JCM
Journals.ASM.org

COMMENTARY

Integration of Whole-Genome Sequencing into Infection Control Practices: the Potential and the Hurdles

Elizabeth Robilotti, Mini Kamboj

Infection Control Program and Infectious Diseases Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA

Microbial whole-genome sequencing (WGS) is poised to transform many of the currently used approaches in medical microbiology. Recent reports on the application of WGS to understand genetic evolution and reconstruct transmission pathways have provided valuable information that will influence infection control practices. While this technology holds great promise, obstacles to full implementation remain. Two articles in this issue of the Journal of Clinical Microbiology (S. Octavia, Q. Wang, M. M. Tanaka, S. Kaur, V. Sintchenko, and R. Lan, J Clin Microbiol 53:1063–1071, 2015, doi:10.1128/JCM.03235-14, and S. J. Salipante, D. J. SenGupta, L. A. Cummings, T. A. Land, D. R. Hoogestraat, and B. T. Cookson, J Clin Microbiol 53:1072–1079, 2015, doi:10.1128/JCM.03385-14) describe the breadth of application of WGS to the field of clinical epidemiology.

The availability of molecular typing methods is essential to aid in outbreak investigations and track evolutionary trends of microbes in the community and health care settings. Currently used methods to determine genetic relatedness between bacteria are based largely on fractional sequencing, e.g., multilocus sequence typing, multilocus variable-number tandem-repeat analysis (MLVA) or gel-based separation and visualization of DNA fragments, e.g., pulsed-field gel electrophoresis (PFGE) or PCR ribotyping. Each of these techniques has its own procedural and analytic limitations, including suboptimal discriminatory power.

Whole-genome sequencing (WGS) has become an important and rapidly accessible tool for microbial identification and pathogenesis and comparative analyses. The clinical impact of this technology on patient care is increasingly discernible (1, 2). From a public health and epidemiologic perspective, robust and higher-resolution genomic analysis provided by WGS has yielded important insights into transmission pathways for several significant pathogens (3–7). Although the majority of these investigations were conducted retrospectively, the findings collectively highlight the potential of WGS as a real-time infection control tool.

Two accompanying articles in this edition of the Journal of Clinical Microbiology (JCM) emphasize the broad array of applications of WGS for epidemiologic investigations (8). The first accompanying paper, by Octavia et al. (9), describes the use of WGS to examine 57 clinical and environmental isolates from five different point source community outbreaks of gastroenteritis caused by endemic *Salmonella enterica* serovar Typhimurium phage type DT170, which accounts for 40% of the human cases in the region studied. Routine implementation of MLVA typing using five variable-number tandem-repeat loci was unable to offer sufficient discriminatory power to distinguish between outbreak and endemic clones. By means of WGS, several additional cases were recognized as outbreak related. Although these additional cases were isolated from the same geographic area and within a 2-week window around the outbreak period, no obvious epidemiologic links had been previously established. In this study, single nucleotide polymorphism (SNP) criteria to determine outbreak thresholds and confirm diversity within the point source were based on a Poisson mutation model. The upper and lower estimates of mutation rates used in the model were derived from previously well-characterized outbreaks, an approach that needs to be corroborated in future studies.

The second accompanying paper, by Salipante and colleagues (10), describes a study that compared WGS to standard PFGE analysis of isolates collected from suspected outbreaks of methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant enterococcus (VRE), and *Acinetobacter baumannii*. Overall, the agreement between PFGE and WGS was highest for VRE and wide discordance was observed for MRSA and *A. baumannii*. Strains of the latter organism that were closely related by PFGE were clonal by WGS. The authors hypothesize that the occurrence of frequent horizontal gene transfer among *A. baumannii* strains could account for this discrepancy, making PFGE prone to produce false-negative results in the event of an actual transmission. The lack of clinical and epidemiologic correlates limits the ability to draw any inference from the proposed classification of genetic relatedness in that study (≤3 SNPs, related; 4 to 12 SNPs, closely related; >13 SNPs, unrelated). Overall, the analysis revealed over- and underidentification of clonality among epidemiologically related samples of three clinically important health care-associated pathogens and provides important information for laboratories transitioning methodologies from PFGE to WGS.

Despite the obvious advantages of WGS, several barriers remain before the technology is adopted for routine use. Technical advances have resulted in a rapid decline in sequencing and equipment costs, overcoming many of the upfront financial challenges (11, 12). The turnaround time needed to retrieve actionable information is also becoming shorter, such that timely implementation

Accepted manuscript posted online 11 February 2015

Citation Robilotti E, Kamboj M. 2015. Integration of whole-genome sequencing into infection control practices: the potential and the hurdles. J Clin Microbiol 53:1054–1055. doi:10.1128/JCM.00349-15.

Editor: Y.-W. Tang

Address correspondence to Elizabeth Robilotti, robilote@mskcc.org, or Mini Kamboj, kambojM@mskcc.org.

doi:10.1128/JCM.00349-15

of control strategies based on WGS data seems plausible in the setting of an outbreak.

In its present state, high-performance computing remains essential for sequence assembly and analysis. The various sequencing platforms have their unique sets of challenges that impact the sensitivity and reproducibility of the data generated, for example, from short-read systems (assembly is complicated) compared to long reads (higher error rate). The requirement of a sophisticated informatic infrastructure for data processing and integration persists as a major obstacle to the widespread implementation of WGS, and the technology remains out of reach for many clinical and public health laboratories (13). Calibration for genetic evolution within and outside the host is essential to establish benchmarks for sample density and variant calling. The studies in this edition of JCM clearly indicate that refinement of the WGS approach in these areas is under way.

In summary, the data presented by Octavia et al. and Salipante et al. reinforce the importance of integrating genomic analysis with epidemiologic information gathered during public health investigations. Rapid access to the whole genome of a bacterial isolate, ideally obtained directly from a clinical specimen within a clinically relevant time frame, holds innumerable opportunities for improving patient care and infection prevention and control. Ultimately, the pace and success of the transition from conventional to WGS-based typing depend upon simplifying analysis and establishing clear criteria for validation and verification, akin to currently accepted standards for the accreditation of routine clinical microbiology laboratories.

ACKNOWLEDGMENT

Grant support for this article was provided by a National Institutes of Health National Institute of Allergy and Infectious Diseases career development award to M.K. (K23 AI083880).

REFERENCES

1. **Köser CU, Bryant JM, Becq J, Torok ME, Ellington MJ, Marti-Renom MA, Carmichael AJ, Parkhill J, Smith GP, Peacock SJ.** 2013. Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. N Engl J Med **369:**290–292. http://dx.doi.org/10.1056/NEJMc1215305.
2. **Roetzer A, Diel R, Kohl TA, Ruckert C, Nubel U, Blom J, Wirth T, Jaenicke S, Schuback S, Rusch-Gerdes S, Supply P, Kalinowski J, Niemann S.** 2013. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. PLoS Med **10:**e1001387. http://dx.doi.org/10.1371/journal.pmed.1001387.
3. **Harris KA, Underwood A, Kenna DT, Brooks A, Kavaliunaite E, Kapatai G, Tewolde R, Aurora P, Dixon G.** 1 December 2014, posting date. Whole-genome sequencing and epidemiological analysis do not provide evidence for cross-transmission of Mycobacterium abscessus in a cohort of pediatric cystic fibrosis patients. Clin Infect Dis http://dx.doi.org/10.1093/cid/ciu967.
4. **Köser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, Hsu LY, Chewapreecha C, Croucher NJ, Harris SR, Sanders M, Enright MC, Dougan G, Bentley SD, Parkhill J, Fraser LJ, Betley JR, Schulz-Trieglaff OB, Smith GP, Peacock SJ.** 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med **366:**2267–2275. http://dx.doi.org/10.1056/NEJMoa1109910.
5. **Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Group NCSP, Henderson DK, Palmore TN, Segre JA.** 2012. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med **4:**148ra116. http://dx.doi.org/10.1126/scitranslmed.3004129
6. **Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TE.** 2013. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. Lancet Infect Dis **13:**137–146. http://dx.doi.org/10.1016/S1473-3099(12)70277-3.
7. **Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, Ip CL, Golubchik T, Batty EM, Finney JM, Wyllie DH, Didelot X, Piazza P, Bowden R, Dingle KE, Harding RM, Crook DW, Wilcox MH, Peto TE, Walker AS.** 2013. Diverse sources of C. difficile infection identified on whole-genome sequencing. N Engl J Med **369:**1195–1205. http://dx.doi.org/10.1056/NEJMoa1216064.
8. **Köser CU, Ellington MJ, Cartwright EJ, Gillespie SH, Brown NM, Farrington M, Holden MT, Dougan G, Bentley SD, Parkhill J, Peacock SJ.** 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog **8:**e1002824. http://dx.doi.org/10.1371/journal.ppat.1002824.
9. **Octavia S, Wang Q, Tanaka MM, Kaur S, Sintchenko V, Lan R.** 2015. Delineating community outbreaks of *Salmonella enterica* serovar Typhimurium by use of whole-genome sequencing: insights into genomic variability within an outbreak. J Clin Microbiol **53:**1063–1071. http://dx.doi.org/10.1128/JCM.03235-14.
10. **Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT.** 2015. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol **53:**1072–1079. http://dx.doi.org/10.1128/JCM.03385-14.
11. **Jünemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, Mellmann A, Goesmann A, von Haeseler A, Stoye J, Harmsen D.** 2013. Updating benchtop sequencing performance comparison. Nat Biotechnol **31:**294–296. http://dx.doi.org/10.1038/nbt.2522.
12. **Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ.** 2012. Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol **30:**434–439. http://dx.doi.org/10.1038/nbt.2198.
13. **Fricke WF, Rasko DA.** 2014. Bacterial genome sequencing in the clinic: bioinformatic challenges and solutions. Nat Rev Genet **15:**49–55. http://dx.doi.org/10.1038/nrg3624.