

Table 2. Fitting the UNTB-HDP model to human gut microbiota.

Taxa	N	S	\tilde{J}	θ	I_i			p_N	p_L
					l	m	u		
Bacteroidetes	231	569	596	148.6	1.5	5.5	13.7	0.0 (0.0)	0.0 (0.0)
Bacteroidaceae	208	224	506	51.4	0.7	3.3	7.6	0.0 (0.0)	0.03 (0.0)
Bacteroides	208	224	506	51.4	0.7	3.3	7.6	0.0 (0.0)	0.03 (0.0)
Firmicutes	277	4770	1009	1382.3	21.4	44.8	81.0	0.0 (0.0)	0.0 (0.0)
Incertae Sedis XIV	87	176	264	39.2	1.7	9.8	27.5	0.0 (0.0)	0.05 (0.004)
Blautia	87	175	264	38.9	1.6	10.1	27.1	0.0 (0.0)	0.06 (0.003)
Lachnospiraceae	164	873	248	262.9	6.5	13.0	21.2	0.0 (0.0)	0.0 (0.0)
Ruminococcaceae	239	1471	409	411.0	4.5	16.1	38.1	0.0 (0.0)	0.0 (0.0)
Faecalibacterium	141	301	297	71.7	1.0	7.5	21.4	0.0 (0.0)	0.004 (0.0)

Results are given for 3% OTUs at different levels, quantities given in the table are: N - the no. of samples with > 150 reads; S - the number of 3% OTUs; \tilde{J} - the median sample size; θ - the fitted biodiversity parameter; I_i - the fitted immigration rates where l, m and u are the lower 2.5%, median and upper 97.5% quantiles respectively; p_N - the proportion of simulated neutral samples exceeding the observed data likelihood; and p_L - the proportion of simulated locally neutral samples exceeding the observed data likelihood. The figures in parantheses give pseudo p-values for the equivalent complete gut microbiome data set randomly sampled down to the same size as the individual taxa.