AWARENESS

Newer Horizons in Human Excellence

ISSUE I | VOL I | JANUARY 2024

SRI SATHYA SAI
UNIVERSITY FOR
HUMAN EXCELLENCE

AWARENESS

Newer Horizons in Human Excellence

Table of Contents

Narrative Review

Approaches for Unraveling Complex Human Genetic Diseases

Nallur B. Ramachandra*, Chavadapur S. Srushti, Kommerahalli R. Meghana

Department of Studies in Genetics and Genomics, University of Mysore, Manasagangotri, Mysore, India.

Abstract: Human genetic diseases significantly burden the family and society. The key concern is understanding the diseases' etiology, prevention, treatment, and management. Even though there are many analytical methods and tools available, overall knowledge of these has yet to reach needy families. Here, we systematically narrated the methods and techniques such as pedigree analysis, the pattern of inheritance, cytogenetic analysis, next-generation sequencing, drug discovery, personalized genomic medicine and artificial intelligence employed to diagnose and treat genetic diseases, particularly complex diseases. These techniques help unravel the high-risk causal genes and their variants, further enabling drug discovery research groups and leading to personalized, targeted therapeutic interventions.

Keywords: Human Complex Genetic diseases, pedigree, inheritance pattern, cytogenetics, next-generation sequencing, drug discovery, genomic medicine.

Corresponding Author: Dr. N.B. Ramachandra; Email: ramachandra@genetics.uni-mysore.ac.in

Citation: Ramachandra NB, Srushti CS, Meghana KR. *Awareness* 2024, 1(1): 49-47.

Academic Editor: KJS Anand.

Received: 12-29-2023
Revised: 01-04-2024
Accepted: 01-06-2024

1. Introduction

Genetics unravel the mysteries of life processes. One-sixth of the world's population suffers from various genetic disorders. Genetic diseases can be seen from birth to later in life [1]. The Indian population has enormous variability due to consanguineous marriages among different cultures, ethnic groups, and caste systems [2]. Thus, genetic disorders of varied inheritance patterns are not uncommon in India and other developing countries.

Studying the origin and mechanism of genetic diseases has become possible with recent scientific advancements. All the diseases studied may have at least one genetic component. Rare and complex diseases are responsible for significant morbidity and mortality, which pose a global threat. Population screening and investigating etiological factors at the phenomics and genomics levels help diagnose and prevent diseases through early intervention and follow-up counselling [3]. The present task is to unravel the genetic component of complex diseases.

2. Genetic diseases

Changes in the genetic material lead to the four main types of genetic diseases as described below [4].

2.1 Single-Gene diseases

Single-gene diseases are caused by a single defective gene in chromosomes with a predictable inheritance pattern. They can be either dominant or recessive. If the disease is dominantly inherited, it requires the presence of one defective copy of the gene from either parent to manifest, whereas recessive diseases require two copies of the affected genes, one from each parent. Currently, the Online Mendelian Inheritance in Man (OMIM) documents 4,434 disease genes that, when mutated, provide the underlying causal basis for 6,306 Mendelian disorders (https://www.omim.org/statistics/geneMap) (OMIM compendium, 2023) [5]. Examples include sickle cell anemia & phenylketonuria.

2.2 Chromosomal Abnormalities

Chromosome abnormalities are characterized by a structural or numerical alteration in single or multiple chromosomes, affecting the autosomes or, allosomes or both. Numerical aberrations include changes in the chromosome number, aneuploidy or whole set of chromosomes, euploidy. Structural aberrations include deletions, duplication, inversion and translocation of chromosomal segments [6, 7]. Examples include Down's syndrome and Prader-Willi syndrome.

2.3 Mitochondrial Diseases

Extranuclear DNA is present in the mitochondria of all eukaryotic cells. Mitochondrial diseases are caused by mitochondrial DNA variants, which are inherited maternally. Therefore, the mutations are transferred from mothers to all their offspring but never from the father. Defects in oxidative phosphorylation identify mitochondrial diseases. These are caused by the changes in mitochondrial and nuclear DNA [8]. Mitochondrial disease can vary in severity and penetrance due to heteroplasmy. Examples include Leber Hereditary Optic Neuropathy (LHON) and Leigh Syndrome.

2.4 Multifactorial Diseases

Multifactorial diseases are caused by the effects of more than one gene in combination with lifestyle and other environmental factors. This is the most frequent disease cluster in families without a clear inheritance pattern. Examples include Diabetes and Hypertension. The prevalence of chromosomal, single-gene and multifactorial diseases are 3.8, 20 and ~600, respectively [9]. The etiology of these diseases can be studied through linkage analysis, DNA makers, and genome sequencing.

3. Pedigree Analysis and Patterns of Inheritance

The pedigree analysis technique is used to study the inheritance patterns of genetic diseases by examining family trees and medical histories. Pedigree diagrams visualize the relationships within families to understand the contribution of genetic factors to various diseases. During the late19th and early 20th centuries, the practice of pedigree studies can be traced back [10].

One of the earliest pedigree analyses is the inheritance of haemophilia, a bleeding disorder prevalent in Europe. The steps involved in pedigree analysis are collecting detailed family information, constructing a pedigree, analyzing the pedigree, interpreting the results, and follow-up testing.

Inheritance is the natural process by which genetic diseases are passed from parent to offspring. The five primary modes of inheritance are autosomal dominant (inheritance probability is 1:2), autosomal recessive (inheritance probability is 1:4), X-linked dominant (negligibly low for males, 100% for females), X-linked recessive (inheritance probability is 1:2 for males or 1:4 for all offspring) and Y- linked inheritance (only in males). Another rare form of inheritance is mitochondrial inheritance, which is a matrilineal inheritance [10]. Not all genetic conditions will follow these patterns. The patterns of inheritance are not easy to define for complex diseases. The chances of developing a complex genetic disease depends on several factors, namely, the number of relatives affected, how closely one is related to the affected individual(s), similarity of the shared environment and lifestyle factors; age, sex, ethnicity, age of onset, and disease severity in the affected relative. Pedigree analysis has its limitations and challenges. These include incomplete information, complex inheritance patterns, small sample sizes, non-paternity or non-maternity, environmental factors, and ethical considerations [10].

Pedigree studies have now become an even more powerful technique for understanding the inheritance patterns of genetic diseases with the advent of DNA markers.

4. Cytogenetic Analysis

During the metaphase of the cell cycle, cytogenetic analysis examines the banded pattern of all chromosomes. Cytogenetic analysis involves cell culture from the source, harvesting of cells, chromosome preparation, banding, microscopic analysis, and the production of karyotypes. Clinicians use cytogenetic studies mainly to investigate a family history, abnormal ultrasound, biochemistry findings, advanced maternal age of >35 years, recurrent miscarriages, abnormal non-invasive prenatal test results and multiple non-syndromic congenital anomalies [11]. Common cytogenetic studies include karyotyping, fluorescence in-situ hybridization, spectral karyotyping, and chromosomal microarray analysis.

4.1 Karyotyping

Karyotyping is the process of ordering and pairing of chromosomes to detect chromosomal abnormalities. Conventional banded karyotyping is the gold standard for diagnosing, classifying, assigning prognosis, and managing complex diseases [12]. Banding patterns can be achieved in different forms, such as Giemsa-banding, Centromere-banding, Quinacrine-banding, and Reverse-banding. Karyotyping allows for studying structural anomalies, such as deletions, duplications, insertions, inversions, and translocations, as well as numerical abnormalities such as trisomies of chromosome 13, 18, or 21, and monosomy, triploidy, or tetraploidy of chromosomes. Cytogenetic testing can be performed prenatally and postnatally.

4.2 Fluorescence in-situ Hybridization (FISH) and Spectral karyotyping (SKY)

Integration of molecular techniques with classic cytogenetic analysis gave rise to molecular cytogenetics. Two essential diagnostic techniques in molecular cytogenetics are FISH and SKY. FISH detects chromosome abnormalities targeting specific DNA sequences using fluorescently labelled probes that attach to complementary DNA. FISH testing involves determining the presence, absence, position, and copy number of DNA segments with the help of fluorescence microscopy. FISH can be performed both in the interphase and metaphase of the cell cycle. FISH can detect the microdeletion or microduplication of genes to assist in diagnosis, prognosis, and therapeutic management of several genetic diseases [13]. SKY, a relatively novel technique, is used to identify all human chromosomes, each labelled with different fluorochromes. This allows us to visualize all the chromosome pairs simultaneously to detect complex rearrangements and translocation within and between chromosomes [14].

4.3 Chromosomal Microarray analysis (CMA)

CMA is used to screen thousands of individual DNA sequences and provide precise information about the location(s) of any identified aberrations in a single experimental run. CMA helps to diagnose subjects lacking an apparent syndromic phenotype and allows the detection of clinically significant microdeletions or duplications. Microarray-based comparative genomic hybridization (aCGH) and single nucleotide polymorphism (SNP) arrays are subtypes of CMA. Like FISH, oligonucleotide probes attached to fluorescent dyes are used to label genomic DNA through hybridization. This helps to detect genomic copy number changes or minor genetic imbalances, such as gain or loss of chromosomal material [15]. Since the whole genome sequencing of several mammals is complete, a crucial next step involves using microarrays to catalogue all transcription units and identify their expression patterns in healthy and diseased individuals.

5. Next-generation Sequencing (NGS)

NGS provides susceptible and accurate high-throughput platforms for large-scale genomic testing. The workflow of NGS technologies is presented in Figure 1. In NGS, multiple genes can be tested in case-control or family-based studies, in parallel across numerous samples, saving time and money over performing multiple individual assays. Therefore, NGS has proved to be a compelling approach for investigating complex diseases quickly and efficiently as a single genetic test to identify and reliably characterize the comprehensive spectrum of genomic variations [16].

5.1 Whole Genome Sequencing (WGS)

WGS identifies all the genes, their variants and pathways that may have subtle phenotypes that are difficult to study biochemically. WGS also helps to understand the genomes' noncoding regions, introns, promoters, and telomers [17]. WGS provides the primary coverage of the entire human genome in regions not covered by other methods but even within regions targeted by other methods [18].

The Human Pangenome Reference Consortium [19] of the National Human Genome Research Institute presents a first draft of the human pangenome reference (GRCh38) assemblies, which have analyzed the genomes of diverse populations to represent the genomic landscape. It identifies about 90 Mb of additional bases derived from structural variations and reduces error to 1 base error per 200,000 assembled bases [19].

The workflow of WGS is given in Figure 1 (a). The essential applications of WGS include: (1) screening of newborn and pediatric diseases, (2) drug trials and pharmacogenomics, (3) regulatory variation and expression of quantitative trait loci (eQTLs), (4) investigation of rare tumor types, (5) family disease pedigrees and (6) large cohorts with extensive phenotyping. Shotgun sequencing is a classic strategy for whole genome sequencing.

5.2 Epigenome Analysis

Epigenomics includes understanding higher-order chromatin folding and covalent modifications of histone tails, attachment to the nuclear matrix, packaging of DNA around nucleosomes, DNA methylation and noncoding RNAs. NGS-based methods for epigenetic analysis include methyl-seq, ChIP-seq, and ATAC-seq. Integration of this massive amount of data promises to revolutionize our understanding of gene-environment interactions and offers unique ways to diagnose and treat complex human diseases [20]. The workflow for epigenome analysis is given in Figure 1 (b).

5.3. Whole Exome Sequencing (WES)

WES is a technique for sequencing all the protein-coding regions of genes in a genome. About 85% of disease-causing mutations lie in the exome [21]. WES is a powerful and reliable tool to identify population-specific, low and rare frequencies of all genetic variations associated with complex disease [22]. The main categories of exome capture technology are solution-based and array-based. In gene- and protein-interaction studies and pathway analyses, WES aids in discerning the genomic variations and reducing the noise generated by voluminous benign variations in complex diseases. The essential applications of WES are rare variant mapping in complex diseases, discovering Mendelian disorders, and clinical diagnostics. The workflow for whole Exome sequencing is given in Figure 1 (c).

5.4 Clinical Exome Sequencing (CES)

CES is a test for identifying disease-causing genomic variants that code for specific proteins. CES is a subset of WES as it covers a limited number of genes. CES is a rapid and common molecular diagnostic technique for those suffering from rare, mendelian, and complex diseases [23]. Examples include Cystic fibrosis, Haemophilia, and Breast cancer. The workflow for CES is given in Figure 1 (d).

5.5. Targeted Gene Sequencing (TGS)

TGS is a sensitive, rapid, more cost-effective, and robust method that delivers accurate, easy-to-interpret results in a single assay for identifying common and rare genomic variants. Examples: KRAS, TP53, BRAF, and EGFR are targeted across various cancer types [24]. It focuses on a panel of genes or targets known to have strong associations with disease pathogenesis and clinical relevance, offering greater sequencing depth and reducing data burden. The workflow for TGS is given in Figure 1 (e).

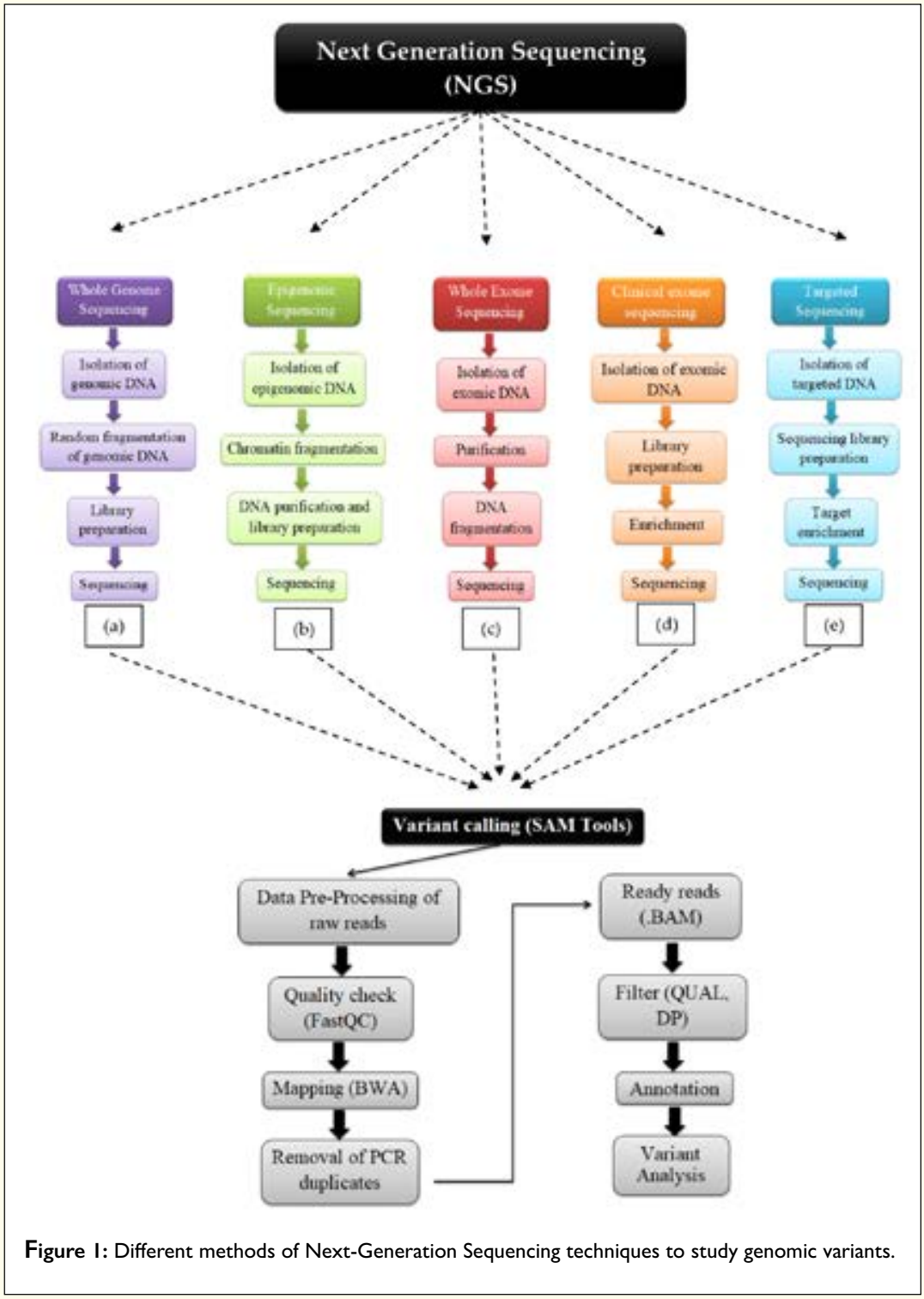

Figure 1: Different methods of Next-Generation Sequencing techniques to study genomic variants.

6. Analysis of Genomic Variations

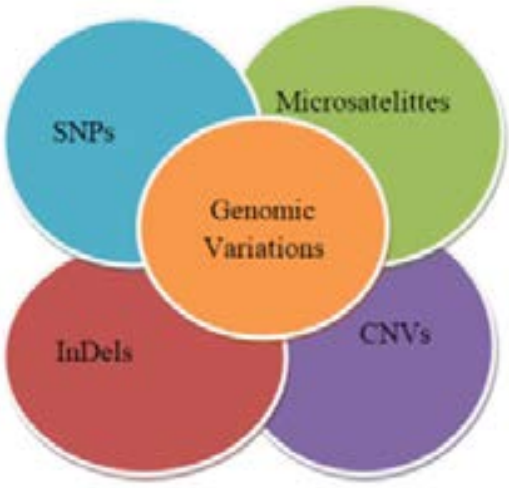

Figure 2: Types of genomic variations

Genomics is a dynamic field in biology that revolutionizes the diagnosis and prognosis of diseases like pathology, radiology, or biochemistry. Although 99.6% of genome sequences are identical in all humans, the 0.4% differences amount to millions of genomic variations in the human population [25]. The genomic variations, namely, single nucleotide polymorphisms (SNPs), Microsatellites, Insertion/deletion (InDels) and Copy number variations (CNVs), are the most powerful tools for the analysis of disease genomes [26, 27] (Figure. 2).

The pathogenicity of identified missense variations can be validated with various prediction tools. Sequence conservation profiling, protein stability changes, homozygosity mapping, and regulatory functional SNPs can be predicted, and different statistical analyses can be employed to analyse the significance of results. NGS allows the detection of common (MAF>0.05), uncommon (MAF<0.05 - >0.01), and rare (MAF<0.01) genomic variations and de novo variants involved in complex and mendelian diseases, thereby facilitating the comprehension of the genome architecture and its function [28-30].

6.1 Single Nucleotide Polymorphism (SNP)

SNP is a nucleotide location with a high substitution rate in individual samples from a population [31]. SNPs are present in about 3% of the coding and noncoding DNA in the genomes. About 10 to 30 million SNPs are identified in the human population. SNPs are highly polymorphic and serve as biological markers to understand population structure and genetic diversity, allowing us to build high-density genetic maps associated with disease [32]. Using DNA microarrays, millions of SNPs can be examined at the same time more effectively [33]. Many diseases are diagnosed with the association of SNPs and verified in the dbSNP and genomAD databases.

6.2 Microsatellites

Microsatellites are tandem repetitive DNA segments of 1 to 6 base pairs scattered throughout the genome. These are inherited in Mendelian inheritance patterns, highly polymorphic and diverse, and act as prominent genetic markers [34]. Microsatellite repeat expansions manifest in several diseases.

6.3 Insertions and Deletions (InDels)

InDels are the short segments of DNA with the size of 100 to 1000 base pairs that have been inserted or deleted from the genome [26]. InDels results in frameshift and non-frameshift variations. These are highly associated with many complex diseases [35] including neurodegenerative conditions like Alzheimer's disease or physical conditions like Cystic Fibrosis.

6.4 Copy Number Variations (CNVs)

CNVs are deletion or addition of 1kb up to several Mb segments of the genomes with >90% identity between the genomes of different individuals. CNVs are present in the population with fixed starting and ending positions. These account for a large portion of human variability [36], tremendously impacting the screening, diagnosis, prognosis, and monitoring of several diseases [37]. CNVs are classified into two main categories, based on sequence length. The first includes copy number polymorphisms (CNPs) – highly prevalent in the general population, with an overall frequency of 1% or more and typically associated with genes that encode proteins for drug detoxification and immunity. The second category of CNVs include variants ranging in size from hundreds of thousands to over 1 million base pairs. These rare variants are observed disproportionately in patients with mental retardation, developmental delay, schizophrenia, and autism.

7. Drug Discovery for Genomic Variants

The enormous human individual sequenced data and genomic variations have expanded the range of therapeutic targets in drug design and drug discovery. NMR spectroscopy, protein purification, and high-throughput crystallography have also made the structural details of protein–ligand and protein complexes available [38, 39]. These techniques lead to molecular docking or computeraided drug design (CADD), which uses the techniques of computational chemistry, molecular modeling, molecular design and rational drug design to design drugs for therapeutics, using ligandbased and structure-based methods [40].

Molecular target-based drug discovery has become the key to identifying potential druggable targets [41, 42]. NGS provides genes, their variants and how they affect the protein structure. A protein target associated with disease pathophysiology is identified using direct biochemical approaches, genetic interactions, and computational inference methods [43].

A chemical library of small molecules is essential in the drug discovery process. The number and diversity of commercially accessible chemical compound libraries are available for the pharmaceutical industry and academic investigators. Small drug molecules account for ~80-90% of marketed treatments and offer several benefits, including well-defined structures, relatively simple manufacturing, oral administration, and primarily non-immunogenic characteristics. These drug molecules can calso ross the blood-brain barrier and reach the central nervous system.

CADD is typically used in three ways in the drug discovery process: (1) to reduce large compound libraries into smaller sets of predicted active compounds, (2) to guide the optimization of lead compounds, and (3) to design novel compounds. The drug screening results are entered into a database and examined with bioinformatics software. For single-concentration screening, primary screening hits are chosen based on criteria such as inhibition >50%, IC50 <5 M and efficacy >70%" [44]. The principal hits consider real chemicals and delete false positives and non-specific chemicals.

8. Personalized Genomic Medicine (PGM)

PGM uses information and data from a patient's genotype and phenotype to stratify disease, select a medication, provide therapy, and initiate preventative measures suited to that patient at the time of administration. This is a new, comprehensive, and integrated approach to disease management and wellness. Most drugs are not effective or partially effective in 60% of the treated patients. Side effects are responsible for millions of deaths and hospitalizations. Predictive toxicology for new drug candidates can predict which individuals will benefit, and those who might be at greater risk for experiencing severe side effects can be identified [45].

The number of targeted treatments in the pipeline for all diseases is increasing dramatically. For instance, the actress Angelina Jolie made headlines with a proactive double mastectomy after tests revealed she carried BRCA1, the same mutations for breast cancer as her mother, who died from the disease. Genes are not destiny, but they indicate suitable decisions about our health and healthcare. As in Jolie's case, this can change our future [46, 47]. Potential benefits of PGM are to predict disease susceptibility, customize disease-prevention strategies, improve disease detection, preempt disease progression, shift the emphasis in medicine from reaction to prevention, prescribe more effective drugs, and avoid prescribing drugs with predictable side effects. PGM helps to reach the point where all medicines are linked to diagnostics [48]. PGM is a promising treatment for improved human welfare.

9. Artificial Intelligence (AI)

AI is the computational simulation of human intelligence. In clinical diagnostics, AI is described as any computer system that can accurately analyze health data, particularly in its original state as viewed by humans. AI can analyze enormous amounts of clinical diagnostic data and genome variation data in various therapeutic situations and contribute to the efficient interpretation of massive and complicated datasets [49, 50].

First, numerous algorithms have been created employing various AI models to discover high-quality genomic variations sensitively and precisely associated with disease. AI also helps to identify long noncoding RNAs, generate protein-coding DNA sequences, and design DNA probes for protein-binding microarrays. Deep learning is the best way to analyze these data sources and complete genomic modelling tasks as the genomic data sets grows exponentially.

Second, AI assists in designing and screening drug molecules and analyzing clinical trials. At the gene level, AI can help to predict the binding affinities of transcription factors, DNA- and RNA-binding proteins, cisregulatory/enhancer elements, DNA methylation sites, histone modifications, chromatin accessibility, transcription start sites, tissueregulated splicing, gene expression and translation efficacies, transcriptome patterns, microRNA precursors and binding targets, variant calling, noncoding and coding variants pathogenicity [51]. Therefore, AI is one of the strategies to address all kinds of clinical and genomic issues to prevent, treat, and manage diseases by potential drug molecules.

10. Discussion and Conclusion

Although efficient technologies are available, we have several limitations to treating and preventing genetic diseases. Some limitations are the need for more public awareness, current knowledge of the technology, cost of the diagnosis, and proper diagnostic skills. Special training on recent technologies must be updated for concerned persons, including the medical community and researchers.

Some of the recommendations for the implementations include: (1) Diagnostic chips from the identified diseasespecific, population-specific variants can be established; (2) Effective integration of genomic, pathological, and clinical data for effortless translation of research data can be enhanced; and (3) These multimodal approaches will effectively assess the effect of variations on protein function through different perspectives and present an opportunity to develop unique biomarker panels for identifying pre-symptomatic individuals with complex diseases.

Finally, understanding the causes and consequences of genetic mechanisms provides meaningful insights into disease genesis and progression. With systematic improvements in the latest software support, NGS data analysis has become an accessible reality to researchers and clinicians. This has paved the way for improved knowledge of the complex interplay of genes and their mechanisms of action in complex human diseases, together with the development of novel treatment and management strategies. Thus, identifying high-risk causal genes and their variants enables research groups to further the discovery process, leading to personalized, targeted therapeutic interventions.

Author Contributions: Conceptualization, Nallur B.Ramachandra; methodology, Chavadapur S. Srushti.; Kommerahalli R. Meghana.; resources, Chavadapur S. Srushti.; Kommerahalli R. Meghana.; data curation, Chavadapur S. Srushti.; Kommerahalli R. Meghana.; writing—original draft preparation, Chavadapur S. Srushti.; Nallur.B.Ramachandra.; writing—review and editing, Chavadapur S. Srushti.; Kommerahalli R. Meghana.; Nallur.B.Ramachandra; supervision, Nallur.B.Ramachandra.; funding acquisition, Department of Atomic Energy (DAE), Government of India, Mumbai Raja Ramanna Chair (RRC) scheme.; Indian Council of Medical Research- Senior Research Fellowship (ICMR-SRF), New Delhi. All authors have read and agreed to the published version of the manuscript.

Funding: None

Institutional Review Board Statement: Not Applicable

Informed Consent Statement: Not Applicable

Data Availability Statement: No new data were created or analyzed during this study. Data sharing is not applicable to this article.

Acknowledgments: We thank the Department of Studies in Genetics and Genomics, University of Mysore; DAE-RRC; and ICMR-SRF.

Conflicts of Interest: There are no conflicts of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

References

1. Jackson, S.L., L. Hart, and F.R. Volkmar, *Preface: Special issue—College experiences for students with autism spectrum disorder*. 2018, Springer. p. 639-642.
2. Sahoo, H., et al., *Changing trends of consanguineous marriages in South India.* Journal of Asian and African Studies, 2022. 57(2): p. 209-225.
3. Elliott, A.M., *Genetic counseling and genome sequencing in pediatric rare disease*. Cold Spring Harbor perspectives in medicine, 2020. 10(3).
4. Oti, M. and H.G. Brunner, *The modular nature of genetic diseases*. Clinical genetics, 2007. 71(1): p. 1-11.
5. Hieter, P., et al., *Highlighting rare disease research with a GENETICS and G3 series on genetic models of rare diseases*. 2023, Oxford University Press US. p. iyad121.
6. Moorthie, S., et al., *Chromosomal disorders: estimating baseline birth prevalence and pregnancy outcomes worldwide*. Journal of community genetics, 2018. 9: p. 377-386.
7. Milani, D.A.Q. and P. Tadi, *Genetics, Chromosome Abnormalities, in StatPearls [Internet]*. 2022, StatPearls Publishing.
8. Gorman, G.S., et al., *Mitochondrial diseases*. Nature reviews Disease primers, 2016. 2(1): p. 1-22.
9. Nussbaum, R., R.R. McInnes, and H.F. Willard, Thompson & Thompson genetics in medicine e-book. 2015: Elsevier Health Sciences.
10. Lewis, R., *Human genetics: The basics*. 2016: Taylor & Francis.
11. Ozkan, E. and M.P. Lacerda, *Genetics, cytogenetic testing and conventional karyotype*. 2020.
12. Kang, J., *Clinical implications of microarray in cancer medicine*. International Journal of Cancer Research, 2015. 11(4): p. 150-158.
13. Dave, B.J. and W.G. Sanger. *Role of cytogenetics and molecular cytogenetics in the diagnosis of genetic imbalances. in Seminars in pediatric neurology*. 2007. Elsevier.
14. Chowdhury, M.R., A. Singh, and S. Dubey, *Role of cytogenetics and molecular genetics in human health and medicine, in Animal biotechnology*. 2020, Elsevier. p. 481-501.
15. Berisha, S.Z., et al., *Cytogenetic and molecular diagnostic testing associated with prenatal and postnatal birth defects*. Birth Defects Research, 2020. 112(4): p. 293-306.
16. Di Resta, C., et al., *Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities*. Ejifcc, 2018. 29(1): p. 4.
17. Clark, M.M., et al., *Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases*. NPJ genomic medicine, 2018. 3(1): p. 16.
18. Meienberg, J., et al., *Clinical sequencing: is WGS the better WES?* Human genetics, 2016. 135: p. 359-362.
19. Liao, W.-W., et al., *A draft human pangenome reference*. Nature, 2023. 617(7960): p. 312-324.
20. Ku, C.S., et al., *Studying the epigenome using next generation sequencing*. Journal of medical genetics, 2011.
21. Rabbani, B., M. Tekin, and N. Mahdieh, *The promise of whole-exome sequencing in medical genetics*. Journal of human genetics, 2014. 59(1): p. 5-15.
22. Amanat, S., T. Requena, and J.A. Lopez-Escamez, *A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders*. Genes, 2020. 11(9): p. 987.
23. Lee, H., et al., *Clinical exome sequencing for genetic identification of rare Mendelian disorders*. Jama, 2014. 312(18): p. 1880-1887.
24. Wala, J.A., et al., *SvABA: genome-wide detection of structural variants and indels by local assembly*. Genome research, 2018. 28(4): p. 581-591.

25. Vollger, M.R., et al., *Segmental duplications and their variation in a complete human genome*. Science, 2022. 376(6588): p. eabj6965.
26. Redon, R., et al., *Global variation in copy number in the human genome*. Nature, 2006. 444(7118): p. 444-454.
27. Ku, C.S., et al., *The discovery of human genetic variations and their use as disease markers: past, present and future*. Journal of human genetics, 2010. 55(7): p. 403-415.
28. Musunuru, K., et al., *From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus*. Nature, 2010. 466(7307): p. 714-719.
29. Ng, S.B., et al., *Exome sequencing identifies the cause of a mendelian disorder*. Nature genetics, 2010. 42(1): p. 30-35.
30. Bilgüvar, K., et al., *Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations*. Nature, 2010. 467(7312): p. 207-210.
31. Ferguson, J.F., et al., *Gene-nutrient interactions in the metabolic syndrome: single nucleotide polymorphisms in ADIPOQ and ADIPOR1 interact with plasma saturated fatty acids to modulate insulin resistance*. The American journal of clinical nutrition, 2010. 91(3): p. 794-801.
32. Hou, H., et al., *An updated systematic review and meta-analysis of association between adiponectin gene polymorphisms and coronary artery disease*. OMICS: A Journal of Integrative Biology, 2017. 21(6): p. 340-351.
33. Khlestkina, E. and E. Salina, *SNP markers: methods of analysis, ways of development, and comparison on an example of common wheat*. Russian Journal of Genetics, 2006. 42: p. 585-594.
34. Morgante, M., M. Hanafey, and W. Powell, *Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes*. Nature genetics, 2002. 30(2): p. 194-200.
35. Lin, M., et al., *Effects of short indels on protein structure and function in human genomes*. Scientific reports, 2017. 7(1): p. 9313.
36. Eichler, E., *Copy number variation and human disease*. Nat Educ, 2008. 1(3): p. 1.
37. Pös, O., et al., *DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects*. biomedical journal, 2021. 44(5): p. 548-559.
38. Bajorath, J., *Integration of virtual and high-throughput screening*. Nature Reviews Drug Discovery, 2002. 1(11): p. 882-894.
39. Jorgensen, W.L., *The many roles of computation in drug discovery*. Science, 2004. 303(5665): p. 1813-1818.
40. Henrich, S., et al., *Computational approaches to identifying and characterizing protein binding sites for ligand design*. Journal of Molecular Recognition: An Interdisciplinary Journal, 2010. 23(2): p. 209-219.
41. Agüero, F., et al., *Genomic-scale prioritization of drug targets: the TDR Targets database*. Nature reviews Drug discovery, 2008. 7(11): p. 900-907.
42. Eder, J., R. Sedrani, and C. Wiesmann, *The discovery of first-in-class drugs: origins and evolution*. Nature Reviews Drug Discovery, 2014. 13(8): p. 577-587.
43. Schenone, M., et al., *Target identification and mechanism of action in chemical biology and drug discovery*. Nature chemical biology, 2013. 9(4): p. 232-240.
44. Inglese, J., et al., *Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries*. Proceedings of the National Academy of Sciences, 2006. 103(31): p. 11473-11478.
45. Roth, S.C., *What is genomic medicine?* Journal of the Medical Library Association: JMLA, 2019. 107(3): p. 442.
46. Juengst, E.T., M.A. Flatt, and R.A. Settersten, *Personalized genomic medicine and the rhetoric of empowerment, in The Ethical Challenges of Emerging Medical Technologies*. 2020, Routledge. p. 177-183.

47. Zeggini, E., et al., *Translational genomics and precision medicine: Moving from the lab to the clinic*. Science, 2019. 365(6460): p. 1409-1413.

48. Xie, W., B. Chen, and J. Wong, *Genomic medicines: the coming waves?* Nature reviews. Drug Discovery, 2023.

49. Dias, R. and A. Torkamani, *Artificial intelligence in clinical and genomic diagnostics*. Genome medicine, 2019. 11(1): p. 1-12.

50. Uddin, M., Y. Wang, and M. Woodbury-Smith, *Artificial intelligence for precision medicine in neurodevelopmental disorders*. NPJ digital medicine, 2019. 2(1): p. 112.

51. Eraslan, G., et al., *Deep learning: new computational modelling techniques for genomics*. Nature Reviews Genetics, 2019. 20(7): p. 389-403.