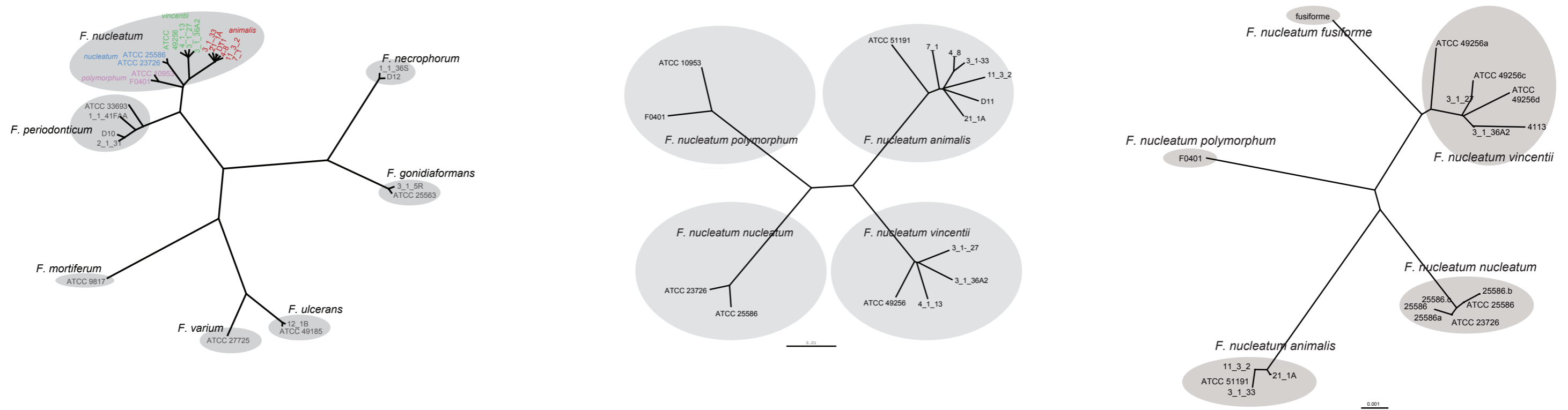


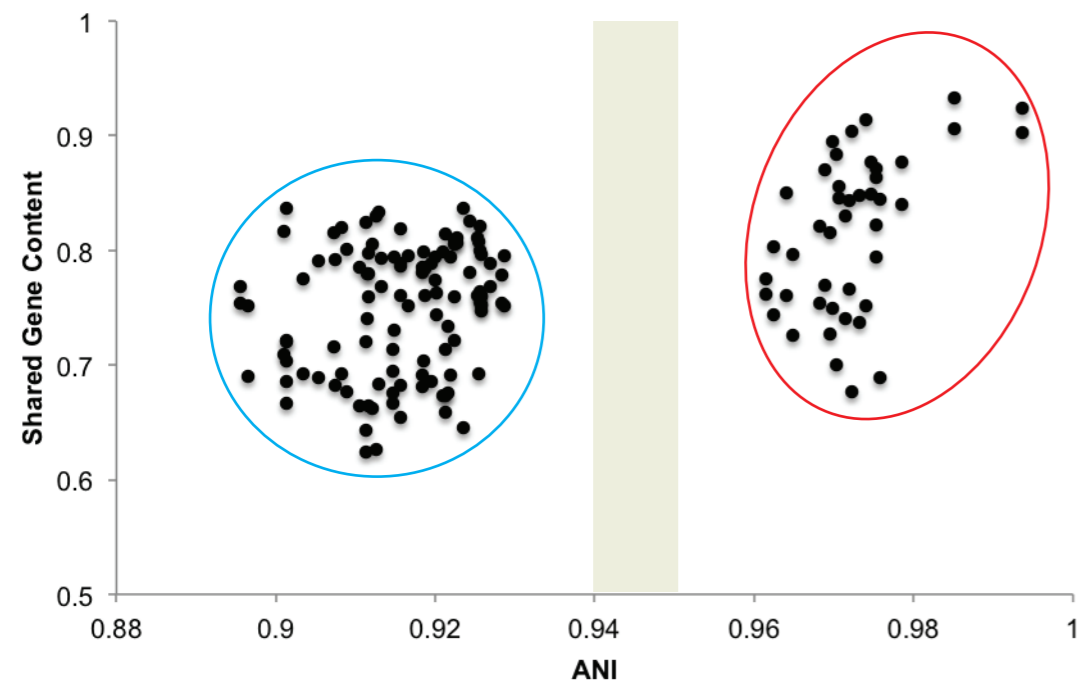
Figure S1.



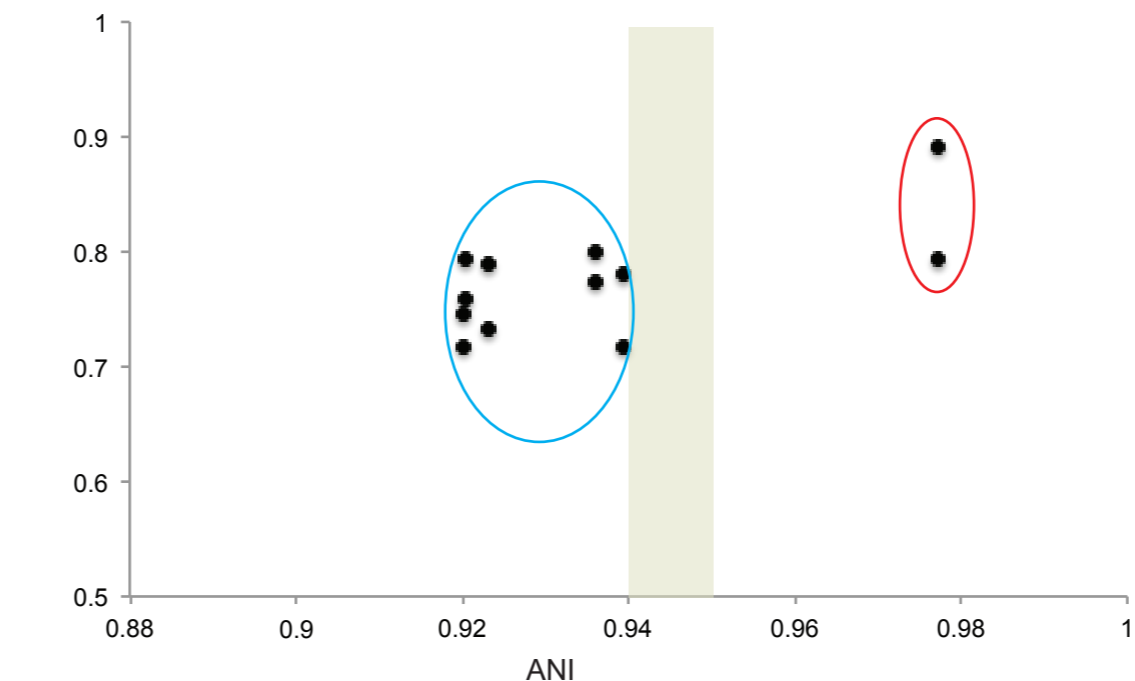
a) Phylogenetic tree used to rename strains. Unrooted tree based on single copy core orthogroups across all 26 Fusobacterium (excluding the outgroup). *F. nucleatum* F0401 clustered closely with *F. nucleatum* subsp. *polymorphum*, so we have renamed this strain *F. nucleatum* subsp. *polymorphum* F0401. *F. sp.* 2\_1\_31 and *F. sp.* 1\_1\_41FAA clustered closely with the two *F. periodonticum* strains, so we have renamed these strains *F. periodonticum*. *F. sp.* D12 clustered closely with *F. necrophorum*; *F. sp.* 3\_1\_5R clustered closely with *F. gonidiaformans*; and *F. sp.* 12\_1B clustered closely with *F. ulcerans*, so we have renamed these accordingly.

b) Unrooted tree based on 994 uniform OrthoMCL orthogroups shared between 15 strains of *F. nucleatum* (including the additional genome *F. nucleatum* subsp. *animalis* ATCC 51191). We found that six of our strains clustered closely with the *F. nucleatum* subsp. *animalis* strain (see Figure S1b). Therefore, we are renaming *F. sp.* 21\_1A, *F. sp.* D11, *F. sp.* 4\_8, *F. sp.* 3\_1\_33, *F. sp.* 11\_3\_2, and *F. sp.* 7\_1 as strains of *F. nucleatum* subsp. *animalis*. It is also clear that *F. sp.* 3\_1\_36A2, *F. sp.* 3\_1\_27, and *F. sp.* 4\_1\_13 are strains of *F. nucleatum* subsp. *vincentii*.

c) Tree of only *F. nucleatum* based on 16S rRNA gene sequences, including both the 16S rRNA gene sequences from the sequenced strain of *F. nucleatum* subsp. *fusiforme*. We observed the same phylogenetic pattern in the 16S rRNA gene-based calculations as we did with the whole-genome orthogroup-based calculations. Our observation that *F. nucleatum* subsp. *fusiforme* clusters most closely with *F. nucleatum* subsp. *vincentii* is in agreement with the Living Tree Project [2]. In the 16S rRNA gene sequence analysis, we observe that none of our strains cluster closely with *F. nucleatum* subsp. *fusiforme*.



d) ANI plot for *F. nucleatum*. Each point represents a pairwise comparison of two *F. nucleatum* genomes. Red circled area: pairwise comparisons between members of the same subspecies. Blue circled area: pairwise comparisons between members of different subspecies. A species threshold of 94-95% ANI is indicated by the green-shaded area.



d) Intra-species ANI plot for species other than *F. nucleatum*. Each point represents a pairwise comparison of two *F. periodonticum* genomes. Red circled area: pairwise comparisons between members of the same subspecies (*F. periodonticum* 2\_1\_31 and *F. periodonticum* D10). Blue circled area: pairwise comparisons between members of different subspecies. A species threshold of 94-95% ANI is indicated by the green-shaded area.