

Table 1: Validation of Ray Communities using NIH Human Microbiome Project mock communities. The mock communities were described in PAPER and the repository and accession numbers are available therein.

Species	16S			Genome			Even mixture			Staggered mixture		
	copies per genome	length (Mb)	16S copies	Expected	Actual nucleotide proportion	16S copies	Expected	Actual nucleotide proportion	16S copies	Expected	Actual nucleotide proportion	16S copies
<i>Acinetobacter baumannii</i> (ATCC 17978)	5	4.02	100000	5.27%	12.68%	12.60%	0.26 %	0.14%	10000	0.26 %	0.14%	10000
<i>Actinomyces odontolyticus</i> (ATCC 17982)	3	2.39	100000	5.23%	2.80%	3.23%	0.02 %	0.00%	1000	0.02 %	0.00%	1000
<i>Bacillus cereus</i> (ATCC 10987)	12	5.43	100000	2.97%	0.17%	0.29%	1.49 %	0.02%	100000	1.49 %	0.02%	100000
<i>Bacteroides vulgatus</i> (ATCC 8482)	7	5.16	100000	4.84%	2.40%	2.14%	0.02 %	0.00%	1000	0.02 %	0.00%	1000
<i>Candida albicans</i> (SC5314)	0	27.56	1120	0.00%	0.00%	0.00%	0.00 %	0.00%	1000	0.00 %	0.00%	1000
<i>Clostridium beijerinckii</i> (ATCC 51743)	14	6.00	100000	2.81%	4.73%	4.00%	1.41 %	0.90%	100000	1.41 %	0.90%	100000
<i>Deinococcus radiodurans</i> (DSM 20539)	3	3.28	100000	7.17%	50.06%	51.95%	0.03 %	0.08%	1000	0.03 %	0.08%	1000
<i>Enterococcus faecalis</i> (ATCC 47077)	4	3.36	100000	5.51%	1.44%	1.59%	0.02 %	0.00%	1000	0.02 %	0.00%	1000
<i>Escherichia coli</i> (ATCC 70096)	7	4.60	100000	4.31%	0.06%	0.09%	21.73 %	0.35%	100000	21.73 %	0.35%	100000
<i>Helicobacter pylori</i> (ATCC 700392)	2	1.66	100000	5.45%	3.56%	3.40%	0.27 %	0.02%	10000	0.27 %	0.02%	10000
<i>Lactobacillus gasseri</i> (DSM 20243)	6	1.89	100000	2.06%	0.00%	0.02%	0.10 %	0.00%	10000	0.10 %	0.00%	10000
<i>Listeria monocytogenes</i> (ATCC BAA-679)	6	2.94	100000	3.21%	1.77%	1.99%	0.16 %	0.00%	10000	0.16 %	0.00%	10000
<i>Methanobrevibacter smithii</i> (ATCC 35061)	2	1.85	100000	6.07%	0.73%	0.69%	30.59 %	5.52%	100000	30.59 %	5.52%	100000
<i>Neisseria meningitidis</i> (ATCC BAA-335)	4	2.27	100000	3.72%	1.00%	1.21%	0.18 %	0.00%	10000	0.18 %	0.00%	10000
<i>Propionibacterium acnes</i> (DSM 16379)	3	2.56	100000	5.60%	0.01%	0.02%	0.28 %	0.00%	10000	0.28 %	0.00%	10000
<i>Pseudomonas aeruginosa</i> (ATCC 47085)	4	6.26	100000	10.27%	0.07%	0.15%	5.17 %	0.42%	100000	5.17 %	0.42%	100000
<i>Rhodobacter sphaeroides</i> (ATCC 17023)	3	4.60	100000	10.06%	3.43%	2.96%	5.07 %	27.23%	100000	5.07 %	27.23%	100000
<i>Staphylococcus aureus</i> (ATCC BAA-1718)	5	2.94	100000	3.86%	4.24%	3.64%	19.45 %	29.25%	100000	19.45 %	29.25%	100000
<i>Staphylococcus epidermidis</i> (ATCC 12228)	5	2.49	100000	3.27%	6.22%	5.58%	1.64 %	24.86%	100000	1.64 %	24.86%	100000
<i>Streptococcus agalactiae</i> (ATCC BAA-611)	7	2.16	100000	2.02%	0.01%	0.06%	10.20 %	0.32%	100000	10.20 %	0.32%	100000
<i>Streptococcus mutans</i> (ATCC 700610)	5	2.03	100000	2.66%	2.34%	2.10%	0.01 %	10.48%	1000	0.01 %	10.48%	1000
<i>Streptococcus pneumoniae</i> (ATCC BAA-334)	4	2.16	100000	3.54%	1.99%	1.60%	1.78 %	0.00%	100000	1.78 %	0.00%	100000