

## GENOME WATCH

# Recombination: genomic mix ‘n’ match

Christine J. Boinett and Amy K. Cain



This month's Genome Watch explores the effect that recombination can have on the interpretation of outbreak investigations, and the far-reaching consequences for genomic diversity in bacterial species.

Recombination is the exchange of genetic material between and within species and, for bacteria, it is a source of genetic variation that can allow adaptation to a new ecological niche, alter virulence or lead to the acquisition of resistance determinants. With the increasing use of whole-genome sequencing (WGS) to study bacterial population structures and evolution, understanding the impact of genomic flux is vital for interpretation of WGS data. In outbreak investigations, in which the population structure is key to determining the source of the outbreak, the impact of recombination can greatly influence the conclusions drawn.

In a recent study, Sanchez-Buso *et al.*<sup>1</sup> used the SOLiD 5500XL platform to sequence 69 strains of *Legionella pneumophila* from Alcoy, Spain, to elucidate the epidemiology of 13 different regional outbreaks that spanned 11 years. *L. pneumophila* genomes have a high degree of plasticity owing to recombination, which makes sequence-based typing schemes hard to interpret. Genomic analysis of the 69 environmental and clinically sourced strains identified a single sequence type (ST578) in all 13 outbreaks. Bayesian evolutionary analysis sampling (BEAST) was used to estimate when certain lineages diverged, and revealed two separate diverging events, one in 1991 and one in 2005, which resulted in sublineages ST578A and ST578B, respectively. Single-nucleotide polymorphism (SNP) analysis of the core genes from the ST578 strains in the collection revealed that >10% of these genes contained regions of high SNP density, which is characteristic of recombination hot spots. Once these hot spots were removed and BEAST analysis was repeated, only 2% of the total variability

between strains could be attributed to nucleotide substitutions; the remainder was due to non-vertical inheritance. The recombination hot spots were mostly found in genes involved in membrane biogenesis, metabolism and virulence or defence mechanisms. ST578A was responsible for the major clinical outbreaks in the early 1990s, but by 2008 it had mostly been eliminated from Alcoy owing to effective public health control measures. A 2009 ST578 outbreak was traced to an asphalt-paving machine located outside the city limits<sup>2</sup>, and the authors hypothesize that this outbreak led to the establishment of the ST578B lineage. The genetic diversity of these chronologically separate outbreaks, stemming from 16 recombination events, led the authors to conclude that recombination in *L. pneumophila* readily occurs in the environment and is responsible for the rapid evolution of the species.

It is clear that recombination has far-reaching consequences for the evolution of bacterial genomes. A study by Chewapreecha *et al.*<sup>3</sup> predicted that the high rates of recombination present in *Streptococcus pneumoniae* would generate enough diversity to make a bacterial genome wide association study (GWAS) possible. The authors carried out a GWAS on two large, geographically diverse data sets (of 3,085 and 616 isolates) to assess the association of SNPs with an antibiotic resistance phenotype. They pinpointed 301 SNPs from 51 loci as being linked to β-lactam resistance, improving on previous estimates that placed resistance

mutants within larger genetic regions. The more deeply sampled data set enabled candidate loci to be narrowed down in both vaccine-targeted and non-vaccine-targeted lineages, which potentially has clinical implications.

Even minor recombination events can shape bacterial genomes, as outlined by another recent study that examined signatures of genetic exchange in 95 *Staphylococcus aureus* strains<sup>4</sup>. Everitt and colleagues found that mobile genetic elements (MGEs) facilitated fine-scale recombination (with 0.5–1 kb exchanged per event) by providing homoplasy hot spots, especially in core regions. An increase in predicted recombination frequency of 2.5-fold at so-called fault lines in the genome occurred within 1 kb of MGE placement, including conjugative transposons and genomic islands.

Collectively, these studies serve as a reminder that recombination is often a misunderstood or underrated driver of evolution. More in-depth studies are required to fully understand its impact on genomic diversity and the consequences for the interpretation of epidemiological data, virulence and resistance.

Christine J. Boinett and Amy K. Cain are at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

e-mail: [microbes@sanger.ac.uk](mailto:microbes@sanger.ac.uk)

doi: 10.1038/nrmicro3388

Published online 10 November 2014

1. Sanchez-Buso, L. *et al.* Recombination drives genome evolution in outbreak-related *Legionella pneumophila* isolates. *Nature Genet.* **http://dx.doi.org/10.1038/ng.3114** (2014).
2. Coscolá, M. *et al.* Legionellosis outbreak associated with asphalt paving machine, Spain, 2009. *Emerg. Infect. Dis.* **16**, 1381–1387 (2010).
3. Chewapreecha, C. *et al.* Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. *PLoS Genet.* **10**, e1004547 (2014).
4. Everitt, R. G. *et al.* Mobile elements drive recombination hotspots in the core genome of *Staphylococcus aureus*. *Nature Commun.* **5**, 3956 (2014).

**Competing interests statement**  
The authors declare no competing interests.

