

Editorial

Whole genome sequencing and metagenomics for public health use in Ethiopia

Eshetu Lemma

The idea of this article is not to address all issues in whole genome sequencing (WGS); rather it is to focus on major areas of this growing science that will fit the existing local needs. Suggestions are also given for what purpose WGS is to be used at the moment even though, this will continue to be updated as the technology continues to evolve at rapid pace. Stress is also made on issues of capacity development in health science research, research institutions and universities to ensure the uptake and sustainability of this cutting edge health research capability for the present and the coming generations.

The genome or genetic material of an organism is made up of DNA. Each organism is composed of four DNA bases (ATCG). Whole genome sequencing is a laboratory procedure that determines the sequence of the bases in the whole genome of an organism whether it is human, animal, microorganism or plant cells is known as WGS.

WGS technique has brought explosive development in medical research in interims of unraveling complex mysteries of life. The technique is comparatively user-friendly that can be of value both in developing and developed countries. Sequencing method was initially developed by Sanger et al. (1977) and over several decades it has gone through substantial changes to result into efficient bench top technology for research and laboratory use (Kwong et al. 2015).

WGS will provide information as never before in detection and characterizing of disease causing agents. Normally, genomic tests which involve human DNA can have substantial ethical ramifications; however tests based on pathogen analysis are free from these. This is because in the laboratory analysis of the genetic test science, it is the pathogen's genetic material in question rather than the patient's DNA which are tested. A traditional culture technique for example, is very slow and labour-intensive. It takes several days and in some cases (in tuberculosis) several weeks and months to identify the cause of an infection and establish which treatments could be effective. WGS could transform the way these laboratories function, helping clinicians to manage patients with suspected infections faster and more effectively. On the other hand, metagenomics refers to both a research technique and

research field and the field can be defined as the genomic analysis of microbial DNA in their natural environment, including skin, mouth, and gut samples as well as soil, water and plant samples to understand the effect of microbial communities in health. Metagenomics tools enable the population analysis of un-culturable or previously unknown microbes. This is important as only around 1-2% of bacteria can be cultured in the laboratory (Wade 2002). The ability to identify microbes without a priori knowledge of what a sample contains is opening new doors in disciplines like microbial ecology, virology, microbiology, environmental sciences and biomedical research. The combination of both WGS and metagenomics has become a powerful tool for generating novel hypotheses in human disease (Kwong et al. 2015).

Using WGS strain typing has been improved with far more discriminatory power provided by culture or by other molecular techniques including rRNA gene based analysis. WGS technique is providing increasingly accessible means for tracking disease outbreaks and sequence-based information of drug resistance as well as virulent gene tracing even though some limitations are observed (Taylor et al. 2015).

Currently, WGS cannot be used in medical laboratories for the detection of disease causing organism for routine purposes. Given the need and the standard of local capacity WGS can have increasing use in Central and Regional/Medical School Laboratories for the purpose of epidemiological and related studies. This approach will initiate the capacity development of laboratory techniques of WGS in institutions together with enhanced manpower development in bioinformatics (Gupta et al. 2014) that can handle the huge amount of sequence data that is generated at national level. Using exiting automated technology of WGS huge sequence data can be generated, however, such large data will not be helpful for public health unless properly analyzed and interpreted by bioinformatics scientists and related experts. Therefore adequate training in bioinformatics both in number and quality is very much essential.

Given the potential of WGS in the control of infectious diseases (and non-infectious diseases for that matter) the wider application of WGS in

Ethiopia is very much within sight if we invest on existing and upcoming young researchers to exploit their avant-garde stamina for the good of the country in earnest.

References

- Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. (2014). ARG-ANNOT, a new Bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents Chemotherapy*, 58: 212–220.
- Kwong JC, McCallum N, Sintchenko V & Howden BP (2015). Whole genome sequencing in clinical and public health microbiology. *Pathology*, 47(3):199–210.
- Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, et al. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, 265:687-695.
- Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ, Medus C, et. Al. (2015). Characterization of food borne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. *Journal of Clinical Microbiology*, 53:3334–3340.
- Wade W (2002). Unculturable bacteria—the uncharacterized organisms that cause oral infections. *Journal of the Royal Society of Medicine*, 95: 81-88.