Supplemental Table S1. Taxonomic diversity of bacterial open reading frames identified from assembled contigs derived from shallow 454 pyrosequencing belonging to the additional wells of interest. The number of 454 reads ranged between 9,000 -10,000 for each sample. ORFs were identified and annotated by FragGeneScan (Rho et al. 2010) and METAREP (Goll et al. 2010), respectively.

Sample	**16S Best Hit**	**Division**	**No ORFs**	**Species with more than 10 matching ORFs***
PO1	*Veillonella dispar*	Firmicutes	78	*Veillonella parvula* (31), *V. dispar* (10), *V.* sp 3_1_44 (24)
A17	*Sphingobacterium spiritivorum*	Proteobacteria	15	*Escherichia coli* (15)
D14	*Sphingobacterium spiritivorum*	Firmicutes	8	-
		Proteobacteria	15	-
		Bacteroidetes	1,330	*Sphingobacterium spiritivorum* ATCC 33300 (623), *Sphingobacterium spiritivorum* ATCC 33861 (542)
D18	*Streptococcus pneumoniae*	Firmicutes	78	*Streptococcus pneumoniae* (43), *S. mitis* (27)
H20	*Leptrotrichia_sp*	Fusobacteria	49	*Leptotrichia hofstadii* F0254 (30), *L. buccalis* C-1013-b (16)
J05	*Streptococcus oralis*	Firmicutes	23	- , (3 hits to *S. oralis*)
K21	*Streptococcus mitis*	Firmicutes	816	*Streptococcus pneumoniae* (403), *S. mitis* (317), S. sp M143 (13)

* excluding ORFS matching suspected MDA contaminants such as *E.coli*