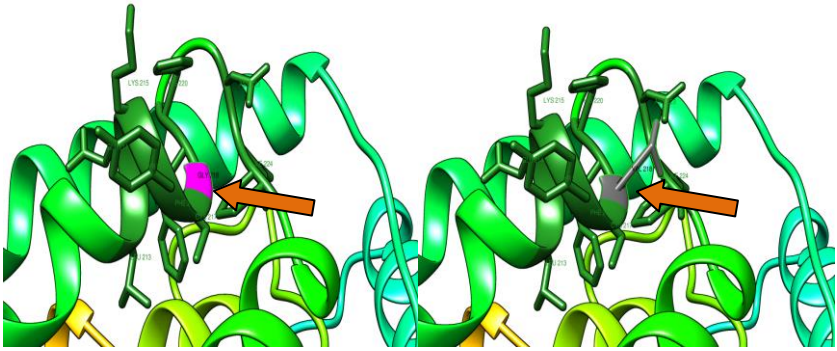
**rs62508595 , P:357, WT:C, NT:G**



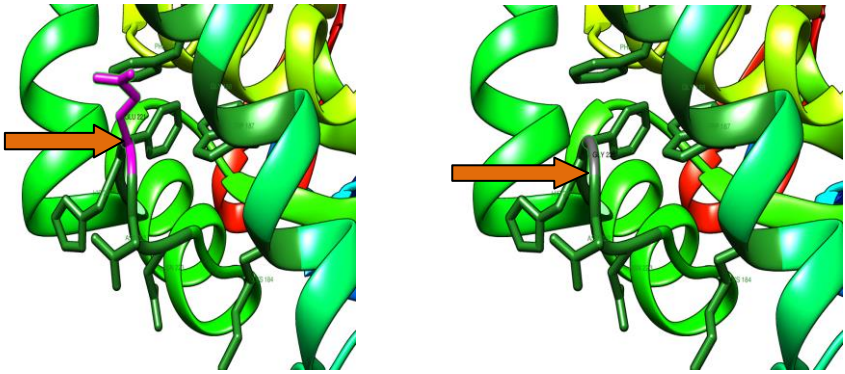
**rs62508695, P:71, WT:R, NT:H**



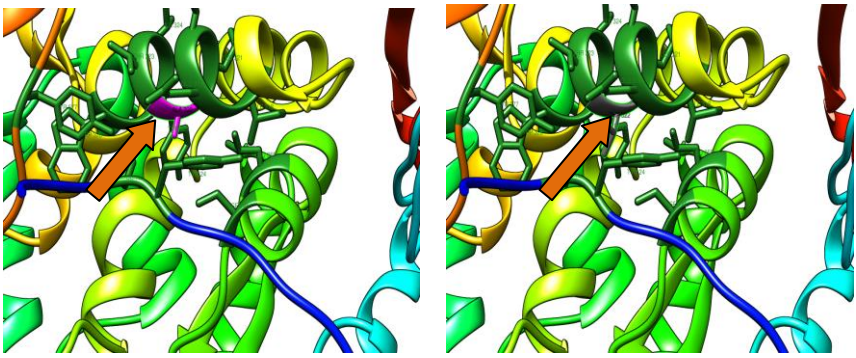
**rs62514931, P:211, WT:P, NT:T**



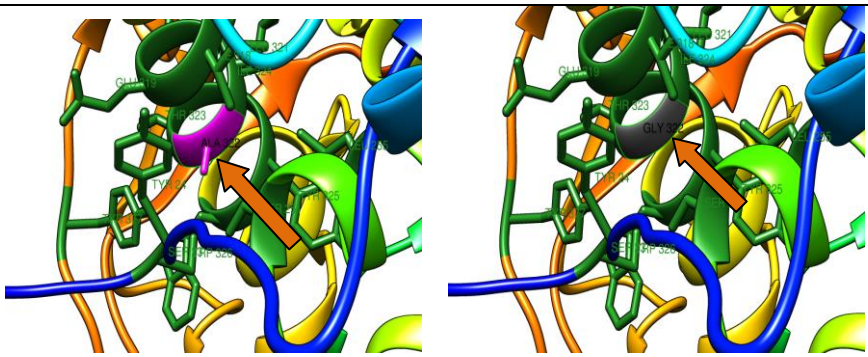
**rs62514933, P:218, WT:G, NT:V**



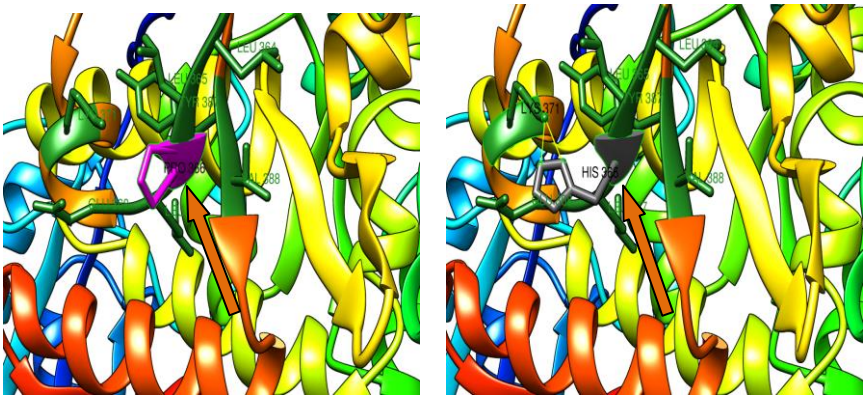
**rs62514934, P:221, WT:E, NT:G**



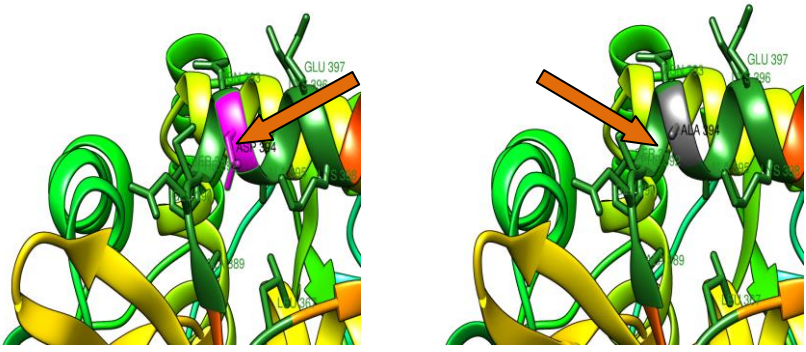
**rs62514957, P:322, WT:A, NT:T**



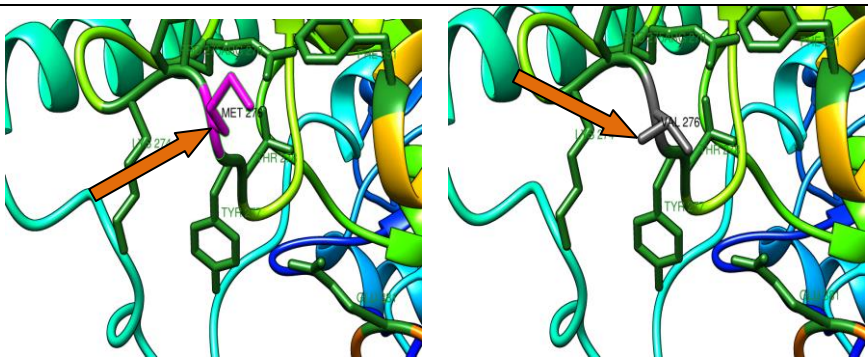
**rs62514958, P:322, WT:A, NT:G**



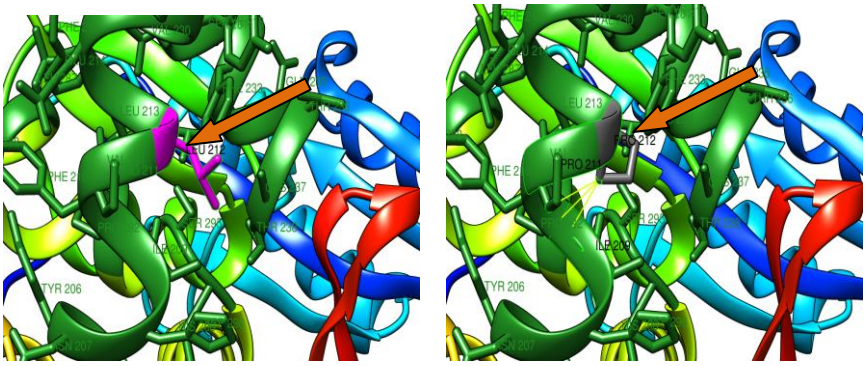
**rs62516098, P:366, WT:P, NT:H**



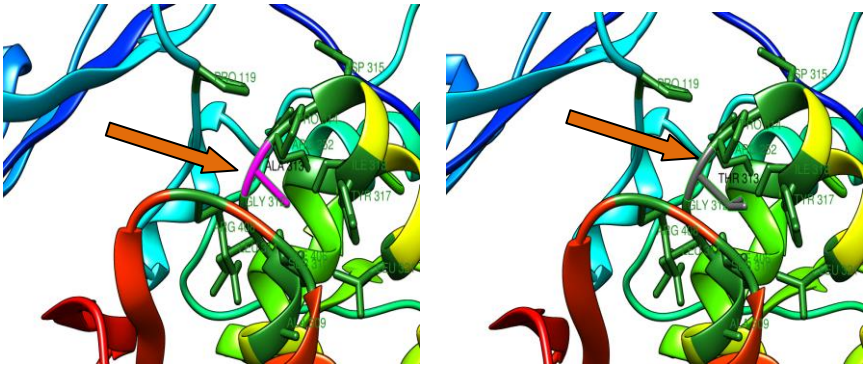
**rs62516102, P:394, WT:D, NT:A**



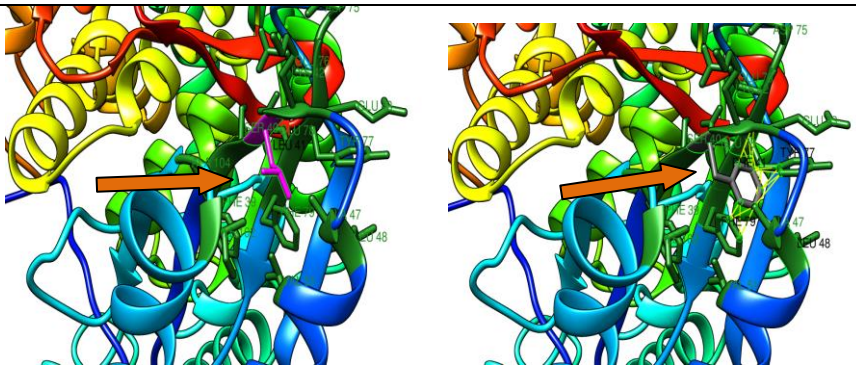
**rs62516149, P:276, WT:M, NT:V**



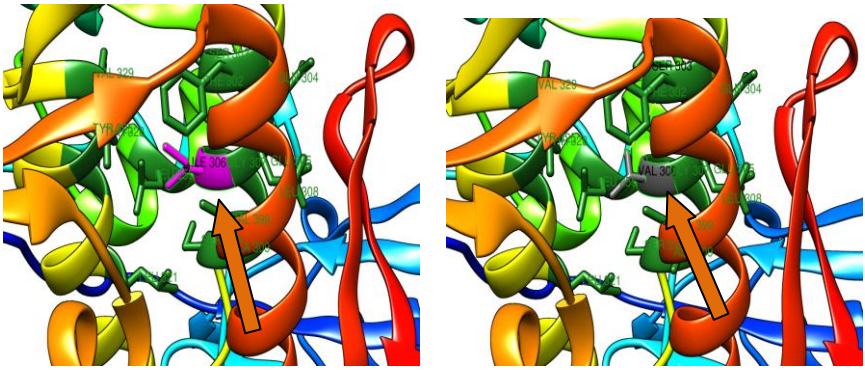
**rs62517198, P:212, WT:L, NT:P**



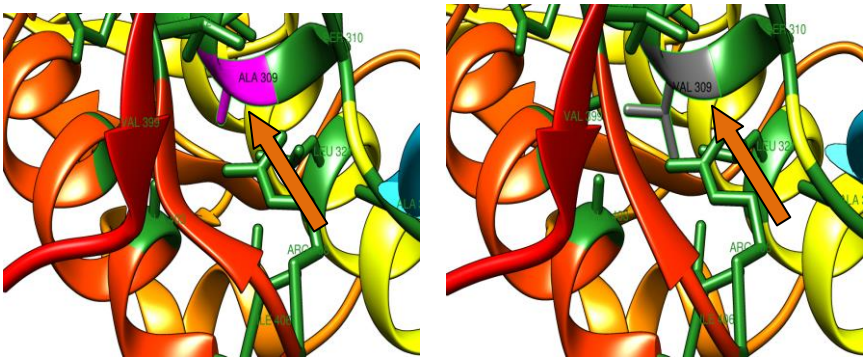
**rs62642912, P:313, WT:A, NT:T**



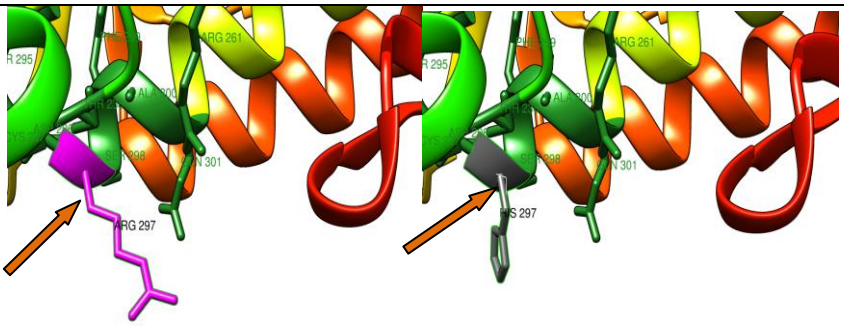
**rs62642928, P:41, WT:L, NT:F**



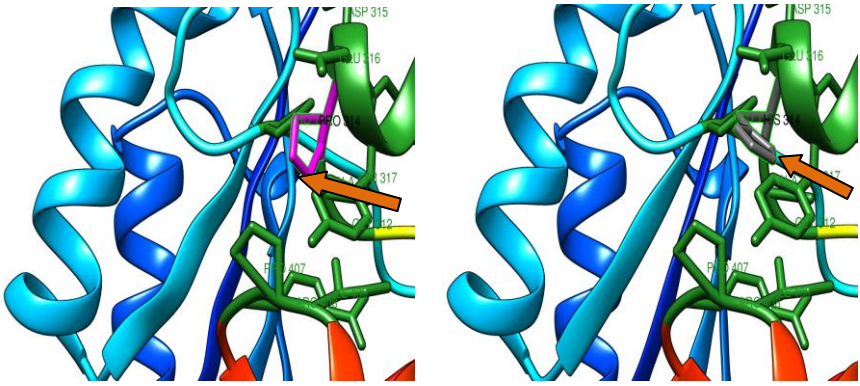
**rs62642934, P:306, WT:I, NT:V**



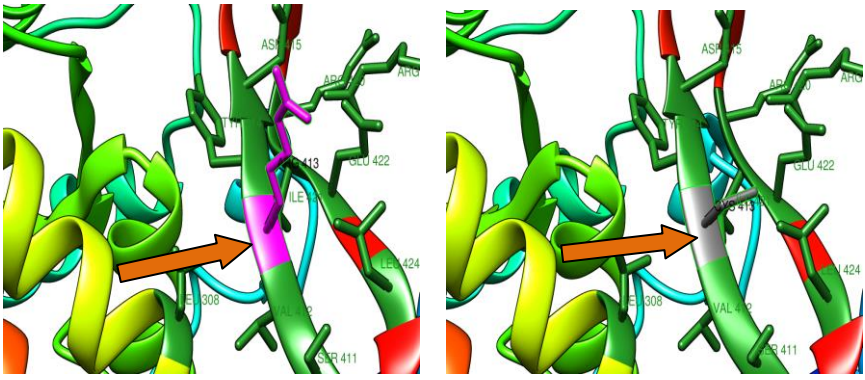
**rs62642935, P:309, WT:A, NT:V**



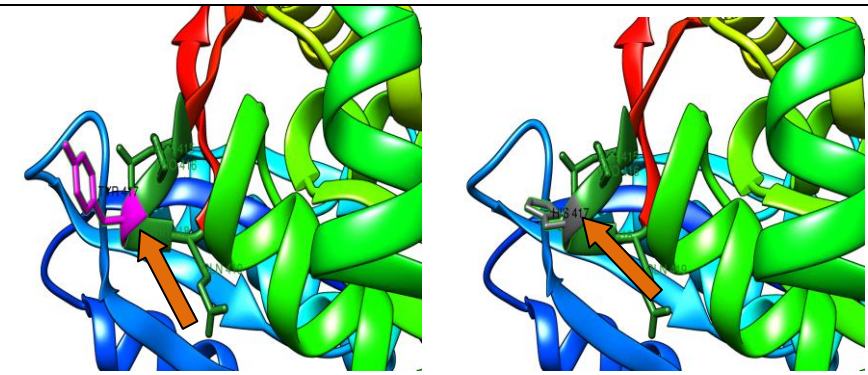

**rs62642939, P:297, WT:R, NT:H**



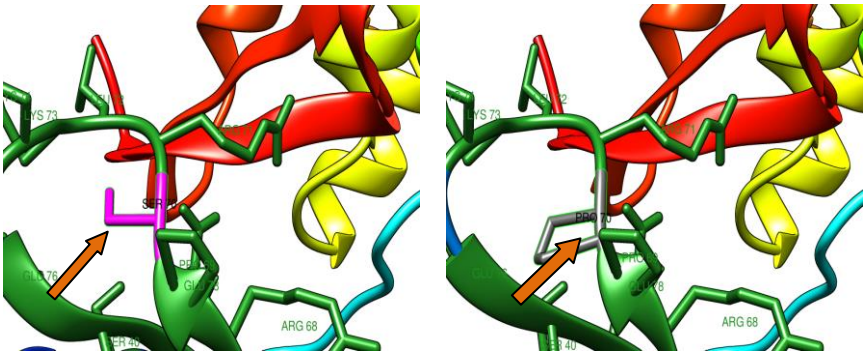
**rs62642940, P:314, WT:P, NT:H**




**rs62644467, P:413, WT:R, NT:C**

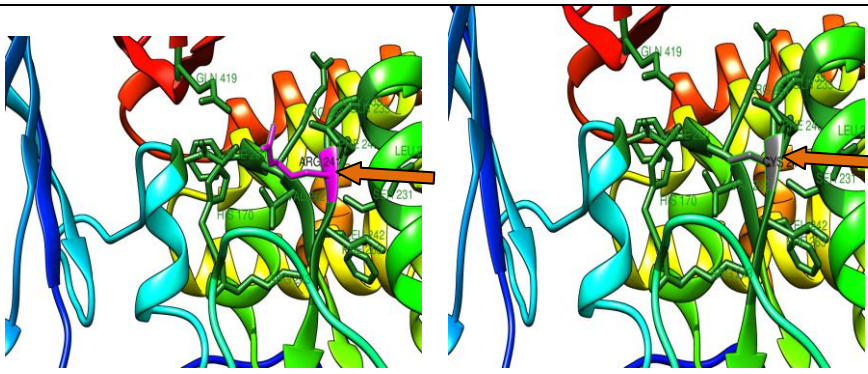


**rs62644471, P:417, WT:Y, NT:H**

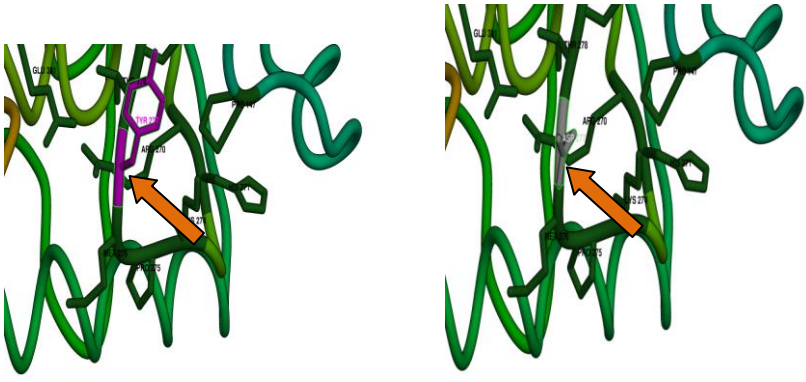


**rs63048261, P:70, WT:S, NT:P**

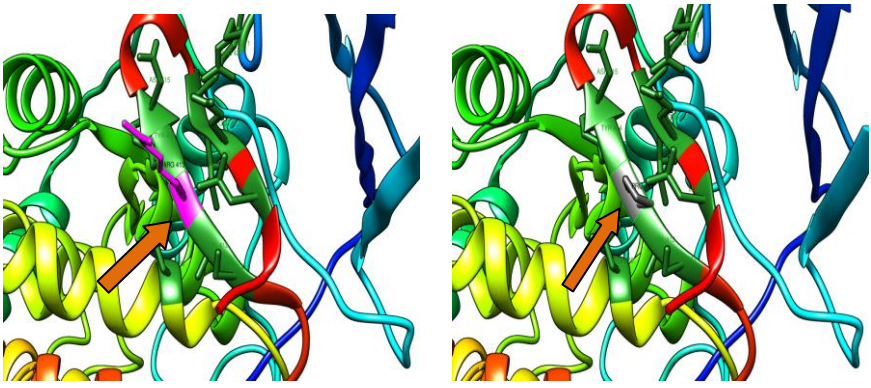
**rs76394784, P:68, WT:R, NT:S**



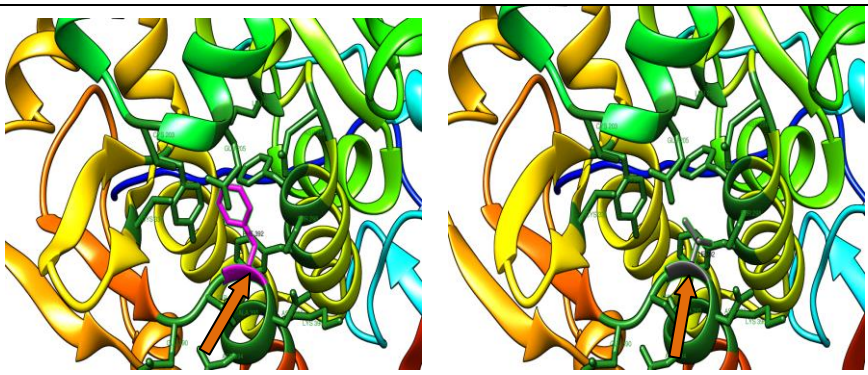
**rs76687508, P:241, WT:R, NT:C**



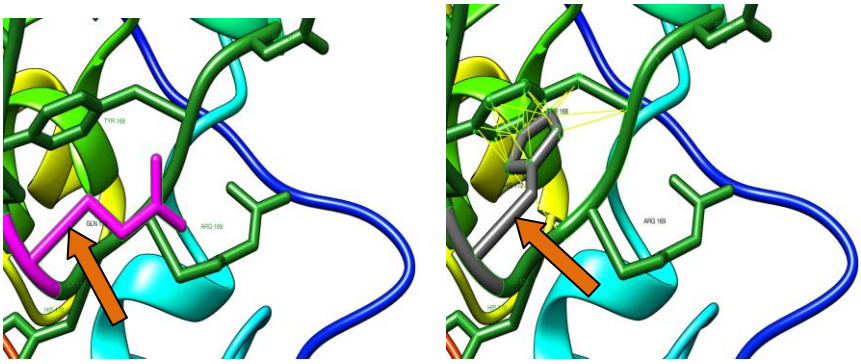
**rs78655458, P:277, WT:Y, NT:D**



**rs79931499, P:413, WT:R, NT:P**



**rs180819807, P: 392, WT:F, NT:I**



**rs192592111, P:172, WT:Q, NT:H**

**Figure 2. Shows position of native (left column) and new protein (right column) residues using Chimera program v1.8. rs: SNP ID , p: protein position, WT: wide type amino acid, NT: new type amino acid.**

*Functional SNPs in 3`UTR predicted by PolymiRTS data base 3.0*

Polymorphisms in miRNA-binding sites not only offer a possibility of novel diagnostic and prognostic markers in PKU disease, but can also help to understand complex regulatory networks of miRNAs in health and disease. Through total 29 SNPs in 3'UTR region of *PAH* gene, six functional SNPs predicted that effected miRNA binding Sites. Table 3.

**Table 3. Polymorphisms in microRNAs target Sites**

| db SNP ID | miR ID | Conservation | miR site | Functional class |
|---|---|---|---|---|
| rs185081299 | hsa_miR_634 | 3 | gcttgGCTGGTAg | C |
| | hsa_miR_1243 | 2 | TCCAGTAtcattt | D |
| | hsa_miR_3654 | 2 | TCCAGTAtcattt | D |
| rs180991851 | hsa_miR_3117_5p | 4 | tccAGTGTCAttt | C |
| | hsa_miR_425_5p | 4 | tccaGTGTCATtt | C |
| | hsa_miR_6757_3p | 2 | tCCAGTGTcattt | C |
| rs141177123 | hsa_miR_1243 | 2 | ctTCCAGTAtcat | D |
| | hsa_miR_3654 | 2 | ctTCCAGTAtcat | D |
| rs184297770 | hsa_miR_3673 | 3 | ATTCCAAttacta | D |
| | hsa_miR_6505_5p | 3 | ATTCCAAttacta | D |
| | hsa_miR_582_3p | 2 | attCCAGTTActa | C |
| | hsa_miR_6128 | 3 | ATTCCAGttacta | C |
| rs189657033 | hsa_miR_5003_3p | 3 | gctAAAAGTAaaa | C |
| Rs1042517 | hsa_miR_4477_b | 7 | aacagCCTTAAAt | D |
| | hsa_miR_7856_5p | 9 | aacagCCTTAAAt | D |
| | hsa_miR_320a | 3 | aaCAGCTTTAaat | C |
| | hsa_miR_320b | 3 | aaCAGCTTTAaat | C |
| | hsa_miR_320c | 3 | aaCAGCTTTAaat | C |
| | hsa_miR_320d | 3 | aaCAGCTTTAaat | C |
| | hsa_miR_4429 | 3 | aaCAGCTTTAaat | C |
| | hsa_miR_9_3p | 6 | aacAGCTTTAAat | C |

miR ID : Link to miRBase, Conservation : Occurrence of the miRNA site in other vertebrate genomes in addition to the query genome, miRSite: Sequence context of the miRNA site. Bases complementary to the seed region are in capital letters and SNPs are highlighted in red, FuncClass : D: The derived allele disrupts a conserved miRNA site (ancestral allele with support >= 2), C: The derived allele creates a new miRNA site.

## Conclusion

Presently, most molecular studies are concentrating on SNPs located in coding, regulatory regions and mirRNA; yet several of these studies have been unable to identify substantial associations between SNPs and disease vulnerability. The assumption is that amino acids conserved across species are more expected to be functionally significant. Therefore, SNPs that alter these amino acids might more probably be related

with disease vulnerability. It has been documented that use of the molecular evolutionary approach may be a potent tool for prioritizing SNPs to be genotyped in upcoming molecular epidemiological studies. Therefore, our analysis will provide useful information in selecting SNPs of PAH gene that are likely to have potential functional impact.

## REFERENCES

1. Agnieszka DK. 2014. MicroRNA-binding site polymorphisms in hematological malignancies. Journal of Hematology & Oncology. 7: 2-6.
2. Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. Cell. 136:215–233.
3. Charles R and Scriver. 2007. The PAH Gene, Phenylketonuria, and a Paradigm Shift. HUMAN MUTATION. 28:831-845.
4. Capriotti EP., Fariselli and R. Casadio. 2005. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research. 33: 306–310.
5. Eric, WS., Tanya B, Dennis AB et al. 2011. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 39: 38–51.
6. Hassan MM., Dowd AA, Faisal I et.al. 2014. In Silico Analysis of Single Nucleotide Polymorphisms (SNPs) in Human HLA-A and HLA-B Genes Responsible for Renal Transplantation Rejection. EUROPEAN ACADEMIC RESEARCH .Vol. II.
7. Lee JE., Choi JH, Lee JH, and MG Lee. 2005. Gene SNPs and mutations in clinical genetic testing: haplotype-based testing and analysis. Mutat Res. 573:195–204.
8. Nielsen M., Lundegaard C, Lund O and TN Petersen. 2010. CPH models 3.2. remote homology modeling using

structure-guided sequence profiles. Nucleic Acids Research. 38:1093.

9. Ng PC and S Henikoff. 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Research. Vol. 31, No. 13.
10. Pietro S and D Concolino. 2014. New Strategies for the Treatment of Phenylketonuria (PKU). Metabolites. 4:1007-1017.
11. Rajith B and GP Doss C. 2011. Path to Facilitate the Prediction of Functional Amino Acid Substitutions in Red Blood Cell Disorders – AComputational Approach. doi: 10.1371/journal.pone. 0024607.
12. Scriver CR., Beaudet AL, Sly WS et al. 1995. The Metabolic and Molecular Bases of Inherited Disease, 7th ed. New York: McGraw-Hill, pp 1015-1075.
13. T. Xi., Jones IM and HW Mohrenweiser. 2004. Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. Genomics. 83: 970–979.
14. Williams RA., Mamotte CD and JR Burnett. 2008. Phenylketonuria: An Inborn Error of Phenylalanine Metabolism. Clin Biochem. 29 :I -31.