OPINION

Sequencing or not sequencing

Ji-Dong Gu
Environmental Science and Engineering Research Group, Guangdong Technion - Israel Institute of Technology, 241 Daxue Road, Shantou, Guangdong 515063, China
Guangdong Provincial Key Laboratory of Materials and Technologies for Energy Conversion, Guangdong Technion - Israel Institute of Technology, 241 Daxue Road, Shantou, Guangdong 515063, China

Abstract: Technology has been playing a very important role in the advancement of science, and big breakthroughs in science are dependent upon technologies available from time to time. Between science and technology, a good understanding is required so that available technologies can be used most efficiently and effectively to solving the targeted questions in scientific research. Genome nucleotide sequencing as an available technology after several generations of the technical improvement is capable of extract the full nucleotide sequences of any single organism or a mixed population of organisms in a given sample. This technology has a major important role in medicine and health science for diagnosis of the microorganisms and early detection of genetic diseases. In microbial ecology and applied microbiology, it is now fashionable to sequence the microbial community, the microbiome, so that the community composition and diversity are described from air, soil, water, sludge to skin and gut of animals, including humans. This routine practice, too frequently used by far too many, neither provide any significant insights of knowledge nor unravel the fundamental science in any way, but it serves as a cosmetic decoration for publication mostly. Because of this, not sequencing or sequencing less is the view expressed here to alert researchers for a better focus on the research questions, and the selection and implementation of the necessary physical science methods to obtain the solid research data that can advance science in a significant way. Publication driven experiments to produce more papers are eroding the essence of science, especially the pure spirit of science.

Keywords: Sequencing, scientific questions, science, publication, spirit of science

Correspondence to: Ji-Dong Gu, Environmental Science and Engineering Research Group, Guangdong Technion - Israel Institute of Technology, 241 Daxue Road, Shantou, Guangdong 515063, China; Guangdong Provincial Key Laboratory of Materials and Technologies for Energy Conversion, Guangdong Technion - Israel Institute of Technology, 241 Daxue Road, Shantou, Guangdong 515063, China; E-mail: jidong.gu@gtiit.edu.cn

Received: August 26, 2022; **Accepted:** December 2, 2022; **Published Online:** December 9, 2022
Citation: Gu, J.-D., 2022. Sequencing or not sequencing. *Applied Environmental Biotechnology*, 7(2): 1-4. http://doi.org/10.26789/AEB.2022.02.001

Discoveries in science can be dependent upon technology advancement, but definitely not every time in the history. After the structure of deoxyribonucleotide acid (DNA) was proposed by Watson and Crick in 1953, there had been no justification to verify such information from every organism or to find any possible slightest differences of DNA among different organisms, large or small. On this general subject, though a new claim was made on the replacement of the phosphate by arsenate on the DNA structure in a bacterium isolated from arsenic contaminated site (Wolfe-Simon et al., 2010), it has been quickly proven to be a false claim by the individual involved, not a major breakthrough in science. The message from this and other similar ones show clearly that fundamental discovery in science is important and will remain so for a long time before a new way of understanding can be established to replace or modify the previous one, e.g., Newton's gravity to Einstein's relativity. This has been discussed in the philosophy of scientific discovery by Thomas Kuhn (1996) and Karl Popper (1959) using different philosophical approaches with different ways of understanding.

After the discovery of 16S rRNA gene as a molecular biomarker for taxonomy and phylogenetic analysis of microorganisms, bacteria and archaea (Woese, 1987), commercial and technological innovations have made it possible to unravel the rich microbial community in any samples containing a mixture of multiple microorganisms not only possible, but also with great depth of the constituting members in the sample. Recently, DNA sequencing from pyrosequencing, Illumina sequencing to nanopore sequencing, coupling with development in bioinformatics for assembling, complex microbial community in various environmental samples can be delineated for composition and diversity analysis, and also metabolic potential (Caporaso et al., 2010). As a result, far too many reports have been published on such descriptive information of the microbiome in different environmental samples to claim the so-called uniqueness of the results, which are based on the samples or locations only. This has become a routine or accepted model in that, if such information or exercise has not been performed, the individuals involved are not doing proper research or science regardless of the significance of this information to the research question involved. By looking into the two ways of doing science mentioned here, it is apparent that commonality of the discovery with universal truth shall be the key or focus of scientific investigation for importance and significance, not the small variation or the locality of the samples (Arias and Omil, 2018; Yuan et al., 2020). It is also clear that technology application may advance the science to a significant extent for the first application, not so afterward in scientific research. Therefore, the question is clear to sequence less and less.

In this article, I would like to continue the discussion from my previous comments about too many insignificant papers are published in my view to the current research landscape and pitfalls with DNA sequencing technologies and applications in research. It seems that the current research practice of using sequencing as a required tool is locked into an accepted pattern, many young and new researchers simply follow the establishment and practice by the majority as the accepted rule without an individual judgement on its true value. This is very alarming for science, specifically in applied and general biological science subjects, e.g., applied microbiology and microbial ecology because the investment into research is not being used wisely and efficiently for the fundamental discoveries and advancement of science as expressed earlier (Gu, 2022).

Technology as a Double-edged Sword

Science has benefit greatly from technology advancement to allow new discoveries to be made. Analytical chemistry and physics serve their fundamental roles quantitatively to advance the scientific hypothesis and questions, and also to establish a new framework or understanding in a scientific way (Gao and Gu, 2021). But, over use of available technology is not rare in the history of science, especially in biological sciences. This phenomenon has been more apparent today due to the competitiveness of the scientific research career and also the nature of scientific investigation is no longer a personal endeavour by choice and passion truly or solely, but a job to secure so to support the living of oneself and the family (Gu, 2022). For example, in an environmental investigation on toxic pollutants in the sediment materials of a selective site, as many as can be analyzed elements from the periodic table are presented as results by including Fe, Na, Ca, Mg, etc. without any justification for choosing some of them with sound reasons. This shows a clear fact in that the authors do not have a good understanding or command on the nature of the issues or pollutants of environmental significance or in the samples, so the experiment, not an investigation, was a collection of everything that can be shown as results from the analysis without any scientific judgement from the beginning to design the study (Gu, 2020). With the available data, the information cannot be synthesized logically, but is presented as they are without any synthesis, selectivity or focus coherently in an article. When such submission has been sent out for reviewing, I can predict the outcome as a rejection for the manuscript.

In a similar way, DNA sequencing as a new technological tool has been used to generate the 16S rDNA sequences of bacteria and archaea or 18S or ITS of fungi for phylogenetic analysis to serve the classification and taxonomy in a productive manner for several decades. With further technological advancement, the same concept and technology can reveal the results of samples containing different microorganisms in natural or impact community of the samples. Currently, a ‘standard’ routine is to present the microbial composition data in terms of the different species with its abundance in the community, so a very long list of so many species including those without any taxonomic names or established knowledge, called novel naively, are described. The simple fact is that almost all environmental samples have the major bacteria as Proteobacteria, Bacteroides, Firmicutes, and Actinobacteria as the most abundant members in the community (Chen and Gu, 2022). We shall feel very uncomfortable with the simple sentence structure on the variations of names or label, and then the corresponding abundance in relative percentage as results and finding. In contrast, a similar trend and practice is also evident in materials science to make new materials through mixing different composition of chemicals, heating them up, and then conducting a suite of routine characterization using the instruments available, physical and chemical. The data are then combined into paper for publication with the higher efficiency in performance comparing with the former using sequencing technology. When science is being conducted in such a manner, this is alarming because this is not the way to train the next generation for science for one and, at the same time, the system of training is eroding itself to disable the future generation of scientists in their own reasoning and competency and also creativity (Gu, 2016). I wonder if any of these authors asks themselves what is the meaning of such publications and the information for? What is the actual contribution of such results to the development of science? Since technological advance has been accelerated in speed, the future generation of sequencing tool is always going to be available within an interval of every 3-5 years, therefore a repetitive analysis of the same or similar samples can be carried out for one’s life time without any exaggeration if a person chooses to do so, not to think.

This and similar practices have a serious eroding effect on the career training of young researchers because they shall be trained with the ways to ask fundamental scientific question and identify the necessary methods, especially the innovative ones to solve the scientific question in a significant way. The on-going practice of collecting data in a mass production manner in the two examples shown above is simply to produce many papers within a short period of time (Gu, 2022) and also to manipulate the citation metrics with the least efforts (Radicchi et al., 2008). It is time to ask the question on what is the true meaning of sequencing and what is it for in any of the research project before it is too late? This phenomenon occurring in applied microbiology or microbial ecology because of lack of a clear hypothesis and the readily availability of sequencing technology is becoming a wide spread one due to the affordable cost to use of the technol-

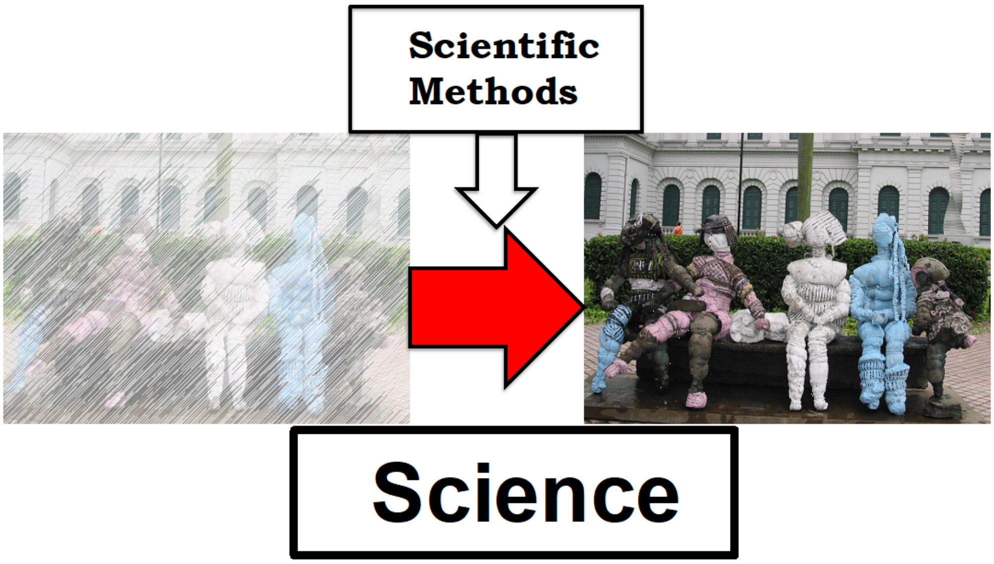

Figure 1. Effective scientific methods play important role in the discovery of science as schematically illustrated by transforming an unknown or less clear understood phenomenon (*left triangle*) to a high clarity (*right*) for human understanding and comprehension as new knowledge

ogy by far too many now. This trend has to be stopped and prevented to allow science to be back on the right track for the fundamental discovery, not peripheral data without any significant contribution to the subject than duplicated data or sometimes generating noises to blur the truth. Population of science with far too many insignificant papers is not a healthy development in science and must be managed by both government agencies and also academic societies collaboratively soon (Gu, 2022).

Statistical analysis is also increasingly used with sequencing data. Inference and correlation analysis of statistical science between sequencing data and the environmental parameters analyzed as an example are another point deserving some words here because they are collectively presented in many of the microbial ecology papers as an accepted formula for a complete paper (Chen and Gu, 2022). Statistical analysis was invented to differentiate the treatment effects of different fertilization schemes at the Rothamsted Experimental Station in England after its establishment in 1843. The method by itself has been invented and established soundly as a subject taught in universities and colleges without difference today, but the results from the analysis, now by software on computers, need validation or further experiments. In other words, this method only provides some guiding information, the fact finding will be solely on the shoulder of the individual to prove or disprove the truth. Currently, the results from such analysis are being used as a proven fact incorrectly and definitely in majority of the reports.

The Future Prospects

Science makes its advance by the participating individuals at different time, especially the talented and passionate ones. The benefits to biology gained from physical sciences and the relationship between them need further discussion to establish a clear and useful framework structure for the development of relevant subjects under the science, especially the former, into the near future. Biology is a subject based on instruments for development to be made in science in contrast to the realistic approaches in physics and chemistry (Rosenberg, 1994). During the course of investigating science, research paradigm shift is common and breakthrough discovery is often the result of a major change in the research framework and thinking to challenge the existing knowledge. This shall be the case for the new emerging environmental science and related subjects to improve the current practice for a better foundation of future research (Gao and Gu, 2021). Because of this, the mass production of papers driven by low quality science and experiments has an eroding consequence to the development of science, but more and more researchers recognize this situation and also its damage to the community and science. More effective means will be implemented in the high education and also academic societies to deal with this from the root of the problem.

New and emerging technology will continue to play its role in the scientific discovery, but its application must be carefully selected and evaluated based on the scientific questions and the specific scheme of research route to tackle the selective scientific issues of high significance. Technology can be a drawback to development because it is neither creative nor innovative to adopt existing technology blindly without careful judgement or a clear understanding of the research hypothesis in one's head. During the current development of nucleotide sequencing, applications of this technology has served our society well in medicine and health science as an early effective diagnostic tool. Its role in science and further discovery of high quality requires a good background and comprehension of the subject knowledge as a prerequisite in science to serve the advancement and the further development into the future. When natural or anthropogenic impacted sites are the topic of research for microbial community information, the novelty shall not be the site or location, but the new knowledge of universal level importance from the investigation. Therefore, my advice is to sequence when the questions is good enough, and not when there is no genuinely anything new.

Acknowledgements

Opinions expressed here are those of the author personally. Research in his laboratory is supported by National Natural Science Foundation of China, but no implications shall be made between the nature of individual statements made here and the research support received from agencies.

Conflict of Interest

Author declares that there is no conflict of interest in the information presented here.

Ethical approval

This article does not contain any studies with human participants or animals performed by the author involved.

References

Arias, A. and Omil, F., 2018. An innovative wastewater treatment technology based on UASB and IFAS for cost-efficient macro and micropollutant removal. *Journal of Hazardous Materials*, 359: 113-120. https://doi.org/10.1016/j.jhazmat.2018.07.042

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pea, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7: 335-336. https://doi.org/10.1038/nmeth.f.303

Chen, J. and Gu, J.-D., 2022. The environmental factors used in correlation analysis with microbial community of environmental and cultural heritage samples. *International Biodeterioration & Biodegradation*, 173: 105460. http://doi.org/10.1016/j.ibiod.2021.105460

Gao, L. and Gu, J.-D., 2021. A unified conceptual framework involving maintenance energy, metabolism and toxicity involved in research on degradation of environmental organic pollutants. *International Biodeterioration & Biodegradation*, 162: 105253. http://doi.org/10.1016/j.ibiod.2021.105253

Gu, J.-D., 2016. Biodegradation testing: so many tests but very little new innovation. *Applied Environmental Biotechnology*, 1 (1): 92-95. http://doi.org/10.18063/AEB.2016.01.007

Gu, J.-D., 2020. On environmental biotechnology of bioremediation. *Applied Environmental Biotechnology*, 5 (2): 3-8. http://doi.org/10.26789/AEB.2020.02.002

Gu, J.-D., 2022. Popularization or population of science and the future. *Applied Environmental Biotechnology*, 7 (1): 1-5. http://doi.org/10.26789/AEB.2022.01.001

Kuhn, T.S., 1996. The structure of scientific revolution (3rd ed.). University of Chicago Press, Chicago. https://doi.org/10.7208/chicago/9780226458106.001.0001

Popper, K.R., 1959. The logic of scientific discovery. Basic Books, New York.

Radicchi, F., Fortunato, S., Castellano, C., 2008. Universality of citation distributions: toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences USA*, 105: 17268-17272. https://doi.org/10.1073/pnas.0806977105

Rosenberg, A., 1994. Instrumental biology or the disunity of science. The University of Chicago Press, Chicago

Woese, C.R., 1987. Bacterial evolution. *Microbiological Reviews*, 51: 221-271.

Wolfe-Simon, F., Blum, J.S., Kulp, T.R., Gordon, G.W., Hoeft, S.E., Pett-Ridge, J., Stolz, J.F., Webb, S.M., Weber, P.K., Davies, P.C., Anbar, A.D., Oremland, R.S., 2010. A bacterium that can grow by using arsenic instead of phosphorus. *Science*, 332: 1163-1166. http://doi.org/10.1126/science.1197258

Yuan, K., Li, S., Zhong, F., 2020. Treatment of coking wastewater in biofilm-based bioaugmentation process: biofilm formation and microbial community analysis. *Journal of Hazardous Materials*, 400: 123117. http://doi.org/10.1016/j.jhazmat.2020.123117