Curr Infect Dis Rep (2017) 19: 15
DOI 10.1007/s11908-017-0570-0

Whole Genome Sequencing—Implications for Infection Prevention and Outbreak Investigations

Kyle J. Popovich[1] · **Evan S. Snitkin**[2]

Published online: 9 March 2017

Abstract
Purpose of Review Whole genome sequencing (WGS) is a laboratory method that has emerged as a promising tool for epidemiologic investigations.
Recent Findings Genomic epidemiology approaches have been utilized in outbreak settings, community settings, within acute care hospitals, and across healthcare facilities to better understand transmission and spread of potential pathogens. These studies have highlighted how essential robust epidemiologic data is in these analyses as well as how results can be translated into clinical practice and infection control and prevention.
Summary Existing studies have highlighted both the promise and challenges of using WGS as an epidemiologic tool in a community and healthcare setting and across a region. Costs for performing and interpreting WGS analyses are decreasing, and availability of and experience with WGS analyses in healthcare epidemiology are increasing. With these favorable trends, this laboratory method soon could emerge as the gold standard for epidemiologic evaluations.

Keywords Whole genome sequencing (WGS) · Genomic epidemiology · Infection prevention · Outbreaks

This article is part of the Topical Collection on *Healthcare Associated Infections*

✉ Kyle J. Popovich
kyle_popovich@rush.edu

[1] Department of Internal Medicine, University Infectious Diseases, Rush University Medical Center, Stroger Hospital of Cook County, 600 South Paulina, Suite 143, Chicago, IL 60612, USA

[2] Department of Microbiology and Immunology, Department of Medicine, Division of Infectious Diseases, Center for Microbial Systems, University of Michigan, 1520D MSRB I, 1150 W. Medical Center Dr., Ann Arbor, MI 48109-5680, USA

Introduction and Background

Healthcare-associated infections remain an important problem for healthcare settings and are associated with increased morbidity, mortality, and cost [1, 2]. Numerous strategies and infection control bundles have been employed to reduce the incidence of infections in healthcare settings [3, 4]. A cornerstone of infection control and hospital epidemiology is the integration of microbiologic and advanced laboratory analysis with epidemiologic data to inform prevention efforts. These laboratory tools have been used in outbreak, outpatient, and hospital settings to identify sources of pathogen spread, thereby allowing for targeting of infection control strategies [5].

In epidemiologic evaluations, laboratory methods can be used to investigate the relatedness of particular healthcare-associated pathogens, which can improve understanding of how pathogens are being acquired and if there is a common source of spread. A downside of several microbiologic methods that have been used (*spa* typing, multilocus sequence typing, pulsed-field gel electrophoresis (PFGE)) is the limited discriminatory power they have for distinguishing among pathogens that are considered endemic in a particular setting [6]. For example, community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) emerged 10 to 15 years ago [7] and that was the most common CA-MRSA strain in the USA as identified by PFGE is USA300 [8]. However, USA300 has become endemic in community, outpatient settings, emergency rooms, and hospitals [9–11], which has made standard molecular methods such as PFGE of limited value in discriminating between USA300 strains. This limitation has made it challenging to fully understand the spread of strains and how best to intervene outside of acute care settings.

Whole genome sequencing (WGS) is a laboratory method that has emerged as a promising tool for epidemiologic investigations due to its capacity to distinguish between endemic

strains [12]. WGS interrogates the entire genetic material of an organism and therefore provides maximal discriminatory power for distinguishing between closely related strains. By considering the number of single-nucleotide variants (SNVs) between two strains in the context of the evolutionary rate of the species, genomic analyses allow for an informed assessment of whether two strains are connected on epidemiologically relevant time scales. Then, by performing phylogenetic analyses to more finely describe the relationships among strains and overlaying epidemiological data, investigators can generate hypotheses for what are the key epidemiological drivers of transmission.

Genomic epidemiology approaches have been utilized in outbreak settings, community settings, within acute care hospitals, and across healthcare facilities to better understand transmission and spread of potential pathogens (Fig. 1). These studies have highlighted how essential robust epidemiologic data is in these analyses as well as how results can be translated into clinical practice and infection control and prevention. Costs for performing and interpreting WGS analyses are decreasing, and availability of and experience with WGS analyses in healthcare epidemiology are increasing. With these favorable trends, this laboratory method soon could emerge as the gold standard for epidemiologic evaluations.

Single-Facility Outbreak Investigations

WGS has been used as an epidemiologic tool to aid in outbreak investigations for a variety of pathogens including MRSA, *Acinetobacter baumannii*, Carbapenem-resistant *Enterobacteraciae* (CRE), *Clostridium difficile*, and tuberculosis (Table 1). In an outbreak of MRSA in a neonatal intensive care unit [26••], WGS along with contact tracing data and surveillance cultures of babies and healthcare workers led to the identification of a colonized healthcare worker who likely contributed to the persistence of the outbreak, despite optimization of infection control efforts. Information supplied with WGS led to the intervention of decolonization and furlough of the specific healthcare worker. Investigators noted that WGS allowed for improved identification of a MRSA transmission network in comparison to conventional infection control methods.

In an adult intensive care unit in a teaching hospital in England, investigators used WGS to identify whether intra-facility *S. aureus* transmission occurred at their institution [13]. Admission and weekly surveillance cultures were performed to identify MRSA carriage and acquisition, and *spa* typing along with patient stay data highlighted potential cases of patient-to-patient transmission. WGS confirmed instances of patient-to-patient transmission within the intensive care unit by identifying genetically similar strains between patients however also disproved three hypothesized patient-to-patient transmissions. In this study, only 18.9% of acquisitions could be attributed to other colonized patients with WGS, demonstrating the importance of robust epidemiologic data when creating genomic transmission networks.

WGS was used in the United Kingdom to examine *Clostridium difficile* transmission over a 3.5-year time period in a specific geographic region [22•]. Isolates in this study were from clinical cases, and therefore, colonization or transmissions from asymptomatic carriers would not be detected or considered in the analysis. The authors observed that only 35% of strains were genetically similar to previously reported isolates, suggesting that for a proportion of individuals, infection was acquired either through another source or an intermediate individual. Furthermore, 36% of patients with highly genetically similar strains did not have an apparent epidemiologic link, highlighting how WGS can illuminate gaps in our understanding and guide future research efforts.

In an outbreak of *Acinetobacter baumannii* at the research-based National Institutes of Health Clinical Center [19], WGS was used to determine the origins a polyclonal outbreak of *Acinetobacter baumannii*. Despite no prior history of *A. baumannii* in the hospital, PFGE typing of outbreak isolates had revealed three distinct clone types contemporaneously causing infections in patients, which typically would suggest multiple simultaneous introductions of *A. baumannii*. However, given the penchant for *A. baumannii* to rapidly alter its genome structure, it was also deemed possible that there was a single importation event, with the clone evolving and changing its PFGE type during the course of the outbreak. In

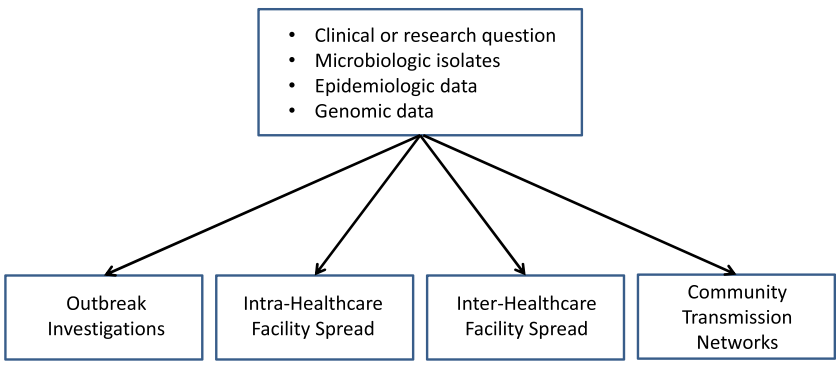

Fig. 1 Overview of how whole genome sequencing can be used as an epidemiologic tool

Table 1 How whole genome sequencing improved the epidemiologic analysis of various pathogens in outbreak, community, and healthcare settings

Pathogen	Epidemiology based on pre-WGS analytics	Epidemiology based on WGS data	Comments
Methicillin-resistant *Staphylococcus aureus* (MRSA)	Adult intensive care unit (ICU): Possible patient-to-patient transmissions of MRSA using *spa* typing and patient stay data [13]. Neonatal ICU (NICU): Persistence of a MRSA outbreak despite optimization of infection control strategies [12]. Community: Colonization and infection with USA300 MRSA, the most common community-associated MRSA strain in the US by pulsed-field gel electrophoresis (PFGE) [8], is associated with certain community exposures, and particular patient populations are disproportionately impacted by USA300 MRSA [14–16].	Adult ICU: Confirmation of patient-to-patient transmission of MRSA however WGS highlighted three instances where transmission was unlikely to have occurred. NICU: WGS led to the identification of a colonized healthcare worker who likely contributed to the persistence of the outbreak. Community: Identified that households likely serve as a reservoir for USA300 MRSA transmission [17] and that transmission networks exist outside the household as well. Individuals in non-household-based transmission networks may share certain epidemiologic features [18].	Work in the ICU has highlighted the need for robust epidemiologic data. Missing epidemiologic links between patients can lead to gaps in the created transmission network. Because USA300 is endemic in many settings, WGS was a useful intervention for differentiating between strains and attempting to identify pathways for transmission.
Acinetobacter baumannii	PFGE analysis showed three distinct clone types contemporaneously causing infections in patients, suggesting multiple simultaneous introductions of *A. baumannii* [19].	Outbreak of *A. baumannii* at the research-based NIH Clinical Center was from two separate importation events. In addition, one of the PFGE types was traced back to a strain present in the hospital a year before the outbreak but the strain had acquired drug resistance elements over time.	This outbreak highlighted that in-depth genomic analysis may be needed in conjunction with well-characterized data sets to fully construct a transmission network.
Carbapenem-resistant *Enterobacteraciae* (CRE)	There is a dense network of patient transfers between healthcare facilities that may have allowed for regional spread of CRE [20].	Confirmation that patient transfers drove the CRE regional outbreak. WGS also allowed for distinguishing between importation and acquisition of CRE at facilities [21].	Because long-term care facilities and acute care settings were felt to be the primary driver of CRE transmission, an analysis of patient sharing with genomic data was able to be conducted.
Clostridium difficile	*C. difficile* clinical isolates from a healthcare system in the United Kingdom were collected over a 3.5-year time period. The prevalence of sequence types and ribotypes was determined [22•].	A proportion of isolates were genetically distinct to those previously reported, highlighting how strains were likely acquired through another source or an intermediate individual.	*C. difficile* can last for prolonged periods in the environment which may complicate a genomic analysis of transmission [23]. In addition, asymptomatically colonized patients without culture data may be the missing links in transmission networks.
Neisseria gonorrhoeae	CDC surveillance *N. gonorrhoeae* isolates were collected from clinics across the USA. Data had suggested spread of antibiotic resistance in gonococcus among men who have sex with men [24].	Confirmation that spread of isolates with reduced susceptibility to cefepime was occurring through sexual networks of men who have sex with men. There were also some introductions into heterosexual networks. Results also demonstrated the geographic spread in the USA (from west coast and has spread eastward).	Large datasets and often collaboration with public health to be able to analyze a large collection of isolates and to have relevant epidemiologic data are needed.
Mycobacterium tuberculosis	An outbreak of tuberculosis occurred from 2006 to 2008 in British Columbia, Canada [25•]; however, investigators could not identify the source using interspersed repetitive unit variable number tandem repeat genotyping with traditional contact tracing.	WGS with social network analysis led to the conclusion that increases in crack cocaine use allowed for the spread of *M. tuberculosis* and that high-risk social networks led to persistence of the outbreak.	Multiple methodologies can be integrated with genomic data, including contact tracing, social network, and epidemiologic exposure data.

contrast to the uncertainty in the PFGE results, WGS with phylogenetic analysis allowed investigators to conclude that the outbreak actually originated from two separate importation events. The first introduction of *A. baumannii* was traced back to a single patient from whom two PFGE types were isolated. The third PFGE type, or second importation, was ultimately traced back to a strain circulating in the hospital in the year prior to the outbreak, which had in the interim acquired multiple-drug-resistant elements. Thus, WGS was able to disentangle this complex outbreak situation and demonstrate how the plasticity of *A. baumannii* allows it to adapt to the healthcare environment over extremely short time frames.

Issues to Consider

There are potential limitations with using WGS in outbreak investigations and examination of intra-facility transmission of a pathogen. First, there can be missing epidemiologic links between people colonized or infected with genetically similar strains and hence gaps in the created transmission network. For example, asymptomatically colonized patients without culture data may serve as missing links in the constructed models of pathogen spread. Second, long-term colonization for a patient with a potential pathogen can lead to greater than expected within host diversity by genomic analysis [27]. This may then complicate comparative genomic analyses between individuals where there is reliance on genetic distance to infer transmission events [28]. Third, directionality is challenging to determine in many of these outbreak investigations, although well-characterized data sets with accurate time stamps can mitigate this limitation. Finally, while WGS provides a wealth of information, how to optimize integration with clinical and epidemiologic data to generate actionable prevention strategies presents an unmet challenge.

Investigations of Inter-facility Pathogen Spread

WGS can also be used beyond a single healthcare facility to examine the transmission of potential pathogens across a healthcare network in a region. Using mathematical modeling, it has been suggested that interventions for antibiotic-resistant bacteria need to extend beyond a single facility to have a significant impact [29]. The frequent movement of patients between particular healthcare facilities in a region likely promotes spread of a potential pathogen and creates a healthcare transmission network. Prevention efforts therefore may need to be across institutions and require public health support to effectively interrupt transmission of the potential pathogen. Genomic data may ultimately play a key role in outlining these healthcare networks of transmission for various multidrug-resistant pathogens.

Among the most compelling studies demonstrating the role of patient movement in regional dissemination of MDROs was an epidemiological investigation of a regional outbreak of *bla-KPC Klebsiella pneumoniae* [20, 21]. In this study, Won and colleagues demonstrated that a dense network of patient transfers connected the facilities affected by this outbreak. In an extension of this work, WGS was applied to both confirm the conclusions of the epidemiological investigation and determine whether WGS had sufficient resolution to dissect a regional outbreak that occurred over the course of 1 year [21]. The results showed a strong concordance between genomically inferred inter-facility links and the patient-sharing network, thereby confirming that patient transfers drove the regional outbreak. Moreover, the genomically inferred network also allowed for the distinction between patients who imported isolates into a facility versus acquiring at the facility in which it was first detected.

Integrating WGS with patient movement data can be used to identify the extent to which patient sharing between healthcare facilities contributes to the spread of potential pathogens. Chang et al. used genome sequencing data to examine the spread of USA300 MRSA across highly connected healthcare facilities [30]. In this investigation, the authors did not observe genetic similarity among patient transfer networks, supporting the idea that for USA300, community transmission networks may be the primary drivers of spread in a region rather than acute care hospitals.

Toleman et al. [31] used WGS and MRSA cultures sent to a microbiology laboratory in the east of England to determine the prevalence and origins of USA300 in the region. The study documented a low prevalence of USA300 in the region. However, with WGS, they were able to identify a few cases of household transmission of USA300 but overall, that USA300 in the region was the result of multiple introductions into eastern England. While similar findings have been seen in the USA with USA300, the authors highlight that in a low prevalence region, WGS and contact tracing could be integrated into public health surveillance to identify emerging pathogens and direct prevention efforts to halt spread.

Issues to Consider

There are potential limitations with using genomic data to understand transmission networks across healthcare centers. First, the use of WGS to improve the understanding of patient sharing between various healthcare settings may need to be limited to certain pathogens that are predominantly healthcare-associated. For example, long-term care facilities and acute care hospitals are felt to be the primary driver of CRE transmission. This characteristic allowed investigators to understand how CRE was spread in a region by linking patient transfer with genomic data. In contrast, for a pathogen such as USA300 MRSA that is largely spread in community settings

and thus does not rely on patient transfers between hospitals for spread, healthcare transmission networks may be more problematic to create [30]. Second, comprehensive sampling of isolates across healthcare facilities and a region is needed to understand inter-facility spread of potential pathogens. This large collection of saved isolates may not be available across a healthcare network which would create gaps in a created transmission network. Therefore, projects seeking to understand inter-facility spread of a potential pathogen would benefit from collaboration with other institutions and public health organizations.

Community Transmission Networks

WGS has also been used as an epidemiologic tool outside of hospitals to characterize the spread of potential pathogens for which community networks appear to be the primary driver of spread. In British Columbia, Canada, an outbreak of tuberculosis occurred from 2006 to 2008 [25•] but investigators were unable to identify the source of the outbreak using mycobacterial interspersed repetitive unit-variable number tandem repeat genotyping and traditional contact tracing. Therefore, WGS was used in conjunction with social network analysis to characterize the spread of this pathogen. With this data, investigators concluded that increased crack cocaine use allowed for the spread of two lineages of *Mycobacterium tuberculosis* and that the outbreak was able to persist because of high-risk social networks. This investigation highlights the importance of robust epidemiologic data and how multiple epidemiologic and laboratory tools can be integrated to fully characterize a community transmission network.

WGS has also been utilized to further delineate the spread of *Neisseria gonorrhoeae* with decreased susceptibility to extended spectrum cephalosporins [24], a pathogen of significant concern given the limits in available therapeutic options. *N. gonorrhoeae* isolates collected as part of the Centers for Disease Control and Prevention Gonococcal Isolate Surveillance Project were collected from clinics across the USA and were sequenced. Phylogenetic analysis was then performed and integrated with location and sexual orientation data. This analysis demonstrated that spread of isolates with reduced susceptibility to cefepime was occurring through sexual networks of men who have sex with men with some introductions into heterosexual networks. Knowledge gained was useful for informing population surveillance efforts and prevention strategies for halting the spread of this pathogen.

Lowy et al. used CA-MRSA colonization, infection, and household environmental isolates for genomic analysis to understand spread of CA-MRSA in a Manhattan community. Their results indicated that households likely constitute an important reservoir for USA300 MRSA transmission [17]. The authors highlight that their results support that prevention strategies may need to target a household and not just the infected patient. However, their results also suggested that networks of MRSA transmission likely exist beyond the household as well.

In Chicago, USA300 MRSA has been observed to disproportionately impact certain patient populations such as those who are HIV-infected, use illicit drugs, have incarceration exposure, and live in certain geographic areas, or those who reside in unstable housing [14–16]. WGS was used to determine whether transmission networks exist in an urban community outside of a household or healthcare setting. WGS was performed on colonizing USA300 isolates from HIV-infected and HIV-negative individuals seeking medical care at the major public hospital in Chicago [18]. Isolates were put into context with strains from other cities and with phylogenetic analysis, it appeared that there were community USA300 MRSA transmission networks characterized by individuals who were predominantly African-American and HIV-infected. Illicit drug use and residence in geographic areas of high detainee release appeared to be associated with inclusion in these identified community transmission networks as well. This study suggests that with a larger sample size, use of genomic data along with community epidemiologic data may allow for the identification of community "epicenters" of USA300 MRSA spread.

Issues to Consider

For certain pathogens, community transmission networks may be valuable to identify points of intervention. However, this type of community analysis can have challenges that may limit the characterization of these networks. First, a set of isolates needs to be available for analysis, and if a large collection of isolates exists, often appropriate sampling strategies are needed to ensure sufficient information is available for creation of a community transmission network. Second, using WGS as an epidemiologic tool requires robust epidemiologic meta-data to properly identify possible transmission events in a community network. Third, creation of a community transmission network often involves an integration of various data sets—microbiologic, genomic, demographic, and epidemiologic—and therefore, informatics as well as epidemiologic expertise is essential.

Future Directions/Conclusion

Existing studies have highlighted both the promise and challenges of using WGS as an epidemiologic tool in a community and healthcare setting and across a region. While the cost of WGS has decreased, it is still costly and not widely available, making widespread use for routine infection control difficult. However, in settings of continued spread or in an outbreak

setting, the costs of WGS may be offset by the benefits of providing the information needed to halt an outbreak and prevent spread of potential pathogens to other patients. Another consideration is what interventions will follow from WGS results. As some studies have highlighted, there are sometimes no epidemiologic links found between highly similar strains making it less clear how actionable large-scale WGS analyses will be. Finally, several studies have nicely demonstrated how WGS can be incorporated into large epidemiologic datasets to gain a more detailed understanding of pathogen acquisition and transmission. With endemic pathogens, such as USA300, such a tool can be highly valuable, particularly if spread of a pathogen continues despite standard infection control strategies. Using WGS for infection prevention shows great promise but there is still no substitute for robust epidemiologic data. A combination of various methodologies will likely have the highest yield for preventing spread of potential pathogens and for infection prevention.

Compliance with Ethical Standards

Conflict of Interest Drs. Popovich and Snitkin declare no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Klevens RM, Edwards JR, Richards Jr CL, et al. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. Public Health Rep. 2007;122:160–6.
2. Roberts RR, Scott 2nd RD, Hota B, et al. Costs attributable to healthcare-acquired infection in hospitalized adults and a comparison of economic methods. Med Care. 2010;48:1026–35.
3. Siegel JD, Rhinehart E, Jackson M, Chiarello L. 2007 guideline for isolation precautions: preventing transmission of infectious agents in health care settings. Am J Infect Control. 2007;35:S65–S164.
4. Calfee DP, Salgado CD, Milstone AM, et al. Strategies to prevent methicillin-resistant Staphylococcus aureus transmission and infection in acute care hospitals: 2014 update. Infect Control Hosp Epidemiol. 2014;35:772–96.
5. Kazakova SV, Hageman JC, Matava M, et al. A clone of methicillin-resistant Staphylococcus aureus among professional football players. N Engl J Med. 2005;352:468–75.
6. David MZ, Taylor A, Lynfield R, et al. Comparing pulsed-field gel electrophoresis with multilocus sequence typing, spa typing, staphylococcal cassette chromosome mec (SCCmec) typing, and PCR for panton-valentine leukocidin, arcA, and opp3 in methicillin-resistant Staphylococcus aureus isolates at a U.S. medical center. J Clin Microbiol. 2013;51:814–9.
7. Fridkin SK, Hageman JC, Morrison M, et al. Methicillin-resistant Staphylococcus aureus disease in three communities. N Engl J Med. 2005;352:1436–44.
8. Tenover FC, McDougal LK, Goering RV, et al. Characterization of a strain of community-associated methicillin-resistant Staphylococcus aureus widely disseminated in the United States. J Clin Microbiol. 2006;44:108–18.
9. Dukic VM, Lauderdale DS, Wilder J, Daum RS, David MZ. Epidemics of community-associated methicillin-resistant Staphylococcus aureus in the United States: a meta-analysis. PLoS One. 2013;8:e52722.
10. Moran GJ, Krishnadasan A, Gorwitz RJ, et al. Methicillin-resistant S. aureus infections among patients in the emergency department. N Engl J Med. 2006;355:666–74.
11. Murphy CR, Hudson LO, Spratt BG, et al. Predictors of hospitals with endemic community-associated methicillin-resistant Staphylococcus aureus. Infect Control Hosp Epidemiol. 2013;34:581–7.
12. Koser CU, Holden MT, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med. 2013;366:2267–75.
13. Price JR, Golubchik T, Cole K, et al. Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of Staphylococcus aureus in an intensive care unit. Clin Infect Dis. 2014;58:609–18.
14. Popovich KJ, Hota B, Aroutcheva A, et al. Community-associated methicillin-resistant Staphylococcus aureus colonization burden in HIV-infected patients. Clin Infect Dis. 2013;56:1067–74.
15. Popovich KJ, Smith KY, Khawcharoenporn T, et al. Community-associated methicillin-resistant Staphylococcus aureus colonization in high-risk groups of HIV-infected patients. Clin Infect Dis. 2012;54:1296–303.
16. Popovich KJ, Weinstein RA, Aroutcheva A, Rice T, Hota B. Community-associated methicillin-resistant Staphylococcus aureus and HIV: intersecting epidemics. Clin Infect Dis. 2010;50:979–87.
17. Uhlemann AC, Dordel J, Knox JR, et al. Molecular tracing of the emergence, diversification, and transmission of S. aureus sequence type 8 in a New York community. Proc Natl Acad Sci U S A. 2014;111:6738–43.
18. Popovich KJ, Snitkin E, Green SJ, et al. Genomic epidemiology of USA300 methicillin-resistant Staphylococcus aureus in an urban community. Clin Infect Dis. 2016;62:37–44.
19. Snitkin ES, Zelazny AM, Montero CI, et al. Genome-wide recombination drives diversification of epidemic strains of Acinetobacter baumannii. Proc Natl Acad Sci U S A. 2011;108:13758–63.
20. Won SY, Munoz-Price LS, Lolans K, Hota B, Weinstein RA, Hayden MK. Emergence and rapid regional spread of Klebsiella pneumoniae carbapenemase-producing Enterobacteriaceae. Clin Infect Dis. 2011;53:532–40.
21. Snitkin ES WS, Lolans K, Whitefield S, Weinstein RA, Popovich KJ, Munoz-Price LS, Lin MY, Hayden MK for the CDC Prevention Epicenters Program. Genomic epidemiology of an early, regional outbreak of blaKPC-encoding Klebsiella pneumoniae (KPC). IDWeek 2015, October 7–11, San Diego CA, abstract 1379.
22. • Eyre DW, Cule ML, Wilson DJ, et al. Diverse sources of C. difficile infection identified on whole-genome sequencing. N Engl J Med. 2013;369:1195–205. **This study highlights how genetically similar strains may not have apparentepidemiologic links, demonstrating the importance of epidemiologic data in theseinvestigations and how WGS can highlight gaps in knowledge**.
23. Kim KH, Fekety R, Batts DH, et al. Isolation of Clostridium difficile from the environment and contacts of patients with antibiotic-associated colitis. J Infect Dis. 1981;143:42–50.
24. Grad YH, Kirkcaldy RD, Trees D, et al. Genomic epidemiology of Neisseria gonorrhoeae with reduced susceptibility to cefixime in the USA: a retrospective observational study. Lancet Infect Dis. 2014;14:220–6.

25.• Gardy JL, Johnston JC, Ho Sui SJ, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med. 2011;364:730–9. **This study demonstrates how WGS can be used as an epidemiologic tool to characterize the spread of a pathogen in community networks.**

26.•• Harris SR, Cartwright EJ, Torok ME, et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis. 2013;13:130–6. **This study outlines the use of WGS in a MRSA outbreak in a neonatal intensive care unit and nicely demonstrates how genomic data with epidemiologic data can lead to actionable infection control interventions.**

27. Lieberman TD, Flett KB, Yelin I, et al. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. Nat Genet. 2014;46:82–7.

28. Worby CJ, Lipsitch M, Hanage WP. Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. PLoS Comput Biol. 2014;10:e1003549.

29. Ciccolini M, Donker T, Kock R, et al. Infection prevention in a connected world: the case for a regional approach. Int J Med Microbiol. 2013;303:380–7.

30. Chang HH, Dordel J, Donker T, et al. Identifying the effect of patient sharing on between-hospital genetic differentiation of methicillin-resistant Staphylococcus aureus. Genome Med. 2016;8:18.

31. Toleman MS, Reuter S, Coll F, et al. Systematic surveillance detects multiple silent introductions and household transmission of methicillin-resistant Staphylococcus aureus USA300 in the East of England. J Infect Dis. 2016;214:447–53.