Clemens Heiderer

A<u>u</u>tomatic annotation of a genome assembly

Brochothrix thermosphacta genome was hybrid assembly with 30 art-illumina coverage and a 60 coverage Bac Bio coverage.

Used versions for assembly: Pbsim 1.04 for Bac Bio
                            Art-illumina: 20160605
Used version: prokka 1.14.6

used scaffold: broc.Msv3.500bp.250bp.30x.pe.pacbio.60.spades/scaffolds.fasta

commands for –prefix broch1: --genus Brochothrix
                             --gram +
                             -- usegenus

Skipped genus-specific proteins as it can't see
/apps/prokka/1.14.6/db/Brochothrix

<u>broch1.tsv file:</u> Tab-separated file of all features:
locus_tag,ftype,len_bp,gene,EC_number,COG,product

Function	gene	EC_number	product
Peptidoglycan modification	oatA		O-acetyltransferase OatA
Intracellular survival:resistance to osmotic stress	prsA1	COG0760	Foldase protein PrsA 1
Synthesis of nucleotides	Prs_1, prs_2	COG0462	Ribose-phosphate pyrophosphokinase
Surface protein anchoring	lgt	COG0682	diacylglyceryl transferase
enzym for mannose 6-phosphate	MngB_1, mngB_2	COG0383	Mannosylglycerate hydrolase
oligopeptide uptake with pleiotropic effects on growth at cold temperatures	OppA_1, oppA_2	COG4166	Oligopeptide-binding protein OppA
Iron Uptake	hbp1	COG5386	Hemin/hemoglobin-binding protein 1

<u>broch1.fnn file:</u>

Nucleotide FASTA file of all the prediction transcripts (CDS, rRNA, tRNA, tmRNA, misc_RNA)

2510 predicted genes

Clemens Heiderer

<u>broch1.faa file:</u> Protein FASTA file of the translated CDS sequences.

2401 CDC genes

875 hypothetical genes
2 dipeptidase
6 glycosyltransferase
8 N-acetyltransferase
57 transporter
3 trans-acting regulator
3 transmembrane protein
23 ABC transporter

2 contigs totalling 2573114 bp:

predicted tRNAs with aragorn: 83 tRNAs

Found 25 rRNAs: 5S ribosomal RNA, 23S ribosomal RNA,23S ribosomal RNA each found 8 times

2402 CDS, there are still 2402 unannotated CDS left,

Total of 107 tRNA + rRNA features

Found 1 CRISPRs: CRISPR1 NODE_1_length_2518277_cov_16.390230 2391868 with 5 spacers

Functional Annotation with sequence similarity search (BLAST):
Fixed 145 colliding /gene names
Assigned 2510 locus_tags to CDS and RNA features

Clemens Heiderer


```
commands for –prefix broch2: --genus Brochothrix
                             --gram +
                             -- usegenus
                             --rfam
```

```
Same:
      Found 83 tRNAs
      Found 25 rRNAs
```

```
Different:

      Found 50 ncRNAs
      Assigned 2560 locus_tags to CDS and RNA features
```