Table S2

Bioinformatics parameters and databases used in the analysis of pyrosequencing results.

|  |  |  |
| --- | --- | --- |
| Step | Parameter Targeted rDNA gene 18S | |
| Preprocessing | Length threshold | 300 |
|  | Number of ambiguities tolerated | 0 |
|  | Detection of proximal primer sequence | Complete and perfect |
|  | Detection of distal primer sequence | Perfect, but potentially incomplete |
| Clustering | Chosen level of similarity (%) | 95 |
|  | Ignoring differences in homopolymer lengths | Yes |
| Filtering | Chosen clustering similarity threshold | 95 |
|  | Used taxonomic database | SILVA (r111) |
|  | Chosen taxonomic level | Phylum |
|  | Similarity or confidence threshold (%) | 85 |
| Taxonomy | Used taxonomic database | SILVA (r111) |
|  | Method or tool of comparison | MegaBLAST |
|  | Similarity or confidence threshold (%) | 80 |
| Analysis | Chosen level of similarity (%) | 95 |
|  | Ignoring differences in homopolymer lengths | Yes |
|  | Computation of a Unifrac distance matrix | Yes |