database providing web based tools for the molecular identification of ectomycorrhizal fungi

Urmas Kõljalg[1*], Karl-Henrik Larsson[2], Kessy Abarenkov[1], Henrik R. Nilsson[2], Ian J. Alexander[3], Ursula Eberhardt[4], Susanne Erland[5], Rasmus Kjøller[6], Ellen Larsson[2], Taina Pennanen[7], Robin Sen[8], Andrew F.S. Taylor[4], Trude Vrålstad[9], Leho Tedersoo[1], Björn M. Ursing[10]

[1] Institute of Botany, University of Tartu, Tartu, Estonia; [2] Systematic Botany, Göteborg University, Sweden; [3] Dept. of Plant and Soil Science, University of Aberdeen, UK; [4] Dept. of Forest Mycology and Ecology, SLU, Sweden; [5] Microbial Ecology, Lund University, Sweden; [6] Botanical Institute, University of Copenhagen, Denmark; [7] Metla, Finnish Forest Research Institute, Finland; [8] Dept. of Biosciences, University of Helsinki, Finland; [9] Dept. of Biology, University of Oslo, Norway; [10] Center for Genomics and Bioinformatics, Karolinska Institutet, Sweden

* Corresponding author: urmas.koljalg@ut.ee

Identification of ectomycorrhizal fungi is often achieved through comparisons of ribosomal DNA internal transcribed spacer (ITS) sequences with accessioned sequences deposited in public databases. A major problem encountered is that annotation of the sequences in these databases is not always complete or trustworthy. In order to overcome this deficiency, we report on UNITE, an open access database comprising well annotated fungal ITS sequences from well-defined herbarium specimens that include full herbarium reference identification data, collector/source, ecological data, etc. UNITE can be searched either by species name or via a sequence similarity search by using program *blastn*. Database incorporates also phylogenetic species recognition. For this purpose the software package *galaxie* – web based tool for the phylogenetic analyses, is implemented into UNITE. *Galaxie* is either using prealigned sequence files to place query sequences inside known taxa or starts out using BLAST to assemble a new dataset that is aligned and then analysed. The test version of the UNITE database is available through the Internet address http://unite.ut.ee/.

Fig.1. Composite of "screenshots" displaying some features of the UNITE. In the home page of UNITE you can BLAST (above left) your ITS sequences against the database or you can run phylogenetic analyses by using galaxie (above right) . In BLAST search and galaxie results pages the best matches and terminal taxa respectively have interactive links to the specimen data (above center). Species name in the last page have interactive link to the species description if available. Species description includes text as well as illustrative material and link to the nomenclatural database *Cortbase*.