

State of The World's Fungi 2018

Fungal Genomes

An initial analysis of all sequenced fungal species across four databases

- <https://www.ncbi.nlm.nih.gov/> NCBI
- <https://genome.jgi.doe.gov/programs/fungi/index.jsf> MycoCosm
- <http://fungidb.org/fungidb/> Fungi DB
- <https://www.ensembl.org/index.html> Ensembl

This revealed that all species within these databases could be found in either NCBI or MycoCosm. We thus focused further analysis only on these databases.

On 31st March 2018, the full list of eukaryotic whole genome sequences on NCBI, including sequence metadata such as release date, assembly size and gene number, was accessed at ftp://ftp.ncbi.nlm.nih.gov/genomes/GENOME_REPORTS/eukaryotes.txt. A full list of fungal whole genome sequences on MycoCosm, also including metadata, was supplied by Igor V. Grigoriev (U.S. Department of Energy, Joint Genome Institute). The two lists were merged and names automatically reconciled against a local download of Index Fungorum by string matching. Where names could not be automatically reconciled, an expert (Paul Kirk, Royal Botanic Gardens, Kew) was consulted for reconciliation.

Following reconciliation, fungal names were summarised to a list of unique species. Names which at this point could not be reconciled were each considered to be novel, undescribed species.

Raw data and the R code used to analyse the data are available from sotwf@kew.org.

