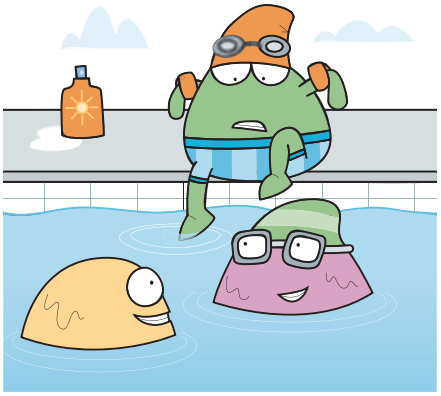
GENOME WATCH

Testing the water: marine metagenomics

Gemma Langridge



Neil Smith

This month's Genome Watch reviews three recent papers that describe metagenomic analyses of marine and freshwater bacteria.

Although bacterial genomes are traditionally sequenced from cultured isolates, metagenomic analyses allow researchers to gain insights into the genomes of unculturable microorganisms. Marine habitats have been a particular target and two recent studies that build on our current knowledge of the 'marine metagenome' have applied the metagenomic principle in different ways[1,2].

Woyke *et al.*[1] used flow cytometry to extract live cells from a sample of coastal water. The DNA from 11 single cells was individually amplified using multiple displacement amplification (MDA) and two cells were identified as the proteorhodopsin-containing *Flavobacteria bacterium* MS024-2A and *F. bacterium* MS024-3C. Following a second round of MDA, DNA from the uncultured flavobacteria was used for both Sanger shotgun sequencing and 454 pyrosequencing. Final assemblies were 1.9 Mb (17 contigs) and 1.5 Mb (21 contigs), which are estimated to represent 91% and 78% of each genome, respectively. Both genomes were smaller in comparison to other sequenced Bacteroidetes, a feature that is thought to be a nutrient and energy conserving adaptation in marine alphaproteobacteria. Both genomes also encode genes for the hydrolysis of allophanate (a by-product of urea breakdown), which is a likely additional source of nitrogen. Unexpectedly, *F. bacterium* MS024-2A is unique among marine flavobacteria, as it encodes the nickel and iron hydrogenase genes *hyaA* and *hyaB*, which suggests that it can use hydrogen as an energy source.

Palenik *et al.*[2] also sampled coastal water but took an alternative approach, using the natural fluorescence of *Synechococcus* spp. with flow cytometry to enrich for fluorescing cells. *Synechococcus*, a cyanobacterial genus, is classified into four clades. The enriched collection of Palenik *et al.* was estimated to contain ~20,000 cells from clade I and ~50,000 cells from clade IV, which are the two clades that are typically present in the coastal environment. Sequence reads were generated by 454 pyrosequencing after whole genome amplification and were then mapped to the four complete *Synechococcus* genomes that are currently available. Although reads mapped to coastal reference strains 20 times more frequently than to open-ocean reference strains, reads tended not to map to regions of atypical nucleotide content, even in the coastal reference strains. This implies that these regions represent horizontally acquired genes, possibly as genomic islands, that are not conserved across *Synechococcus* spp. In particular, a large region of atypical nucleotide content in the *Synechococcus* sp. CC9902 reference genome is absent in the metagenome, indicating that this region is a possible hot spot for recombination of acquired DNA.

Three families of plasmids that represent genetic elements previously unseen in marine cyanobacterial genomes were also reconstructed from the sequence data. This raises the possibility that plasmids have a role in gene transfer in coastal *Synechococcus* genomes, which contain none of the genomic islands that are characteristic of the phage seen in open-ocean genomes such as *Synechococcus* sp. WH 8102. Multiple mechanisms for horizontal gene transfer might therefore be active in the *Synechococcus* genus.

Alongside these investigations into the marine metagenome, a collaborative study in Mexico has tackled the microorganisms living in freshwater 'microbialites', or matrices of exopolymeric substances[3]. A diverse range of heterotrophs and autotrophs were identified, and members of the cyanobacterial order Chroococcales (of which *Synechococcus* is a genus) were commonly found in morphologically distinct samples. However, the metagenome was enriched for genes involved in phosphorus metabolism and the establishment and development of biofilms, features that distinguish these freshwater microbialite communities from their marine counterparts.

Metagenomic analyses were initially used to gain a broad understanding of which species were present in a microbial community, and this remains the most common form of investigation. However, as described above, marine metagenomics is expanding the approaches that can be taken to gain a better understanding of unculturable organisms.

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

1. Woyke, T. *et al.* Assembling the marine metagenome, one cell at a time. *PLoS ONE* **4**, e5299 (2009).
2. Palenik, B., Ren, Q., Tai, V. & Paulsen, I. T. Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. *Environ. Microbiol.* **11**, 349–359 (2009).
3. Breitbart, M. *et al.* Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Cienegas, Mexico. *Environ. Microbiol.* **11**, 16–34 (2009).

DATABASES

Entrez Genome Project: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj
Flavobacteria bacterium MS024-2A | *Flavobacteria bacterium* MS024-3C | *Synechococcus* sp. CC9902 | *Synechococcus* sp. WH 8102

ALL LINKS ARE ACTIVE IN THE ONLINE PDF