

Analysis of selected processes influencing the architecture of fungal genomes and proteomes

Nowadays whole genomes have been sequenced for many forms of life, what means that the whole genetic information of a given organism has been read. Genomes are not only sets of genes, but also serve as the environment into which genes are immersed. The landscape of the genome is shaped by various factors including the way how an organism lives, reproduces and gets food. There are big genomes with lots of additional elements gearing how genes are used simultaneously in diverse cellular processes and simplified miniature genomes with only the most indispensable tools (genes) left. The adaptation of the organism to a particular lifestyle can shape the composition of genes, their spatial arrangement, the share of non-coding sequences and the eagerness to use foreign sources of genomic information. Genomes are dynamic structures with constant cycles of death and birth of genes and virus-like sequences called transposons. The latter are able to move in the genome and change the context of neighbouring genes. The genomes are also shaped by horizontally acquired genes which can be a source of new traits acquired faster than through random walk of evolution driven by infidelity in copying and transferring genetic material from parents to descendants. However, genomes are strictly protected from gene invasion, which uncontrolled, would blur host's identity.

The main idea of our project is to find how the way organism lives, together with its genome features (how big it is, how many genes are there, how densely packed they are) affects the abundance of horizontally transferred genes and transposons. We will collect all sequenced fungal genomes and extract their characteristic features, predict transposons, gene clusters and horizontally transferred genes. We will also describe the organisms with tags describing their preferred lifestyle (for example: pathogen, soil inhabiting). This will result in a big dataset spanning many features for each organism which will be analysed simultaneously with statistical tools. All this data together may shed new light on how we understand the relationship between organisms lifestyle and genome organisation. On the practical side we may find new gene clusters with useful features for biotechnological application.