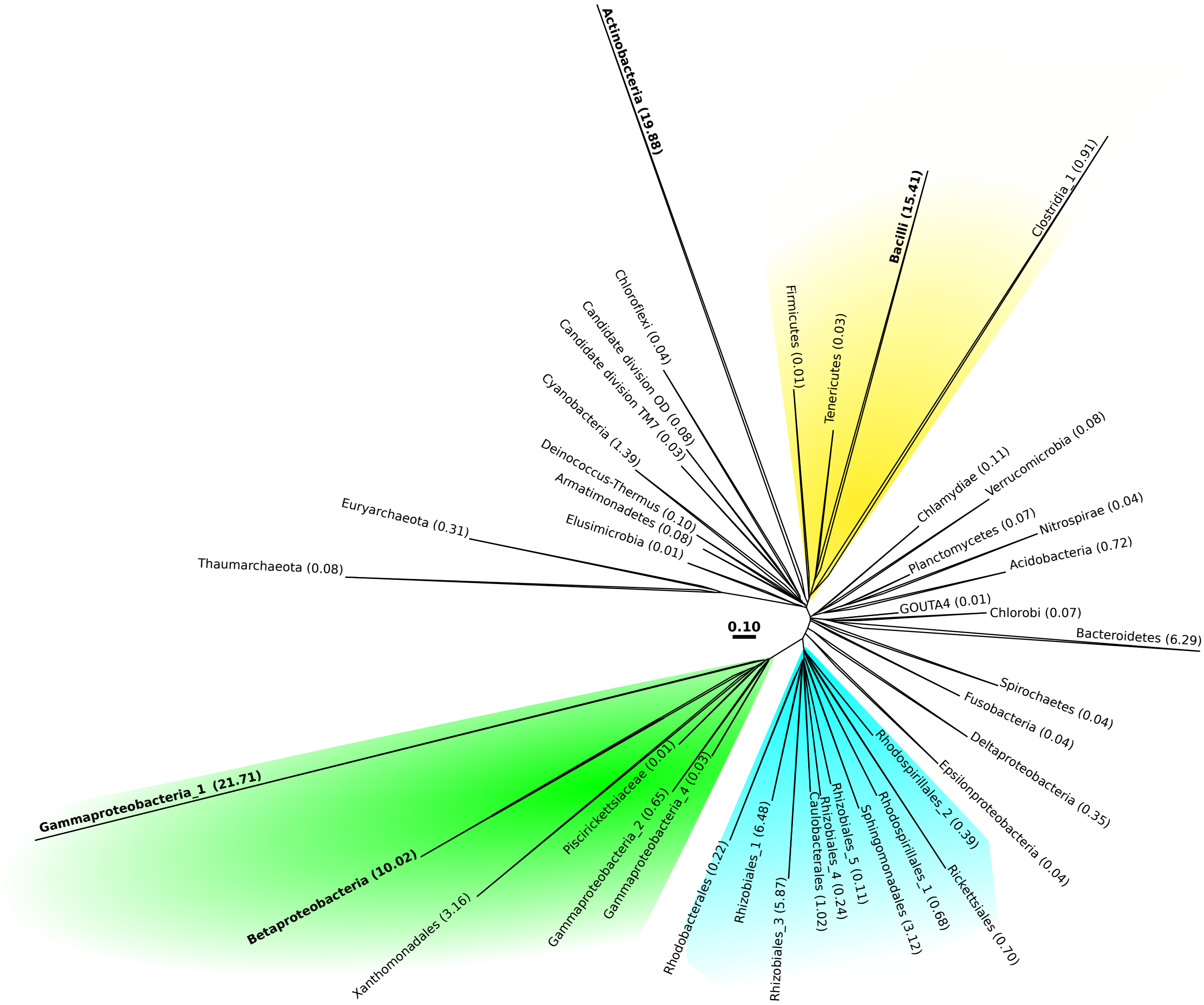

Fig. S1. Phylogenetic analysis of 16S rRNA genes from prokaryotic endophyte dataset. Names outside wedges of clustered sequences are phylum/subphylum/order taxa with their relative abundance indicated as percentages within round brackets. Taxa from endophytes with more than 10% of the total dataset are shown in bold. The (sub)phyla Gamma-/Betaproteobacteria, Alphaproteobacteria and Firmicutes are shown within green, blue and yellow shaded area, respectively. Scale bar indicates 10% sequence divergence.

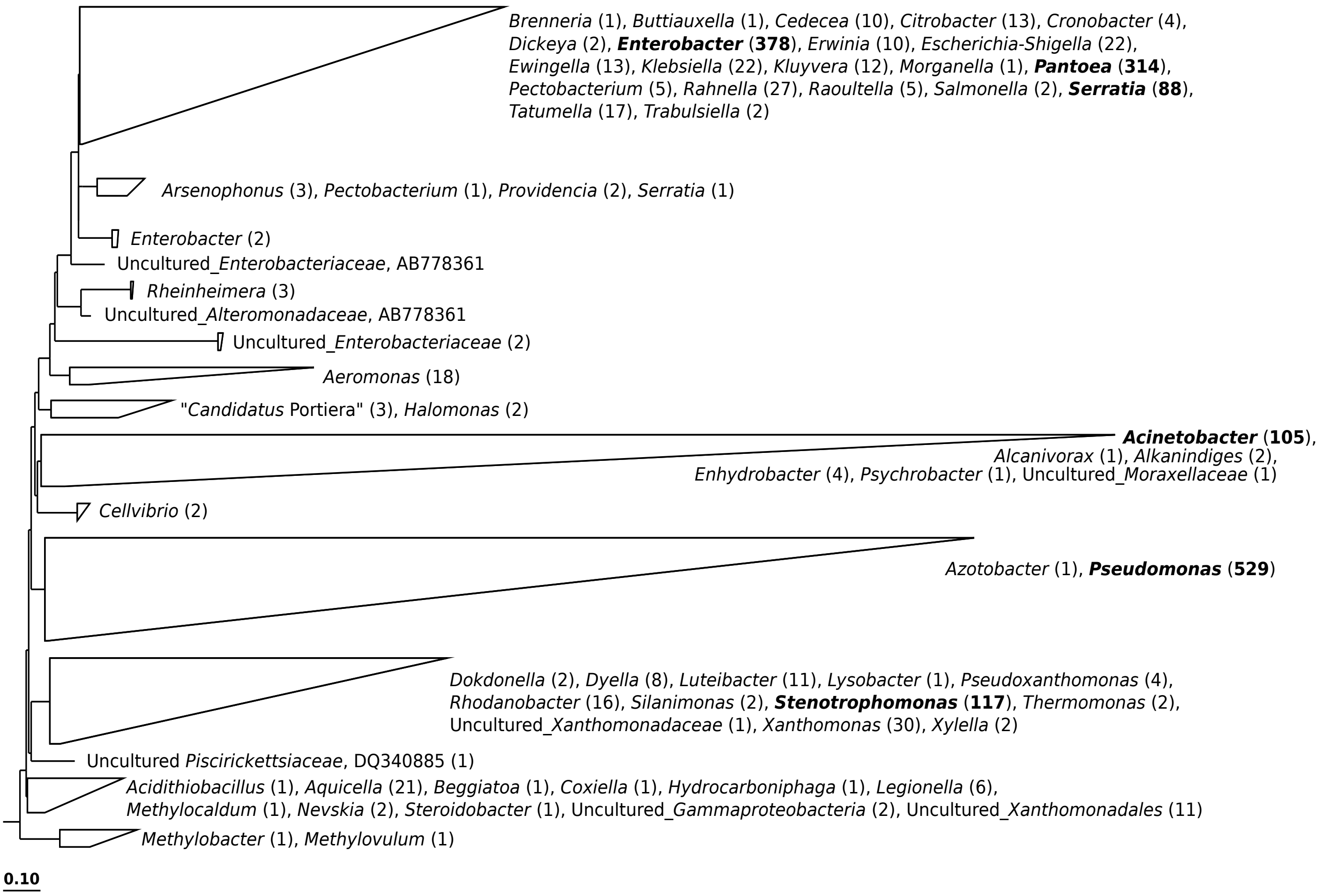

Fig. S2. 16S rRNA gene-based phylogeny analysis of Gammaproteobacteria endophytes.
Names outside wedges of clustered sequences are genus/taxon from the curated endophyte dataset
The total number of sequences observed in each taxa is displayed within brackets.
Highly abundant taxa are shown in bold.
The outgroup (not shown) consisted of a range of sequences representing several other bacterial phyla.
Scale bar indicates 10% sequence divergence.

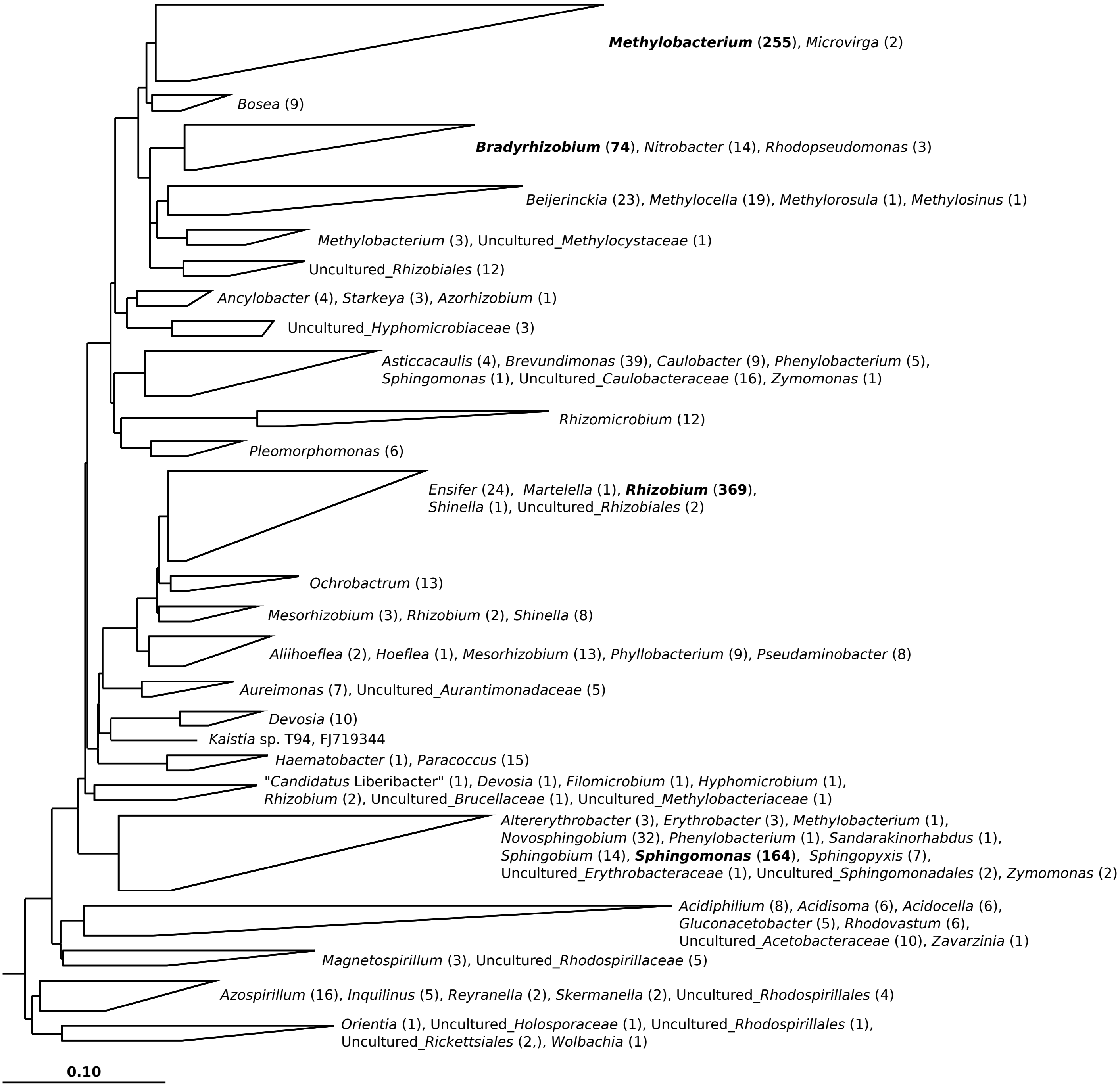

Fig. S3. 16S rRNA gene-based phylogeny analysis of Alphaproteobacteria endophytes. Details are the same as those provided for Fig. S2.

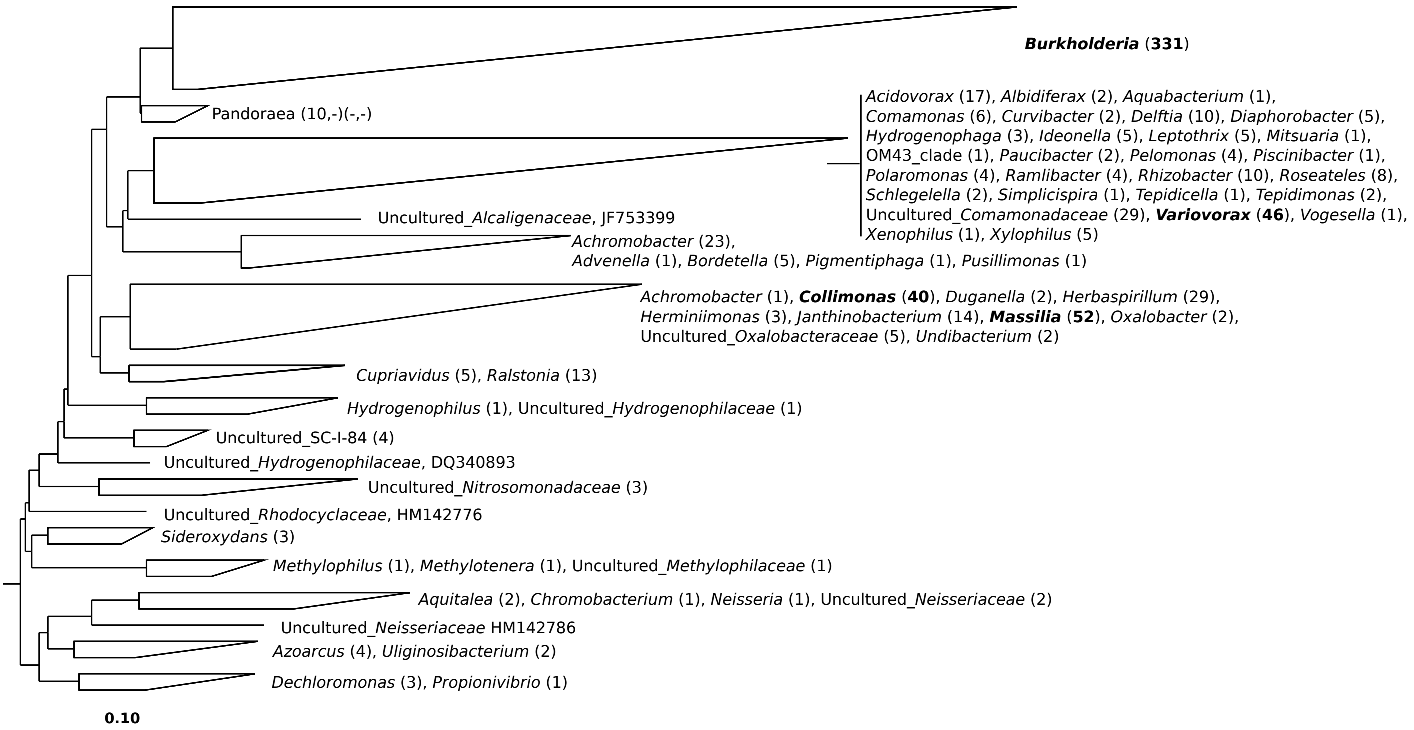

Fig. S4. 16S rRNA gene-based phylogeny analysis of Betaproteobacteria endophytes. Details are the same as those provided for Fig. S2.

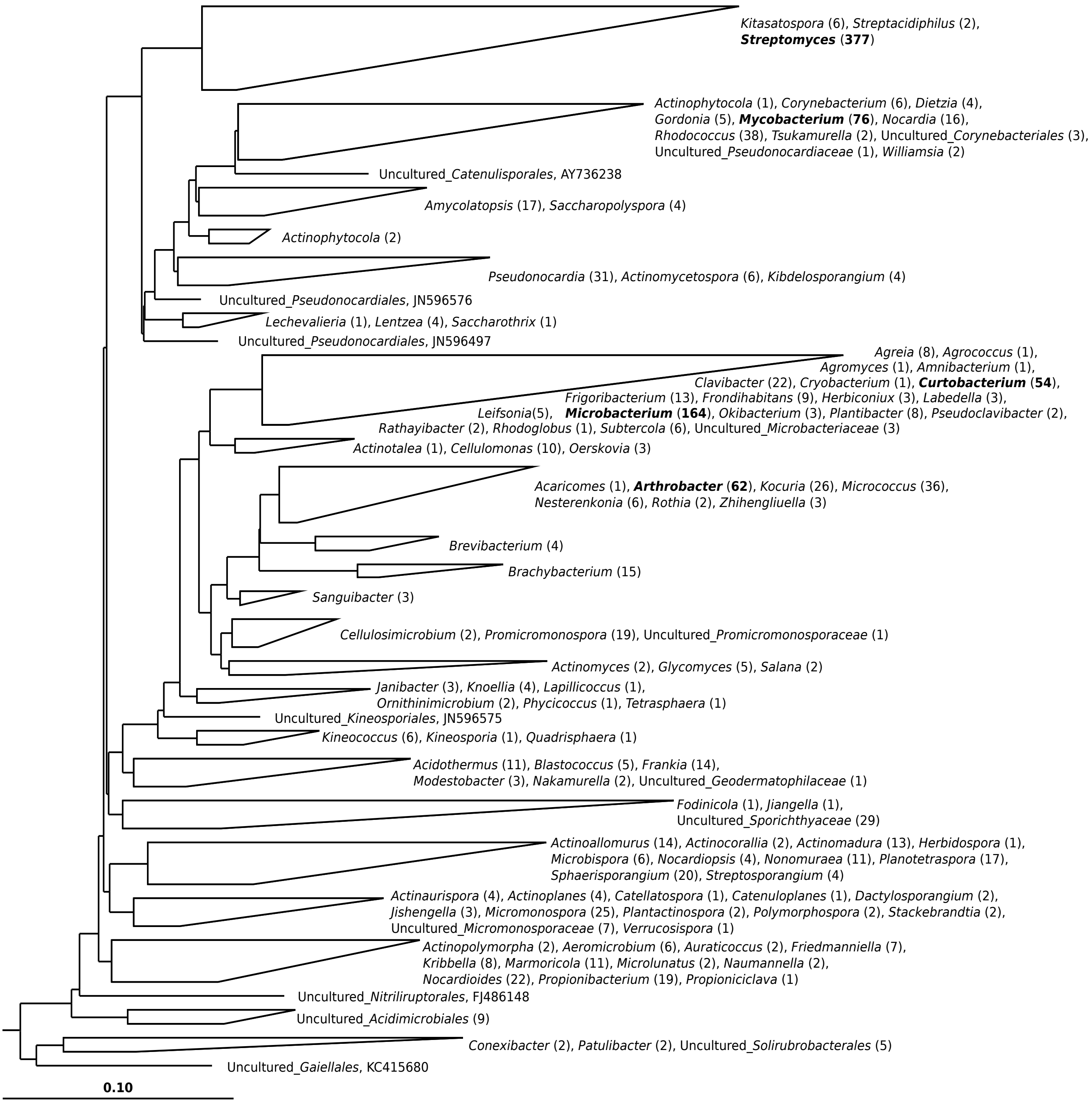

Fig. S5. 16S rRNA gene-based phylogeny analysis of Actinobacteria endophytes. Details are the same as those provided for Fig. S2.

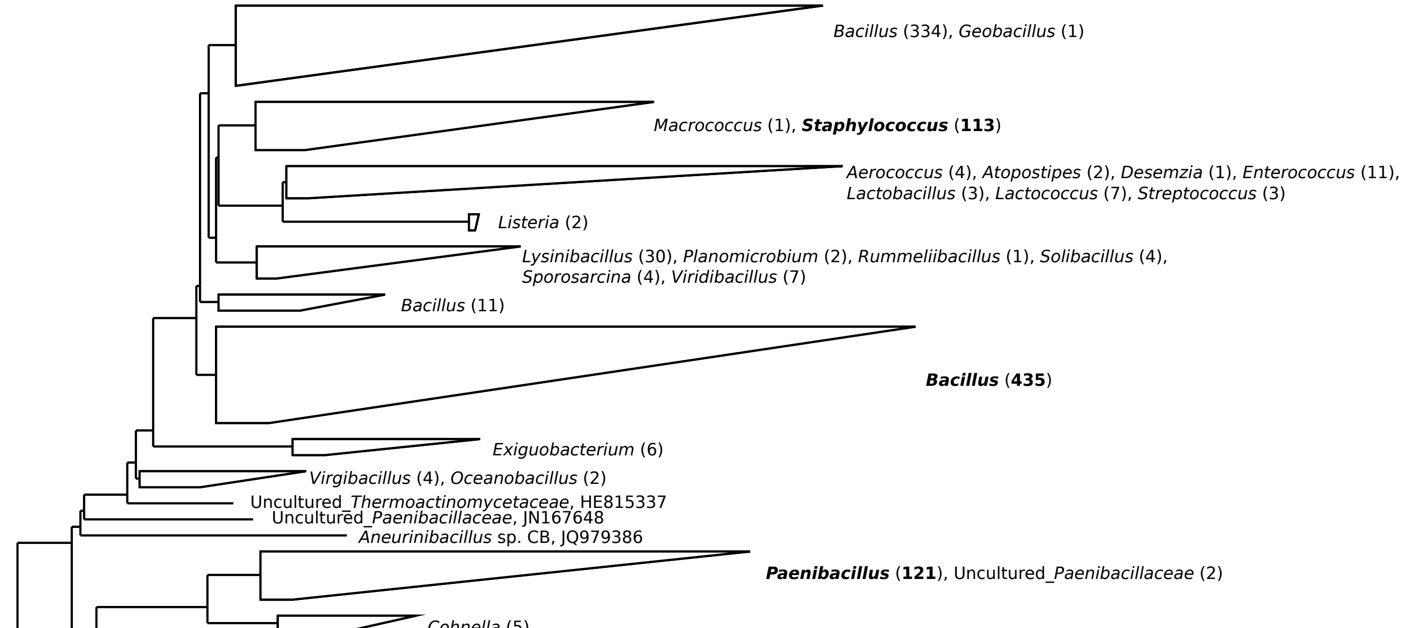

Fig. S6. 16S rRNA gene-based phylogeny analysis of Bacilli endophytes. Details are the same as those provided for Fig. S2.