Toward the Genomic-Information Society

Takashi Gojobori[✉]

Computational Bioscience Research Center (CBRC), King Abdullah,
University of Science and Technology (KAUST),
Thuwal 23955-6900, Kingdom of Saudi Arabia
takashi.gojobori@kaust.edu.sa

Abstract. In life science, strong needs for big data analysis and its related data-oriented methodologies such as AI (Artificial Intelligence) are acutely increasing in the recent years. These needs are recognized in not only medical sciences but also other disciplines in which genomic information plays a crucial role in understanding phenomena of interest and in conducting their innovative application to practical usages in the society. It eventually leads us to significant implication that our society will have undergone totally new experiences with enormous benefits from various advancements by research and development of genome information. On the basis of such observations, I would like to propose a future vision of our society as a "Genome-Information-Oriented Society (G-Society)" or "Genomic Information Society".

Keywords: Big data · AI · Metagenomics

1 Metagenomics as an Example

Taking a methodology called "metagenomics" as an example, I would like to discuss how our society changes with advancements of big data analyses of genomic information.

Metagenomics is an approach of understanding the whole features of microbes existing in a given sample by directly conducting extensive sequencing almost all the DNA fragments of microbial genomes without any microscopic observations. In particular, this approach can be applied to all the microbes including non-culturable bacterial species.

When this approach can be used for seawater samples, we will know all the microbial communities in that seawater, which will be useful to monitor marine environments and further to explore fishing grounds. This may be called as "Marine Metagenomics".

When this approach is conducted for soil samples, it will be very useful for improvements of agricultural efficiencies and conservation of ecological habitation. It has huge potential to change the way of conventional agriculture. This may be called as "Soil Metagenomics" or "Agricultural Metagenomics".

This approach can be applied for air samples, by which we can identify pathogens such as detrimental bacteria and viruses present in an air sample. In particular, immediate application to the sand storms may lead to understanding of etiological

agents that can be brought into a human society. This is called as "Airborne Metagenomics".

Moreover, metagenomics has been applied to human cavities such as human guts. For example, intestinal bacteria have been understood in a variety of aspects in an enormous speed. It is getting known that some bacterial species are deeply associated with human disease such as not only diseases associated with adult lifestyle habits but also cancers and even mental diseases.

Although we have already known how beneficial human complete genome information is for curing the diseases and maintaining our health, the state-of-art methodologies of metagenomics are surely changing our society, as stated above.

2   Big Data Analyses as Crucial Tools

The data produced from genomic approaches including metagenomics is really huge: It is literally "big data". In order to extract useful knowledge from such big data as genomic information, it is obvious that informatics of big data analyses is crucial. Database construction, functional annotation with appropriate ontology of controlled vocabularies, and AI (artificial Intelligence) such as machine learning, text mining and deep learning will be key elements of big data analyses.

How usefully we can utilize genetic information depends heavily upon successful developments of those IT (Information Technology) elements. It follows that our future vision of the so-called "G-Society" will come true only if such big data analyses can be applied, in a timely and appropriate fashion, to genome information that will be hugely produced further in the coming years.

3   Conclusion

In the situation that a huge amount of genomic information has been produced by the recent advancements of genomic approaches such as metagenomics, it is pointed out that our human society may be faced by new experiences including benefits from useful knowledge extracted from big data of genomic information. This vision of new human society may be called "Genome-Information-oriented Society (G-Society)" or "Genome Information Society". To make this vision come true, developments of various IT elements for big data analyses must be crucial.