Possibilities and obstacles in recovery of genomes from elusive microbes in complex metagenomes
Karst, Søren Michael; Albertsen, Mads; Nielsen, Jeppe Lund; Nielsen, Kåre Lehmann; Nielsen, Per Halkjær

Publication date:
2014

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Possibilities and obstacles in recovery of genomes from elusive microbes in complex metagenomes

Søren M. Karst, Mads Albertsen, Jeppe L. Nielsen, Kåre L. Nielsen and Per H. Nielsen

Center for Microbial Communities, Aalborg University, Aalborg, Denmark

Introduction

Representative genomes provide an entry point for understanding a given microbial ecosystem. The genomes give insights into the metabolic potential and possible roles of the bacteria, as well as being essential when applying other -omics based techniques. For elusive, unculturable bacteria, metagenomics can be a useful tool to extract their genomes directly from their environments. However the approach has limitations when the community complexity is high and microdiversity is present.

In this study, we explored the potentials and obstacles faced when assembling genomes from complex metagenomes using activated sludge as a model system.

Aims

- To investigate the impact of microdiversity and community complexity on metagenomic assembly, binning and genome extraction.
- To investigate whether short-term enrichment can mitigate the adverse effects of the above.

Methods

Short-term Enrichment

Seed sample
- Fullscale ETPR WWTP

Batch Incubations
- Casein, Glucose or Butyrate
- Aerobic or Anaerobic
- Incubations at 20 °C
- Incubation for 7 days

Screen Incubations
- V4 16S rRNA amplifications
- 14 of 50 samples chosen for metagenome sequencing

Metagenome Binning

Sequencing and Assembly
- Illumina HiSeq 250/200 bp (500 Gbp)
- De novo assembly (SPAdes) and read mapping using DILLC genomics workflow

Coverage Binning
- Read coverage profiles from 14 samples
- Taxon relative abundance
- Essential genes
- Fasta and read connections
- Tetranucleotide frequencies

Re-Assembly and Finishing

- Bio-specific read de novo re-assembled in SPAdes
- Manual finishing using TRIMMER, Circos and R-Binning

Microdiversity Simulation

Simulate diversity
- Simulate reads
- De novo assembly
- Visualize

Differential coverage binning and re-assembly

Results

Conclusions

- Enrichments decreased overall complexity, changed relative abundance and facilitated better coverage binning and allowed for numerous high quality genome bins to be extracted.
- Despite the short-term enrichments microdiversity was still present, which compromised complete recovery of genomes from many species.
- The simulation underlines the devastating impact of micro-diversity on assembly and binning.

Differential coverage binning and re-assembly

- Points represent scaffolds, scaled by their length and colored based on the presence of essential single copy genes. (A) is a coverage plot of a metagenome assembly of all data. Circled numbers indicate extracted high quality genome bins. In this example bin 5 (A) was targeted for extraction. Bin 5 was found to be enriched in the casein incubation and a sample from this was therefore used for targeted metagenome re-assembly. Bin 5 was binned again from the casein sample metagenome (B & C), and subsequently refined with re-assembly and scaffolding (D). The extracted genome belongs to phylum Bacteroidetes and genus Ferribacteriaceae.