

EDITORIAL

Next-generation pathogen genomics

George M Weinstock^{1*} and Sharon J Peacock^{2,3}

Editorial

In the early 1990s, one of us was involved in one of the first projects to sequence a bacterial genome, the meager 1.1 Mb chromosome of *Treponema pallidum*, the causative agent of syphilis. Completing the project ultimately took about seven years (until published in 1998 [1]), over US\$1.8 million in National Institutes of Health grants (R01AI031068 and R01AI040390) [2], and required pooling forces with The Institute for Genomic Research. Recently, that original *T. pallidum* strain was re-sequenced to get a 'perfect' sequence, a process that took a few days and cost only hundreds of dollars [3]. The original sequencing was performed with the dideoxy-chain termination technique using slab gel electrophoresis instruments. Newly developed software was used for genome assembly and data management and analysis. The latter re-sequencing was performed with next-generation sequencing (NGS) technology and mature software tools. Such is the enormous progress in microbial genome sequencing in the last 20 years.

The mind-boggling evolution of DNA sequencing and bioinformatics technologies is driving a new era of pathogen research. Recent studies of old, well-scrutinized pathogens are now greatly extended based on the sequencing of thousands of strains from collections [4,5]. This increased density of genetic data for individual species allows new insights and definition of mechanisms, just as an aerial photograph gives a clearer picture of the landscape as the pixel density increases. Such large-scale studies, now possible with the increased throughput and lower cost of sequencing, allow a more comprehensive picture of a species' gene pool (the pan-genome), population genetic and/or evolutionary analyses, and more accurate insights into epidemiology, to name a few advances. In the realm of epidemiology, NGS of pathogens is now pushing into the applied genomics area of the clinic, with, for example, studies of clinical outbreaks that can now precisely define complex transmission

chains [6,7]. Perilous clinical challenges posed by new antibiotic-resistant organisms benefit from NGS which can identify mutations, thereby defining mechanisms by which resistance is acquired [8,9], as well as discerning new threats from resistance genes found in whole genome sequences [10].

It is in this context of a new era in pathogen genomics that this special issue of *Genome Biology* and *Genome Medicine* on the Genomics of Infectious Diseases has been assembled. It coincides with an exhilarating time for pathogen genomics research and covers a broad range of bacterial, viral, and parasitic pathogens. Genomic analysis, and sequencing in particular, is agnostic, and applies equally well to the diverse types of pathogens studied in this special issue. Pathogen genomics continues to be an area of some urgency. We need look no further than the current challenges of containing Ebola virus outbreaks or the emergence and expansion of new antibiotic-resistant bacteria, such as carbapenemase-producing *Klebsiella pneumoniae*, to be reminded that infectious disease is not, and will never be, a solved problem. Rather, only by dramatic technological innovation, such as offered by NGS, can we keep up with the pathogenic population.

Genome sequencing continues to advance and provide new tools and applications for pathogen research. Sequencing can now be performed on hundreds of strains in parallel in overnight instrument runs, and this drives forward the data density for the description of genomes and gene expression patterns. Metagenomic application of NGS is another bright new area, affording new culture-independent detection of pathogens in clinical samples as well as illuminating interactions between the pathogen and resident microbiome. One looks forward to future applications of this information to combat infection and restore health, possibly with reduced dependence on antibiotics.

Abbreviation

NGS: Next-generation sequencing.

Competing interests

The authors declare that they have no competing interests.

* Correspondence: george.weinstock@jax.org

¹The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut, USA
Full list of author information is available at the end of the article

Author details

¹The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut, USA.

²Department of Medicine, University of Cambridge, Box 157 Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK. ³The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Published online: 19 November 2014

References

1. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, Gwinn M, Hickey EK, Clayton R, Ketchum KA, Sodergren E, Hardham JM, McLeod MP, Salzberg S, Peterson J, Khalak H, Richardson D, Howell JK, Chidambaram M, Utterback T, McDonald L, Artiach P, Bowman C, Cotton MD, Fujii C, Garland S, Hatch B, Horst K, Roberts K, Sandusky M, *et al*: **Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.** *Science* 1998, **281**:375–388.
2. NIH RePORTER [http://projectreporter.nih.gov/reporter.cfm?def=1]
3. Pětrošová H, Pospíšilová P, Strouhal M, Čejková D, Zobaníková M, Mikalová L, Sodergren E, Weinstock GM, Smajs D: **Resequencing of *Treponema pallidum* ssp. pallidum strains Nichols and SS14: correction of sequencing errors resulted in increased separation of syphilis treponeme subclusters.** *PLoS One* 2013, **8**:e74319.
4. Chewapreecha C, Marttinen P, Croucher NJ, Salter SJ, Harris SR, Mather AE, Hanage WP, Goldblatt D, Nosten FH, Turner C, Bentley SD, Parkhill J: **Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes.** *PLoS Genet* 2014, **10**:e1004547.
5. Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long SW, Kristinsson KG, Gottfredsson M, Vuopio J, Raisanen K, Caugant DA, Steinbakk M, Low DE, McGeer A, Darenberg J, Henriques-Normark B, Van Beneden CA, Hoffmann S, Musser JM: **Evolutionary pathway to increased virulence and epidemic group A *Streptococcus* disease derived from 3,615 genome sequences.** *Proc Natl Acad Sci U S A* 2014, **111**:E1768–1776.
6. Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Group NCS, Henderson DK, Palmore TN, Segre JA: **Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing.** *Sci Transl Med* 2012, **4**:148ra116.
7. Koser 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: **Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak.** *N Engl J Med* 2012, **366**:2267–2275.
8. Tran TT, Panesso D, Gao H, Roh JH, Munita JM, Reyes J, Diaz L, Lobos EA, Shamoo Y, Mishra NN, Bayer AS, Murray BE, Weinstock GM, Arias CA: **Whole-genome analysis of a daptomycin-susceptible *Enterococcus faecium* strain and its daptomycin-resistant variant arising during therapy.** *Antimicrob Agents Chemother* 2013, **57**:261–268.
9. Tran TT, Panesso D, Mishra NN, Mileykovskaya E, Guan Z, Munita JM, Reyes J, Diaz L, Weinstock GM, Murray BE, Shamoo Y, Dowhan W, Bayer AS, Arias CA: **Daptomycin-resistant *Enterococcus faecalis* diverts the antibiotic molecule from the division septum and remodels cell membrane phospholipids.** *MBio* 2013, **4**:e00281–13.
10. Rossi F, Diaz L, Wollam A, Panesso D, Zhou Y, Rincon S, Narechania A, Xing G, Di Gioia TS, Doi A, Tran TT, Reyes J, Munita JM, Carvajal LP, Hernandez-Roldan A, Brandão D, van der Heijden IM, Murray BE, Planet PJ, Weinstock GM, Arias CA: **Transferable vancomycin resistance in a community-associated MRSA lineage.** *N Engl J Med* 2014, **370**:1524–1531.

doi:10.1186/s13059-014-0528-6

Cite this article as: Weinstock and Peacock: Next-generation pathogen genomics. *Genome Biology* 2014 **15**:528.