**S2 Table. Taxonomic profiling with MG-RAST statistics.**

|  |  |  |
| --- | --- | --- |
|  | Sample 1 | Sample 2 |
| Sequences pre-screened with 70% identity | 9,771,866  | 6,845,168  |
| *De novo* clusters with 97% identity | 462506  | 462506  |
| Number of BLAT queried sequences with a hit | 22543  | 22543  |
| Number of BLAT hits | 33457  | 28667 |