

Evolution of the eukaryotic cell (containing nucleus) is one of the most topical issues in evolutionary biology, and endosymbiosis is a key element of the process. The first successful endosymbiosis occurred about 2 billion years ago and involved acquisition of an α -proteobacterium by the ancestors of all eukaryotes. The acquired endosymbiont was transformed into the mitochondrion, the cellular powerhouse, which in fact initiated the origin of eukaryotes. About 500 million years later, one of the eukaryotic lineages acquired a cyanobacterium, which was transformed during the so called primary endosymbiosis into the primary plastid (e.g. chloroplast). This resulted in the first autotrophic eukaryotes capable of performing photosynthesis. Their descendants are terrestrial plants and various algae groups, such as green algae, red algae and glaucophytes. Some representatives of these lineages or their plastids were engulfed by heterotrophic eukaryotes during secondary endosymbioses, and some of those were subsequently enslaved by other eukaryotes in higher order endosymbioses. As a result, many of the original heterotrophic eukaryotes became autotrophs. In this way, plastids and the process of photosynthesis have virtually penetrated the whole eukaryotic tree of life, including some animal groups such as *Elysia* marine gastropods.

The transformation of an endosymbiont, e.g. the cyanobacterium, into the cellular organelle was gradual. In order to better manage the genetic information contained in the newly created organelle, many endosymbiont genes were transferred to the host nuclear genome, which resulted in the necessity for a protein import system that would allow proteins to be transported back to the organelle. Therefore, proteins targeted to the organelle acquired appropriate targeting signals, and the organelle evolved appropriate transporters for the proteins in its membranes.

The aim of this project is to create the first database that will contain both genetic, morphological and biochemical information about plastids, with a user-friendly interface, and a set of tools for interactive data browsing, searching, sorting, comparing and downloading. As part of this project, we will create a webserver that will recognize plastid targeting sequences in order to facilitate plastid proteins identification. This is especially important in seeking new targets for the pharmaceutical industry as many parasites, such as the malaria agent, have reduced plastids that may become the target for new therapies. Information contained in the database and the predictor will be used to create an evolutionary model of plastid presequences and their translocation machineries. Based on our database and phylogenetic analyses, we will also recreate how plastids have been evolving for over 1.5 billion years. This project is also a part of one of the greatest challenges to modern evolutionists since the publication of Darwin's theory of evolution, i.e. the construction of the Tree of Life. The Tree of Life is supposed to depict the phylogenetic relationships between all known organisms, and one of the aims of this project is to reveal the relationships between poorly investigated but dominating groups of eukaryotes classified by many authors to the historical kingdom of Protista.