Phylogenomic insights into the evolution of function in fungi and plants - datasets, outcomes and biases

Grzegorz Koczyk

Thanks to innovations in both semiconductor sequencing and sequencing-by-synthesis over the last 15 years, the decline of costs for whole genome sequencing of eukaryotes outpaced Moore's law. This has led to the extant situation where the ability to interpret quickly changing data, rather than the availability of data itself, becomes a severely limiting factor. In many applied scenarios (from molecular diagnostics to selection of causal candidate genes in a novel organism), this problem is compounded by prohibitive costs of large scale experimental characterisation of homologous genes.

Phylogenomic methods facilitate understanding molecular and biological functions of genes in context of the evolution of hosting species. These analyses (1) take into account large, non-redundant body of genomic data and relatively sparse functional annotations (2) employ phylogenetic methods in a (semi)automated, repeatable manner. An explicit reconciliation of gene family histories in context of species history constitutes a frequent part of the process.

In this lecture, I broadly discuss two example cases where phylogenomic analyses led us to novel insights: biosynthesis of secondary metabolite compounds in fungi (polyketides and terpenes) and symbiotic nitrogen fixation in plants. Last but not least, general pitfalls and rules of thumb in phylogenomic data analysis are outlined in brief.