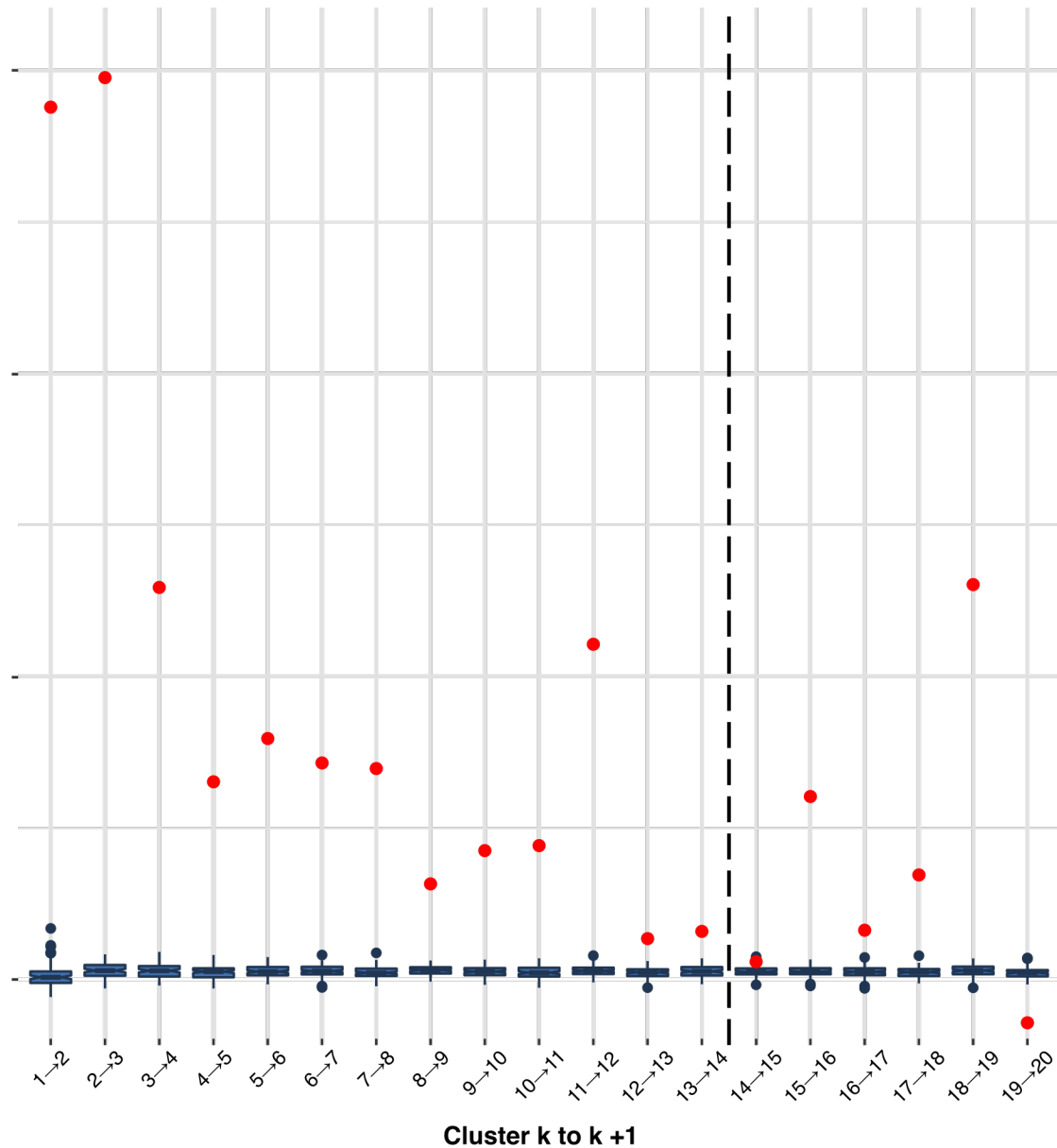
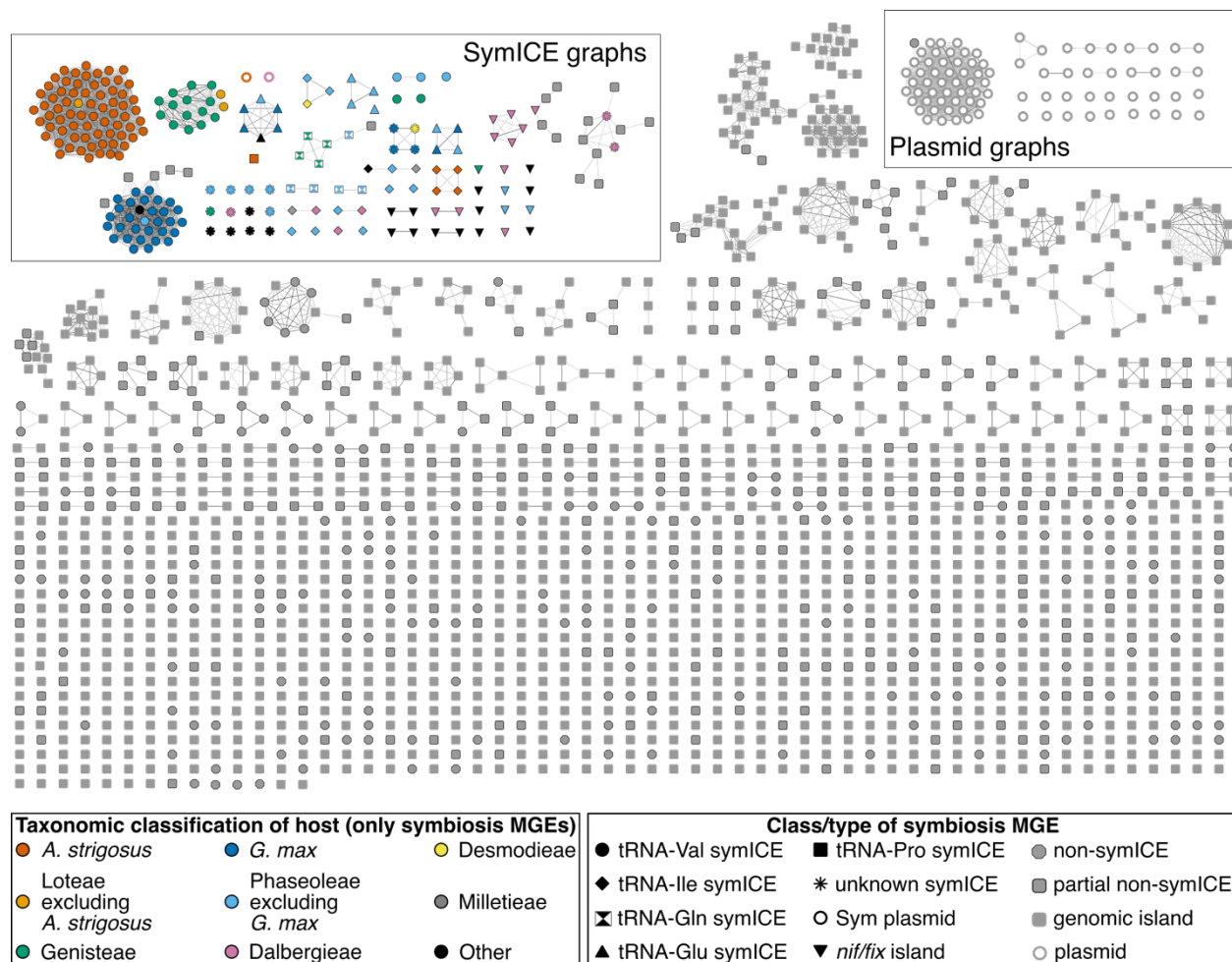


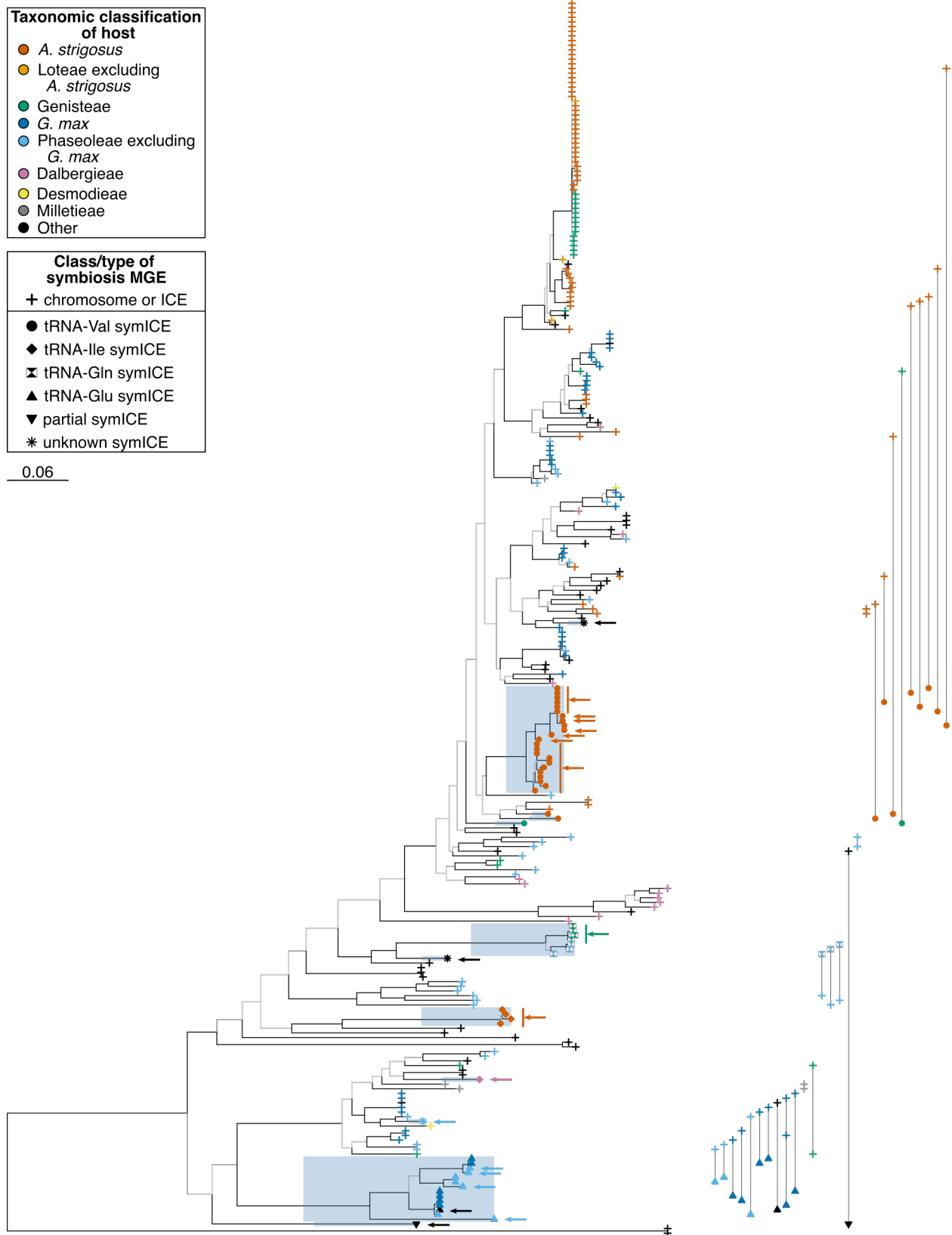
**Extended Fig. S1. The two Sym plasmids originated from different *repABC* plasmids.** ML phylogenies of translated *repC* gene sequences. The two symbiosis plasmids are indicated. The tree is midpoint rooted. Branches colored in black exceed 95% UFBoot and 80% SH-aLRT support. The scale bar indicates average number of substitutions per site. Tip points of the tree are colored according to the taxonomic classification of plant hosts and shaped according to the genomic location of the gene.



**Extended Fig. S2. Clustering of 170 trees of symbiosis MGE genes into 14 sets.** A total of 170 genes in  $\geq 40\%$  of symbiosis MGEs were identified. Their phylogenetic trees were clustered into sets based on similarities in topology. The difference in likelihood sums (y-axis) was calculated for each increase in cluster number ( $k = 1 - 20$ ; x-axis). The vertical dashed line indicates the cut-off for the minimum supported value, the  $k$  in which the increase in likelihood score for the samples (red points) is not significantly greater than that of the permuted dataset phylogenies (whisker-box plots; 100 permutations per  $k$ ).



**Extended Fig. S3. *Bradyrhizobium* has diverse regions with signatures of horizontal gene transfer.** Weighted undirected network of genomic islands and mobile genetic elements. These include all elements identified from hybrid or finished genome sequences, plasmids from metapopulation strains with draft genome sequences, and symbiosis MGEs from draft genome sequences. ICEs are defined based on having integrase gene(s) and/or a type IV secretion system (T4SS)-encoding locus. Plasmids have *repABC* genes and conjugative plasmids also carry a T4SS locus. Symbiosis MGEs additionally have *nod*, *nif*, and *fix* genes. Nodes are shaped according to the class, type, or subtype of genetic elements. Symbiosis MGEs are colored according to the taxonomic classification of plant hosts, other non-symbiosis MGEs are colored light gray. Darker edges indicate greater Jaccard similarity of *k*-mer signatures.



**Extended Fig. S4.** The *fix* genes encoding a *cbb*<sub>3</sub>-type cytochrome oxidase are

**exchanged between chromosomes and symICEs.** ML phylogeny of the *fixO* gene, which encodes a membrane-associated monoheme cytochrome *c*. Boxes shaded blue indicate clades of *fixNOQP* and *fixGHIS* gene clusters present in symICEs. Colored arrows indicate strains with the cluster only in a symICE. Vertical lines link strains that have two sets of *fixNOQP* and *fixGHIS*. The strain with three clusters is USDA76. The tree is midpoint rooted. Black colored branches have 70% or more bootstrap support. Gray colored branches have less than 70% bootstrap support. The scale bar indicates average number of substitutions per site. Tip points are colored according to the taxonomic classification of plant hosts and shaped according to the genomic location of the *fixO* gene.