# Division of Biological Sciences

Phylogenetic Inference on the Universal Tree of Life

The central focus of our research is to gain insight into origin and early evolution of eukaryotes, the most important open problem in evolutionary biology.One of the goal of this research is to reconstruct the Universal Tree of Life including diverse organisms on the Earth, based on molecular phylogenetic approach. Since a reliable phylogenetic inference simultaneously requires numerous genes and species, a large scale data analysis using high performance computers plays a key role in our study.

Maximum-likelihood Tree of the "24-gene" data set

Phylogenetic analyses was performed by using 24 protein-coding genes with >= 10,000 amino acid positions. The tree was selected from 945 possible pre-defined trees for 7 taxonomic groups. The numbers on nodes indicate bootstrap percent support values.

Although the sisterhood between opisthokonts and amoebozoa is well accepted, the 24-gene analyses connected opisthokonts and other eukaryotic group, diplomonads+parabasalids with a strong support value.
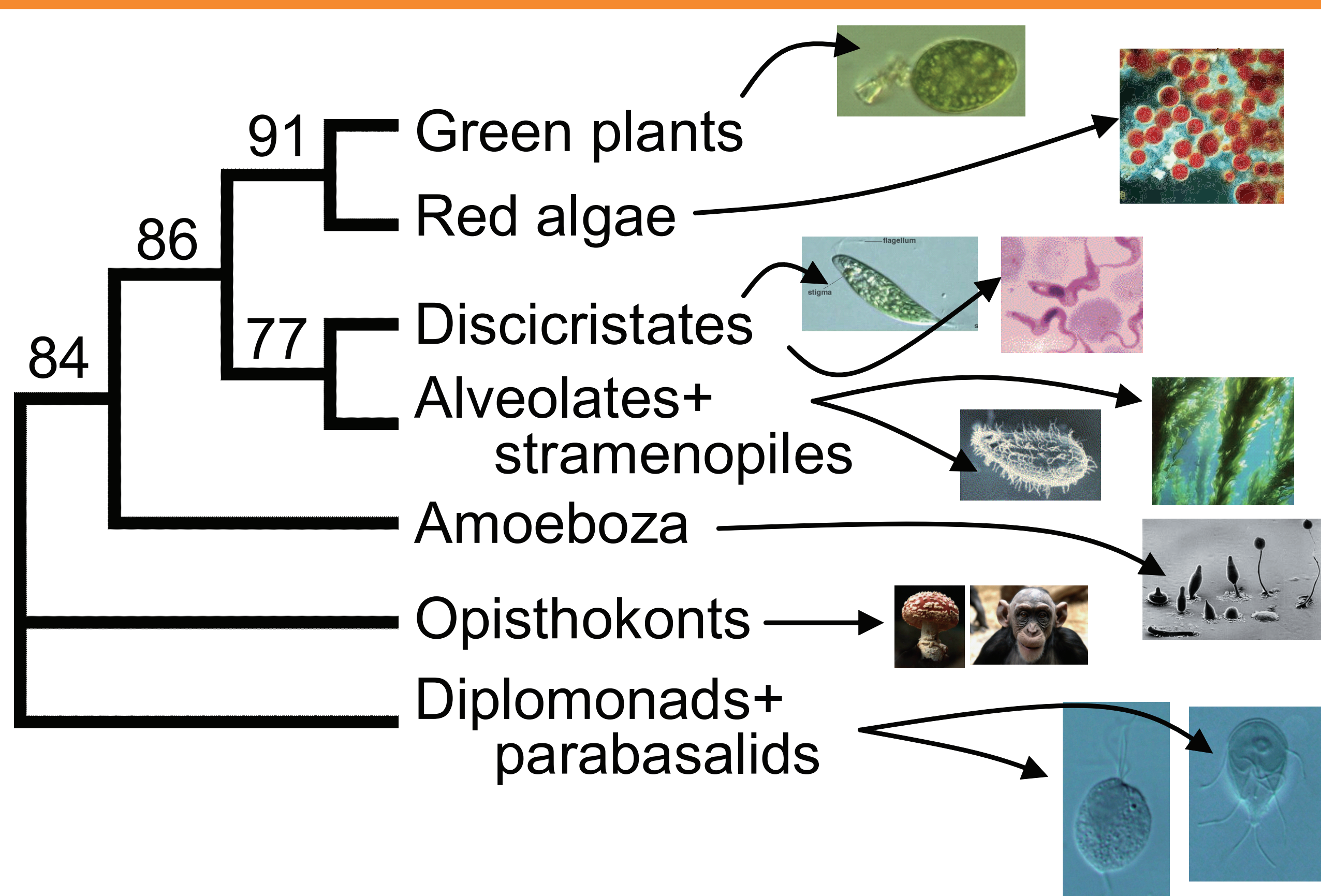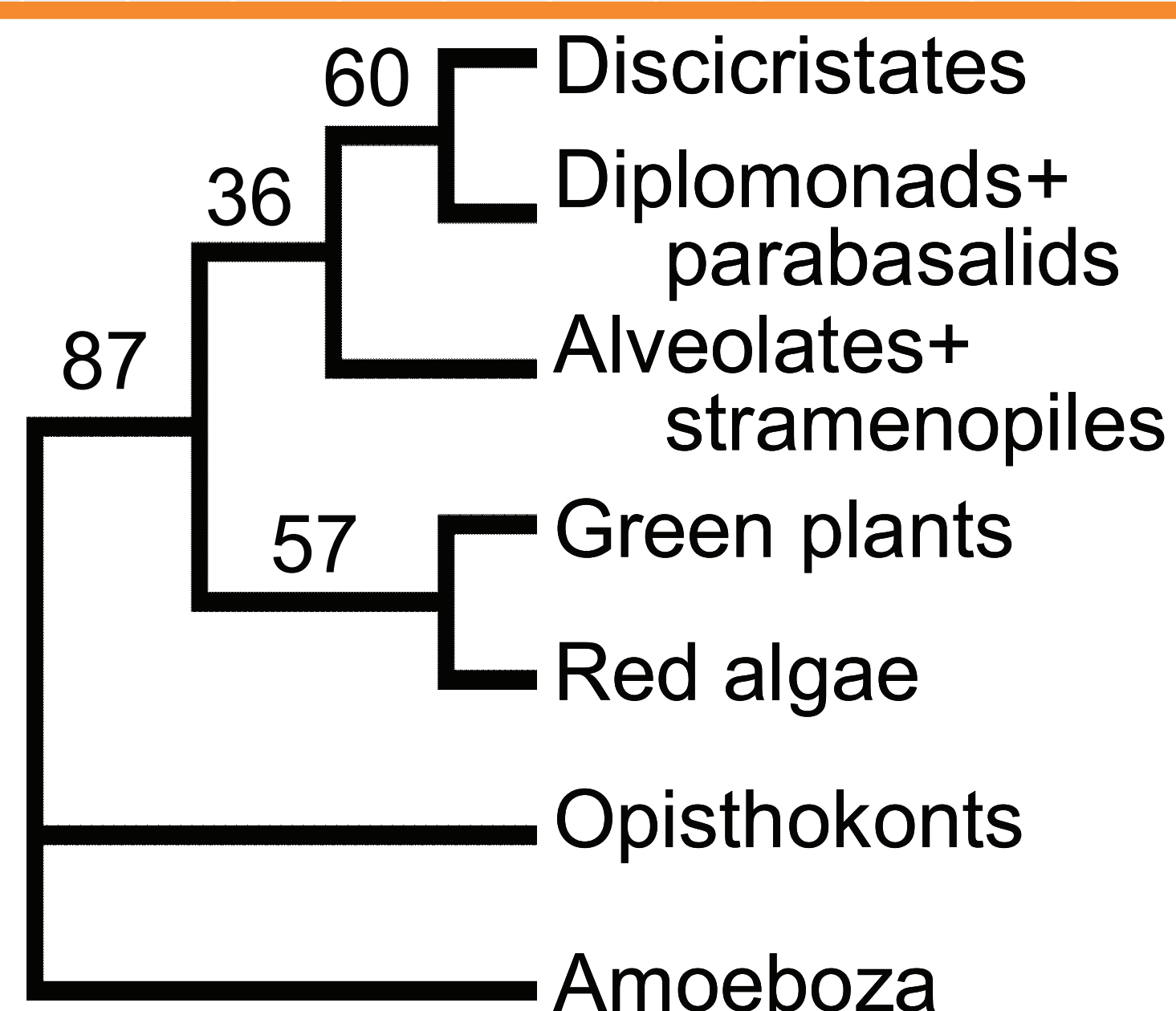

Phylogenetic signals in the individual genes considered
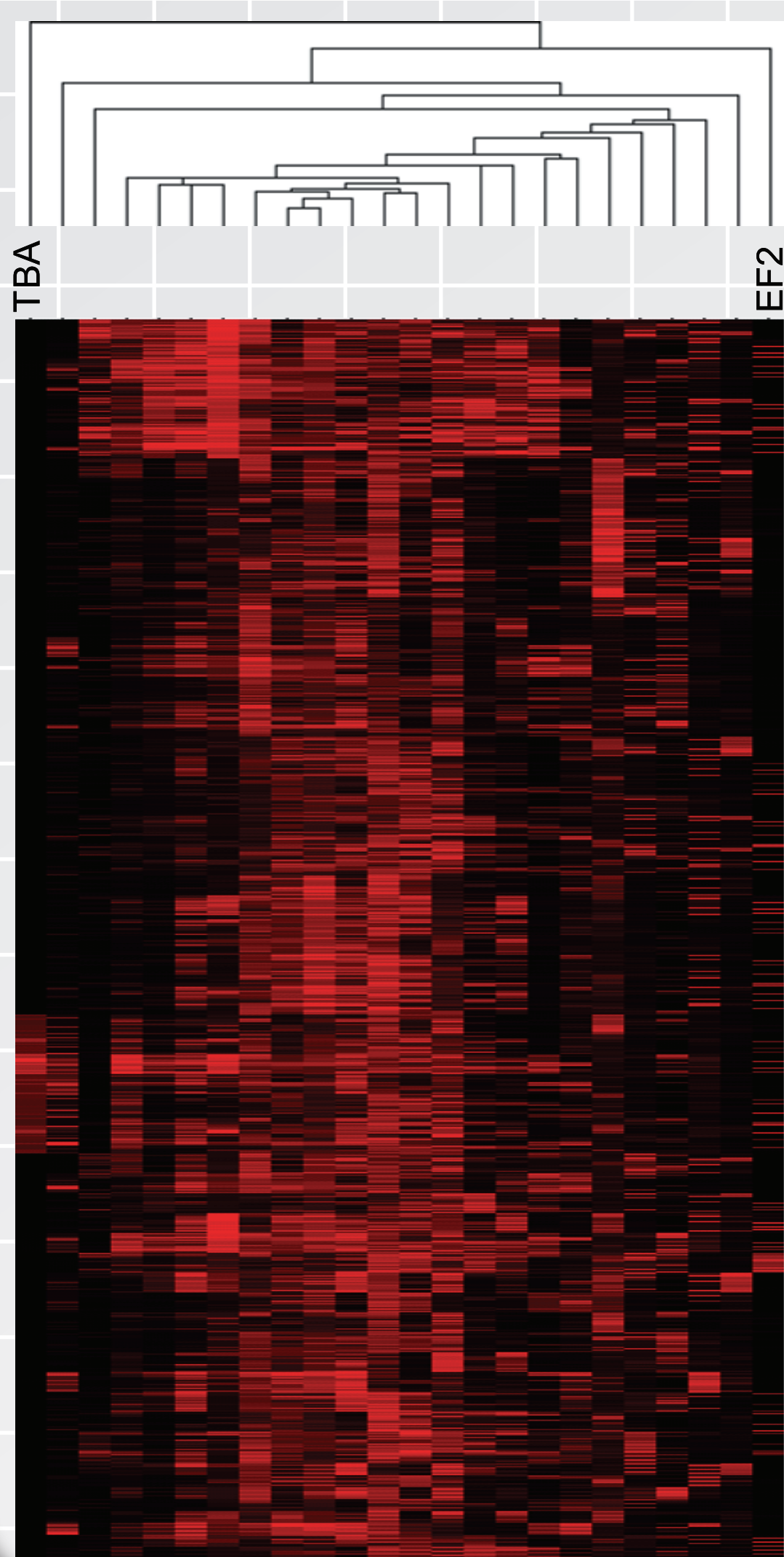
Phylogenetic signal of 24 protein-coding genes are color-coded from black/weak to red/strong.

Hierarchical clustering indicates two genes, alpha-tubulin (TBA) and translation elongation factor 2 (EF2) potentially possess the signals that are distinctive from other genes considered.

Phylogenetic analyses was performed excluding TBA and EF2 (22 protein-coding genes in total). The details of the analyses are same as described.

By exclusion of TBA, the sisterhood between opisthokonts and amoebozoa was successfully recovered. Two eukaryotic groups, Diplomonads+parabasalids and discicristates, which supposed to belong a hypothetical supergroup Excavata, now clustered. The phylogentic "noise" in TBA most probably masked the genuine historical "signal" in the 24-gene data set.