Application of Whole Genome Sequencing in Medical and Pharmaceutical field

Awofisayo O.A[1] *. and Eseyin O.A[1].

[1]Department of Pharmaceutical and Medicinal Chemistry, University of Uyo, Nigeria
Submitted:6th May, 2022; Accepted:26th August, 2022; Available online: 31st August, 2022

Doi: https://doi.org/10.54117/jcbr.v2i4.9
*Corresponding author: E-mail: oladojaawofisayo@uniuyo.edu.ng

Abstract

Whole Genome Sequencing (WGS) is a high throughput process utilized to obtain the nucleotide bases that constitutes a genome. The Sanger Technique was used to sequence the first human genome; however the limitations include inability to sequence more than few genes per time and it cannot detect structural rearrangement. The method also has far reaching cost implications. The New Generation Sequencing Platform has further lend credence to WGS, as sequencing can be carried out in a shorter period of time with accuracy hence a great improvement in the clinical setting for diagnostic purposes. This also needs to correspond with effective data analysis and interpretation. This method has proved useful as it can be used to compare bacterial strains that have a divergence at one nucleotide base, transmission, Virulence, phenotypic characteristics can be deduced. Various strains can be analyzed via a single process and it is an essential tool in clinical diagnosis and in monitoring the outbreak of a new disease. This is important in handling multidrug resistance strains of bacteria. In addition, it equally helps in the identification of drug response markers on drug target. Moreover, drug metabolism or disease pathway can enhance development of drug tailored to specific individuals for specific purposes, hence the ability to define a population genetically.

Keywords: Whole Genome Sequencing, genome, drug metabolism, disease pathway, phenotypic characteristics

Introduction

The completion of the human genome in 2001 gave rise to the field of genomics (Nerlich et al., 2002). Whole Genome Sequencing is the aspect of genomics that allows for the determination of genetic variation in humans and microorganisms at the level of the nucleotide basis (Gibbs, 2022). The genetic basis of various diseases of public health importance has been discovered, and this has led to mapping of diseases to the associated genes (Ng et al., 2010). The advent of whole genome sequencing has greatly allowed the genetic predictors of human diseases to be deduced . Whole Genome Sequencing NGS is a laboratory technique that sequences the Whole genome of an organism bacteria, viral or man (Tam et al., 2019). The genes, exons and introns are thereby elucidated and deciphered (Plenge, 2013). The genome encodes the total information required for the existence of the organism (Griffiths, 2001). This information includes enzymes required for metabolism, membrane proteins

and cell signaling proteins. The presence of an enzyme in the metabolic pathway confirms the presence of the enzyme in the genome (Zhao et al., 2013).

Methodology

Laboratory and Computational Methods are applied. The computational method involves the use of artificial vectors to subclone regions of the genome where DNA inserts are present and this may be up to 200,000 base pairs in humans. Physical mapping is done by different experimental methods like Direct Method, Shortgun Method and Primer Walking (Zhao et al., 2013).

This advent of New Generation Sequencing NGS has greatly improved the capacity of genomic analysis at low cost and higher efficiency. The gene cannot act alone but works alongside promoters, regulatory elements, response elements, silencers and enhancers. NGS has been able to identify epigenetic changes in the genome (Davey, 2011).

Sequencing the whole genome provides the most comprehensive SV analysis. Breakpoints of CNV and copy-neutral SV are detectable by paired-end reads that have discordant mappings to the reference genome. Complex genomic structures identified by WGS may require FISH or chromosome banding to place rearranged segments.

Other forms of genetic variation include Single Nucleotide Polymorphism SNPs, Indel, Insertion and Deletions, Structural variations duplication, deletion and inversion. They are used to identify mendelian diseases of which there are more than 7000 distinct diseases in the MIM (Schuster, 2008)

The advantages include:

1. Identification and Treatment of multidrug resistant tuberculosis. Sequencing of the clinical isolates and mutations were identified on the M gene The strains used were related to the East Asian strain.
2. Identification of Human Genetc Variations.
3. Adverse drug reactions to current medications have been documented and found to be alarming.
4. Development of a personalized drug which will be suitable for an individual since his genetic make-up has been identified.
5. Development of novel drugs that can target specific part of the genome.
6. Identification of novel targets for a particular disease with novel ligands being developed.
7. The molecular basis of diseases can be better understood: early detection allows application of novel therapies to combat the disease.
8. Early identification of biomarkers of certain diseases via genome sequencing can pave way for a proactive means of prevention via change of lifestyle which may prevent the disease from progressing.
9. Ancestry of individuals can be determined.
10. Forensic analysis is made possible and solution of crimes enhanced.
11. Individual responses to particular drugs can be catalogued.

12. Genetic screening must be done before administration of personalized medications.
13 Genomics and WGS can be applied to decipher individuals at risk of hereditary diseases like leukaemia, beta thallasemia.

Clinical application of whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases

Identification and Treatment of multidrug resistant tuberculosis.
The limitation in the availability of drugs for TB has made the treatment of MDR-TB very challenging. Drug resistance have been developed by microorganisms to variety of drug. I. For example in 2017 (Witney et al, 2015), Multi drug resistance tuberculosis MDR-TB have been discovered with solutions being proffered by WGS and second line anti TB drugs have been developed. However, based on WGS, there were 558,000 new cases of Rifampicin Resistant TB RR-TB (Witney et al, 2015).
2. Identification of Human Genetc Variations.
The genetic differences between and among populations is such that a particular gene may have a number of variants. This phenomenon is known as polymorphism (Higasa et al., 2016). Genome Sequencing can thus be done in healthy humans and can give genetic details (Frazer et al., 2009) .Phenotypic variations exist in the color of the eyes, skin, hair, facial traits and body build. However, scientists now have a wealth of molecular and genomic data at their disposal to work with. Humans variation is such that 85 percent exists inter population, while 15 percent occurs intra population (Frazer et al., 2009). The human genome is 99% the same in all individuals hence human DNA sequences are the same in every individual. However, interspecies variability in the DNA sequences occur and is used to distinguish one individual from another. The variations can occur in places not known to code for proteins –noncoding regions unless they are monozygotic twins (single zygote divided into two) (Deplancke et al., 2016). DNA profiling uses respective sequences that are highly variable called Variable Number Tandem Repeats (VNTRs), particularly Short Tandem Repeats (STRs) (Madsen et al., 2010). VNTR loci are very short and similar between closely related humans, but so variable that unrelated individuals are extremely different (Linacre et al., 2014).). Approximately 25% of neurodevelopmental disorders are due to SNVs (Kaminsky EB, et al). Single Nucleotide Polymorphism occurs when there is a change in a single base pair of a gene (Udogadi et al., 2020). They occur at the rate of 1 in every 300 nucleotides and 30 percent of variations have been due to SNPs. Other forms of polymorphism include Simple Sequence Length Polymorphisms (SSLPs), Restriction Fragment Length Polymorphism (RFLP) (Botstein et al., 1980). SSLPs include variations like variant minisatelite which are repeat sequences like Variable Number Tandem Repeats (VNTRs) (Bakhtiari et al., 2021) and microsatellite example of which is the Simple Tandem Repeats (STRs). In STRs, pieces of sequence are repeated several times in a row (Madsen et al., 2010). These variations do not affect cellular functions adversely and have no effect; however, they can alter

individual responses to drugs. Examples of diseases associated with single-gene disorders are Sickle Cell Disease (SCD), Cystic Fibrosis, Duchenne disorder and Hungtinton's disease. Others like diabetes, Cancer, Congestive Heart Disease, psychiatric disorder like schizophrenia and bipolar disease are polygenic or multifactorial due to interaction of the genes with the environment. Individuals with about 15–20% of those with intellectual disability and autism spectrum disorders have been discovered to have clinically relevant SV.

The variations can be continuous or quantitative.

Mechanism of SV formation

Human Structural Variation: mechanisms of chromosome rearrangements Brooke Weckselblatt and M. Katharine Rudd

Breakpoint junction sequencing reveals the mutational mechanisms and complexity of Structural Variation (SV),

Most human SV is generated via non-homologous end-joining and microhomology-mediated break-induced replication.

Non-allelic homologous recombination between segmental duplications, LINEs, and HERVs drives recurrent SV.

Chromothripsis is usually de novo, arises on paternal alleles, and in some cases is transmitted maternally.

3. Drug Reactions to current medications have been documented and found to be alarming.

ADR is defined as unpleasant, unintended, uncommon effects of a drug used in the prevention, diagnosis and treatment of a disease. It occurs in 2 to 7 percent of hospital patients. It is the fourth leading cause of death in the USA apart from heart disease, cancer and stroke (Hallberg et al., 2021).

The World Health Organization(WHO) has defined adverse drug reactions (ADRs) as "a response to a drug that is noxious and unintended and occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease or for the modification of physiological function." There are two traditional pharmacologic classifications. Type A generally referred to as "side effects," are dose-dependent and predictable reactions that account for 85–90% of all ADRs (Osanlou *et al*., 2018). Type B reactions generally referred to as "idiosyncratic" or "allergies," are not dose dependent and are unpredictable and account for approximately 10–15% of all ADRs. Patients sometimes misinterpret some side effects as allergies (e.g., diarrhea with amoxicillin/clavulanate), which may be perpetuated through the patient's medical record (Mockenhaupt, 2012). Majority of drugs have humans reacting to them.However, some have far reaching grave consequences than others (Lonjou *et al*, 2006). In most cases, it is difficult to identify such ADRs due to its similarities with viral infections and other dermatological incidences (Valeyrie-Allanore *et al*., 2007). The reactions are immune-mediated reaction to innocuous antigens produced in the presence of drugs. The skin being the most affected aside; other organs as the liver gives rise to reactions such as Stevens-Johnson Syndrome and

toxic epidermal necrolysis which are characterized by blisters. Serious cutaneous ADR Type B being the most severe and may cause hospitalization of individuals (Hung *et al*, 2006).

The ability to analyze the whole genome in an unbiased fashion as well as molecular typing in genomic studies have detected novel associations for Type B ADR reactions with Class I and II Human leukocyte Antigen HLA alleles as important biomarkers and predictors of drug induced liver injury DILI as well as susceptibility to hypersensitivity skin rash and myotoxicity (Kullak-Ublick., 2017). It is also possible to prevent ADRs with Pharmacogenetic screening (Plumpton *et al*, 2016).

4. Development of a personalized drug which will be suitable for an individual since his genetic makeup has been identified.

Existing drugs have been found to have a response rate of 40% and the detection of the disease is known via symptoms; therefore, diagnosis oftentimes is feasible after a disease has progressed and there are biochemical changes in the body (Hely *et al*, 2008). Whole Genome Sequencing however allows the prediction of diseases in the future based on the genomic composition (Lohmann et al., 2014). The ability to understand the molecular basis of these diseases, determine the 3 percent of protein coding genes, determine the variations in these genes will ultimately help individuals to take informed decision about their health with respect to change of lifestyle or diet (Varsan, 2006). Candidate genes have been used before; however, a lot of other things have to be taken into cognizance for instance the family and the genetic makeup of the other family members. Multidisciplinary approach is essential in the interpretation of genomic data for an individual and clinicians must understand the role of genomics in their areas of specialty to enable precise diagnosis (Ormondroyd *et al*, 2018). Examples include Tolvarptan used in the treatment of autosomal dominant polycytic kidney disease from the germline. Olaparib used in the treatment of platinum sensitive, BRCA mutation and peritoneal cancer (Brittain *et al*, 2017).

Development of novel drugs that can target specific part of the genome.

Identification of suitable targets has become pertinent due to a lot of factors that borders on improper target selection (Payne *et al*, 2007). Target selection is a complex process that involves adequate understanding of the association between the drug and target, the biological function of the target, druggability and appropriate determination of its use and relevance in the clinical setting (Knowles *et al*, 2003). The issue of mutant variants or development of multidrug resistance strains has jeopardized the efficacy of existing drugs and calls for selection of novel targets. Selected novel drug targets must meet with the criteria of inhibiting a major protein or pathway that is essential to the organism and thus selectively toxic but safe in the human host. (Crowther *et al*, 2010)The target must also be validated as follows (i) Biologically relevant that is the genes encodes for proteins that function as an enzyme, receptor, or transporter (ii) Druggable

attributes should be known or predicted (iii) Potential therapeutic agents should be predicted. The rationale is that conserved genes often fulfill essential functions (4). A conserved single copy gene that lacks close matches in the same genome is likely to be indispensable. WGS therefore allows identification of protein coding genes that are useful and can act as possible novel targets discovery for novel drugs (Gashaw *et al*, 2011).

6. Identification of novel targets for a particular disease with novel ligands being developed.

7. Molecular basis of diseases can be better understood; early detection allows application of novel therapies to combat the disease (Hopkins, 2008)

8. Early identification of biomarkers of certain diseases via genome sequencing can pave way for a proactive means of prevention via change of lifestyle which may prevent the disease from progressing (4). WGS enables identification of biomarkers which can be used to predict the future occurrence of a disease like cancer in oncology for example ECFR exon 19 deletions or exon 21 (L858R) substitution(4) mutations is an indication of NSCLC or KRAS mutation negative or KRAS and NRAS mutation negation negative is an indication of colorectal cancer. ERBB2 amplification on HER2 expression is also an indication of breast cancer while ERBB2 amplification can be an expression of gastric cancer.

9. Ancestry of individuals can be determined.

WGS has made it possible to enable individuals who have been adopted into families to trace their ancestry via the submission of their biological samples where their DNA can be extracted, analysed and compared to a database to find matches . Families have been united and siblings have been able to locate themselves via the WGS .

10. Forensic analysis is made possible and solution of crimes enhanced.

The deciphering of the human genome in the Human Genome Project and the use of whole genome sequencing has been able to unravel the mysteries that surround the genome (Ng et al., 2010). Humans have been found to be different despite the fact that 99.5% of their genomes are similar. The coding regions which are interspersed by the non coding regions have been deciphered (Schneider, 1997). Individual differences have been found in the Short Tandem Repeats (STRs) of the non coding regions of the genome (Wyner, 2020). These regions have served as markers in low quantities of DNA and degraded DNA hence its applicability in the field of forensics to identify a criminal from the DNA left at a crime scene or the victim of a crime (Udogadi *et al*, 2020). Single Nucleotide Polymorphism SNPs can also be used when there is degraded DNA as in the case of fire disaster victims (Linacre *et al*, 2014). Y-chromosome STR can also be used in the case of sexual assault to identify the victim. The limitations of STRs and SNPs can be overcome by DNA phenotyping or ancestry (Linacre *et al*, 2014).

13. Genomics and WGS can be applied to decipher individuals at risk of hereditary diseases

WGS has been useful to identify individuals at risk of heredictary diseases like Diabetes, early onset dementia, cancer, leukaemia, beta thallasemia and many rare diseases (Diamandis, 2014). Preventive measures can be made early when detected and early intervention can help to prevent the onset of the diseases (Prokopenko *et al*, 2022).

As sequencing costs continue to decrease and computer resources expand, WGS analysis for cancer genome research and clinical utilities will become more common and more sophisticated. Cancer WGS provides abundant information to understand the biology underlying the cancer genome and the function of unexplored noncoding regions and SV in the human genome. There is much potential for transcriptional or functional consequences of SV and noncoding mutations, and they should be further explored by integrative analysis of RNA-Seq and multi-omics analysis with DNA methylation data,78 protein expression data, and chromatin structure85 or epigenome data to interpret mutational consequences and to understand the biology and immunology of cancer. Taking into account the diversity of cancer genomes and phenotypes, interpretation of the mutational data from cancer WGS will also require the analysis of much more WGS data and integration with multi-omics data, functional data, immuno-genomic data and clinic-pathological data in a larger sample set.

Conclusion

The availability of Whole Exome and Whole Genome Sequencing has drastically impacted genetic diagnostics, and the clinical genetics specialty is undergoing rapid development. Genetic diagnostics was until recently limited to investigations of chromosome aberrations by karyotyping or array analysis, and gene by gene sequencing. Consequently, a strong focus has been on conditions like unclear malformation syndromes and intellectual disability, together with selected monogenic disease groups where a limited number of underlying disease genes have been defined. Genomics fundamentally changes this scenario. Around 4200 different monogenic disease genes are currently known, causing conditions that present across all clinical disciplines, at all ages, and ranging from insidious, chronic, to dramatically acute diseases. The possibility to incorporate WGS in the diagnostic workup across these vastly different clinical situations provides tremendous opportunities, but also poses challenges.

WGS's importance in the discovery of disease gene has opened up development of novel drugs in the rare disease area via recombinant enzymes/vaccines, identification of drug targets, novel development of small molecules as drugs, sense/antisense techniques, gene therapy, and genome editing/cell therapies implying the genetic basis of rare familiar disease 4-6 and explains novel disease biology. Informed health decisions can be made by seemingly healthy individuals whose genome are sequenced by WGS. The genetic findings can enable individuals guard against the onset of future disease.

Therefore, accurate diagnosis through WGS can improve therapeutics such that personalized medicine can be developed in the near future.The limitations are centered on ethical considerations.

References

Bakhtiari, M., Park, J., Ding, Y. C., Shleizer-Burko, S., Neuhausen, S. L., Halldórsson, B. V.,Bafna, V. (2021). Variable number tandem repeats mediate the expression of proximal genes. *Nature communications*, 12(1): 1-12

Barton, N. H., Keightley, P. D. (2002). Understanding quantitative genetic variation. *Nature Reviews Genetics*, 3(1): 11-21.

Botstein, D., White, R. L., Skolnick, M., Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, 32(3): 314.

Brittain, H. K., Scott, R.,Thomas, E. (2017). The rise of the genome and personalised medicine. *Clinical medicine*, 17(6): 545–551.

Crowther GJ, Shanmugam D, Carmona SJ, Doyle MA, Hertz-Fowler C, Berriman M, Nwaka S, Ralph SA, Roos DS, Van Voorhis WC, Agüero F (2010). Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. *PLoS neglected tropical diseases*, 4(8):804.

Dancey, J. E., Bedard, P. L., Onetto, N., Hudson, T. J. (2012). The genetic basis for cancer treatment decisions. *Cell*, 148(3), 409-420.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7): 499-510.

Diamandis, E. P. (2014). Present and future of cancer biomarkers. Clinical Chemistry and Laboratory Medicine (CCLM), 52(6):791-794.

Eichler, E. E., Nickerson, D. A., Altshuler, D., Bowcock, A. M., Brooks, L. D., Carter, N. P., and Human Genome Structural Variation Working Group. (2007). Completing the map of human genetic variation: A plan to identify and integrate normal structural variation into the human genome sequence. *Nature*, 447(7):141- 161.

Fernald GH, Capriotti E, Daneshjou R. Bioinformatics challenges for personalized medicine. *Bioinformatics* 27 (13): 1741-8, 2011.

Fernandes, G. F. D. S., Man Chin, C., Dos Santos, J. L. (2017). Advances in drug discovery of new antitubercular multidrug-resistant
compounds. *Pharmaceuticals*, 10(2):51.

Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4): 241-251.

Gashaw, I., Ellinghaus, P., Sommer, A., & Asadullah, K. (2011). What makes a good drug target? *Drug discovery today*, 16(23-24):1037-1043.

1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, 526(7571): 68-75.

Gibbs, R. A. (2020). The human genome project changed everything. *Nature Reviews Genetics*, 21(10): 575-576.

Griffiths, P. E. (2001). Genetic information: A metaphor in search of a theory. *Philosophy of Science*, 68(3): 394-412.

Hallberg, P., Yue, Q. Y., Eliasson, E., Melhus, H., Ås, J., & Wadelius, M. (2020). SWEDEGENE—a Swedish nation-wide DNA sample collection for pharmacogenomic studies of serious adverse drug reactions. *The pharmacogenomics journal*, 20(4): 579-585.

Higasa, K., Miyake, N., Yoshimura, J., Okamura, K., Niihori, T., Saitsu, H and Matsuda, F. (2016). Human genetic variation database, a reference database of genetic variations in the Japanese population. *Journal of human genetics*, 61(6): 547-553

Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature chemical biology*, 4(11): 682-690.

Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, Moreno-De-Luca D, Moreno-De-Luca A, Mulle JG, Warren ST, Richard G. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genetics in Medicine*, 13 (9):777–784

Knowles, J., & Gromo, G. (2003). Target selection in drug discovery. *Nature Reviews Drug Discovery*, 2(1): 63-69.

Linnarsson, S. (2010). Recent advances in DNA sequencing methods–general principles of sample preparation. *Experimental cell research*, 316(8):1339-1343.

MacConaill, L. E. (2013). Existing and emerging technologies for tumor genomic profiling. *Journal of Clinical Oncology*, 31(15):1815.

Madsen, B. E., Villesen, P., & Wiuf, C. (2010). Short tandem repeats and genetic variation. In Genetic Variation (pp. 297-306). Humana Press, Totowa, NJ.

Mullighan CG (2013) Genomic characterization of childhood acute lymphoblastic leukemia. *Semin Hematol* 50 (4): 314-24.

Mullighan CG, Goorha S, Radtke I, et al.: Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446 (7137): 758-64, 2007.

Mullighan CG, Su X, Zhang J, et al.: Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. (2009). *N Engl J Med* 360 (5): 470-80.

Myler, P. J. (2008). Searching the Tritryp genomes for drug targets. Drug Targets in Kinetoplastid. *Parasites* 133-140.

Nerlich, B., Dingwall, R., Clarke, D. D. (2002). The book of life: How the

completion of the Human Genome Project was revealed to the public. *Health*, 6(4): 445-469.

Ng, P. C., Kirkness, E. F. (2010). Whole genome sequencing. Genetic variation pp 215-226.

Osanlou, O., Pirmohamed, M.,Daly, A. K. (2018). Pharmacogenetics of adverse drug reactions. *Advances in Pharmacology*, 83, 155-190.

Payne, D. J., Gwynn, M. N., Holmes, D. J., & Pompliano, D. L. (2007). Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature reviews Drug discovery*, 6(1): 29-40.

Petty, N. K., Zakour, N. L. B., Stanton-Cook, M., Skippington, E., Totsika, M., Forde, B. M., and Beatson, S. A. (2014). Global dissemination of a multidrug resistant Escherichia coli clone. *Proceedings of the National Academy of Sciences*, 111(15): 5694-5699.

Plenge R.M., Scolnick E.M., Altshuler D (2013). Validating therapeutic targets through human genetics. *Nat. Rev. Drug Discov*.,12:581–594.

Prokopenko, D., Hecker, J., Kirchner, R., Chapman, B. A., Hoffman, O., Mullin, K., Tanzi, R. E. (2020). Identification of novel Alzheimer's disease loci using sex-specific family-based association analysis of whole-genome sequence data. *Scientific reports*, 10(1): 1-9.

Reshmi SC, Harvey RC, Roberts KG, et al.: Targetable kinase gene fusions in high-risk B-ALL: a study from the Children's Oncology Group. *Blood* 129 (25): 3352-3361, 2017. [PUBMED Abstract]

Rizzo JM, Buck MJ: Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer Prev Res (Phila)* 5 (7): 887-900, 2012. [PUBMED Abstract] 11

Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74 (12): 5463-7, 1977. [PUBMED Abstract] 9

Schneider, P. M. (1997). Basic issues in forensic DNA typing. *Forensic science international*, 88(1): 17-22.

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature methods*, 5(1): 16-18.

Sirot, E. J., van der Velden, J. W., Rentsch, K., Eap, C. B., Baumann, P. (2006). Therapeutic drug monitoring and pharmacogenetic tests as tools in pharmacovigilance. *Drug Safety*, 29(9):735-768.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, 20(8): 467-484.

Wang, F., Huang, S., Gao, R., Zhou, Y., Lai, C., Li, Z., Liu, L. (2020). Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell discovery*, 6(1): 1-16.

Witney, A. A., Gould, K. A., Arnold, A., Coleman, D., Delgado, R., Dhillon, J., Hinds, J. (2015). Clinical application of

whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases. *Journal of clinical microbiology*, 53(5): 1473-1483.

Witney, A. A., Cosgrove, C. A., Arnold, A., Hinds, J., Stoker, N. G., Butcher, P. D. (2016). Clinical use of whole genome sequencing for Mycobacterium tuberculosis. *BMC medicine*, 14(1): 1-7.

Wyner N., Barash M. Mc Nevin D. (2020). Forensic Autosomal Short tandem Repeats and their Potential Association with Phenotype. *Frontiers in Genetics,* 11(6): 64-80

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., Eng, C. M. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *New England Journal of Medicine*, 369(16): 1502-1511.

Zhao, S., Kumar, R., Sakai, A., Vetting, M. W., Wood, B. M., Brown, S., Jacobson, M. P. (2013). Discovery of new enzymes and metabolic pathways by using structure and genome context. *Nature*, 502(7473): 698-702