Bioinformatics Technology In Clinical And Public Health Microbiology Applying Computational Methods

Hyunjo Kim

Abstract
The role of clinical genomics in infectious disease diagnostics and public health microbiology is the topic of discussion during a recent decade. Although much of this work is aimed at describing the structure of outbreak communities, the methodology works equally well to identify pathogens in clinical samples. Clinical genomics is the exploitation of genome sequence data for diagnostic, therapeutic, and public health purposes. Central to this field is the high-throughput DNA sequencing of genomes and metagenomes. The key concept in using clinical genomics methodology is that detection of microbes is independent of culture and is not limited to targets used for in-depth PCR assays. Rather, it is a process of generating large-scale sequence data sets that adequately sample a specimen for microbial content and then of applying computational methods to resolve the sequences into individual species, genes, pathways, or other features.

Keywords
clinical microbiology, infectious diseases surveillance, public health, clinico genomics, whole-genome sequencing, bioinformatics pipeline

Introduction
Infectious diseases are disorders caused by organisms such as bacteria, viruses, fungi or parasites. Many organisms live in and on our bodies are normally harmless or even helpful, but under certain conditions, some organisms may cause disease. Infectious diseases currently contribute about 20% of the global annual death causes and public health priorities in this area include antimicrobial resistance (AMR), vaccine preventable diseases, tuberculosis, influenza, and sexually transmitted infections. Addressing this challenge, the key role of clinical microbiologists (CMs) in improving appropriate use of antimicrobials in human medicine has been reaffirmed, notably by ensuring timely production and communication of critical diagnostic results and standardized drug susceptibility testing reports in accordance with local treatment guidelines. The provision of facility-specific cumulative susceptibility reports for bacterial pathogens against antibiotics on the formulary also forms an essential part of CM work. Additionally, CMs provide daily counseling to clinicians on etiological infection diagnoses and management, including correct sampling for and interpretation of test results, and targeted therapy of difficult-to-treat resistant pathogens and complicated infections. As members of the hospital antimicrobial stewardship team, CMs take on responsibilities that include coordination, planning of infection control activities, post-prescription review, and feedback. Hence, the explosion of microbiome research is driven by high-throughput DNA sequencing, so-called next-generation sequencing (NGS), technologies that allow the genomic content of entire microbial communities (bacterial, viral, and eukaryotic organisms) to be described.

Materials and Methods

Collection and processing of meta-analysis samples
Preparation of quantitative determination As in other disciplines in medicine, systematic reviews (SR) and meta-analyses (MA) in infectious diseases are an important aid to clinical decision-making. In the present article we review features important to SRs in infectious diseases that should be addressed in most SRs and Mas. We stress the need to include in the SR analysis all patients that were randomized; all studies that were performed. Authors of SRs should choose one main outcome that matters to patients, and base their conclusions mainly on this outcome. Resistance as an outcome is a topic that should be addressed in all SRs of antibiotic treatment. Ethical aspects and especially patients’ safety should be addressed in SRs.

Optimization of analysis conditions for outbreak investigation Public health microbiology laboratories (PHLs) are on the cusp of unprecedented improvements in pathogen identification, antibiotic resistance detection, and outbreak investigation by using whole-genome sequencing (WGS). However, considerable challenges remain due to the lack of common standards. Here, we describe the validation of WGS on the Illumina platform for routine use in PHLs according to Clinical Laboratory Improvements Act (CLIA) guidelines for laboratory-developed tests (LDTs). The following objectives were accomplished: (i) the establishment of the performance specifications for
WGS applications in PHLs according to CLIA guidelines, (ii) the development of quality assurance and quality control measures, (iii) the development of a reporting format for end users with or without WGS expertise, (iv) the availability of a validation set of microorganisms, and (v) the creation of a modular template for the validation of WGS processes in PHLs. The validation panel, sequencing analytics, and raw sequences could facilitate multilaboratory comparisons of WGS data. Additionally, the WGS performance specifications and modular template are adaptable for the validation of other platforms and reagent kits.

Results and Discussion

In systematic reviews data from original research are assembled in their entirety, critically appraised, and sometimes combined through meta-analysis to answer a clear clinical question as shown in Figure 1-1 and Figure 1-2.

A study protocol should be written in advance, completed literature searches should be performed, and studies selected in a reproducible and objective fashion. Biologic plausibility should be addressed. Reported findings should be interpreted with caution, taking into account the limitations of methodologic aspects of risk factors studies. Figure 3 showed that Dual-use research (DUR) is research conducted for legitimate purposes that generates knowledge, information, technologies, and/or products that could be utilized for both benevolent and harmful purposes.
threat with broad potential consequences to public health and safety, agricultural crops and other plants, animals, the environment, materiel, or national security. The seven viruses of concern listed include circulating, synthetically generated, and eradicated agents. Although spanning all of microbiological research, much of the DURC debate has focused on virology, where selection processes of circulating and synthetically generated agents have been used to enhance transmission [in Table 1].

Although selection is a mainstay of microbial genetic research and gain and loss of function experiments are commonplace if the DURC list is strictly adhered to, there is a very small number of DURC projects. Furthermore, gain of function is often accompanied by loss of function and vice versa. Alterations in phenotype which give cause of concern can include loss and gain of function. This generic figure depicts the flow of information associated with a patient diagnosed with gastroenteritis who is tested for Salmonella. Although human input or interpretation may be required at various nodes, information is often generated, transmitted, and received digitally (in Figure 4).

Figure 4. Agents of notifiable infectious diseases and their associated data often travel through layers of agencies, including clinics, laboratories, and public health registries.

Various terminology and messaging standards, such as SNOMED, LOINC, and HL7, are employed to ensure the standardization, efficiency, and security of the process. However, many clinical laboratories still use paper to transmit information.14,15 The more often information is exchanged between two nodes, the more incentive exists to create a robust electronic interface between these nodes. Therefore, information between the patient and clinician is often exchanged verbally, although electronic communication with patients is likely to become more common. As shown in Figure 5, the WGS assays evaluated in order to deduce the corresponding performance parameters are shown in green boxes.

Figure 5. Summary of WGS validation. The estimated performance parameters are shown in blue boxes. The components of WGS accuracy determined in this study are shown in purple boxes.

Percentages alongside the boxes represent values measured during this validation for the corresponding parameters. The bar within a group shows the percentage of strains in that MCG group carrying a given virulence gene. Six genes encoding immune proteins are marked in red (Figure 6).

Figure 6. Distribution of known putative virulence genes in the seven core genome groups.

The horizontal dotted blue line demarcates the virulence genes that are present in all isolates. The encoded protein names are shown on the left.16–18

Conclusion

The clinical microbiology laboratory has responsibilities ranging from characterizing the causative agent in a patient’s infection to helping detect global disease outbreaks. All
Table 1. Online molecular typing databases and Currently available software for the analysis of typing results

<table>
<thead>
<tr>
<th>Method</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLST</td>
<td>MLST.net&lt;br&gt;Pubmlst.org&lt;br&gt;Institut Pasteur MLST&lt;br&gt;European Working Group for Legionella Infections Sequence-based typing database&lt;br&gt;Environmental Research Institute, University College Cork&lt;br&gt;MLVAbank&lt;br&gt;Groupe d’Etudes en Biologie Prospective</td>
</tr>
<tr>
<td>MLVA</td>
<td>MLVApplus&lt;br&gt;Institute Pasteur MLV: MLV-NET&lt;br&gt;MLV.net</td>
</tr>
<tr>
<td>ccrB typing</td>
<td>Staphylococci ccrB sequence typing</td>
</tr>
<tr>
<td>dru typing</td>
<td>dru typing database</td>
</tr>
<tr>
<td>spa typing</td>
<td>Ridom Spa Server</td>
</tr>
<tr>
<td>CRISPR typing</td>
<td>CRISPRdb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application</th>
<th>Software</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gel analysis</td>
<td>GelCompare II&lt;br&gt;Phoretix 1D&lt;br&gt;Gel-Pro Analyzer 4.5&lt;br&gt;Lasergene CLCbio workbench&lt;br&gt;Geneious</td>
<td>Commercial</td>
</tr>
<tr>
<td>Sequence assembly and analysis</td>
<td>Lasergene&lt;br&gt;CLCbio workbench&lt;br&gt;MEGA 5&lt;br&gt;Geneious</td>
<td>Freeware</td>
</tr>
<tr>
<td>Phylogenetic inference</td>
<td>PHYLOViZ 1.0&lt;br&gt;Structure 2.3.3&lt;br&gt;BAPS 5.4&lt;br&gt;ClonalFrame 1.2&lt;br&gt;Ridom Epicompar&lt;br&gt;Comparing Partitions</td>
<td>Freeware</td>
</tr>
<tr>
<td>Typing methods comparison</td>
<td>RDP3&lt;br&gt;Mauve</td>
<td></td>
</tr>
</tbody>
</table>

of these processes are increasingly becoming partnered more intimately with informatics. Effective application of informatics tools can increase the accuracy, timeliness, and completeness of microbiology testing while decreasing the laboratory workload, which can lead to optimized laboratory workflow and decreased costs. Informatics is poised to be increasingly relevant in clinical microbiology, with the advent of total laboratory automation, complex instrument interfaces, electronic health records, clinical decision support tools, and the clinical implementation of microbial genome sequencing. This review research discusses the diverse informatics aspects that are relevant to the clinical microbiology laboratory, including the following: the microbiology laboratory information system, decision support tools, expert systems, instrument interfaces, total laboratory automation, telemicrobiology, automated image analysis, nucleic acid sequence databases, electronic reporting of infectious agents to public health agencies, and disease outbreak surveillance. The breadth and utility of informatics tools used in clinical microbiology have made them indispensable to contemporary clinical and laboratory practice. Continued advances in technology and development of these informatics tools will further improve patient and public health care in the future.

Acknowledgement
The authors are grateful to the Department of Bioinformatics, Gachon University, Seongnam city, South Korea for laboratory facility and to the Department of Medical informatics, CHA medical school for the financial support respectively.

References
3. Buehler SS, Madison B, Snyder D, JH C, NE S, A M. Effectiveness of practices to increase timeliness of providing


