For a better understanding about the biological interaction between the microsporidia and the host species as well as their position in the tree of life, it is necessary to investigate the ancestor of these eukaryotic parasites (Keeling and Fast, 2002). Thus, we carried out this study to examine the fungal related origin of microsporidia and to explore the metabolic network of their last common ancestor. In Chapter 2, " The estimation of the microsporidian last common ancestor protein set ", we used a parsimony approach to estimate the protein set of microsporidia last common ancestor (LCA), which was then the basic data for the downstream analyses. In order to analyze the phylogenetic distribution of the microsporidian LCA proteins in an effective and informative manner, we developed an phylogenetic visualization and analysis tool named PhyloProfile, which is introduced in Chapter 3, "PhyloProfile: an interactive visualization tool for exploring complex phylogenetic profiles". In Chapter 4, "Distribution analysis of microsporidian LCA proteins", we applied PhyloProfile to the protein set of the microsporidian LCA to measure the evolutionary ages of those sequences. The last chapter, " Metabolic pathway analysis of the microsporidian LCA proteins", describes a novel approach to assign the functional annotations to the microsporidia LCA proteins and based on those annotations, the microsporidian LCA metabolic network was compared with those of the extant species.

The estimation of the microsporidian last common ancestor protein set

# Introduction

The analysis of species phylogeny can give insight into the evolutionary history of those species, such as what is the systematic relationship between them and the others in the phylogenetic tree of life, or how their pathways evolved across species and time (Futuyma, 2005). Since the evolutionary process of microsporidia is still poorly understood, a comparative analysis between the contemporary microsporidia and their ancestor is required (Keeling and Fast, 2002). For this reason, in this chapter we describe an orthology-based approach for estimating the microsporidian last common ancestor (LCA) protein set, which was served as an initial data for further analyses.

Discussion

Proportion of orthologous and lineage specific proteins (or the reduction and expansion in microsporidian genome)

We analyzed the proportion of orthologous and lineage specific proteins in 11 microsporidia species.

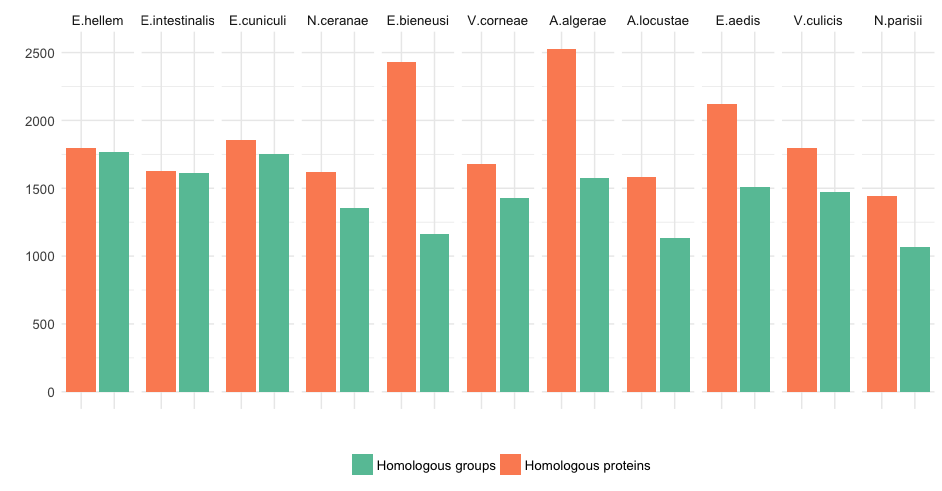


Figure 1‑1: The distribution of number of homologous proteins (orange) and number of homologous groups (green) in each microsporidia taxon.

Figure 1‑1 shows the relative number of homologous proteins and number of homologous groups in each microsporidia species. In some species, such as E.bieneusi or A.algerae, the number of homologous genes is substantially higher than the number of corresponding homologous groups. We check the number of in-paralogs for each microsporidia taxon in the homologous group. The result in Figure 1‑2 shows that there is no evidence for whole genome duplication in any species. But there are some instances where the homologous groups contain more than 10 co-orthologs for one microsporidia species showing the effect of gene dosage (Kondrashov and Koonin, 2004).

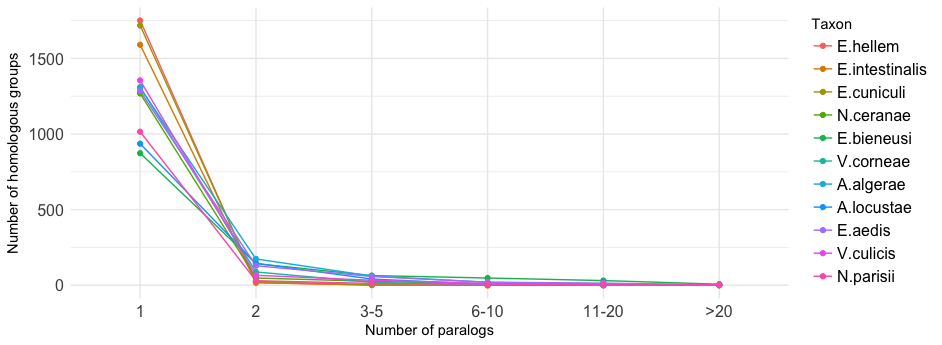


Figure 1‑2: The distribution of number of homologous groups as a function of number of in-paralogs. Colors denote different microsporidia taxa.