Lawrence Berkeley National Laboratory
Recent Work

Title
Metagenomic Bacterial Finshing at the JGI

Permalink
https://escholarship.org/uc/item/6c65v3s7

Authors
Lowry, Stephen
Clum, Alicia
Goltsman, Eugene
et al.

Publication Date
2008-05-26
Phylopythia binning of Thalassiosira species (from McHardy et al.(2))

This figure shows the binned species of Phylopythia at the level of genus and 32% more at phylum level, with only 3% unclassified. The binned species constitute the largest chal of gammaproteobacteria in an Australian waste treatment sludge community studied for its Accumulibacter species in parallel with the US species sequenced at the JGI.

Desulfodrillus audaxviator

Described as a 'one organism ecosystem' (4), which is to say it has been environmentally enriched. The representations in Orchid (available from the Stanford Genome Center) at right show clearly the low complexity of this community. (Orchid displays an arc file distributed around a circle, the heavy blue sections being continuous sequence.) The left hand image is a representation of the contigs in the earliest assembly arc-file, while the right is the final. Only about 100 bp remain uncorporated out of 28000.

Korarchaeum cryptofilum OPF8

From a more complex thermophilic community, it was maintained as a community in the lab, and enriched by differential lysis. This organism was finished by conventional means, its small size was helpful.

Accumulibacter phosphatis Type IIA

A more difficult, and unfortunately more typical situation was presented by A. phosphatis. It was the first community subjected to Phylopythia binning, reducing its number of reads by more than half and which eventually produced the assembly illustrated at right. One can see that after completion about 40% of the initial reads have fused into the final consensus.

There is moderate polymorphism. It was our good fortune that 454 sequencing came on the scene as we considered finishing the genome. It allowed the bridging of many small gaps, and confirmed other only moderately supported regions.

Endomicrobium trichonymphae

A more challenging situation exists with E. trichonymphae. The sequence obtained from the combined hind-guts of a group of termite nest-mates was biased targeting a bacterial symbiote of a symbiotic flagellate protist. The enrichment reduced the read number by about 60%. The pertinent portion of the assembly at this point seems to consist of a number of strains considerably more distant from each other than the polymorphic members of the previous projects. This can be illustrated by the cartoons on the right. In it, contigs with clear BLAST related sequences is a very closely related organism are aligned against that genome (AP009510.3), which is represented by the left vertical bar. The coloring is arbitrary to make contigs distinct.

There are certainly limits to the genomic resolution of communities, but luck and ingenuity have not been exhausted yet.