Problem-based Learning for Bioinformatics: Culture-independent pathogen detection for public health and patient care

Instructor: William Hsiao

Jan 22, 2019

While British Columbians enjoy high standards of healthcare, infectious diseases remain a serious and constant threat. The goal of microbiological diagnostic tests in a hospital is to identify and characterize the pathogen(s) infecting a patient as quickly as possible to aid treatment decisions. The challenge is to identify the culprit in a mixed sample containing host cells and a microbiome. Traditionally, microbiological tests are done by conducting biochemical tests and microscopic examinations. To enrich the pathogen(s) from the background cells, these tests often require the microbes to first be grown on Petri dishes to produce a “pure culture” (called a colony or an isolate). This is time and labour intensive. The power of these tests to differentiate microbes is also low. More and more microbiological diagnostic tests are now PCR-based molecular tests (often called Nucleic Acid Amplification Tests or NATTs). These NAATs are fast and can be done on primary clinical samples (culture-independent assays), but they require the amplification targets to be known. Moreover, when amplification fails, one can not rule out that it’s due to the absence of the target rather than some genetic mutations rendering the test ineffective. Lastly, as NAATs do not produce cultured isolates, further characterization (called sub-typing tests) of the pathogen for the purpose of disease surveillance and outbreak investigation is not possible. As public health agencies around the world are increasingly relying on genomic sequencing for tracking the spread of pathogens, the switch to culture-independent testing in hospital labs has a major impact on public health operations.

You are the co-owner and scientific director of an established biotechnology company contracted by the Canadian government to figure out how to **implement a new solution that would take advantage of the speed and flexibility culture-independent assays for patient care yet allowing public health to conduct disease surveillance using the test results**. Moreover, the Canadian health care budget is tight so your solution has to be cost-effective and can be implemented across Canada without undue burden on the health care personnel and budget. Your company has expertise in designing targeted and non-targeted assays using high throughput sequencing (HTS) technologies, bioinformatics analysis, software engineering, and workflow development. Here are the formal requirements:

1) The solution does not require pure culture and can be performed from primary clinical samples
2) Be able to enrich and characterize the genomes of potential pathogens in the presence of host cells and microbiome - primarily focused on bacterial pathogens; bonus if can also detect viruses.
3) Have fast “turn-around-time” so the test results can inform patient care (typically this means within 24 hours)
4) Produce sufficient data for public health to be able to conduct disease surveillance and analyze transmission using the test results. The public health response time frame is typically longer (ideally within 72 hours). (hint: this means your solution could be staged where you first produce results sufficient for patient care before you generate results for public health use)
5) Design a bioinformatics workflow that is easy enough for untrained hospital staff to use to produce test results within 24-72 hours - including sequencing time.

Discussion Questions:

1. What are the roles of public health vs. clinical care? What are the priorities

2. What are some culture-independent microbiological assays besides HTS? What are the advantages and disadvantages of these approaches for clinical care and for public health?

3. For HTS based approaches, discuss the difference between targeted (amplicon-based) vs. non-targeted (metagenomics) sequencing. Keep in mind that pathogens often are often present as a small fraction of the microbiome and host genetic material.

4. What are some laboratory methods to enrich genetic material from pathogens and deplete host or non-pathogenic microbial genetic material so you can have more “on-target” sequences when performing metagenomic sequencing?

5. After sequencing, how can you differentiate between normal microbiota and pathogen? What reference databases would you use for the sequence search? Are there database-independent methods to perform the search? Keep in mind that some pathogens are very similar to commensal organisms (think about *E. coli*)

6. Knowing that rapid turn-around time is essential (1 day), what kind of bioinformatics workflow would you perform to satisfy the clinical care needs?

7. Knowing that rapid turn-around time is desirable (3 days), what kind of bioinformatics workflow would you perform to satisfy the public health surveillance needs?

8. How would you design and deploy your bionformatics workflow so the hospital and public health personnel won’t need years of bioinformatics education to use your software?

9. How would you ensure that your bioinformatics analysis results are trustworthy and high-quality so patients receive the right treatment or care? How can you ensure the results are reproducible and can be examined later if necessary?

10. What procedures may need to be put in place to ensure that your bioinformatics workflows and analysis results are up to date and correct?

11. How would you share your results between hospitals and public health agencies? Across different provinces? Keep in mind that each province has its own health care system.

12. What can you do to reduce the overall cost of operating your laboratory and bioinformatics workflow for pathogen detection?