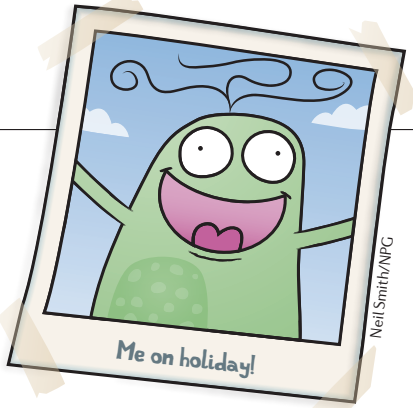

Neil Smith/NPG

GENOME WATCH

TB or not TB? Genomic portraits provide answers

Kate S. Baker and Matthew J. Ellington

This month's Genome Watch highlights the application of bacterial whole-genome sequencing in public health microbiology and epidemiological profiling.

Previous studies using whole-genome sequencing (WGS) of bacteria for the investigation of public health issues reported the identification and drug-susceptibility genotyping of slow-growing bacteria, such as *Mycobacterium tuberculosis*, and the epidemiological investigation and surveillance of a range of pathogens[1], including *Shigella flexneri* and *M. tuberculosis*. Two recent studies highlight the potential of WGS to be used in public health microbiology to investigate the evolution and epidemiology of multidrug-resistant (MDR) bacteria[2,3].

In a very large WGS study, 1,000 clinical *M. tuberculosis* isolates were sequenced to obtain a region-wide snapshot of the prevalence, transmissibility and drug resistance profiles of this pathogen in Samara, Russia[2]. Comparing the genomes of these isolates with previously sequenced *M. tuberculosis* genomes revealed that the pandemic Beijing lineage predominates in Samara. Owing to short genetic distances between some of the isolates, together with the patient location data, the authors concluded that such strains were probably spread from a common source or among patients who share a household. Mapping of drug resistance mutations indicated that 48% of isolates were MDR and 16% of these were extensively drug resistant (XDR). By relating drug resistance SNPs to population phylogeny, SNPs that may exacerbate spread by conferring resistance were identified, such as promoter mutations in the *eis* locus. Some SNPs that confer resistance and that appeared early in the evolution of MDR strains had spread throughout the population and, in some cases, occurred in combination with epistatic mutations that compensated for the fitness costs of resistance. For example, mutations in *rpoA*, *rpoB* and *rpoC* that compensated for resistance-conferring *rpoB* alleles were found. Together, these data suggest that the suite of resistance and compensatory mutations enabled the pathogen to retain clinical resistance without any loss in transmission.

In the second study, the Chinese Centre for Disease Control and Prevention used WGS to analyse the emergence of a novel epidemic serotype (Xv) of *S. flexneri*, which is the major cause of shigellosis in developing countries. A phylogenetic tree that was generated from 59 *S. flexneri* isolates and available reference sequences revealed the emergence of a highly MDR ST91 *S. flexneri* serotype X (REF. 3). Resistance loci that are associated with MDR seemed to result from horizontal gene transfer (for example, a transposon that encodes multiple resistance genes) or were of chromosomal origin (for example, *gyrA* mutations that confer quinolone resistance). The data indicated that the clone emerged in the 1990s and that it acquired a plasmid that converted it from serotype X to serotype Xv in around the year 2000. By 2006, the plasmid was present in three distinct lineages that were epidemic in China. The data that were obtained by WGS were then used to develop a novel SNP-typing scheme, which was applied to 380 *S. flexneri* isolates from three Chinese provinces. Although the sublineages of the epidemic clone clustered geographically, which indicated local evolution of the pathogen, there was also evidence of inter-regional transmission.

Both studies demonstrate the cumulative value of WGS data sets to investigate the evolution and transmission dynamics of MDR epidemic bacterial strains. These studies also show that transmission occurs at household and regional levels and that WGS is useful for providing details of antimicrobial resistance genes and compensatory mutations. Both reports correlated genotypes with laboratory-detected phenotypes (for example, the antibiotic resistance profiles of *M. tuberculosis* and the serotype of *S. flexneri*), which has important implications for updating treatment recommendations, identifying infection sources and epidemiological tracking. However, despite their commonalities, the studies applied WGS in distinct ways; one study leaned towards a 'big data' approach, using statistical analyses to identify molecular mechanisms that are associated with pathogen success, whereas the other study used a smaller data set to develop an SNP-typing assay for further screening of a larger collection of isolates. The commonalities and differences between these two studies exemplify some of the ongoing discussions about how WGS should be applied to different organisms and questions in public health microbiology.

Kate S. Baker and Matthew J. Ellington are at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro3273
Published online 7 May 2014

1. Koser, C. U. *et al.* Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog.* **8**, e1002824 (2012).
2. Casali, N. *et al.* Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nature Genet.* **46**, 279–286 (2014).
3. Zhang, N. *et al.* Genomic portrait of the evolution and epidemic spread of a recently emerged multidrug-resistant *Shigella flexneri* clone in China. *J. Clin. Microbiol.* **52**, 1119–1126 (2014).

Competing interests statement
The authors declare competing interests: see Web version for details.

© 2014 Macmillan Publishers Limited. All rights reserved