Acta Scientific Veterinary Sciences (ISSN: 2582-3183)

Volume 4 Issue 9 September 2022

Review Article

Bioresearch Highlights on Single Nucleotide Polymorphisms (SNPs), and their Applications

Bhawanpreet Kaur[1] and CS Mukhopadhyay[2]*

[1]Ph.D. Scholar, College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana, Punjab, India

[2]Senior Scientist, Department of Bioinformatics, College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana, Punjab, India

***Corresponding Author:** CS Mukhopadhyay, Senior Scientist, Department of Bioinformatics, College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana, Punjab, India.

DOI: 10.31080/ASVS.2022.04.0504

Received: August 05, 2022

Published: August 29, 2022

© All rights are reserved by **Bhawanpreet Kaur and CS Mukhopadhyay.**

Abstract

Single nucleotide polymorphisms (SNPs) are an important type of genetic variation that occurs due to differences in one nucleotide between individuals. SNPs are distributed throughout the genome and can occur in both coding as well as non-coding regions. Different SNPs databases and tools for predicting protein-altering variants are available with huge information. For detecting the synonymous and nonsynonmous SNPs for genome-wide studies online and offline tools are accessible. In bioinformatics, SNPs are crucial to understanding the molecular mechanisms of sequence evolution, association with human diseases, and variation between individuals in response to drugs and can be used to make personalized medicines. Therefore, in this systematic review, we also discuss the vital role of SNPs in genetic mapping, pharmacogenomics, gene discovery, pharmacogenetics, diseases, evolution, nutrigenetics, nutrigenomics, and strong biological markers.

Keywords: Single Nucleotide Polymorphism; Genome; Markers; Databases; Mutation

Introduction

Single nucleotide polymorphism (SNP) is the most common form of polymorphism. The variations include insertions and deletions called INDELS, microsatellites, copy number variants (CNV), and epigenetic markers are vital to consider and can impact disease. Usually, SNPs are present in the promoter's region, exons, and introns as well as untranslated regions (UTRs) (5'- and 3'). Their presence affects the expression of genes by different mechanisms which depend on the role of the genetic elements in which the individual SNPs are located. In Humans and Eukaryotic genomes, SNP is the single nucleotide polymorphism occurring between DNA samples concerning only one base. Therefore, even if there is a polymorphism identified when two homoinsertion/deletions of a single base, those changes cannot be recognized as SNPs. DNA fragments have the same sequence except for a specific point where it differs.

Single nucleotide polymorphisms (SNPs) and their types

Throughout the whole genome, SNPs occur at a frequency of 1 in 1,000 base pairs (bp). SNP markers allow to automatize and maximize as ten folds the efficiency of genotype analysis. SNPs are excellent markers for studying complex genetic traits analysis and understanding genome evolution. SNPs are usually bi-allelic, even though in principle any of the four (A, T, C, G) nucleotides can be present at any location in a stretch of DNA strand. This is due to the low mutation frequency that leads to new SNPs. According to

their location in the genome, SNPs can be classified as synonymous (silent) SNP and non-synonymous (missense) SNP (nsSNPs). SNPs are the most ubiquitous genetic variations in the human genome. SNPs occur in the coding region and non-coding regions of the genome. Approximately 50% of SNPs lie in the non-coding regions named as ncSNPs, 25% lie in the coding region or called cSNPs and 25% are the silent mutations also named synonymous SNPs (sSNP they do not change encoded amino acids) while on the other hand, nonsynonymous SNPs (nSNPs or nsSNP change-encoded amino acids). Figure 2 shows the types of SNPs. The low mutation rate and simplicity of SNPs also make them excellent markers for studying complex genetic traits and a tool for understanding genome evolution, analyzing genetic variations responsible for differential expression of several economically important traits, and susceptibility to complex diseases. Due to lower mutation rates, SNPs are highly abundant and more stable than STRs (Short tandem repeats). The basic genetic variations that occur at the genic level between two individuals can be tracked through single nucleotide polymorphisms (SNPs). However, genome-wide scanning of SNPs and genotyping of all the existing SNPs in experimental samples is time and resource-intensive.

Single Nucleotide Polymorphism

Polymorphism is defined as the occurrence of two or more forms of a gene (allelomorphs) so that the least frequent allele should have a frequency of 1% and that cannot be maintained by recurrent mutation. Point mutation is the main contributor to the existence of SNPs [4].

Nucleotide substitutions can take place both in coding and non-coding regions of the nuclear/mitochondrial/plastid DNA. When there is a base change in the coding region of the DNA strand, it may either modify the encoded amino acid which is also called a non-synonymous mutation, or the resulting amino acid may not change due to degeneracy of codon. The latter is called a synonymous mutation. Some of the salient features of the mono-nucleotide variation or SNPs are, that SNPs are assumed to be biallelic, and the distribution patterns of the SNPs are heterogeneous over the genome. Besides, SNPs are more frequent in the non-coding region and introns. The biological implications of the SNPs vary with the position in the genome. If the SNP is located in the coding region, it may result in non-sense or missense mutation or even no immediate impact on the encoded amino acid. However, if the SNP is located in the non-coding region, it may affect the gene splicing, binding of the transcription factors, or sequence and activity of encoded non-coding RNA (like, miRNA).

Example of SNP

Sickle cell anemia is a recessive autosomal disease, which is completely genetic. This disease is prevalent in humans of the sub-Saharan region. A single SNP leads to alteration of the sixth amino acid (the "Methionine" at the beginning is being omitted from the counting) of the hemoglobin B (HB-B) chain. A transversion of "Adenine" to "Thymine" implies mutation of the codon "GAG" (encoding hydrophilic Glutamate) to "G**T**G" (encoding hydrophobic Valine). The detail has been shown below in figure 1.

```
>HB A: ATG GTG CAT CTG ACT CCT GAG GAG AAG
>HB S: ATG GTG CAT CTG ACT CCT GTG GAG AAG
 Match:*** *** *** *** *** *** *** *** ***
        Met  Val  His Leu  Thr  Pro  Glu Glu  Lys
                                      to
                                     Val
Position: 0    1    2   3   4    5   6    7   8
```

Figure 1: Sickle cell anemia showing an example of SNP.

Multinucleotide polymorphism (MNP)

In contrary to an SNP, a multi-nucleotide polymorphism (MNP) refers to the substitution of multiple bases one after another. The number of such contiguous bases may vary from two to five or even more. Example: CCC/GGG. MNPs could be bi-, tri, or tetra-allelic polymorphisms. MNPs are comparatively rare in the eukaryotic genome, as compared to SNPs.

Characteristics of SNPs

- SNPs are bi-allelic (only two alleles).
- In genomic regions, SNPs are not homogenous but rather inconsistently heterogeneous.
- Point mutation is the major contributing factor to the creation of SNP.
- SNPs are more frequently found within the non-coding regions as compared to the exons, (Barreiro., *et al.* 2008).

Applications of SNPs in GWAS

Genome-wide association study (GWAS) has gained popularity since 2002, with the research work on susceptibility to myocardial infarction, reported by [37]. In subsequent years, Klein and co-workers investigated age-related macular degeneration (AMD) patients. This ground-breaking research could identify two SNPs significantly associated with the disease. Thereafter hundreds of GWAS work has been reported in animals [16] and humans [44,27]. GWAS refers to the scrutinizing of detectable single-nucleotide

Types	Functions
cSNP (coding SNP)	SNP present within the coding region of a gene
ncSNP (noncoding SNP)	SNP present within a noncoding region
eSNP(expression SNP)	SNP that affects the gene expression
rSNP (regulatory SNP)	SNP present in a gene regulatory region
sSNP (synonymous or "silent" SNP)	SNP lies within a protein, codon that does not result in an amino acid change
nSNP, nsSNP (nonsynonymous SNP)	SNP present inside a protein, codon that results in an amino acid change
srSNP (structural regulatory SNP)	SNP that affects the mRNA gene products

Table 1: Different types of SNPs and their functions.

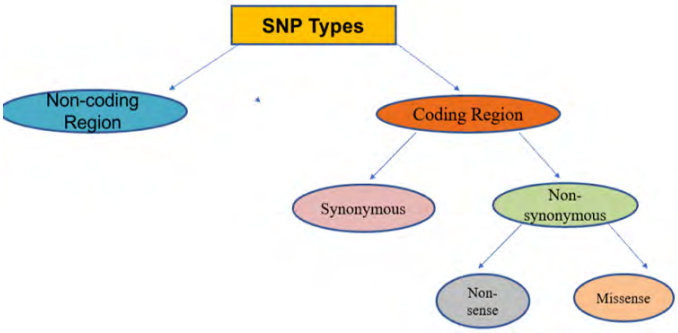

Figure 2: Types of SNP.

polymorphisms among different groups of individuals who are being screened to identify which SNP-variant(s) is/are associated with the trait of interest. Thus, the DNA profiles of individuals belonging to two distinct groups (viz. diseased vs. Healthy) are compared to statistically determine the causal loci contributing to the varying phenotype. SNPs are identified after subjecting the DNA samples of the individuals of those two contrasting groups to microarray analysis or genotyping by sequencing (or ddRAD) or whole genome sequencing study (DNAseq to identify SNPs) for detecting which more prevalent SNPs (hence, higher frequency) in any one group. It is highly recommended to all countries to maintain a DNA bank for maintaining and conserving rare repositories of human and animal genomes for future initiatives to identify causal genetic variants or SNPs.

Thus SNP databases (of different species) are recommended to be maintained and regularly updated to allow researchers to identify the causal genetic contributors to diseases or production variation (in animals), map economically or clinically important traits to the genome vis-a-vis identify QTL underlying many common, complex diseases.

Cis- and trans-acting QTLs

It is imperative to discuss a bit about the quantitative trait loci (QTLs) that are considered of much importance in animal breeding experiments. In the genome-wide association studies, the mRNA abundance is taken as a trait. The term expression-QTLs (eQTLs) has now been incorporated into regular molecular breeding experiments scientists associate the trait of interest with a region on the genome. eQTLs are positioned with the transcript in any one of the possible two different ways

- **Cis-acting eQTL**: the eQTL is positioned in such a way that it overlaps the location of the affected gene (i.e., transcript) due to sequence variation in the regulatory region of the gene. The cis-eQTL can be used in the candidate gene approach for the validation of quantitative trait genes (QTG).
- **Trans-acting eQTL**: alternatively, when more than one gene (or transcript) is mapped to the same SNP (or eQTL), it is called trans-acting eQTL or "eQTL hotspots". It results from the abundance of transcription factor(s) resulting in enhanced gene expression [12].

SNP databases

A huge number of databases are available that give information about SNPs and their characteristics. It includes a single nucleotide polymorphism database (dbSNP) [41], SNPSTR a database for Single nucleotide polymorphism (SNP) Short tandem repeats or mic-

rosatellites [1], PicSNP, a database of non-synonymous mutation in human genomes [8] HapMap [45], UCSC Browser [25], SNP2NMD a Database of human SNPs causing nonsense-mediated mRNA decay (NMD) [20]. Online Mendelian Inheritance in Man (OMIM) database [19], Human Genome Variation database (HGVbase) [14]. Table 2 shows the important databases and their links.

SNP detection tools

There are enlisted several tools as shown in table 3 for SNP detection which works in offline and online mode. Polyphred, novoSNP, SNPMap, SNP chip, and SNP harverster follow the offline mode for SNP detection and on another side, SNPselector, QuickSNP, SNP vista, PupaSuite, SNPbox, SNPsFinder are the online tools.

Database	Features	Link/ References
dbSNP	comprehensive repository for SNPs substitutions, short deletion, and insertion.	http://www.ncbi.nlm.nih.gov/snp
SNPSTR	A database of compound microsatellite-SNP markers in human, dog, mouse, rat, and chicken	http://www.sbg.bio.ic.ac.uk~inocgi-bin/SNPSTRdatabase.html- SNPSTR database
PicSNP	Catalog of non-synonymous SNP in the human genome	http://plaza.umin.ac.jp/~hchang/picsnp/
HapMap	The HapMap Consortium in which >1.1 million SNPs were genotypedhich	http://www.hapmap.org/
UCSC Browser	University of California, Santa Cruz	http://genome.ucsc.edu/cgi-bin/hgTables
SNP2NMD	Database of human SNPs causing nonsense-mediated mRNA decay (NMD)	http://variome.kobic.re.kr/SNP2NMD/
OMIM	Powerful, comprehensive, and widely used database.	http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim
HGVbase	High-quality, and nonredundant database. provides both neutral polymorphisms and disease-related mutations.	http://hgvbase.cgb.ki.se
GVS	dbSNP and HapMap SNPs access	http://gvs.gs.washington.edu/GVS
Human Cytochrome P450 (CYP) Allele Nomenclature Committee	Human cytochrome P450 genes	http://www.cypalleles.ki.se
Cytokine Gene Polymorphism	Gene polymorphisms in human disease	http://www.nanea.dk/cytokinesnps/
SNP500Cancer	Genes associated in cancer	http://snp500cancer.nci.nih.gov/home_1.cfm
PharmGKB	Genes associated in drug metabolism	http://www.pharmgkb.org
Perlegen Genotype Data	Whole genome SNP in three populations	http://genome.perlegen.com/

Table 2: SNP databases and their web addresses.

Name of tool	Mode of action	Link or Reference
Polyphred	Offline mode	http://droog.gs.washington.edu/polyphredpoly_get.htm
novoSNP	Offline mode	http://www.molgen.ua.ac.be/bioinfo/novosnp/download.html?email
SNPselector	Online mode	http://snpselector.duhs.duke.edu/hqsnp36.html
QuickSNP	Online mode	http://bioinformoodics.jhmi.edu/quickSNP.pl
SNP vista	Online mode	http://hazelton.lbl.gov/~teplitski/dtree/
PupaSuite	Online mode	http://bioinfo.cipf.es/pupasuite/www/index.jsp
SNPbox	Online mode	http://www.snpbox.org/
SNPsFinder	Online mode	http://snpsfinder.lanl.gov/
SNPMap	Offline	http://cran.r-project.org/
SNP chip	Offline	http://bioconductor.org/packages/2.6/bioc/html/SNPchip.htm
DataBins	Online	http://www.mrl.ubc.ca/who/who_bios_scott_tebbutt.shtml
SNP harverster	Offline	http://bioinformatics.ust.hk/SNPHarvester.html

Table 3: SNP detection tools and their web addresses for genome-wide studies.

SNPs Tools for predicting protein-altering variants

Many tools have been developed to prioritize a given amino acid substitution and many analyses have been applied to understanding the effects of nsSNPs and mutations that are not included in the tools below. These tools are all supervised, that is, they use a training set of positive and negative examples to "learn" sites. The tools are named Polymorphism Phenotyping or synonymous with PolyPhen. It includes the protein structure [39]. PhD-SNP [7]. They usually use features based on sequence, structure, or known function. The initially published method is named Sorting Intolerant from Tolerant or SIFT, It uses the multiple sequence alignment (MSA) [34], LS-SNP [24], Predicting the Amino Acid Replacement Probability or Parepro, [46], and Protein Analysis Through Evolutionary Relationships or PANTHER [32], nsSNPAnalyzer is the non-synonymous single nucleotide polymorphism (nsSNP) is phenotypically neutral or disease-associated [3]. Mupro [6,9] CanPredict is a computational tool for predicting cancer-associated missense mutations [22], PROVEAN analyzes the consequence of sequence variation on the function. It indicates that the amino acid variant has a damaging effect or neutral effect on the protein [10]. Table 4 describes the description of the tools and their online available websites.

Application of SNPs

1. SNPs in Genetic Mapping

SNPs play a vital role in genetic mapping for the improvement of plants genetically. SNPs act as powerful markers for generating new maps. Linkage maps based on SNPs have been constructed in different species like *Brassica* [29] and rice [49], etc.

SNPs in pharmacogenomics

In pharmacology studies, it plays an important role in two ways one is the gene approach and the other is LD or Linkage disequilibrium mapping. The prior knowledge of disease pathogenesis to identify genes relevant to disease. SNPs found in the genes are tested for statistical association with disease in patients enrolled. The occurrence of diseases in the family, case-control, or cohort studies. These "susceptibility genes" are hypothesized to directly influence an individual's likelihood of developing the disease and the other concept is LD mapping where the SNPs, most are located in the vast non-coding DNA regions between genes and play no obvious role in drug response. Linkage disequilibrium (LD) acts as the anonymous marker that can be used to identify a region of the genome that may harbor a susceptibility gene.

SNPs in pharmacogenetics

In the future perspective, the major challenge is to understand the genetic variance response to medicine, and interaction with environmental factors. To replace the error-based selection of medicine in the future, it helps clinicians genetically profile individual patients according to their genetic makeup. But the studies do not unambiguously prove the clinical value of pharmacogenetic testing

Name of tools	Description	Link / Reference
PolyPhen	Generates MSA, multiple attribute classifier	http://genetics.bwh.harvard.edu/pph/
PhD-SNP	Multiple attribute SVM classification support vector machines (SVMs) which predict whether a point mutation is a neutral polymorphism	http://gpcr2.biocomp.unibo.it/cgi/predictors/PhD-SNP/PhD-SNP.cgi
SIFT	SIFT predicts an amino acid substitution affects protein function	http://blocks.fhcrc.org/sift/SIFT.html
LS-SNP	1. It isthe large-scale annotation of coding non-synonymous SNPs	http://modbase.compbio.ucsf.edu/LS-SNP/
Parepro	method of identifying which non-synonymous single base	http://www.mobioinfor.cn/parepro/
PANTHER	Database for the evolution of a system.	http://www.pantherdb.org/tools/csnpScoreForm.jsp
nsSNPAnalyzer	Generates MSA, classifier, accessibility,	http://snpanalyzer.utmem.edu/
Mupro	Multiple attribute SVM classification	http://www.ics.uci.edu/~baldig/mutation.html
I-mutant	Protein stability changes	http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant2.0/I-Mutant2.0.cgi
CanPredict	Cancer mutation classification	http://www.cgl.ucsf.edu/Research/genentech/canpredict/
PROVEAN	analyzes sequence variation	http://provean.jcvi.org/

Table 4: Different SNP for predicting protein-altering variants.

[13]. Figure 3 shows the role of polymorphisms in pharmacodyamics and pharmacogenetics.

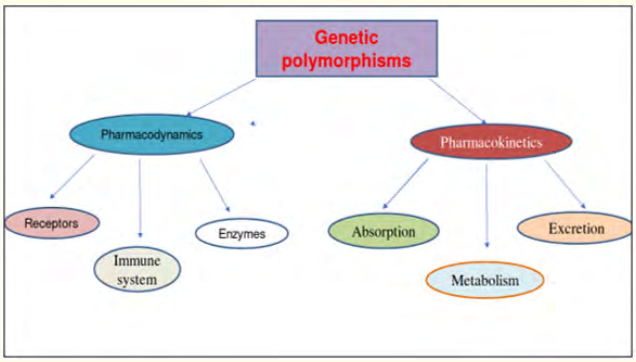

Figure 3: Role of Polymorphisms in pharmacodynamics and pharmacogenetics.

SNPs in gene discovery

SNPs are present frequently throughout the genome. By an association study (SNPs act as strong markers to identify disease-causing genes [18]. The two closely located alleles (gene and marker) are inherited together. Therefore, a simple comparison of patterns of genetic variations between patients and normal individuals may provide a method of identifying the loci responsible for disease susceptibility [21].

SNPs in diseases and genetic studies

Role of SNPs in diseases, they act as strong biomarkers as they are the genes associated with various complex diseases such as cardiac diseases, diabetes, cancer, schizophrenia, blood pressure, migraine, and Alzheimer's. These SNPs are mostly located within a gene or in a regulatory region near a gene and can affect the gene's function to play a more direct role in disease. It helps to under-

stand the molecular mechanisms of disorders by identifying new genetic variants and is key to developing new therapeutic strategies in the future. Below table 5 shows the various disorders, the genes associated, and the responsible SNPs.

Name of Diseases	Locus SNPs	Gene associated	References
Migraine	rs1835740 *rs1801133 and* *rs1801131* *rs10504*61	MTDH MTHFR MMP16	[47] [17] [40]
Breast Cancer	rs1045485 rs889312	*CASP8* *MAP3*K1	[11] [15]
Cervical cancer	rs1048943	*CYP1A1*	[22]
Alzheimer's Disease	rs121918399	APOE	[31]
Asthma	*rs11078927* *rs2786098*	*GSDMB* *CRB1*	[33] [42]
coronary heart disease	rs7025486	*DAB2IP*	[5]
Brain tumor	rs1136410	*PARP-1*	[26]
Atopic dermatitis	rs612529	VSTM1	[28]

Table 5: Enlisted the different diseases and responsible SNPs.

SNPs in evolution

Genetic evolution in part depends upon a balance between natural selection and environmentally driven mutation. By natural selection, deleterious mutations that affect the biological functions of proteins are effectively eliminated from the gene pool. SNPs are present at all levels of evolution, they can be used to study sequence variation among species.

1. SNPs in nutrigenetics and nutrigenomics

Several monogenic disorders like phenylketonuria are associated with the interaction between genes variant and nutrients. Several recent studies are based on the population and intervention of this gene support–nutrient interaction [43]. Nutrients play an important role in modulating gene expression. The diet-related disease risk is more efficiently common in some multifactorial disorders [36]. This can be reduced by better understanding the individual genetic make-up for the development of a personalized diet.

SNPs as biological markers

SNPs occur in coding as well as non-coding regions of the genome. Depending on the potential application, SNP markers might be selected outside of exons or the direct cause of a genetic mutation. It helps to map the history of populations by identifying the distribution of SNPs among present and past populations. SNP markers are used for studying the association between the markers and a particular trait or disease. Table 6 shows the SNPs as powerful markers used for the genotyping methods in recent association studies.

Genotyping method	Susceptibility gene	Disease	Genotyping scale	References
MassARRAY	β-Chemokine gene in 17q11	Multiple sclerosis (MS)	232 SNPs, 1369 subjects	[48]
MassEXTEND™	*NuMA* in 11q13	Breast cancer	> 25,000 SNPs, 522 subjects	[23]
TaqMan	*APOE*	Alzheimer's disease	60 SNPs, 220 subjects	[30]
TaqMan	*ELAVL4*	Parkinson's disease	9 SNPs, 1223 subjects	[35]
AS-PCR	*Il4r*	Rheumatoid arthritis	2 SNPs, 842 subjects	[38]
SBE-FRET/FP	*PPARγ*	Type 2 diabetes	2 SNPs, > 3000 subjects	[2]

Table 6: SNP markers used for Genotyping methods in recent association studies.

Conclusion and Future Directions

SNPs are the most abundant genetic polymorphisms and are strong molecular genetic markers in disease genetics studies and research. Two basic strategies of SNP genotyping detection are allele discrimination and allele detection. These are sequence alterations that occur more frequently in the human genome and a wide range of studies. They have been used as molecular markers for drug designing and development for individualized therapy and personalized medicine. The success of such studies can be greatly limited by technical aspects of SNP genotyping, specifically accuracy, cost-effectiveness, and throughput. Further, the genotyping technologies will transform the health care industry and contribute to the advancement of biology and medical science.

Bibliography

1. Agrafioti I., *et al.* "SNPSTR: a database of compound microsatellite-SNP markers". *Nucleic Acids Research* 35 (2017): D71-D75.

2. Altshuler D., *et al*. "The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes". *Nature Genetics* 26 (2000): 76-80.

3. Bao L., *et al.* "nsSNPAnalyzer: identifying disease-associated nonsynonymous single nucleotide polymorphisms". *Nucleic Acids Research* 33 (2005): W480-W482.

4. Barreiro., *et al*. "Natural selection has driven population differentiation in modern humans". *Nature Genetics* 40 (2000): 340-345.

5. Beaney KE., *et al. "*Clinical utility of a coronary heart disease risk prediction gene score in UK healthy middle-aged men and in the Pakistani population". *PLoS ONE* 10.7 (2015): e0130754.

6. Capriotti E., *et al.* "I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure". *Nucleic Acids Research* 33.2 (2005): W306-W310.

7. Capriotti E and Fariselli P. "PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants". *Nucleic Acids Research* 45.W1 (2017): W247-W252.

8. Chang H., *et al.* "PicSNP: a browsable catalog of nonsynonymous single nucleotide polymorphisms in the human genome". *Biochemical and Biophysical Research Communications* 287.1 (2001): 288-291.

9. Cheng J., *et al.* "Prediction of Protein Stability Changes for Single Site Mutations Using Support Vector Machines". *Proteins* 62.4 (2006): 1125-1132.

10. Choi Y., *et al.* "Predicting the functional effect of amino acid substitutions and indels". *PLoS One* 7.10 (2012): e46688.

11. Cox A., *et al.* "Common coding variant in CASP8 is associated with breast cancer risk". *Nature Genetics* 39 (2007): 352-358.

12. Drake T., *et a*l. "Integrating genetic and gene expression data: application to cardiovascular and metabolic traits in mice". *Mammalian Genome* 17 (2006): 466-479.

13. Evans WE and Johnson JA. "Pharmacogenomics: the inherited basis for interindividual differences in drug response". *Annual Review of Genomics and Human Genetics* 2 (2001): 9-39.

14. Fredman D., *et al.* "HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources". *Nucleic Acids Research* 30.1 (2002): 387-391.

15. Garcia-Closas M., *et al.* "Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics". *PLoS Genet* 4 (2008): 1000054.

16. George AW., *et al*. "Eagle for better genome-wide association mapping". *G3 (Bethesda)* 11.9 (2001): 204.

17. Goyette P., *et al.* "Human methylenetetrahydrofolatereductase isolation of cDNA, mapping and mutation identification". *Nature Genetics* 7 (1994): 195-200.

18. Gray IC., *et al.* "Single nucleotide polymorphism as tools in human genetics". *Human Molecular Genetics* 9 (2000): 2403-2408.

19. Hamosh A., *et al.* "Online Mendelian Inheritance in Man (OMIM), aknowledgebase of human genes and genetic disorders". *Nucleic Acids Research* 33 (2005): 514-D517.

20. Han., *et al.* "SNP2NMD: a database of human single nucleotide polymorphisms causing nonsense-mediated mRNA decay". *Bioinformatics* 23.3 (2007): 397-399.

21. Hirschhorn JN and Daly MJ. "Genome-wide association studies for common diseases and complex traits". *Nature Reviews Genetics* 6 (2005): 95-108.

22. Joshua S., *et al.* "CanPredict: a computational tool for predicting cancer-associated missense mutations". *Nucleic Acids Research* 35.2 (2007): W595-W598.

23. Kammerer S., *et al. "*Association of the NuMA region on chromosome 11q13 with breast cancer susceptibility". *Proceedings of the National Academy of Sciences of the United States of America* 102 (2005): 2004-2009.

24. Karchin R., *et al.* "LS-SNP: large-scale annotation of coding non-synonymous SNPs based on multiple information sources". *Bioinformatics* 21 (2005): 2814-2820.

25. Karolchik D., *et al*. "The UCSC Genome Browser Database". *Nucleic Acids Researc*h 31.1 (2003).

26. Khan AU., *et al*. "Modulation of brain tumor risk by genetic SNPs in PARP1gene: Hospital based case control study". *Plos one* 14.10 (2015): e0223882.

27. Kim SK., *et al*. "A genome-wide association study for shoulder impingement and rotator cuff disease". *Journal of Shoulder and Elbow Surgery* 30.9 (2021): 2134-2145.

28. Kumar D., *et al.* "A functional SNP associated with atopic dermatitis controls cell type specific methylation of the *VSTM1* gene locus". *Genome Medicine* 9 (2017): 18.

29. Li F., *et al.* "A *Brassica rapa* linkage map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits". *DNA Research* 16.6 (2009): 311-323.

30. Martin ER., et al. "SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease". *The American Journal of Human Genetics* 67 (2000): 383-394.

31. Masoodi TA., *et al.* "Screening and Evaluation of Deleterious SNPs in APOE Gene of Alzheimer's Disease". *Neurology Research International* 2012 (2012): 8.

32. Mi H., *et al.* "The PANTHER database of protein families, subfamilies, functions and pathways". *Nucleic Acids Research* 33 (2005): D284-D288.

33. Moffatt MF., *et al.* "A large-scale, consortium-based genome-wide association study of asthma". *The New England Journal of Medicine* 363.13 (2010): 1211-1221.

34. Ng PC and Henikoff S. "SIFT: predicting amino acid changes that affect protein function". *Nucleic Acids Research* 31 (2003): 3812-3814.

35. Noureddine MA., *et al.* "Association between the neuron-specific RNA-binding protein ELAVL4 and Parkinson disease". *Human Genetics* 117 (2005): 27-33.

36. Ordovas JM and Mooser V. "Nutrigenomics and nutrigenetics". *Current Opinion in Lipidology* 15 (2004): 101-108.

37. Ozaki K., *et al*. "Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction". *Nature Genetics* 32.4 (2002): 650-654.

38. Prots I., *et al.* "Association of the IL4R single-nucleotide polymorphism I50V with rapidly erosive rheumatoid arthritis". *Arthritis and Rheumatology* 54 (2006): 1491-1500.

39. Ramensky V., *et al*. "Human non-synonymous SNPs: server and survey". *Nucleic Acids Research* 30 (2002): 3894-3900.

40. Rozanov DV., *et al.* "The low-density lipoprotein receptor-related protein LRP is regulated by membrane type-1 matrix metalloproteinase (MT1-MMP) proteolysis in malignant cells". *Journal of Biological Chemistry* 279 (2004): 4260-4268.

41. Sherry ST., *et al*. "DbSNP: the NCBI database of genetic variation". *Nucleic Acids Research* 29.1 (2001): 308-311.

42. Sleiman PM., *et al*. "Variants of DENND1B associated with asthma in children". *The New England Journal of Medicine* 362.1 (2010): 36-44.

43. Subbiah MT. "Nutrigenetics and nutriceuticals: the next wave riding on personalized medicine". *Translational Research* 149 (2007): 55-61.

44. Tegelberg P., *et al. "*A genome-wide association study for shoulder impingement and rotator cuff disease". *PBMC Oral Health* 21.1 (2021): 611.

45. Thorisson GA., *et al.* "The International HapMap Project Web site". *Genome Research* 15.11 (2005): 1592-1593.

46. Tian J., *et al.* "Predicting the phenotypic effects of non-synonymous single nucleotide polymorphisms based on support vector machines". *BMC Bioinformatics* 8 (2007): 450.

47. Van den Maagdenberg AM., *et al.* "Migraine: gene mutations and functional consequences". *Current Opinion in Neurology* 20 (2007): 299-305.

48. Vyshkina T and Kalman B. "Haplotypes within genes of beta-chemokines in 17q11 are associated with multiple sclerosis: a second phase study". *Human Genetics* 118 (2005): 67-75.

49. Xie W., *et al.* "Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing". *Proceedings of the National Academy of Sciences of the United States of America* 107.23 (2010): 10578-10583.