

Magnitude of Human Genetic Variations: A Mini Review

Prem Chandra Suthar^{1,2*}, Pulakes Purkait^{1*}, Rakshit Ameta², B.N. Sarkar³

¹Anthropological Survey of India, Western Regional Centre, Udaipur, India

²Faculty of Science, Pacific Academy of Higher Education and Research University, Udaipur, India

³Anthropological Survey of India, Eastern Regional Centre, Kolkata, India

ABSTRACT

Genetic variations are the genetic differences found among individuals of the same species. These variations are adaptive i.e. they ensure flexibility and survival of a population in the changing environment over time. It is quite interesting to know how and when these genetic variations offer to a population the survival benefit under the altered situations and at what extent over time. Whether there is any significant difference among male and female in expression of such variations. Further if the population subgroups and demographical factors offer any selective advantage to these variations. Moreover we often find that one drug which is effective on an individual may not work well with other individual. This interindividual variation of drug effects is due to the genetic variation of drug metabolism. The variations of xenobiotic or drug metabolizing enzymes at genetic level bring about the evolution of diverse and specific subgroups in the population that offers a selective ability to perform the biotransformation reactions. With the advent of advance DNA sequencing technologies we are now able to find and characterize most of the natural variations such as single nucleotide polymorphisms, structural variations like insertion/deletion, inversion, translocations, tandem repeat polymorphism, complex chromosomal rearrangements etc. that are linked with human evolution, migration, disease susceptibility, adaptation to diverse geographical regions etc. This mini review focus on basic genetic variations and their possible associations with human health, forensics, environmental stress management, various diseases and other relevant aspects.

Keywords: Genetic variation, Adaptation, polymorphism, population, disease, structural variation, DNA sequencing.

INTRODUCTION

The source of natural genetic variations is mutation in the genes which cause altered expression or activity of these enzymes. The variant allele frequency is always relatively higher in the population. Many comparative studies have been undertaken on ethnically diverse human populations to find patterns and the evolutionary origins of genetic diversity and their implications for the mapping of complex traits, including disease susceptibility. Also to understand the genetic basis of adaptations that has evolved in response to diverse climates and diets, as well as exposure to infectious disease.^{1,2}

Mutation is the Source

DNA sequences vary at a large extent at different regions across the genome. And mutation is the ultimate source of these variations in all organisms. These variations in populations occur at individual level in DNA sequences or at gene level. The frequency with a gene changes from the wild type to a mutant type is known as mutation rate. Mutation rate vary considerably for different genes, different individuals of same species and different sex which can be determined in terms of frequency distribution. The incidence of a specific type of mutation within a group of individual organism, expressed as a percentage of the total size of study population is termed as mutation frequency. The average mutation rate was estimated to be $\sim 2.5 \times 10^{-8}$ mutations per nucleotide site or 175 mutations per diploid genome per generation.³ Mutation rates in humans vary across the genome from about 10^7 to 10^9 per nucleotide base per generation. These rates are roughly equivalent to 10^4 – 10^6 per gene per generation.⁴ There are 6×10^9 nucleotides in the human genome so every time human DNA is passed from one generation to the next it accumulates 100–200 new mutations, according to a DNA-sequencing analysis of the Y chromosome. This number the first direct measurement of the human mutation rate is equivalent to one mutation in every 30 million base pairs.⁵

The frequency of any mutation is affected by many factors such as frequency of primary changes in DNA, DNA stability, probability of repair, probability of recognition, physiological condition, the behavior of genetic elements⁶, reproductive fitness, human population history, chromosomal location and recombination rates. These factors determine the ultimate consequence of mutations among the populations. These factors influence in such a way that mutations become conserved at some regions of genome and transient at other locations. This is because some part of the genome is more susceptible to mutation than the other. The new variants may be beneficial if they increase the biotic potential of the population and thus selected through variability while the deleterious effects present challenges for it. Demographic factors affect the pattern of this variability across the genome in a population. A favorable mutation on average has long period of persistence in a large population than in small population.

Large extent of mutations influencing a small portion of genome are ‘selectively neutral’ because they do not have any effect on the reproductive fitness. This may be due to the fact that only 1.5% of the genome has protein coding genes which may give rise to synonymous substitution i.e. no change in amino acid or nonsynonymous substitutions i.e. change in amino acid but have no effect on reproductive fitness. Some variants influence physiological, morphological and pathological variation in the human population and provide the genetic substrate for evolutionary change in response to physiological or environmental stress. These variants influence diverse traits in human at a variable extent.

When the inheritance of a trait is controlled by a single genetic locus with two alleles and the least common allele has a frequency of about 1% or greater, it is known as genetic polymorphism. Genetic polymorphisms have been studied extensively by many researchers as causal agent for various life threatening diseases, risk assessment marker, evolutionary relationships, drug metabolism, forensic detections, and for others. Some of the polymorphisms/ variations are briefly reviewed here.

Single Nucleotide Polymorphisms (SNPs)

Undoubtedly the most common type of mutation is single nucleotide substitution. Change of one purine base to another purine base is called a transition, whereas change of a purine to a pyrimidine base or vice versa is called a transversion. A large number of polymorphisms in human genomes have been identified and classified as single nucleotide polymorphisms (SNPs).^{7, 8,9,10} SNPs have the additional property of being common in the population as a ‘polymorphism’ i.e. has a population frequency of at least 1%. There are an estimated 3–10 million SNP variants in the human genome with a frequency higher than 1%. Most detected SNP loci show two common alleles with frequencies of at least 10–20%. SNPs that lie in noncoding regions are likely to be nonfunctional and selectively neutral, but there are important subgroups that can influence multifactorial diseases or traits. A selectively neutral SNP is predicted to occur at every 500bp for those with allele frequencies in the 1–10% range, every 1500bp for those in the 10–20% range, and every 3000bp for those in the 40–50% range. Characterizing the nature of gene variation and their association with particular diseases are the major goals of human genetics. 1000 Genome Projects has provided the location, allele frequency, and local haplotype structure of ~15 million SNPs. Use of SNPs in drug designing has offered wide applications in medical and health.^{9, 11,12,13,14}

Structural Variation (SV)

Structural variation is generally defined as a region of DNA approximately 1 kb and larger in size¹⁵ and can include microscopic and submicroscopic variants, which include duplications, insertions, inversions and balanced translocations or genomic deletions, commonly referred to as copy number variants (CNVs). These CNVs frequently overlap with segmental duplications and if present at >1% in a population a CNV may be referred to as copy number polymorphism (CNP). In a broad way SV incorporate variants as small as 50 bp in length including some with the potential of disrupting gene function or regulation and that can be detected with improved technologies and analysis strategies.^{16,17,18,19} Structural variants can comprise millions of nucleotides of heterogeneity within every genome, and are likely to make an important contribution to individual’s genetic makeup, human diversity and disease susceptibility.^{20,21,22} A number of genetic disorders have been linked with CNVs^{23,24} are also involved in environmental response like perception and immunity.^{25,26} A full range of structural variation can now be detected from NGS data through long read sequencing platforms such as PacBio and Oxford Nanopore.²⁷ The Database of Genomic Variants (DGV) has annotated

18.8% of the euchromatic human genome as CNVs but one study²⁸ suggest that the actual CNV content of the human genome will be less than 12%. The variable extent of CNV in the phenotypically normal human genome allows studying the association of demographic variation in diverse human populations. A new form of polymorphic insertions, templated sequence insertion polymorphism (TSIP) which is inserted in the genome through reverse transcription of an RNA intermediate may be important in revealing more about human ancestry, migrations, diversity and genetic disorders.²⁹

Inversions

Balanced rearrangements change the chromosomal gene order but do not remove or duplicate any of the DNA of the chromosomes. The two simple classes of balanced rearrangements are inversions and reciprocal translocations. An inversion is a rearrangement in which an internal segment of a chromosome has been broken twice, flipped 180 degrees, and rejoined. A reciprocal translocation is a rearrangement in which acentric fragments of two nonhomologous chromosomes trade places.³⁰ Small sequence inversions are not uncommon but most have no functional consequences unless they occur in a gene. Inversion of large chromosomal segment is less common but occasionally gives rise to disease. Polymorphic inversions have been characterized to uncover the functional variants that drive phenotypes and disease susceptibilities, among individuals for population studies, or tissue types for biomarker discovery.³¹

INDELS

Short insertions and deletions (INDELS) are the second most abundant form of human genetic variation, but relatively fewer studies have been conducted to identify these alternative forms of natural genetic variation thus there is a huge research gap which needs attention. Human populations are expected to collectively harbor at least 1.6–2.5 million INDEL polymorphisms. Over 415,000 unique INDELS identified in the genomes of 36 diverse humans using DNA sequencing traces that initially were generated for SNP discovery projects.³² Recently a set of 1.6 million indels have been identified from 179 individuals representing 3 diverse human populations. The rates of indel mutagenesis are highly heterogeneous, with 43-48% of indels occurring in only 4.03% of the genome classified as indel hotspots, while in the remaining 96% their prevalence is 16-times lower than that for SNPs.¹⁰ One of the most important INDELS is Alu insertion elements, a family of SINEs that are ancestrally derived from 7SL RNA gene by a process called ‘retroposition’ process.^{33,34} Alu insertion polymorphisms identify the patterns of human genetic diversity and history, race determination, gender identification, personal identification, paternity testing. They are generally located in non-coding regions and have an important application in phylogenetic analyses of human populations.^{35, 36} Compared to SNPs, rates of indel mutagenesis are known to be variable across the genome,³⁷ with high rates in microsatellites.³⁸ Polymerase slippage across the genome is thought to be responsible for the large majority of indel mutations¹⁰ There may be general hotspots of natural genetic variation where both SNPs and INDELS make their effect. INDELS have unusual levels of genetic variations in the population i.e. at some regions alleles are maintained under selective pressure in response to the variable environment whereas at some other regions there is no functional selection. Deletion of short repetitive sequences occurs in and between genes but deletion of large genomic segments containing many genes occurs in relatively gene-poor regions of the genome. High-throughput sequencing investigation has a scope of the incidence of INDELS in coding regions of human genes to identify INDELS associated to diseases for a better clinical outcome.³⁹

Segmental Duplications

Segmental Duplications include both relatively large DNA segments of 10–400kb that are duplicated a few times throughout the genome, and a large class of short (1–5 bp), dispersed and tandemly repeated segments. Segmental duplications have originated from the duplicative transpositions of small portions of chromosomal material, often containing intron-exon structure of known genes, and tend to be localized in pericentromeric and subtelomeric regions.^{40,41} The enzymes that copy DNA sometimes slip extra copies of a gene into a chromosome, and scientists estimate that such genetic replicas make up about 5% of the human genome.⁴² They share a high degree of sequence identity (>90%) with duplicated region. Tandemly repeated segments show polymorphic variation due to a relatively high rate of insertion or deletion, with different individuals showing different numbers of short tandem repeat (STR) sequences. This class of repetitive sequence is important in human

genetics.⁴ Both duplications and deletions have been implicated in shaping human characteristics, such as bigger brains. An explosion of DNA duplication in the common ancestor of humans, chimpanzees and gorillas may be responsible for many of the differences among the species.⁴³ Duplications supply additional genetic material capable of evolving new functions and segmental duplications have experienced extraordinary rates of evolutionary turnover, which result in considerable structural change and rapid gene innovation in the genomes of man and great apes.^{41,44} Bigger, better human brains may be the result of a double dose of a gene that helps brain cells move around. Extra copies of SRGAP2 gene may account for humans' thicker brain cortex, the brain's gray matter where thinking takes place; a study suggests (Dennis et al 2012).⁴⁵ Human whole-genome duplications (WGD) are more divergent and entails the adaptation of WGDs to novel and important functions that consequently lead to their evolutionary conservation in the course of evolution.⁴⁶

Tandem Repeat Polymorphisms

Tandem repeat are a very common class of polymorphism in which copies of DNA sequences lie adjacent to each other in the same orientation (direct tandem repeats) or in the opposite direction to each other (inverted tandem repeats). Tandem repeats have been linked many hereditary disorders in humans, including fragile X syndrome, myotonic dystrophy, Huntington's disease, spinal and bulbar muscular atrophy and Friedreich's ataxia⁴⁷ and play an important role in cancer biology⁴⁸. Advance tools have been designed to search and analyze them like Tandem repeat finder,⁴⁹ tandem repeat meta-search engine, Treads,⁵⁰ ProGeRF: Proteome and Genome Repeat Finder,⁵¹ pSTR Finder.⁵²

DNA Repeat Expansion

This is an increase in number of repeats of a genomic, tandemly repeated DNA sequence from one generation to the next.

Trinucleotide Repeat Expansion- An increased number of contiguous trinucleotide repeats in the DNA sequence from one generation to the next. The presence of these regions is associated with diseases such as Fragile X syndrome and Myotonic dystrophy. Some chromosome fragile sites are composed of sequences where trinucleotide repeat expansion occurs.

Satellite DNA

Highly repetitive DNA sequences found in heterochromatin, mainly near centromeres. They are composed of simple sequences (very short) repeated in tandem many times to form large blocks of sequence. Additionally, following the accumulation of mutations, these blocks of repeats have been repeated in tandem themselves. The degree of repetition is on the order of 1000 to 10 million at each locus. Loci are few, usually one or two per chromosome. They were called satellites since in density gradients, they often sediment as distinct, satellite bands separate from the bulk of genomic DNA owing to a distinct base composition.

Inverted Repeat Sequences

Inverted repeat sequences are copies of nucleic acid sequence that are arranged in opposing orientation. They may lie adjacent to each other (tandem) or be separated by some sequence that is not part of the repeat (hyphenated). They may be true palindromic repeats, i.e. read the same backwards as forward, or complementary which reads as the base complement in the opposite orientation. Complementary inverted repeats have the potential to form hairpin loop or stem-loop structures which results in cruciform structures when the complementary inverted repeats occur in double stranded regions.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - These are repetitive nucleic acid sequences that are principal components of the archaeal and bacterial CRISPR-CAS systems, which function as adaptive antiviral defense systems.

Microsatellite Repeats

A variety of simple repeat sequences that are distributed throughout the genome provides the necessary information to measure the genetic distance between the marker and its associated trait, as is required for genetic mapping by linkage analysis. They are also known as short tandem repeats (STRs). Sometimes referred to as simple sequence repeats (SSRs), are accordion-like stretches of DNA containing a short repeat unit of 2-8 basepairs that is tandemly repeated up to 100 times.⁵³

Although the human genome contains thousands upon thousands of STR markers, only a small core set of loci have been selected for use in forensic DNA and human identity testing.⁵⁴ Current forensic DNA analysis predominantly involves identification of human donors by analysis of short tandem repeats (STRs) using Capillary Electrophoresis (CE). Recent developments in Massively Parallel Sequencing (MPS) technologies offer new possibilities in analysis of STRs since they might overcome some of the limitations of CE analysis.⁵⁵ Expression STRs (eSTRs) are enriched in conserved regions, colocalize with regulatory elements and may modulate certain histone modifications. STRs significantly contribute to the genetic architecture of quantitative human traits and various clinically relevant conditions.⁵⁶

Dinucleotide Repeats- The most common of the microsatellite tandem repeats dispersed in the euchromatic arms of chromosomes. They consist of two nucleotides repeated in tandem; guanine and thymine, (GT)_n, or (CA)_n is the most frequently seen which is present at an estimated 50 000–100 000 sites per genome. The (CA)_n repeats occur on average every 30–60 kb, of which about 8000 have been identified and are listed in human genome databases.⁴

Trinucleotide Repeats- Microsatellite repeats consisting of three nucleotides dispersed in the euchromatic arms of chromosomes. A few STRs are neither functionally nor selectively neutral because they occur in functionally important regions of the human genome and give rise to human disease. These trinucleotide repeat diseases generate abnormally large alleles that interfere with the expression or function of associated genes.

Minisatellite Repeats

Tandem arrays of moderately repetitive, short (10-60 bases) DNA sequences which are found dispersed throughout the genome, at the ends of chromosomes (telomeres), and clustered near telomeres. Their degree of repetition is two to several hundred at each locus. Loci number in the thousands but each locus shows a distinctive repeat unit. Also known as VNTRs are highly informative markers owing to the very large number of alleles in the population and particularly important in forensic work.

Complex Chromosomal Rearrangements

Complex chromosomal rearrangements (CCRs) are constitutional structural rearrangements involving three or more chromosomes or having more than two breakpoints. CCRs preferentially occur during spermatogenesis and are transmitted in families through oogenesis and have important role in fertility consequences. Complex genomic rearrangements are thought to be derived from a single catastrophic event rather than by several incremental steps^{57, 58, 59} Holland and Cleveland⁶⁰ suggested the term chromoanagenesis or chromosomal rebirth to group these one-step catastrophic events together.⁵⁹ They have been studied for the differential transmission of CCRs in males and females, the incidence of pregnancy outcomes of CCR carriers, genetic counseling and prenatal diagnosis. Despite the increasing understanding of the mechanisms involved in their genesis, CCRs arise as unique, complex events for which the genetic and reproductive counseling of carriers remains a challenge.^{61, 62} They are rare in humans but their clinical identification is important. Chromothripsis, a new mechanism has been proposed to explain the presence of the CCRs in cancer cells and patients with mental retardation and/or congenital abnormalities, including phenotypically normal subjects as the CCR carriers can display various phenotypes.⁶³

Rare variants

Rare variants are those which occur in coding regions and give rise to a change in amino acid in the protein product and implicated in severe monogenic disorders. These coding SNPs (cSNPs) are capable of directly influencing gene expression and have potential relevance to morphological, physiological and pathological traits thereby influence reproductive fitness. Rare variants are found with a low frequency in population. The contribution of rare variants to genetically complex, multifactorial traits is likely to be significant. Various studies have been performed to identify trait- and disease-associated rare variants^{64, 65} and with the rise of sequencing technologies, it is now feasible to assess the role rare variants play in the genetic contribution to complex trait variation.⁶⁶

Future prospects

Natural genetic variations have presented before us the challenge to uncover the mysterious facts of the human genome at large. As the next generation sequencing technologies are becoming affordable

to scientists working even at small laboratories more input is being given for micro level details of the genomic regions responsible for the genetic variations. Some time I think that this task of associating the DNA to various diseases, adaptations, evolution, anatomy, physiology and on is so big that more and more skilled manpower with more advanced and reliable technologies are required to get the in-depth research to reveal the hidden truths. Although significant progress have been made in last decade through robotics, advanced programming and ease in accessibility of newer tools which has made the scientific community more confident but more mining of the genome is essential for untouched natural genomic variations.

REFERENCES

- [1] Campbell MC and Tishkoff SA. African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. *Annu Rev Genomics Hum Genet.* 2008; 9: 403–433.
- [2] Gomez F, Hirbo J, Tishkoff SA. Genetic Variation and Adaptation in Africa: Implications for Human Evolution and Disease. *Cold Spring Harb Perspect Biol.* 2014; 6(7): a008524.
- [3] Nachman MW and Crowell SL. Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics.* 2000; 156: 297–304.
- [4] Wright AF. Genetic Variation: Polymorphism and Mutation. *Encyclopedia of Life Sciences & 2005, John Wiley & Sons, Ltd., doi: 10.1038/npg.els.0005005*
- [5] Dolgin E. Human Mutation Rate Revealed. *Nature News,* 27 August 2009, doi:10.1038/news.2009.864.
- [6] Woodruff RC and Thompson JN Jr. Genetic Factors that Affect Rates of Spontaneous Mutation and Chromosome Aberrations In *Drosophila melanogaster*. *Cytogenet Cell Genet.* 1982; 33(1-2):152-159.
- [7] Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, et al. LINE-1 Retrotransposition Activity in Human Genomes. *Cell* 2010; 141:1159–1170.
- [8] Huang CR, Schneider AM, Lu Y, Niranjan T, Shen P, Robinson MA et al. 2010. Mobile Interspersed Repeats are Major Structural Variants in the Human Genome. *Cell* 141:1171–1182.
- [9] 1000 Genomes Project Consortium, et al. 2012. An Integrated Map of Genetic Variation from 1,092 Human Genomes. *Nature* 491:56–65.
- [10] Montgomery SB, Goode D, Kvikstad E, Albers CA, Zhang Z, Mu XJ *et al.* The Origin, Evolution, and Functional Impact of Short Insertion-Deletion Variants Identified in 179 Human Genomes. *Genome Res.* 2013; 23: 749–761.
- [11] Manolio TA, Brooks LD, Collins FSA. A HapMap Harvest of Insights into the Genetics of Common Disease. *J Clin Invest* 2008; 118:1590–1605.
- [12] Hardy J and Singleton A. Genomewide Association Studies and Human Disease. *N Engl J Med* 2009; 360: 1759–1768
- [13] Buchanan CC, Torstenson ES, Bush WS, Ritchie MD. A Comparison of Cataloged Variation between International Hapmap Consortium and 1000 Genomes Project Data. *J Am Med Inform Assoc* 2012; 19: 289–294.
- [14] Fareed M and Afzal M. Single Nucleotide Polymorphism in Genome-Wide Association of Human Population: A tool for broad spectrum service. *Egypt. J. Med. Hum. Genet.* 2013; 14, (2): 123–134.
- [15] Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM et al. Copy Number Variation: New Insights in Genome Diversity. *Genome Res.* 2006; 16(8): 949-961.

- [16] Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C et al. Mapping Copy Number Variation By Population-Scale Genome Sequencing. *Nature* 2011; 470:59-65.
- [17] Sudmant P, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An Integrated Map of Structural Variation in 2,504 Human Genomes. *Nature* 2015; 526:75-81.
- [18] Zarrei M, MacDonald JR, Merico D, Scherer SW (2015) A Copy Number Variation Map of The Human Genome. *Nature Review Genetics* 16:172-183.
- [19] Mak AC, Lai YY, Lam ET, Kwok TP, Leung AK, Poon A, et al. Genome-Wide Structural Variation Detection by Genome Mapping on Nanochannel Arrays. *Genetics*. 2016; 202(1):351-62.
- [20] Sachidanandam, R. Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A Map of Human Genome Sequence Variation Containing 1.42 Million Single Nucleotide Polymorphisms. *Nature* 2001; 409: 928-933.
- [21] The International HapMap Consortium. The International HapMap Project. *Nature* 426, 789796 (2003).
- [22] Feuk L, Carson AR, Scherer SW. Structural Variation in the Human Genome. *Nature Reviews Genetics* 2006; 7:85-97.
- [23] Koolen, D. A. Vissers LE, Pfundt R, de Leeuw N, Knight SJ, Regan R, et al. A New Chromosome17q21.31 Microdeletion Syndrome Associated with a Common Inversion Polymorphism. *Nat Genet* 2006; 38: 999-1001.
- [24] Shaw-Smith, C. Pittman AM, Willatt L, Martin H, Rickman L, Gribble S et al. Microdeletion Encompassing MAPT at Chromosome 17q21.3 is Associated with Developmental Delay and Learning Disability. *Nat Genet* 2006; 38: 1032-1037.
- [25] Cooper GM, Nickerson DA, Eichler EE. Mutational and Selective Effects on Copy-Number Variants in the Human Genome. *Nat Genet* 2007; 39: S22-S299.
- [26] McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, et al. Common deletion polymorphisms in the human genome. *Nat Genet* 38, 86-92 (2006).
- [27] Ye K, Hall G, Ning Z. Structural Variation Detection from Next Generation Sequencing. *Next Generat Sequenc & Applic*. 2016; S1:007.
- [28] Perry GH, Tchinda J, McGrath SD, Zhang J, Picker SR, Caceres AM et al. Hotspots for Copy Number Variation in Chimpanzees and Humans. *Proc Natl Acad Sci USA* 2006; 103: 8006-8011.
- [29] Onozawa M, Goldberg L, Aplan PD. Landscape of Insertion Polymorphisms in the Human Genome. *Genome Biol. Evol.* 2015; 7(4):960–968.
- [30] Griffiths AJF, Gelbart WM, Miller JH, Lewontin RC. *Modern Genetic Analysis*. New York: W. H. Freeman; 1999.
- [31] Sanders AD, Hills M, Porubsky D, Guryev V, Falconer E, Lansdorp PM. Characterizing Polymorphic Inversions in Human Genomes by Single Cell Sequencing. *Genome Res*. 2016. doi:10.1101/gr.201160.115.
- [32] Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS. An Initial Map of Insertion and Deletion (INDEL) Variation in the Human Genome. *Genome Res*. 2006; 16:1182–1190.
- [33] Ullu E, Murphy S, Melli M. Human 7S RNA Consists of a 140 Nucleotide Middle Repetitive Sequence Inserted in an Alu Sequence. *Cell* 1982; 29: 195–202.
- [34] Rogers J. Retroposons Defined. *Nature* 1983; 301: 460.

- [35] Chinniah R, Vijayan M, Thirunavukkarasu M, Mani D, Raju K, Ravi PM, et al. Polymorphic Alu Insertion/Deletion in Different Caste and Tribal Populations from South India. PLoS ONE 2016; 11(6): e0157468.
- [36] Laybourn S, Akam EC, Cox N, Singh P. and Mastana SS. Genetic Analysis of Novel *Alu* Insertion Polymorphisms in Selected Indian Populations. Am. J. Hum. Biol. 2016; doi: 0.1002/ajhb.22881
- [37] Kvikstad EM, Tyekucheva S, Chiaromonte F, Makova KD. A Macaque's-Eye View of Human Insertions and Deletions: Differences in Mechanisms. PLoS Comput Biol. 2007; 3(9): 1772-1782.
- [38] Kelkar YD, Tyekucheva S, Chiaromonte F, Makova KD. The Genome-Wide Determinants of Human and Chimpanzee Microsatellite Evolution. Genome Res. 2008; 18(1): 30-38.
- [39] Wajnberg G and Passetti F. Using High-Throughput Sequencing Transcriptome Data for INDEL Detection: Challenges for Cancer Drug Discovery. Expert Opin Drug Discov. 2016; 11(3):257-68.
- [40] Eichler EE. Recent Duplication, Domain Accretion and the Dynamic Mutation of the Human Genome. Trends Genet. 2001; 17: 661–669.
- [41] Samonte RV, and Eichler EE. Segmental Duplications and the Evolution of the Primate Genome. Nat. Rev. Genet. 2002; 3: 65–72.
- [42] Callaway E. Human Brain Shaped by Duplicate Genes. Nature News. 03 May 2012
- [43] Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW et al. A Burst of Segmental Duplications in the Genome of the African Great Ape Ancestor. Nature 2009; 457 (7231):877-881.
- [44] Nahon JL. Birth of “Human-Specific” Genes during Primate Evolution. Genetica 2003; 118(2): 193–208.
- [45] Dennis MY Nuttle X, Sudmant PH, Antonacci F, Graves TA, Nefedov M, et al. Human-Specific Evolution of Novel *SRGAP2* Genes by Incomplete Segmental Duplication. Cell. 2012; 149(4):912-22.
- [46] Acharya D and Ghosh TC. Global Analysis of Human Duplicated Genes Reveals the Relative Importance of Whole-Genome Duplicates Originated in the Early Vertebrate Evolution. BMC Genomics 2016; 17:71.
- [47] Gatchel JR and Zoghbi HY. Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. Nat Rev Genet. 2005; 6(10):743-55.
- [48] Galindo CL, McCormick JF, Bubb VJ, Abid Alkadem DH, Li LS,McIver LJ. A Long AAAG Repeat Allele in the 5' UTR of the *ERR-γ* Gene is correlated with Breast Cancer Predisposition and Drives Promoter Activity in MCF-7 Breast Cancer Cells. Breast Cancer Res Treat 2010; 130(1):41-8.
- [49] Gary Benson. Tandem Repeats Finder: A Program to Analyze DNA Sequences. Nucl. Acids Res. 1999; 27 (2):573-580.
- [50] Pellegrini M, Renda ME, Vecchio A. Tandem Repeats Discovery Service (TReaDS) Applied to Finding Novel Cis-Acting Factors in Repeat Expansion Diseases. BMC Bioinformatics. 2012;13 Suppl 4:S3.
- [51] Lopes RS, Walas Jhony Lopes Moraes, Thiago de Souza Rodrigues, Daniella Castanheira Bartholomeu ProGeRF: Proteome and Genome Repeat Finder Utilizing a Fast Parallel Hash Function. BioMed Research International Volume 2015 (2015), Article ID 394157, 9 pages

- [52] Lee JC, Tseng B, Ho BC, Linacre A. pSTR Finder: a Rapid Method to Discover Polymorphic Short Tandem Repeat Markers From Whole-Genome Sequences. *Investigative Genetics* 2015;6:10
- [53] Butler JM. *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. 2nd ed. Elsevier Academic Press, New York. 2005
- [54] Butler JM. Genetics and Genomics of Core STR Loci Used in Human Identity Testing. *J. Forensic Sci.* 2006; 51:253-265.
- [55] Van der Gaag KJ, de Leeuw RH, Hoogenboom J, Patel J, Storts DR, Laros JF, et al. Massively Parallel Sequencing of Short Tandem Repeats-Population Data and Mixture Analysis Results for the PowerSeq™ system. *Forensic Sci Int Genet.* 2016; 24:86-96.
- [56] Gymrek M, Willems T, Guilmatre A, Zeng H, Markus B, Georgiev S. Abundant Contribution of Short Tandem Repeats to Gene Expression Variation in Humans. *Nature Genetics*, 2016; 48, 22–29.
- [57] Forment JV, Kaidi A, Jackson SP. Chromothripsis and Cancer: Causes and Consequences of Chromosome Shattering. *Nat Rev Cancer.* 2012; 12:663–670.
- [58] Maher CA, Wilson RK. Chromothripsis and Human Disease: Piecing Together the Shattering Process. *Cell.* 2012; 148:29–32.
- [59] Tubio JM, Estivill X. Cancer: When Catastrophe Strikes A Cell. *Nature.* 2011; 470:476–477.
- [60] Holland AJ, Cleveland DW. Chromoanagenesis and Cancer: Mechanisms and Consequences of Localized, Complex Chromosomal Rearrangements. *Nat Med.* 2012; 18:1630–1638.
- [61] Patsalis PC. Complex Chromosomal Rearrangements. *Genet Couns.* 2007; 18(1):57-69.
- [62] Pellestor F, Anahory T, Lefort G, Puechberty J, Liehr T, Hedon B, et al. Complex Chromosomal Rearrangements: Origin and Meiotic Behavior. *Hum. Reprod. Update* 2011, 17(4): 476-494.
- [63] Del Rey J, Santos M, González-Meneses A, Milà M, Fuster C. Heterogeneity of a Constitutional Complex Chromosomal Rearrangement in 2q. *Cytogenet Genome Res.* 2016; 148(2-3):156-64.
- [64] Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *Am J Hum Genet.* 2014 Jul 3; 95(1): 5–23.
- [65] Fang S, Zhang S, Sha Q. Literature Reviews on Methods for Rare Variant Association Studies. *Human Genet Embryol.* 2016; 6:133.
- [66] Kosmicki JA, Churchhouse CL, Rivas MA, Neale BM. Discovery of Rare Variants for Complex Phenotypes. *Hum Genet* 2016; 135: 625.

Citation: Prem Chandra S, Pulakes P, Rakshit A, Sarkar B. *Magnitude of Human Genetic Variations: A Mini Review. International Journal of Research Studies in Medical and Health Sciences.* 2017;2(2):15-23.

Copyright: © 2017 Prem Chandra S et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.