**Supplementary Table S1.** Pipeline employed for bioinformatics analysis.

|  |  |  |
| --- | --- | --- |
| Step | Command | Note |
| Implemented in QIIME |
| Joining paired-end MiSeq reads | join\_paired\_ends.py | Using SeqPrep algorithm |
| Demultiplexing and quality filtering | split\_libraries\_fastq.py | Phred quality score 19 |
| OTU picking strategy | pick\_open\_reference\_otus.py  | OTU-picking method UCLUST, taxonomy - GreenGenes v13.5 database |
| Negative OTUs elimination | filter\_otus\_from\_otu\_table.py | Exclusion of all OTUs associated with negative samples |
| Samples summary | biom summarize-table | Samples with less than 200 counts were excluded |
| Individual samples alpha diversity | alpha\_rarefaction.py thenadd\_alpha\_to\_mapping\_file.py | Using Shannon, chao1 and Simpson index. Rarefaction implemented at depth of 200 sequences per sample |
| Alpha diversity between samples | alpha\_rarefaction.py thencompare\_alpha\_diversity.py | Comparison with the usage of non-parametric t-test |
| Beta diversity | beta\_diversity\_through\_plots.py then make\_2d\_plots.py | unweighted Unifrac distance |
| Beta diversity with data normalization | filter\_otus\_from\_otu\_table.py, thennormalize\_table.py, thenbeta\_diversity\_through\_plots.py and make\_2d\_plots.py | Include all features that are at least in 3 samples. Normalization with CSS algorithm; unweighted Unifrac distance |
| ANOSIM | compare\_categories.py | Number of permutations: 999 |
| Implemented in Galaxy |
| Normalization | Normalize By Copy Number |  |
| Metagenome predictions | Predict Metagenome |  |
| Function annotation | Categorize by function |  |
| Implemented in R |
| Construction of MRobject | newMRexperiment | MetagenomeSeq-package |
| Filtering | filterData | MetagenomeSeq-package;Include all features which are at least in 3 samples |
| Normalization | cumNorm | MetagenomeSeq-package (CSS normalization algorithm);normalization factor = 0.5 |
| Searching of differences in taxonomy between groups | aggTaxfitZig | Taxonomic aggregation was performed on each level |
| Heatmap construction | plotMRheatmap | Heatmap on phylum level |
| Bubble-plot construction | ggplot | Bubble-plot on phylum level |
| Searching of differences in functional annotations (KEGG) between groups | wilcox.testglm | logistic regression was used for glm model |

.