

Table S1. Number of 2 441 787 high quality sequences in the RDP database matching different primers.

Number of mismatches allowed	0	1	2	3
27F&1492R	35926	57086	75361	119395
27F&1492R but not 63F&M1387R	30051	36813	34059	14477
63F&1387R	52363	152042	259616	450640
63F&M1387R	52740	153122	263574	453724
63F&M1387R but not 27F&1492R	46865	132849	222272	348806