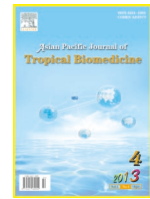




Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Biomedicine

journal homepage: www.elsevier.com/locate/apjtb

Document heading

doi:

© 2013 by the Asian Pacific Journal of Tropical Biomedicine. All rights reserved.

Utilization of multiple "omics" studies in microbial pathogeny for microbiology insights

Viroj Wiwanitkit

Wiwanitkit House, Bangkhuae, Bangkok Thailand

Hainan Medical University, China

Joseph Ayobabalola University, Nigeria

Faculty of Medicine, University of Nis, Serbia

KEYWORDS

Omics
Microbiology
Pathogeny

ABSTRACT

In the present day, bioinformatics becomes the modern science with several advantages. Several new "omics" sciences have been introduced for a few years and those sciences can be applied in biomedical work. Here, the author will summarize and discuss on important applications of omics studies in microbiology focusing on microbial pathogeny. It can be seen that genomics and proteinomics can be well used in this area of biomedical studies.

1. Introduction

In the present day, the complete of genome project leads to many new findings and applications in biomedical society^[1–5]. With the heap of data, the manipulation of the information for medical usage becomes a challenging thing. The answer for manipulating of numerous data is the using of computational technology. With help of computer, the process becomes simple and possible. Hence, it can be seen that this approach is a real bridging among sciences. This is the basic on the new science namely "bioinformatics"^[6,7]. In bioinformatics, three important parts should be mentioned^[7]. First, the basic biomedical research within the basic biomedical laboratory using basic biomedical techniques serves the data for further manipulation. Second, the computational research leads the new software and computational tool for manipulating via computational engineering and informatics technique. Third, this is an actual important point. The way to make an interrelationship between biomedical and computational technique for answering the complex biomedical question

is the critical process of bioinformatics. Bioinformatics can be presently divided into many categories. In general, it is usually divided into steps classified based on the process from generation to protein expression in genetics. Hence, many "omics" science are available and those omics sciences can be useful in biomedical study. Here, the author will summarize and discuss on important applications of omics studies in microbiology. Indeed, microbiology is a classical biomedical science. In the past, the approach was usually on two important concepts, *in vitro* and *in vivo* studies. For sure, those classical old approaches take time and highly cost. With implementation of new bioinformatics, the use of computational application, *in silico* study, for answering the research question is possible^[7]. This is an actual advanced microbiology and the general practitioner in microbiology must realize and know those new technologies. As noted, bioinformatics can be well applied for studying in microbiology. Here, a special focus is on microbial pathogeny. It can be seen that genomics and proteinomics can be well used in this area of biomedical studies.

*Corresponding author: Professor Viroj Wiwanitkit, M.D. Wiwanitkit House, Bangkhuae, Bangkok Thailand.
Tel: 6624132436
E-mail: wviroj@yahoo.com

Article history:
Received 13 Jan 2013
Received in revised form 17 Jan, 2nd revised form 23 Jan, 3rd 31 Jan 2013
Accepted 20 Feb 2013
Available online 28 Apr 2013

2. Genomics for microbial pathogen study

After the end of genome projects, emerging of the first omics science occurred. That pioneer omics science is called “genomics”. Simply, genomics is the specific omics science in bioinformatics that deals with genome. At present, genomics technique becomes widely used in biomedicine. As already mentioned, the actual root of genomics is genetics, which is important concern in microbiology. There is no doubt that genomics can be applied in microbiology study. The use of genomics technique for studying microbial pathogen can be in several ways as the followings:

2.1. Genomics database search

The database is an important tool in bioinformatics. Database is a kind of tool collecting several data. Searching on database is usually the first step in bioinformatics approach[7]. Many databases are available at present and some can be applied for medical microbiology study. The most basic and widely known database is provided by PubMed. The practitioner can search for gene and protein that corresponding to the pathogenesis of the disease via access PubMed. In microbiology, the well-known database for bacteria is Bacteriome.org. This is a resource for the *Escherichia coli* interactome[8]. The tool helps identify and visualize functional, evolutionary and structural relationships between groups of focused interacting proteins[8]. For fungi, the reference database is e-Fungi[9]. Generally, the e-Fungi database posed data for more than 30 fungal genomes[9]. Publicly available genome data, functional annotations, and pathway information could be searched from e-Fungi[9]. The database can be freely accessible via <http://www.e-fungi.org.uk>[9]. Apart from the group of pathogen, there are also some interesting databases that are specific for specific pathogens. The examples are coliBASE, ShiBASE and CandidaDB[10–12]. The coliBASE is a specific database focusing on *Escherichia coli*, *Shigella* and *Salmonella* comparative genomics[10]. Comparative data which include the whole genome alignments and lists of putative orthologous genes can be derived from searching coliBASE[10]. The ShiBASE is a specific database focusing on comparative genomics of *Shigella*[11]. Data on comparative analysis on DNA sequences, gene orders, metabolic pathways and virulence factors can be derived from searching shiBASE[11]. In addition, ShiBASE also allow the alignment of any query sequence with the reference sequences in the database[11]. For CandidaDB, it is a specific database for *Candida* species and related Saccharomycotina[12]. CandidaDB covers number of fully or partially completed genome sequences of related fungal species and can be used for further comparative genomics study[12].

2.2. Comparative genomics

Comparative genomics is the use of computational technique

to help compare among genomes. Basically, comparison between genomes can imply the similarity and this is widely used concept in microbiology. Starting to perform a comparative genomics study, searching of the database is the basic thing as already mentioned. Apart from those already mentioned database, another important database that is useful is xBASE2[13]. The xBASE2 is a genome database that help manipulate the bacterial genome sequence data for further comparative genomics work[13].

Based on the concept of comparative genomics, the comparison becomes an important process in this omics science. The simple concept is comparing the puzzled derived genome with the standard referencing genomes, which are usually collected in standard database. The comparison will result in similarity. The similarity searching can be very useful in microbiology. The similarity of genomes usually means the similar genome function and further subsequent proteins and biological processes. Similarity usually occurs in the same or closely group of pathogen. This helps classification of the new detected pathogens. For example, the use of comparative genomics was applied to clarify the nature of the newly emerging influenza virus in 2009[14]. The origin of the new swine flu virus was discovered by the comparative genomics approach[14]. Finally, the virus was found to be a new type generated by mixing between human, swine and avian viruses[14]. One might imagine that this kind of approach might be limitedly useful for taxonomy, however, it is actually useful for realize of the new emerging pathogen. Apart from the already mentioned case of new virus in 2009, it can be applied in other situation. The best example is the comparative genomics tools applied to bioterrorism defence[15].

In addition to phylogenetics usefulness, similarity of the queried genome to reference virulent genome or drug resistant genome might imply that the queries genome might pose virulence or drug resistance properties. This data is very important when one studies a new unknown microbial genome. A recent interesting report on this application is the use of comparative genomics to study the genome sequence of *Salmonella enterica* that cause typhoid fever in Malaysia and Papua New Guinea[16,17]. Comparing between fatal and non-fatal isolates genomes could show genetic clue of virulence of the pathogen[18]. Further implication of the similarity data can be used in searching of new drug. The homology search can reveal drug target that can be safe to use in human beings. The example of the study is the report on new drug searching for *Mycobacterium leprae* by Wiwanitkit[18].

On the other hand, similarity search can also show the difference. Indeed, the different genomes with significant differences in sequences can be the sign of different species. This is important for grouping. It should be noted that the concept to assess the difference between genomes of species has been used in biomedical science for years and it is called

phylogenetics. There is no doubt that the phylogenetics study can make use of comparative genomics approach.

2.3. Phylogenomics

As already noted, the phylogenetics focuses on assessment for difference among species. The specific application of genomics on this area is called phylogenomics. There are many computational tool to perform phylogenomics study. The good example is BPhyOG, which is an interactive server allowing genome-wide inference of bacterial phylogenies based on overlapping genes^[19]. BPhyOG uses two main phylogenetic tree-reconstruction methods, Neighbor Joining and Unweighted Pair-Group Method using Arithmetic averages and it can be public available via <http://cmb.bnu.edu.cn/BPhyOG>^[19]. The recent well-known report that used the phylogenomics approach to study the pathogenesis of microbial is on Microsporidia. Capella-Gutiérrez *et al.* found from their phylogenomics work and proposed Microsporidia as the earliest diverging clade of sequenced fungi^[20].

2.4. Functional genomics

Functional genomics focuses the interest of the study on the function of the genome. Basically, the function of gene is protein creation. Therefore, functional genomics is strongly related to the functional proteomics. For computational approach, ESTree db is a good tool^[21]. More discussion will be provided in the next heading on functional proteomics.

3. Proteomics for microbial pathogeny study

In bioinformatics, there are two main important omics sciences. One is the genomics that is already mentioned. The other is proteomics, which is the specific omics science focusing on the protein which is the product of the gene. Conceptually, gene generates protein and protein expresses. Gene shows genotype which is not overt while protein shows phenotypes which is overt. Hence, the thing that we see or determine in biomedical test is usually the phenotype.

Similar to genomics, proteomics is also very useful and can be applicable in microbiology. Indeed, the classical microbiology deals mostly on protein. The gene has just been known and investigated for few decades. It is also no doubt that proteomics can be applied in microbiology study. The use of genomics technique for studying microbial pathogeny can be in several ways as the followings:

3.1. Proteomics database search

As noted, database is basic requirement in bioinformatics. Similar to genomics tool, PubMed is also the simple database for searching on protein. Another well-known proteomics database is Swiss-Prot protein knowledgebase^[22]. It provides the way

to search for the the function of proteins^[22]. Recently, its new presentation namely UniProt was created and it contained data from Swiss-Prot, Translation of the EMBL nucleotide sequence (TrEMBL), and the Protein Information Resource-Protein Sequence Database (PIR-PSD)^[22]. Specifically to the field of microbiology, there are also other interesting databases. Of those, the 2Dbase is an interesting database for proteome of *E. coli*^[23]. It contain data from two dimensional polyacrylamide gel electrophoresis and mass spectrometry and it can be accessed via <http://2dbase.techfak.uni-bielefeld.de>^[23]. Another interesting database is MannDB^[24]. The MannDB was created to support searching of proteins to be targets for pathogen or protein toxin detection^[24].

3.2. Structural proteomics

Protein structure is very important when one studies in an organism. Since protein is the actual thing that plays roles and presents phenotype. Structural proteomics is the specific “omic” science to clarify and predict structure of protein. Prediction of the protein structure from the nucleic acid sequence is possible and is the basic study in structural proteomics. Conceptually, protein structure prediction can reveal the pathogenesis process of the pathogen. The good example is the study on the structure of influenza virus by Wiwanitkit^[25]. According to this study, it is concluded that “Structural aberration in HA associated with adaptation of human influenza A H3N2 virus in embryonated chicken eggs”^[25]. Finally, it should also be noted that the advance technology helps generating predict protein structure and it can further allow manipulation to interact with other molecules. Hence, it can be used in studying effect of pathogen on cells and interaction with drug. The good example is the study Wiwanitkit on immune complex formation from dengue. Prediction of the interacted molecule between dengue proteins and immune protein was done in this work and further implied for the pathogenesis of renal failure^[26].

3.3. Comparative proteomics

As already noted, comparison between two things including to proteins can imply the similarity and this is widely used concept in microbiology. Comparative proteomics is the use of computational technique to help compare among proteomes. Searching of the database to get protein sequence then comparing by computational tool is the basic thing in comparative proteomics. The good examples on applied comparative proteomics to study the pathogenesis of microbial pathogens are the serial reports on non-structural proteins of dengue and mechanism of bleeding by Wiwanitkit^[27–29].

4. Conclusion

Based on the present advanced bioinformatics technique, the use of “omics” science can be helpful in performing study

in microbiology. The omics techniques can be used for answering the question on microbial pathogeny. Using genomics, comparison of the puzzled microbial genome to standard genomes in database helps identify the similarity and can predict the microbial properties such as virulence and resistance. Using proteomics, structure clarification and prediction can be done and this can be helpful in further assessment on the pathogenesis process and prediction on interaction with host cells and drugs.

Conflict of interest statement

I declare that I have no conflict of interest.

References

- [1] Trgovcevic Z. 50 years of molecular biology. *Lijec Vjesn* 1994; **116**: 315–318.
- [2] Portin P. Historical development of the concept of the gene. *J Med Philos* 2002; **27**: 257–286.
- [3] Csako G. Present and future of rapid and/or high-throughput methods for nucleic acid testing. *Clin Chim Acta* 2006; **363**: 6–31.
- [4] Watson JD, Cook–Deegan RM. Origins of the human genome project. *FASEB J* 1991; **5**: 8–11.
- [5] Lewin R. National academy looks at human genome project, sees progress. *Science* 1987; **235**: 747–748.
- [6] Yu U, Lee SH, Kim YJ, Kim S. Bioinformatics in the post–genome era. *J Biochem Mol Biol* 2004; **37**: 75–82.
- [7] Wiwanitkit V. *Medical biochemoinformatics*. New York: Nova Publisher; 2008.
- [8] Su C, Peregrin–Alvarez JM, Butland G, Phanse S, Fong V, Emili A, et al. Bacteriome.org an integrated protein interaction database for *E. coli*. *Nucleic Acids Res* 2008; **36**(Database issue): D632–D636.
- [9] Hedeler C, Wong HM, Cornell MJ, Alam I, Soanes DM, Rattray M, et al. e–Fungi: A data resource for comparative analysis of fungal genomes. *BMC Genomics* 2007; **8**: 426.
- [10] Chaudhuri RR, Khan AM, Pallen MJ. coliBASE: an online database for *Escherichia coli*, *Shigella* and *Salmonella* comparative genomics. *Nucleic Acids Res* 2004; **32**(Database issue): D296–D299.
- [11] Yang J, Chen L, Yu J, Sun L, Jin Q. ShiBASE: an integrated database for comparative genomics of *Shigella*. *Nucleic Acids Res* 2006; **34**(Database issue): D398–D401.
- [12] Rossignol T, Lechat P, Cuomo C, Zeng Q, Moszer I, d’Enfert C. CandidaDB: a multi–genome database for *Candida* species and related Saccharomycotina. *Nucleic Acids Res* 2008; **36**(Database issue): D557–D561.
- [13] Chaudhuri RR, Loman NJ, Snyder LA, Bailey CM, Stekel DJ, Pallen MJ. xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Res* 2008; **36**(Database issue): D543–D546.
- [14] Holmes EC. Evolution in health and medicine Sackler colloquium: The comparative genomics of viral emergence. *Proc Natl Acad Sci USA* 2010; **107**(Suppl 1): 1742–1746.
- [15] Slezak T, Kuczumski T, Ott L, Torres C, Medeiros D, Smith J, et al. Comparative genomics tools applied to bioterrorism defence. *Brief Bioinform* 2003; **4**: 133–149.
- [16] Yap KP, Teh CS, Baddam R, Chai LC, Kumar N, Avasthi TS, et al. Insights from the genome sequence of a *Salmonella enterica* serovar typhi strain associated with a sporadic case of typhoid fever in Malaysia. *J Bacteriol* 2012; **194**: 5124–5125.
- [17] Baddam R, Thong KL, Avasthi TS, Shaik S, Yap KP, Teh CS, et al. Whole–genome sequences and comparative genomics of *Salmonella enterica* serovar typhi isolates from patients with fatal and nonfatal typhoid fever in Papua New Guinea. *J Bacteriol* 2012; **194**: 5122–5123.
- [18] Wiwanitkit V. Analysis of *Mycobacterium leprae* genome: *in silico* searching for drug targets. *Southeast Asian J Trop Med Public Health* 2005; **36** (Suppl 4): 225–227.
- [19] Luo Y, Fu C, Zhang DY, Lin K. BPhyOG: an interactive server for genome–wide inference of bacterial phylogenies based on overlapping genes. *BMC Bioinformatics* 2007; **8**: 266.
- [20] Capella–Gutiérrez S, Marcet–Houben M, Gabaldón T. Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. *BMC Biol* 2012; **10**: 47.
- [21] Lazzari B, Caprera A, Vecchiotti A, Stella A, Milanese L, Pozzi C. ESTree db: a tool for peach functional genomics. *BMC Bioinformatics* 2005; **6**(Suppl 4): S16.
- [22] Schneider M, Tognolli M, Bairoch A. The Swiss–Prot protein knowledgebase and Expasy: providing the plant community with high quality proteomic data and tools. *Plant Physiol Biochem* 2004; **42**: 1013–1021.
- [23] Vijayendran C, Burgemeister S, Friehs K, Niehaus K, Flaschel E. 2DBase: 2D–PAGE database of *Escherichia coli*. *Biochem Biophys Res Commun* 2007; **363**: 822–827.
- [24] Zhou CL, Lam MW, Smith JR, Zemla AT, Dyer MD, Kuczumski TA, et al. MannDB—a microbial database of automated protein sequence analyses and evidence integration for protein characterization. *BMC Bioinformatics* 2006; **7**: 459.
- [25] Wiwanitkit V. Structural aberration in HA associated with adaptation of human influenza A H3N2 virus in embryonated chicken eggs. *Turk J Vet Anim Sci* 2009; **33**: 437–438.
- [26] Wiwanitkit V. Immune complex: does it have a role in pathogenesis of renal failure in dengue infection? *Ren Fail* 2005; **27**: 803–804.
- [27] Wiwanitkit V. A study on functional similarity between dengue non structural protein 1 and platelet integrin/adhesin protein, CD61. *J Ayub Med Coll Abbottabad* 2006; **18**(3): 13–16.
- [28] Wiwanitkit V. Platelet CD61 might have an important role in causing hemorrhagic complication in dengue infection. *Clin Appl Thromb Hemost* 2005; **11**: 112.
- [29] Wiwanitkit V. Weak binding affinity of immunoglobulin G, an explanation for the immune mimicking theory in pathophysiologic findings in the recovery phase of dengue. *Nanomedicine* 2005; **1**: 239–240.