Table S2. Statistics of sample sequences

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **SampleType** | **Number of Samples**  | **Sequence Site** | **Maximum** | **Minimum** | **Mean** | **Standard Deviation** | **Source** |
| Gastric juice | 28 | 16S rRNA gene V4 region | 40445 | 26062 | 34976  | 3814  | Illumina Miseq sequencing |
| Saliva | 248 | 16S rRNA gene V4 region | 27558 | 3004 | 6355  | 3432  | Downloaded from HMP and cut from 16S rDNA V3-V5a |
| Stool | 271 | 16S rRNA gene V4 region | 76075 | 3018 | 8343  | 6645  | Downloaded from HMP and cut from 16S rDNA V3-V5a |

a: In Human Microbiota Project (HMP), V4 sequences of 16S rRNA gene are not included, but V3-V5 sequences of 16S rRNA gene are available. As a result, V3-V5 sequences of saliva and stool were downloaded from HMP and V4 sequences were cut from V3-V5 sequences using a program designed by Perl.