Proceedings of the 2nd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/32/20240802

Current research progress of whole genome sequencing in practical

Z Huang

Queen Mary University of London, IQ East Court, 450 Mile End Road, E1 4GG.

hzssally@gmail.com

Abstract. Gene sequencing, a vital technique nowadays, detects patients' diseases and helping with diagnosis. Whole genome sequencing (WGS) is one member of this family, which mainly functions in analysing participant's gene sequences in clinical diagnosis for the purpose of indicating therapeutic interventions. This essay compares WGS and other sequencing methods, such as whole exome sequencing (WES), which comprises both positive and negative sides of this sequencing test for evaluation. And point out the help of WGS in certain diseases. This paper serves as a synopsis of some WGS-related themes as well as potential directions in the future for technical advancement.

Keywords: whole genome sequencing; whole exome sequencing; tuberculosis.

1. Introduction

WGS differentiates it from the other forms of sequencing because it analyses the entire gene and provides complete information instead of partial information. Another distinction is that it has the greatest resolution rate. It can be used to examine a microbe's potential for pathogenicity and antibiotic resistance, as well as for clinical diagnosis. It used to be enormously expensive compared to the other methods because it examined an organism's entire genome group rather than a small piece sample. However, with the rapid pace of technological development in recent decades, it is now usable not only in institutions with abundant resources but also in constrained contexts. WGS can identify risk exposure factors based on pathology; one such component is pathogen closeness, which can be used to anticipate and prevent epidemics. It might also discover the existence of genes that code for antimicrobial resistance (AMR) [1]. WGS is being used as a baseline for assessing data in fields like epidemiology and front-line clinical practise. However, WGS is not yet the ideal option; there are still certain areas that can be improved. Both the positive and negative aspects of WGS must be considered in order to enhance it or combine it with other techniques that will serve science more effectively, which calls for increased effort and scientific cooperation in the future, if possible.

2. Comparison of whole genome sequencing

2.1. Whole Genome Sequencing and Whole Exome Sequencing

The exome capture technology family includes whole exome sequencing (WES), which is solution based rather than array based. Biotinylated oligonucleotide probes (baits) and shattered DNA fragment

DOI: 10.54254/2753-8818/32/20240802

samples are the two main components contributing to WES working. They are chosen to hybridise into genome-targeted zones to facilitate the operation principle of WES. The biotinylated probe is intended to connect with magnetic streptavidin beads before refining the targeted DNA specimen and expanding the result by polymerase chain reaction (PCR). The material will then be sequenced and prepared for bioinformatic examination to take place. It is the sequencing technique that is now clinically accessible and differs from WGS by concentrating solely on regions that encode proteins and exons. Exons comprise only 1-2% of the genetic material in genome, compared to over 20,000 genes, where the percentage varies depending on the species. As a result, utilising WES is typically regarded as being more cost-effective than using WGS. Additionally, WES might be preferred for utilisation by the extremely constrained intronic portion of the genome [2].

Exome sequencing (ES) is described at some length in a scientific research paper. The passage discusses the positive and negative aspects of ES. Because it can identify disease-gene variants in samples without associating them with prior diseases, ES makes it easier for clinicians to provide comprehensive treatment plans because they are no longer limited to individual gene sequencing on a case-by-case basis. Due to this, ES is now the primary diagnostic technique for inborn metabolic diseases (IMD) and rare genetic diseases (RGD) [3]. However, ES has several drawbacks as well, including data cooperation and repeated analyses of continuous statistics. For instance, it may have taken 1-3 years for a reanalysis to improve the diagnostic by 3-15% because there were not enough clues for the initial analysis in the first place [4]. On a technical level, regulatory and splice sites, as well as the complex DNA rearrangements persist to limit ES. Short sequences with proximal matching sites are common in ES analyses, although in some circumstances translocation elements fail to map with the actual location of the genome. Frameshifting insertions or gene regulation due to mobile elements could be the case that results in migration and alters the genome's copy frequency [5]. Despite certain problems, ES has significantly increased the population accuracy and diagnosis, becoming the front-tier clinical tool for use [6].

One study compared WES with WGS and showed that while both can be affected by reading coverage and callers, which result in varying repeatability, WES is additionally influenced by the size of the insert fragment, the grade of the global imbalance, and the content of the genomic copy. Particularly, WES has superior coverage performance. However, coverage at its targeted locations was inconsistent. Likewise, WES exhibits more artefacts as a result of lab processing, which results in greater differences between runs. Because WGS offers more uniform coverage and is, therefore, more reproducible, WES is less active when comparing the results of different runs. They have also discovered intracenter variance, but WES dominates intercenter variation [7].

Additionally, since only a tiny portion of the genome has been sequenced in the examination, some considerable modifications in structure may go undetected when making a comparison with WES. On 2520 patients with carcinomas, 2399 of whom have various metastatic tumours, researchers in the Netherlands used WGS. Up to 70 million cancer genome alterations can be identified, including both minor localised modifications and significant structural abnormalities such as short insertions, deletions, point mutations, and chromosomal regions, were found in this study [8].

2.2. Whole genome sequencing and other comparison

In one investigation, two alleged nosocomial explosions are invested. They originated from E. faecium, which is vancomycin-resistant with E. cloacae; and it is available to resist carbapenem. WGS was then compared with other conventional microbiology assays. The outcome demonstrates that WGS is more advanced than conventional diagnostic methods in identifying carbapenem resistance in a number of gram-negative bacteria that found in machinery, and that WGS could reliably discriminate between outbreak and non-outbreak isolates. They replace multiple current tests with just a single WGS test that examines both the forward gearbox and machinery of resistance. Their research indicates the ability to detect carbapenem or other resistance as well as genotype and phenotype in results. Through comparison, the WGS data demonstrate prior comprehensive and accurate information intended to refute/validate the outbreak in society and the medical community. Additionally, the WGS platform can identify some

DOI: 10.54254/2753-8818/32/20240802

species involved in the spread of multidrug resistance, which aids in diagnostics and enhances public health. It is more proactive and crucial to discover abnormal isolates faster before they produce a catastrophic epidemic across the community since the switch to employing WGS could give the ability to react in real-time when the virus is quickly spreading across [9].

A research project compared WGS with standard-of-care (SOC) diagnoses from 1302 samples with metastatic cancer and 1200 consecutive tumour patients. Two results were contrasted and discussed side by side. WGS has a 70% rate of accuracy and is effective in treating those patients who have metastatic disease as part of routine clinical care because it successfully profiles the majority of whole genomes. Of these, 1216 of them had cancer cells. In the study, researchers discovered that they were becoming increasingly comfortable with the additional diagnostic WGS results and their interpretation, particularly in a number of complex diagnoses. 49 previously unknown pathogenic mutations in the germline of cancer susceptibility were found by WGS, demonstrating the added utility of germline diagnosis. Then, WGS can facilitate additional treatment optimisation. Additionally, WGS with 71% acknowledges numerous therapeutic choices and the distribution of valuable biomarkers from clinical trials. The majority of biomarkers, including the population that was specifically targeted in the patients' sequencing panels, were not identified by the SOC's current diagnostic methodology [10]. Further, WGS provides a strong foundation for the development of the healthcare system and is necessary to demonstrate precise therapy at its best. Tumour thorough characterisation in the genome of merging clinical phenotyping information [11].

3. Whole Genome Sequencing in tumour and tuberculosis issue

Here are a few examples of when WGS has been employed to recognise and classify illnesses. The Mycobacterium Tuberculosis Complex (MTBC) bacteria cause around 10 million new cases of tuberculosis (TB) each year, making them the most lethal infectious illness in the world. WGS, in this case, was built adequately for the individual small MTBC chromosomal genome (4.4Mb) and may lead the full component of regulating TB from assessment to source research [12], and cure by reconstruction of one organism's total DNA genome sequence utilising the platform of DNA sequencing [13]. The application of WGS-based TB treatment methods for the general community followed a swift transition from the research to the therapeutic phase. The WHO has started using WGS to track drug resistance and evaluate the efficacy of drug susceptibility testing (DST) technology for analysing the drug resistance-related mutations [14]. Additionally, WGS has advanced to the point that it can be used for diagnostic purposes in hospitals and health promotion in communities, and in certain instances, it has taken the place of phenotypic assessment for first-line drugs. The main area of strain type research, molecular epidemiology, is steadily establishing itself as the benchmark for WHO drug resistance monitoring [15].

4. Whole Genome Sequencing in cancer

Next-generation sequencing (NGS), in combination with other tests like cytogenetics for clinical oncological information, is a current laboratory established method used in detecting the pathology of genomics for cancer patients. Genome analysis is crucial throughout their entire process including therapy and diagnosis. However, using this method necessitates a variety of tests and proof. WGS is capable of characterising a variety of kinds in a thorough and objective manner (the best comprehensive characterization of both CNVs and structural variants, or SVs, which are crucial in the somatic change of cancer genes) [16].

5. Risk and difficulties in practical

There have already existed some excellent instances of WGS application, thanks to major technological developments that have made it feasible to use in clinics as therapeutic procedures for treating complicated disorders such as breast cancer. However, there are still certain problems that need to be resolved in the areas of analysis, outcome interpretation, quality control, and other areas in order to improve the accuracy of WGS adoption in hospitals. In order to improve the standard of patient care,

DOI: 10.54254/2753-8818/32/20240802

WGS's potential realisation in clinics should be strictly implemented before being applied to the general population [17]. WES and targeted arrays are able to assess CNV using similar read counts, but WGS has the resolution to provide unbiased analysis for microdeletions/microamplifications yet WES continues to be constrained by the depth of sequencing bias challenge. As a result, the outcomes of CNV calls are less accurate and the proportion of false detection is higher [18]. By identifying new genetic drivers of cancer biology, WGS has significantly enhanced the results for patients and our understanding of medical conditions. WGS has become the clinical standard for properly and comprehensively cataloguing CNVs and SVs. Patients were subjected to a single wet bench test, a targeted panel test, chromosomal analysis, and fluorescence in situ hybridization (FISH) in the study. However, for improved detection of new fusion partners, WGS displays breakpoints with higher resolution.

The accuracy of the patient's genome base-calling, which requires rigorous optical sensor platform base recognition, is a challenge in the sequencing process. The platform features a custom programme that the vendor uses to deliver base-calling together with a different programme from a different source. Data can be integrated more accurately into a consensus sequence by resequencing the DNA samples because this lowers the fraction of error and improves coverage, but the cost is higher.

WGS's shift from concept to practise is also challenging. Despite the fact that WGS seems to be helpful in tackling a number of essential issues like infection control, researching drug resistance, and promoting the best possible care for patients. However, the implementation in hospitals may take some time since several fundamental issues must be resolved before WGS can be used for clinical microbiological diagnosis. For instance, M. tuberculosis is a clonal species that, as it undergoes horizontal gene transfer (HGT), may project a rough natural categorization phylogenetically [19]. Organisms with normal HGT and additional genomes cannot be included in the existing classification system because WGS yields information that is too specific and each individual genome is unique. The current classification system would be overturned, creating a significant problem. Furthermore, occasionally the test results from WGS are comparable to those from other conventional phenotyping techniques rather than being superior [20]. S. aureus and M. tuberculosis are the most feasible and early applications, with few auxiliary genomes characterised; this may not be a coincidence. High similarities can also be shown in other WGS cases, such as streptococcal pathogens [21]. One benefit of WGS is tracking epidemics, but this is often retrospective and takes place about 5 years after the outbreak was first noticed. It still takes a while to inform people about current events, and most of them are in the academic community, despite the fact that its analytical report can be issued in a shorter amount of time [22]. It is not clear when and how WGS will replace the existing universal principles in clinical microbiology. It faces obstacles like high expenses, a lack of bioinformatics expertise, and inadequate facilities seen in most hospitals. And crucial for establishing guidelines for bioinformatics protocols. The possibility of losing current knowledge in microbiology following a WGS shift is another danger to microbiologists [23].

6. Conclusion

In conclusion, investigations have shown that WGS-based diagnosis in the routine pathology practise is viable and could help with hospital decision-making [24]. With high accuracy, WGS finds pathogens in the sample population that were originally missed using traditional testing procedures, assisting in a more detailed diagnosis of the condition and improving the patient's treatment and rehabilitation even more. WGS has shown its benefits in comparison with other sequencing methods and illustrates practical uses in real-life diseases, showing its feasibility in use as a technique applied to participants. However, there are still some drawbacks to use this gene sequencing test method, which would cause some inconvenience. Therefore, for future evolution, a combination of various tests, for example, could be considered based on the development in technical field for better exploring the genes in many organisms. With the growth, we can see that the development is to achieve the aim that the test result could function maximally in the form of helping patients.

DOI: 10.54254/2753-8818/32/20240802

References

[1] Purushothaman S, Meola M, Egli A. Combination of Whole Genome Sequencing and Metagenomics for Microbiological Diagnostics. *International Journal of Molecular Sciences.* 2022 Aug 30;23(17):9834.

[2] Warr A, Robert C, Hume D, Archibald A, Deeb N, Watson M. Exome Sequencing: Current and Future Perspectives. *G3: Genes|Genomes|Genetics.* 2015 Jul 2;5(8):1543–50.

[3] Quentin T, Vitobello A, Frédéric Tran Mau-Them, Yannis Duffourd, Agnès Fromont, Giroud M, et al. High efficiency and clinical relevance of exome sequencing in the daily practice of neurogenetics. *Journal of Medical Genetics.* 2021 Mar 5;59(5):445–52.

[4] Basel-Salmon L, Orenstein N, Markus-Bustani K, Ruhrman-Shahar N, Kilim Y, Magal N, et al. Improved diagnostics by exome sequencing following raw data reevaluation by clinical geneticists involved in the medical care of the individuals tested. *Genetics in Medicine: Official Journal of the American College of Medical Genetics [Internet].* 2019 Jun 1;21(6):1443–51. Available from: https://pubmed.ncbi.nlm.nih.gov/30377382/

[5] Kim J, Hu C, Moufawad El Achkar C, Black LE, Douville J, Larson A, et al. Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease. *New England Journal of Medicine.* 2019 Oct 24;381(17):1644–52.

[6] Wortmann SB, Oud MM, Alders M, Coene KLM, van der Crabben SN, Feichtinger RG, et al. How to proceed after "negative" exome: a review on genetic diagnostics, limitations, challenges and emerging new multi‐omics techniques. *Journal of Inherited Metabolic Disease.* 2022 May 4;

[7] Xiao W, Ren L, Chen Z, Fang LT, Zhao Y, Lack J, et al. Toward best practice in cancer mutation detection with whole-genome and whole-exome sequencing. *Nature Biotechnology.* 2021 Sep;39(9):1141–50.

[8] McPherson A. Whole genome sequencing drives progress in cancer [Internet]. *Genomics Education Programme.* 2020. Available from: https://www.genomicseducation.hee.nhs.uk/blog/whole-genome-sequencing-drives-progress-in-cancer

[9] Reuter S, Ellington MJ, Cartwright EJP, Köser CU, Török ME, Gouliouris T, et al. Rapid Bacterial Whole-Genome Sequencing to Enhance Diagnostic and Public Health Microbiology. *JAMA Internal Medicine [Internet].* 2013 Aug 12;173(15):1397. Available from: https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/1726980

[10] Cobain EF, Wu YM, Vats P, Chugh R, Worden F, Smith DC, et al. Assessment of Clinical Benefit of Integrative Genomic Profiling in Advanced Solid Tumors. *JAMA Oncology [Internet].* 2021 Apr 1;7(4):525–33. Available from: https://jamanetwork.com/journals/jamaoncology/fullarticle/2776760

[11] Priestley P, Baber J, Lolkema MP, Steeghs N, de Bruijn E, Shale C, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature.* 2019 Oct 23;575(7781):210–6.

[12] Satta G, Lipman M, Smith GP, Arnold C, Kon OM, McHugh TD. Mycobacterium tuberculosis and whole-genome sequencing: how close are we to unleashing its full potential? *Clinical Microbiology and Infection.* 2018 Jun;24(6):604–9.

[13] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature.* 1998 Jun;393(6685):537–44.

[14] Zignol M, Cabibbe AM, Dean AS, Glaziou P, Alikhanova N, Ama C, et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study. *The Lancet Infectious Diseases.* 2018 Jun;18(6):675–83.

[15] Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues. *Nature Reviews Microbiology [Internet].* 2019 Jun 17;17(9):533–45. Available from: https://www.nature.com/articles/s41579-019-0214-5

[16] Cortés-Ciriano I, Lee JJK, Xi R, Jain D, Jung YL, Yang L, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nature Genetics [Internet]*. 2020;52(3):331–41. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7058534/

[17] Chrystoja CC, Diamandis EP. Whole Genome Sequencing as a Diagnostic Test: Challenges and Opportunities. *Clinical Chemistry*. 2013 Nov 13;60(5):724–33.

[18] Zhao EY, Jones M, Jones SJM. Whole-Genome Sequencing in Cancer. *Cold Spring Harbor Perspectives in Medicine*. 2018 May 29;9(3):a034579.

[19] Coll F, McNerney R, José Afonso Guerra-Assunção, Glynn JR, João Perdigão, Viveiros M, et al. A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. *Nature Communications*. 2014 Sep 1;5(1).

[20] Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, et al. Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. *Nature Communications [Internet]*. 2015 Dec;6(1). Available from: http://www.nature.com/articles/ncomms10063

[21] Metcalf BJ, Chochua S, Gertz RE, Hawkins PA, Ricaldi J, Li Z, et al. Short-read whole genome sequencing for determination of antimicrobial resistance mechanisms and capsular serotypes of current invasive Streptococcus agalactiae recovered in the USA. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases [Internet]*. 2017 Aug 1;23(8):574.e7–14. Available from: https://pubmed.ncbi.nlm.nih.gov/28257899/

[22] Kwong JC, Lane CR, Romanes F, Gonçalves da Silva A, Easton M, Cronin K, et al. Translating genomics into practice for real-time surveillance and response to carbapenemase-producing Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak. *PeerJ*. 2018 Jan 3;6:e4210.

[23] Balloux F, Brønstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, et al. From Theory to Practice: Translating Whole-Genome Sequencing (WGS) into the Clinic. Trends in Microbiology [Internet]. 2018 Dec;26(12):1035–48. Available from: https://www.sciencedirect.com/science/article/pii/S0966842X18301768

[24] Samsom KG, Schipper LJ, Roepman P, Bosch LJ, Lalezari F, Klompenhouwer EG, et al. Feasibility of whole-genome sequencing-based tumor diagnostics in routine pathology practice. *The Journal of Pathology [Internet]*. 2022 Oct 1;258(2):179–88. Available from: https://pubmed.ncbi.nlm.nih.gov/35792649/