Supporting Information

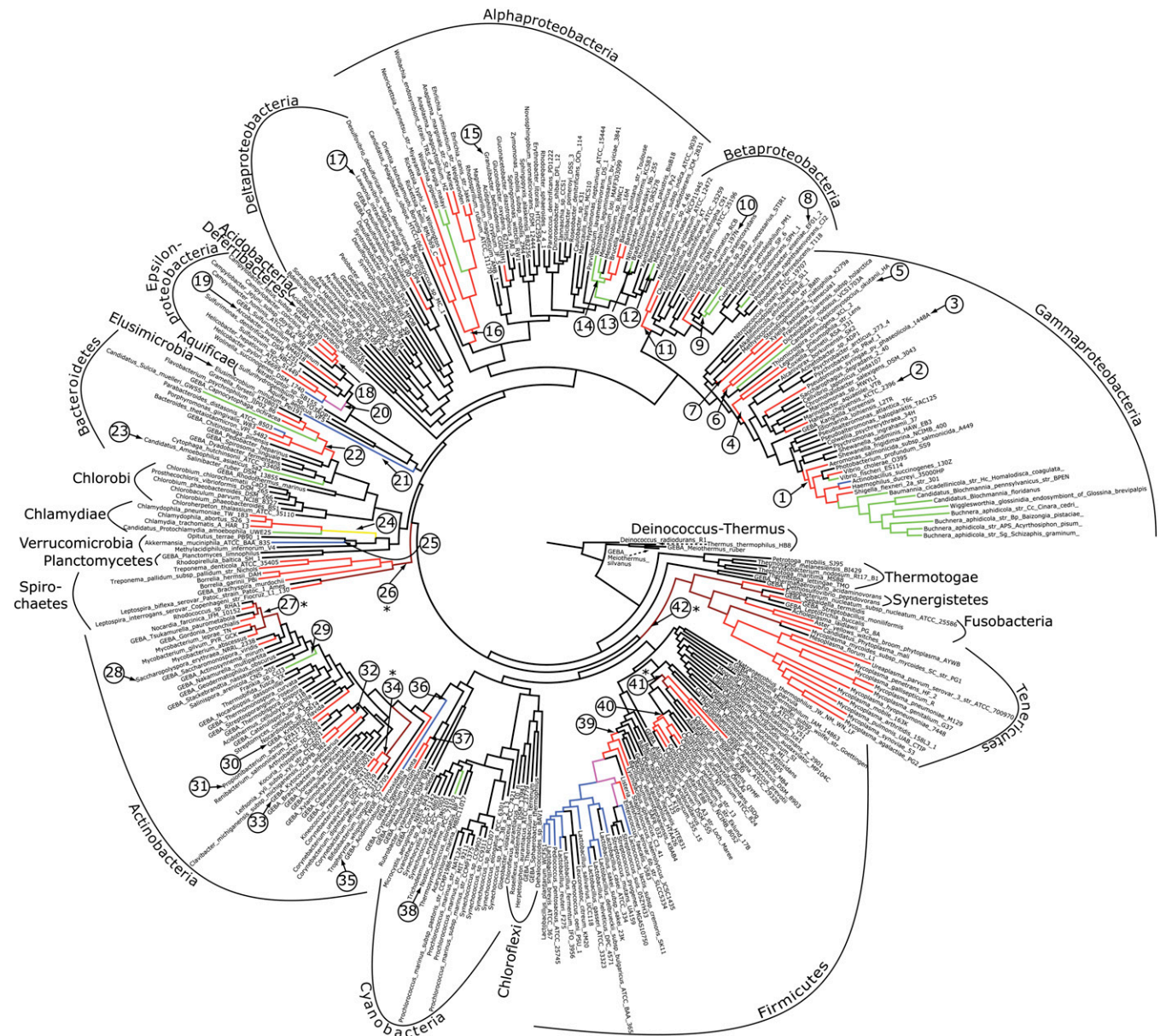## Sachs et al. 10.1073/pnas.1100304108

Fig. S1. Inferred evolutionary history of bacterial host association with taxon labels. Ancestral states are inferred on a domain-level bacterial phylogeny, a modified version of a tree produced in a previous study (1). The tree is a maximum-likelihood reconstruction of a concatenated set of 31 single copy genes from 350 bacterial species chosen to optimize phylogenetic sampling. Phyla and proteobacterial classes are labeled. Branch colors represent host-associated traits on the tips of the tree and inferred states on ancestral nodes (black, environmental; blue, commensal; green, mutualist; red, parasite). Host association traits were obtained from an earlier review (2). We used parsimony in Mesquite (version 2.74) (3) to infer ancestral character states and transitions in host-associated phenotypes. We inferred a minimum of 42 origins of host association (labeled 1–42). Origins at five nodes had equivocal parsimony reconstructions, noted with asterisks. Equivocal ancestral states are represented by blended branch colors (brown, environmental or parasite; purple, parasite or commensal; yellow, parasite or mutualist). Additional origins are equally parsimonious at these nodes and provide an upper bound for global origins at 52. [Adapted from ref. 2.]

1. Wu DY, et al. (2009) A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462:1056–1060.
2. Toft C, Andersson SGE (2010) Evolutionary microbial genomics: Insights into bacterial host adaptation. *Nat Rev Genet* 11:465–475.
3. Maddison WP, Maddison DR (2010) Mesquite: A modular system for evolutionary analysis. Version 2.74. Available at http://mesquiteproject.org.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)