BIOINFORMATIC APPROACHES TO INFECTIOUS DISEASES

Taijiao Jiang

Institute of Basic Medical Sciences-CAMS, China

Abstract

Due to the event of high throughput sequencing technologies, large-scale sequencing of pathogens, like influenza viruses and HIV, has not only become a routine add surveillance of infectious diseases but also become indispensable in the identification of newly emerging communicable diseases in their outbreaks. Besides, analyses of those large-scale sequence data of pathogens will significantly enhance our understanding of the evolution of diseases, promoting new strategies for the prevention and control of those viruses. This research mainly focuses on modeling the evolution of the communicable disease from gene data and epidemiological data. Many computational methods are developed during this lab to know the origin and evolution of influenza viruses. They were demonstrated to be helpful in the influenza vaccine, estimating death rate and disease lethality, sourcing newly emerging influenza viruses, and so on. They have made a significant influence in the field of bioinformatics and informatics of infectious diseases. Among them, together with China CDC, the new methods developed for vaccine recommendation of human influenza viruses have now been in use in China for assisting flu vaccine strains selection.

The application of Bioinformatics tools and techniques in analyzing the increasing data generated in molecular biology, genomics, transcriptomics, and proteomics is gaining momentum. Moreover, the amount of information gleaned in the form of databases and literature for generating molecular profiles and for collecting data related epidemiology of pathogens has been also mounting. Therefore, the use of Bioinformatics tools and techniques in pathogen identification and typing, identifying markers for early diagnosis and treatment, enabling personalized interventions, and predicting patient outcomes is imperative. Bioinformatics aided next-generation sequencing data analysis are promising to spot clinically relevant viruses from a spread of specimen types. Similarly, bacterial pathogens such as Francisella tularensis and Leptospira santarosai were successfully identified using culture-Independent NGS identification from primary human clinical specimens. The application of Bioinformatics techniques in the surveillance of pathogen outbreaks in the fight against infectious diseases is also essential. Thus, this review documents available bioinformatics resources and databases that are employed by medical microbiology scientists and physicians to regulate emerging infectious pathogens.

Bioinformatics tools are extensively used in the identification, characterization, and typing of all kinds of pathogens. This followed the widespread use of genomic approaches in the diagnosis and management of viral, bacterial, and fungal infections. Applications of bioinformatics have been used in pathogen identification, detection of virulence factors, resistome analysis, and strain typing. Next-generation sequencing (NGS) technology supported by bioinformatics, phylogenetic, and patho-genomics analyses helped in the identification of the causative agent were a Clostridium haemolyticum isolate. This isolate possesses virulence factors necessary to establish an infection and cause all the observed symptoms. Thus, NGS holds considerable potential for pathogen identification isolated from human specimens using whole-genome sequencing (WGS) assisted by powerful bioinformatics tools. The application of Bioinformatics tools in analyzing WGS and Ribosomal (rRNA) gene sequencing data for the identification of both bacterial and fungal pathogens is becoming routine in recent years. The need for advanced yet improved bioinformatics tools in the analysis of NGS-rRNA sequencing data is emerging in microbiome studies.

The available Bioinformatics tool for microbiome studies does the detection and removal of the amplification-derived chimeric sequence. These could be classified as novel organisms if not properly identified as an anomalous score. Thus, the removal of chimeric sequences is an essential step in microbiome analysis. In addition to the above-mentioned tools, there are automated pipelines dedicated to analyzing both processed data and raw sequences such as QIIME, It also provides a pipeline for extended processing and analysis of high-throughput sequencing data, including single-strand and paired-end reads.

Genomics and bioinformatics are contributing to our understanding of infectious diseases caused by bacterial pathogens and parasites. This ranges from development for research investigations of disease outbreaks and pathogenesis, host and pathogen genomic variation, and host immune evasion mechanisms to the identification of potential diagnostic markers, vaccine targets, and vaccine developments. High throughput genomics data generated from pathogens and animal models are often combined with the host genomics and patients' health records to offer advice on diagnosis options also as potential drug and vaccine interactions. However, despite accounting for the very best burden of infectious diseases, Africa has rock bottom research output on communicable disease genomics. Here we review the contributions of genomics and bioinformatics to the management of infectious diseases of great public health concern in Africa including tuberculosis, dengue, malaria, and filariasis. Furthermore, we discuss how genomics and bioinformatics are often applied to spot drug and vaccine targets. We conclude by identifying challenges to genomics research in Africa and highlighting how these are often overcome where possible.

Herein we highlight how genomics and bioinformatics have contributed to our understanding of infectious diseases of serious health concern, starting from bacterial and viral to parasitic infections, also as their applications to drug and vaccine target identification. This ranges from understanding pathogenesis, host systemic responses, and host-pathogen interactions to the identification of prognostic and diagnostic markers. However, in Africa, despite the high morbidity and mortality thanks to infectious diseases, there's limited expertise within the field of bioinformatics and hence limited bioinformatics research output in terms of publications. Thus, there's a requirement to strengthen training and capacity building in bioinformatics in Africa to enhance communicable disease genomics and host-pathogen genomics on the continent. This can be achieved through the establishment of well-structured courses, mentorship for junior and trainee bioinformaticians, and better career prospects to take care of trained bioinformaticians on the continent. Even though there is more advancement in the field of science still needs to improve in various aspects. Several Bioinformatics tools are available for analyzing data for combating and control of infectious diseases as discussed in this review. However, there are several bioinformatics tools for drug resistance testing, pathogen-host interaction, infection, and treatment outcomes. Nonetheless, the need to facilitate and incorporate bioinformatics tools and applications in clinical microbiology and infectious diseases through training of personnel and by developing simple yet robust user-friendly bioinformatics.

Reference:

1. Zhaozhong Zhu, Zheng Zhang, Wenjun Chen, Zena Cai, Xingyi Ge, Haizhen Zhu, Taijiao Jiang, Wenjie Tan, Yousong Peng (2018) The neuropathology of alcohol-related braindamage Predicting the receptor-binding domain usage of the coronavirus based on kmer frequency on spike protein. Infect Genet Evol 2018 Jul 4;61:183-

184.

2. Taijiao Jiang et al (2018). Genome-wide RNA-Seq identifies Fas/FasL-mediated tumoricidal activity of embryonic stem cells Int J Cancer 2018 May 20;142(9):1829-1841

3. Yousong Peng, Xiaodan Li, Hongbo Zhou, Aiping Wu, Libo Dong, Ye Zhang, Rongbao Gao, Hong Bo, Lei Yang, Dayan Wang, Xian Lin, Meilin Jin, Yuelong Shu, Taijiao Jiang (2018) Continual Antigenic Diversification in China Leads to Global Antigenic Complexity of Avian Influenza H5N1 Viruses. Sci Rep 2018 May 31;8:46994.

4. Xiao Ding, Jiejian Luo, Lijun Quan, Aiping Wu, Taijiao Jiang (2017) Evolutionary genotypes of influenza A (H7N9) viruses over five epidemic waves in China. Infect Genet Evol 2017 Nov 22;55:269-276.