GENOME WATCH

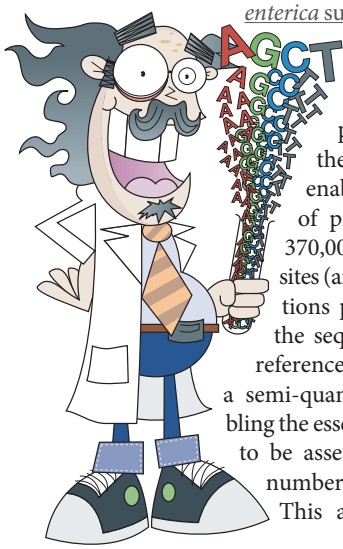
Taming the next-gen beast

Stephen Bentley

This month's Genome Watch discusses how alternative approaches to using second-generation sequencing technologies are powerful tools for the analysis of common pathogenic bacteria.

Almost 5 years have passed since the emergence of fundamentally novel approaches to DNA sequencing, which promised huge increases in data output. Progress has been rapid, and the machines that were initially described as 'next generation' are now described as 'second generation', with the 'third generation' waiting in the wings. Probably the greatest impact so far has been the variation revealed by re-sequencing of the human genome. For smaller microbial genomes, the design scale represents a challenging and tantalising opportunity; however, innovation will be necessary to fully exploit the potential of these machines.

One example of an alterative use for these sequencing technologies is a new method known as TraDIS (transposon-directed insertion site sequencing). Using this approach, Langridge and colleagues[1] analysed a transposon library of more than 1 million *Salmonella enterica* subsp. *enterica* serovar Typhi mutants by sequencing out from the transposons directly into the insertion site. This enabled the mapping of precise locations for 370,000 unique insertion sites (an average of 80 insertions per gene). Mapping the sequence reads onto a reference genome allows for a semi-quantitative assay, enabling the essentiality of each gene to be assessed relative to the number of insertion sites. This analysis found that 356 genes were essential for growth under standard conditions, and a further 274 were designated 'advantageous for growth', owing to the decreasing number of insertions into these genes following multiple liquid culture passages. To survive in the gall bladder, *S.* Typhi requires a high tolerance to bile, leading Langridge *et al.* to screen their mutant library after growth in bile-supplemented broth. They identified 169 genes with significantly fewer insertions, confirming previously known bile tolerance genes and identifying many others. Second-generation sequencing has enabled the rapid screening of much larger mutant libraries than was previously possible, providing a high-resolution view of the conditional essentiality of many genes, and this has been exploited by three other recently published approaches[2,3,4].



Until recently, solving the genome of a bacterial isolate using some second-generation machines was inefficient owing to the high depth of sequence coverage that is generated. However, Harris and colleagues[5] have recently shown how pooling individually tagged shotgun libraries can tame the second-generation sequencing power for use on small genomes.

The genomes of 63 methicillin-resistant *Staphylococcus aureus* (MRSA) isolates were sequenced in parallel, and reads were mapped onto a reference sequence to look for variation. The isolates were all from an MRSA lineage known as *S. aureus* ST239, which is commonly found in health care-associated infections, and were indistinguishable by multilocus sequence typing. The analysis generated a high-resolution phylogenetic view of the lineage with geographically distinct clusters, providing evidence of recent intercontinental transmission. Twenty of the isolates had been collected from a single hospital over 7 months. Although multiple sub-lineages of *S. aureus* ST239 were shown to be present in the hospital, 5 samples taken from one ward over a few weeks had 14 differentiating single-nucleotide polymorphisms (SNPs), demonstrating the potential of this approach for tracing the spread of infection in hospitals.

The SNP data was also used to determine a distance from the phylogenetic root for all isolates, which was then used to calculate that the *S. aureus* ST239 clone emerged sometime between 1965 and 1970. This correlates with the rise in clinical use of antibiotics, a likely driving force behind the expansion of the clone. SNP analysis also showed that there was little evidence for homologous recombination between members of the clone but that mutations known to confer resistance to antibiotics occurred on multiple branches of the tree, further emphasizing the influence of widespread antibiotic use on clone evolution.

These studies suggest that the unprecedented recent expansion in microbial genomics is set to continue as second- (and third-) generation sequencing technologies allow researchers to scale up current sequencing efforts, and further innovation should give rise to many novel approaches.

Stephen Bentley is at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK.
e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro2322

1. Langridge, G. C. *et al.* Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* **19**, 2308–2316 (2009).
2. Goodman, A. L. *et al.* Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* **6**, 279–289 (2009).
3. van Opijnen, T., Bodi, K. L. & Camilli, A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods* **6**, 767–772 (2009).
4. Gawronski, J. D., Wong, S., Giannoukos, G., Ward, D. V. & Akerley, B. J. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc. Natl Acad. Sci. USA* **106**, 16422–16427 (2009).
5. Harris, S. R. *et al.* Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327**, 469–474 (2010).

Competing interests statement
The authors declare no competing financial interests.