GENOME WATCH

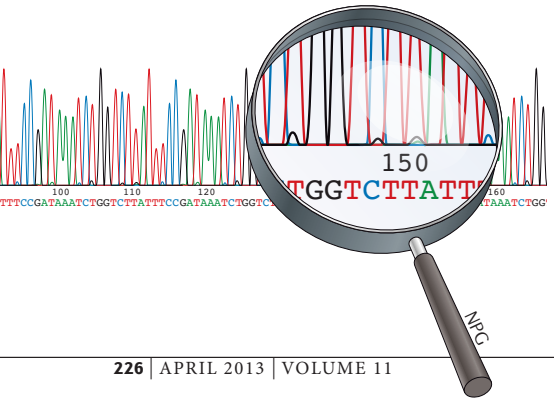
A CRISPR view of genome sequences

Amy K. Cain and Christine J. Boinett

This month's Genome Watch explores recent applications of the CRISPR immune system for bacterial phylogenetic analysis and genome editing.

The CRISPR–Cas (clustered regularly interspaced short palindromic repeats–CRISPR-associated proteins) adaptive immune system is widespread in bacteria and archaea and provides heritable protection against disruptive mobile genetic elements (MGEs), such as bacteriophages and plasmids. CRISPR loci contain a series of repetitive DNA motifs separated by spacer sequences; these spacers are derived from MGEs and incorporated after exposure to each new foreign element. The CRISPR transcript is processed into small CRISPR RNAs, which are displayed on Cas protein complexes, enabling RNA-guided degradation of the foreign DNA by Cas nucleases.

Recent studies have shown that distinct CRISPR profiles occur within different species and strains, which has facilitated their use as a typing tool. For example, in strains of *Salmonella* spp., CRISPR spacer content was found to be strongly correlated with both serotype and multilocus sequence type (MLST), leading to the development of a comprehensive typing scheme[1]. Another study combined CRISPR analysis with MLST typing and pulsed-field gel electrophoresis, yielding increased discriminatory power for typing of the food-borne pathogen *Salmonella enterica* subsp. *enterica* serovar Enteritidis[2].



Another recent study highlights the use of CRISPR loci in providing insight into the phylogenetic relationship between two closely related tuberculosis-causing bacteria: *Mycobacterium tuberculosis* and *Mycobacterium canettii*[3]. The latter species, which belongs to the diverse group of smooth tubercle bacilli (STB), is restricted geographically to East Africa, whereas *M. tuberculosis,* from the relatively homogeneous *M. tuberculosis* complex (MBTC) group, is globally distributed. The authors sequenced the genomes of five STB strains and compared these sequences to those of four previously sequenced MBTC strains. Compared with the MBTC strains, which shared one conserved CRISPR locus, both distinct and additional CRISPR loci were found in the STB strains. Although some STB CRISPR loci occupied the same genomic position as, and encoded similar Cas proteins to, the MBTC locus, the STB loci contained some different sets of spacers, suggesting that these CRISPR systems were acquired before divergence of the two groups and that the different spacers, as well as the additional CRISPR loci, emerged in STB strains after divergence of the two groups. The variation in CRISPR systems, in combination with other genomic differences detected in the analysis, led the authors to suggest that *M. tuberculosis* emerged from an ancestral STB-like pool of mycobacteria.

CRISPR–Cas systems have also been cleverly exploited in the laboratory to edit the genomes of *Streptococcus pneumoniae* and *Escherichia coli*[4]. *S. pneumoniae* cells were transformed with both a modified CRISPR construct designed to target specific *S. pneumoniae* chromosomal loci, and an editing template that recombined with the target sequence to introduce mutations. Transformants that failed to recombine the editing sequence into the genome were susceptible to CRISPR attack and killed, whereas those containing the mutated gene avoided recognition and were spared Cas-mediated cleavage. This approach eliminates the need for selectable markers when generating site-specific mutations in bacterial genomes. The highly recombinogenic nature of *S. pneumoniae* allowed for the use of DNA constructs, whereas plasmids containing the CRISPR loci were successfully used in *E. coli*, which means that the approach could be applied to a range of microorganisms. Similar techniques have been used for editing multicellular eukaryotic genomes, including human cell lines, mice and zebrafish[5–7].

The flexibility and specificity of genome editing using CRISPR loci enables the efficient generation of mutated genotypes in diverse species. Furthermore, as CRISPR loci show strain-specific conservation at the nucleotide level, they are proving to be valuable markers for typing studies and, in conjunction with whole-genome sequencing, can provide insights into the phylogenetic relationships between different bacteria.

Amy K. Cain and Christine J. Boinett are at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro2997
Published online 11 March 2013

1. Fabre, L. *et al.* CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. *PLoS ONE* **7**, e0036995 (2012).
2. Shariat, N. *et al.* The combination of CRISPR-MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. *Food Microbiol.* **34**, 164–173 (2013).
3. Supply, P. *et al.* Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis*. *Nature Genet.* **45**, 172–179 (2013).
4. Jiang, W. *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotech.* 29 Jan 2013 (doi:10.1038/nbt.2508).
5. Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotech.* 29 Jan 2013 (doi:10.1038/nbt.2501).
6. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
7. Cho, S. W. *et al.* Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature Biotech.* 29 Jan 2013 (doi:10.1038/nbt.2507).

Competing interests statement
The authors declare no competing financial interests.