**Supplementary Table S3.** Assessment effects in the discovery cohort all over the whole experiment process\*.

|  |  |  |
| --- | --- | --- |
| Assessment effects | ROC curves# AUC (95% CI) | *P* - values |
| Alpha diversity (Shannon index) | 0.98 (0.95 — 1.00) | < 0.0001 |
| Beta diversity& |  |  |
| Unweighted UniFrac | 0.98 (0.95 — 1.00) | < 0.0001 |
| Weighted UniFrac | 0.89 (0.72 — 1.00) | < 0.0001 |
| Relative abundance of major phyla |  |  |
| *Proteobacteria* | 0.79 (0.60 — 0.97) | 0.010 |
| *Firmicutes* | 0.61 (0.37 — 0.84) | 0.345 |
| *Bacteroidetes* | 0.87 (0.69 — 1.00) | 0.001 |
| *Fusobacteria* | 0.98 (0.94 — 1.00) | < 0.0001 |
| *Actinobacteria* | 0.69 (0.49 — 0.88) | 0.096 |
| *Spirochaetes* | 0.82 (0.66 — 0.98) | 0.005 |
| *Thermi* | 0.90 (0.77 — 1.00) | < 0.0001 |

ROC, Receiver Operating Characteristic curve; AUC, Area Under the Curve; CI, Confidence intervals.  
\*The 16S rDNA tags were clustered using USEARCH (v9.1) after filtering (low qulaility reads, N reads) and overlapping.

#ROC curves and P-values were analysed according to Wilson/Brown method recommended by GraphPad Prism v8.0.2. Good reproducibility is showed by AUC results.

&Beta diversity used the first principal coordinate from weighted UniFrac and unweighted UniFrac distance matrix