



GENOMICS IN DIAGNOSIS: AIDS IN MEDICINE

Dr. Anil Batta*

Professor and Head, Department of Biochemistry, Muzaffarnagar Medical College, Muzaffarnagar.



*Corresponding Author: Dr. Anil Batta

Professor and Head, Department of Biochemistry, Muzaffarnagar Medical College, Muzaffarnagar.

Article Received on 13/05/2024

Article Revised on 02/06/2024

Article Accepted on 23/06/2024

ABSTRACT

The sequencing of complete human genome revolutionized the genomic medicine. However, the complex interplay of gene-environment-lifestyle and influence of non-coding genomic regions on human health remain largely unexplored. Genomic medicine has great potential for diagnoses or disease prediction, disease prevention and, targeted treatment. However, many of the promising tools of genomic medicine are still in their infancy and their application may be limited because of the limited knowledge we have that precludes its use in many clinical settings. In this review article, we have reviewed the evolution of genomic methodologies/tools, their limitations, and scope, for current and future clinical application.

KEYWORDS: Genomic medicine, medical genetics, Gene sequencing, DNA sequencing, RNA sequencing, Clustered regularly interspaced short palindromic repeat, Gene based therapy, Genomic tools, Genome editing.

INTRODUCTION

Remarkably, the human genome and the closest related species chimpanzees differ in single nucleotide alterations by a mere 1.23% and in deletions, insertions, and copy number variations by 3%.^[6] In humans, the genomes of any two individuals are about 99.9% identical. However, a mere 0.1% variation allows for changes in a massive number of nucleotides because the human genome has approximately 30 billion base pairs (3.3×10^9).^[7] The human genome project, which was completed in 2003, revolutionized the understanding of the human genome and served as a turning point to fast forward the genomic methodologies. However, the clinical application of findings from these genomic studies is still in its infancy. This is largely because we still have not understood or made complete sense of the available information. That is, the sequence data have been difficult to correlate to functional outcomes, making it difficult to understand the genetic basis of diseases and the complex gene-lifestyle-environment influences or their interaction. Moreover, most of the initial focus of the research had been on coding regions of DNA which comprises approximately 2% of the DNA and the knowledge about specific implications of non-coding DNA regions (98% of DNA) are largely unknown.^[4,5]

**GENOMIC TOOLS AND THEIR EVOLUTION
DNA SEQUENCIES**

It relied on the template DNA strand and had limited capacity for sequencing gene panels. Subsequently, with commercial production of high throughput technologies

or next-generation sequencing (NGS) revolutionized the DNA sequencing by 2007.^[10] Also called as massively parallel sequencing, NGS does parallel sequencing of millions of small DNA fragments. Each DNA fragment is fixed at a unique location on the solid support. While the sample of the patient's DNA which serves as a template in NGS is amplified and fragmented, the third-generation sequencing uses single DNA molecules rather than the amplified DNA as a template thus eliminating errors from DNA amplification processes. The NGS can be used for whole-genome sequencing, exome sequencing, or targeted gene panels comprising tens to hundreds of genes.

SINGLE NUCLEOTIDE POLYMORPHISM

Single nucleotide polymorphism (SNP) is the variation in genetic sequence by a single nucleotide. It is the most common type of genetic variation in man.^[11] It was detected in the 1980s using restriction enzymes.^[12] With application of the microarray technology to SNPs, the scope of SNP in clinical practice has widened, especially in oncology. The first SNP array analysis was done in 1998 and the first application of SNP array analysis in cancer was done in 2000.^[13] SNP array analysis is used to determine loss of heterozygosity, allelic imbalance, genomic copy number changes, frequency of homozygous chromosome regions, uniparental disomy, DNA methylation alterations and linkage analysis of DNA polymorphisms in cancer cells.^[13,14]

DNA AMPLIFICATION

Kary Banks Mullis successfully demonstrated polymerase chain reaction (PCR) in 1983.^[15] PCR is a cost-effective method that can amplify a single DNA exponentially.^[16] It is a rapid, highly specific, and extremely sensitive method. PCR is being used in SNP genotyping, detection of rare sequences, insertion-deletion variants, and structural variants like copy-number variants.

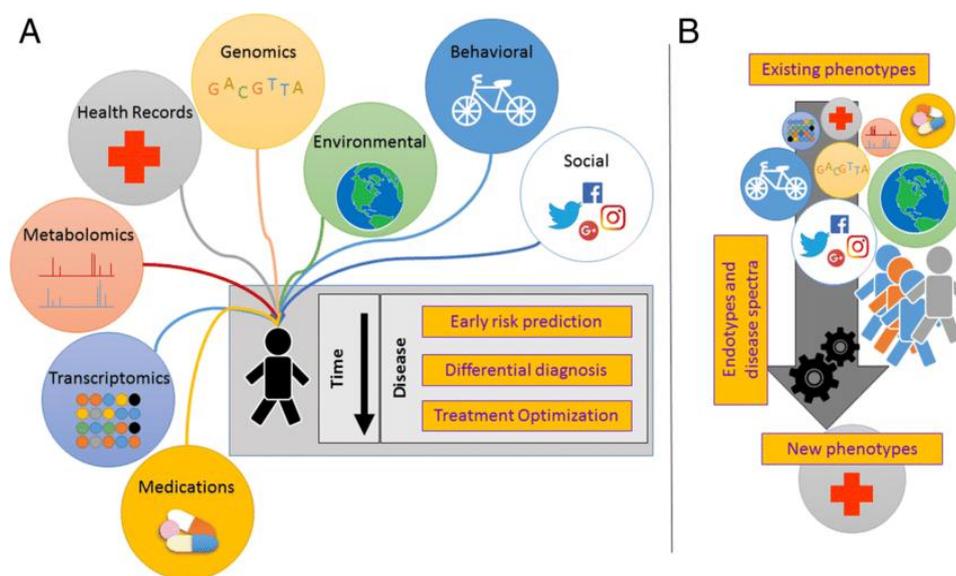
LINKAGE AND ASSOCIATION ANALYSIS

Linkage studies have been used for mapping Mendelian traits with high penetrance in families and relatives.^[20] They are especially useful to identify rare alleles that are present in a small number of families^[21], for disease genes with weak effects and polygenic diseases, linkage disequilibrium association mapping has proved to be more useful. In genome-wide association studies

(GWAS), genotyping of hundreds or thousands of SNPs is done in cases and control populations and their association with heritability is analysed. A combination of linkage and association methodologies helps to identify and characterize the wider range of disease-susceptibility variants.^[22]

HAPMAP AND 1000 GENOME PROJECTS HAVE CREATED A CATALOG OF SNPS

The HapMap project was started in 2002 to develop a haplotype map of the human genome. It can also describe the common patterns of human genetic variation.^[24] The 1000 Genomes Project comprised a total of 26 diverse population set in which whole-genome sequencing was performed. It also used deep exome sequencing and dense microarray genotyping to give a comprehensive description of common human genetic variation.^[25]



TARGETED GENOME EDITING OR GENOME ENGINEERING

It involves modification of the genome at a precise, prespecified locus using programmable nucleases. Examples of some of the programmable nucleases include zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR)-Cas (CRISPR-associated) system. These programmable nucleases are designed to impart site-specific double-strand breaks (dsBs) in chromosomal DNA. The cell is therefore forced to use one of the endogenous DNA repair mechanisms — homologous recombination or homology-directed repair (HDR) and nonhomologous end-joining (NHEJ). This enables targeted genetic modifications during the repair process in the living cells (in vivo) (Table (Table11)).^[26] ZFNs and TALENS recognize the target sequence through protein-DNA interaction. CRISPR-Cas nucleases recognize target sequences through RNA and DNA base pairing.^[26]

DISCUSSION

The newer genomic technology and tools have broadened the scope and pushed the time limits for development of new diagnostic kits, preventive strategies like vaccines, therapeutic strategies like gene modulation and gene therapy. A lot is yet to be studied in terms of the complex interaction of gene-environment-lifestyle-disease. Knowing the impact of genomics on disease pathophysiology and response to medications, expands the scope of research and clinical application. While genome editing holds promise to correct the defective genome in vivo, therapies can also be designed to alter the gene expression without altering the genomic code (example exon skipping, or inclusion discussed above). The newer genomic editing tools have showed great potential and promise but they need to be studied extensively before clinical application. Also, uniform international ethical guidelines and guiding principles need to be established so that these genomic technologies are not misused. It is very important to include diverse populations and to represent minority population in the genomic studies, so that results could be generalized and

more accurate diagnostic, predictive and therapeutic tools can be developed. Genomics in medicine is indeed a new era in medicine. Even the control of coronavirus disease 2019 pandemic has just begun at the time of writing of this article with gene-based therapies eliciting immune response against severe acute respiratory syndrome coronavirus 2 spike proteins. A unified international collaboration is needed to continue expanding gene therapy use in opening new frontiers for fight against novel infections and disease.

CONCLUSION

Genomic medicine holds great promise for providing insight into disease pathophysiology, provide better diagnostic or disease predictive tools, preventive therapies and finally for targeted treatment of diseases. Although some of the newer tools (like CRISPR system) have great potential, more research is needed before these tools can be unleashed to clinical use. Hence there is great need for studies to unravel the mystery of complex interaction of both coding and noncoding genomic regions with environment and lifestyle influences on disease occurrence and management. The newer genomic technology and tools have broadened the scope and pushed the time limits for development of new diagnostic kits, preventive strategies like vaccines, therapeutic strategies like gene modulation and gene therapy. A lot is yet to be studied in terms of the complex interaction of gene-environment-lifestyle-disease. Knowing the impact of genomics on disease pathophysiology and response to medications, expands the scope of research and clinical application. While genome editing holds promise to correct the defective genome *in vivo*, therapies can also be designed to alter the gene expression without altering the genomic code (example exon skipping, or inclusion discussed above). The newer genomic editing tools have showed great potential and promise but they need to be studied extensively before clinical application. Also, uniform international ethical guidelines and guiding principles need to be established so that these genomic technologies are not misused. It is very important to include diverse populations and to represent minority population in the genomic studies, so that results could be generalized and more accurate diagnostic, predictive and therapeutic tools can be developed. Genomics in medicine is indeed a new era in medicine. Even the control of coronavirus disease 2019 pandemic has just begun at the time of writing of this article with gene-based therapies eliciting immune response against severe acute respiratory syndrome coronavirus 2 spike proteins. A unified international collaboration is needed to continue expanding gene therapy use in opening new frontiers for fight against novel infections and disease.

REFERENCES

1. Watson JD, Crick FH. The structure of DNA. *Cold Spring Harb Symp Quant Biol*, 1953; 18: 123–131. [PubMed] [Google Scholar]
2. McKusick VA. Mendelian Inheritance in Man and its online version, OMIM. *Am J Hum Genet*, 2007; 80: 588–604. [PMC free article] [PubMed] [Google Scholar]
3. Johns Hopkins University. OMIM Gene Map Statistics. [cited 20 December 2020]. In: Johns Hopkins University [Internet]. Available from: <https://www.omim.org/statistics/geneMap>.
4. Ling H, Vincent K, Pichler M, Fodde R, Berindan-Neagoe I, Slack FJ, Calin GA. Junk DNA and the long non-coding RNA twist in cancer genetics. *Oncogene*, 2015; 34: 5003–5011. [PMC free article] [PubMed] [Google Scholar]
5. Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. *Exp Mol Med*, 2018; 50: 1–8. [PMC free article] [PubMed] [Google Scholar]
6. Suntsova MV, Buzdin AA. Differences between human and chimpanzee genomes and their implications in gene expression, protein functions and biochemical properties of the two species. *BMC Genomics*, 2020; 21: 535. [PMC free article] [PubMed] [Google Scholar]
7. Gyles C. The DNA revolution. *Can Vet J.*, 2008; 49: 745–746. [PMC free article] [PubMed] [Google Scholar]
8. Maxam AM, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci U S A.*, 1977; 74: 560–564. [PMC free article] [PubMed] [Google Scholar]
9. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977; 74: 5463–5467. [PMC free article] [PubMed] [Google Scholar]
10. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. *Mol Cell.* 2015; 58: 586–597. [PMC free article] [PubMed] [Google Scholar]
11. Shen LX, Basilion JP, Stanton VP Jr. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci U S A.*, 1999; 96: 7871–7876. [PMC free article] [PubMed] [Google Scholar]
12. Gray IC, Campbell DA, Spurr NK. Single nucleotide polymorphisms as tools in human genetics. *Hum Mol Genet*, 2000; 9: 2403–2408. [PubMed] [Google Scholar]
13. Mao X, Young BD, Lu YJ. The application of single nucleotide polymorphism microarrays in cancer research. *Curr Genomics*, 2007; 8: 219–228. [PMC free article] [PubMed] [Google Scholar]
14. Sato-Otsubo A, Sanada M, Ogawa S. Single-nucleotide polymorphism array karyotyping in clinical practice: where, when, and how? *Semin Oncol*, 2012; 39: 13–25. [PubMed] [Google Scholar]
15. Fairfax MR, Salimnia H. Diagnostic molecular microbiology: a 2013 snapshot. *Clin Lab Med*, 2013; 33: 787–803. [PMC free article] [PubMed] [Google Scholar]

16. Wages JM. Polymerase Chain Reaction. In: Worsfold P, Townshend A, Poole C. Encyclopedia of Analytical Science. Oxford: Elsevier, 2005; 243-250. [Google Scholar]
17. Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. An Introduction to Genetic Analysis - NCBI Bookshelf. 7th ed. New York: W. H. Freeman, 2000. [Google Scholar]
18. Dawn Teare M, Barrett JH. Genetic linkage studies. *Lancet*, 2005; 366: 1036–1044. [PubMed] [Google Scholar]
19. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet*, 2015; 16: 275–284. [PMC free article] [PubMed] [Google Scholar]
20. Pulst SM. Genetic linkage analysis. *Arch Neurol*, 1999; 56: 667–672. [PubMed] [Google Scholar]
21. Hinrichs AL, Suarez BK. Incorporating linkage information into a common disease/rare variant framework. *Genet Epidemiol*, 2011; 35 Suppl 1: S74–S79. [PMC free article] [PubMed] [Google Scholar]
22. Ott J, Kamatani Y, Lathrop M. Family-based designs for genome-wide association studies. *Nat Rev Genet*, 2011; 12: 465–474. [PubMed] [Google Scholar]
23. Hu L, Ru K, Zhang L, Huang Y, Zhu X, Liu H, Zetterberg A, Cheng T, Miao W. Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine. *Biomark Res*, 2014; 2: 3. [PMC free article] [PubMed] [Google Scholar]
24. National Human Genome Research Institute. The International HapMap Project. Updated 06/04/2012. [cited 10 December 2020]. In: National Human Genome Research Institute [Internet]. Available from: <https://www.genome.gov/11511175/about-the-international-hapmap-project-fact-sheet>.
25. 1000 Genomes Project Consortium. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. *Nature*, 2015; 526: 68–74. [PMC free article] [PubMed] [Google Scholar]
26. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. *Nat Rev Genet*, 2014; 15: 321–334. [PubMed] [Google Scholar]
27. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013; 339: 819–823. [PMC free article] [PubMed] [Google Scholar]