**Supplementary figures**

****

**Figure S1:** A) The “Classical” sequence based phylogenetic approach. This approach determines phylogenetic relationships based on changes in amino acid (see Figure 2A, S2, S3) and/or nucleotide sequence (see Figure 2B S5, S6). Fewer changes generally indicate closer relationship between species. B) The presence/absence approach determines phylogenetic relationship based on the presence or absence of genes in the genome (see Figure 2, S7). This approach assumes that closer related species share more common genes (blue=present, red=absent). C) The synteny-based analysis determines phylogenetic relationship based on the conservation of gene order in the genome (see Figure 4). The farther species are from one another the lower the conservation of gene order due to gene duplications, losses and genome rearrangements.



**Figure S2:** Phylogenies inferred from concatenated amino acid sequences under LG+I+G+F using (A) RAxML and (B) Mrbayes. Nodes show bootstrap values in (A) and posterior probabilities in (B).



**Figure S3:** BI phylogenies inferred from concatenated amino acid sequences using (A) CAT-LG, (B) CAT-GTR and (C) CAT-Poisson models in Phylobayes. Nodes show posterior probabilities.



**Figure S4:** 50% majority rule consensus tree based on the 1,021 best ML single-gene trees calculated by RAxML (–J MR). Nodes show bootstrap values.



**Figure S5:** Phylogenies inferred from concatenated nucleotide sequences under GTR+ I+G using (A) RAxML and (B) Mrbayes. Nodes show bootstrap values in (A) and posterior probabilities in (B).

****

**Figure S6:** Bayesian phylogenies inferred from concatenated nucleotide under (A) CAT-Poisson, and (B) CAT-GTR models in Phylobayes. Nodes show posterior probabilities.



**Figure S7:** Bayesian phylogenies inferred from gene presence/absence data using the binary F81-like model in Mr. Bayes. Genes present in fewer than two species were remvoved and nodes show posterior probabilities.

****

**Figure S8:** Phylogenetic analysis based on the number of paralogs per species and ortholog/paralog group. The nodes supports are AU (Approximately Unbiased) p-values and BP (Bootstrap Probability) values. Distances were calculated using the “canberra” method and clustered using the “average” Hcluster method and 100 bootstraps.