

Results Report

Autor: Mike Burton

Date: 11/03/2018

The software Abyss assembler was use to analyze the file from SRR4149288 (<https://www.ebi.ac.uk/ena/data/view/SRR4149288>) and then select the software Prokka to make the annotation in both software it was evaluated two k-mer different to compare the metrics of assembler and metrics for annotation

Table 1. Metrics of assembler and annotation to the library SRR4149288.

| K-mer32 | | Assembly | | | | | | | | | Annotation | |
|---------|-------|----------|-----|-------|-------|--------|--------|--------|---------|----------------------------|------------|------|
| n | n:500 | L50 | min | N80 | N50 | N20 | E-size | max | sum | name | Contigs | CDS |
| 3957 | 672 | 79 | 500 | 6873 | 19577 | 38241 | 26263 | 131144 | 5229144 | azotobacter32-unitigs.fa | | |
| 3256 | 392 | 48 | 506 | 13696 | 33053 | 64277 | 43483 | 190775 | 5345967 | azotobacter32-contigs.fa | 3256 | 4954 |
| 3201 | 350 | 39 | 506 | 16798 | 41430 | 71450 | 51781 | 190775 | 5345609 | azotobacter32-scaffolds.fa | | |
| K-mer64 | | Assembly | | | | | | | | | | |
| n | n:500 | L50 | min | N80 | N50 | N20 | E-size | max | sum | name | Contigs | CDS |
| 1300 | 424 | 39 | 500 | 15490 | 37348 | 85338 | 53186 | 198383 | 5425105 | azotobacter64-unitigs.fa | | |
| 1040 | 305 | 35 | 505 | 19713 | 46208 | 98936 | 59648 | 200118 | 5566651 | azotobacter64-contigs.fa | 1040 | 5137 |
| 978 | 262 | 28 | 505 | 23322 | 54067 | 117122 | 73513 | 240210 | 5565150 | azotobacter64-scaffolds.fa | | |

With the annotation found in the assembler k=32, 4954 CDS meanwhile the assembler k=64 found 5137 CDS. Analyzing the annotation oh both assembler found one protein histidine kinase (the six genes cording to the domains barA1-6)