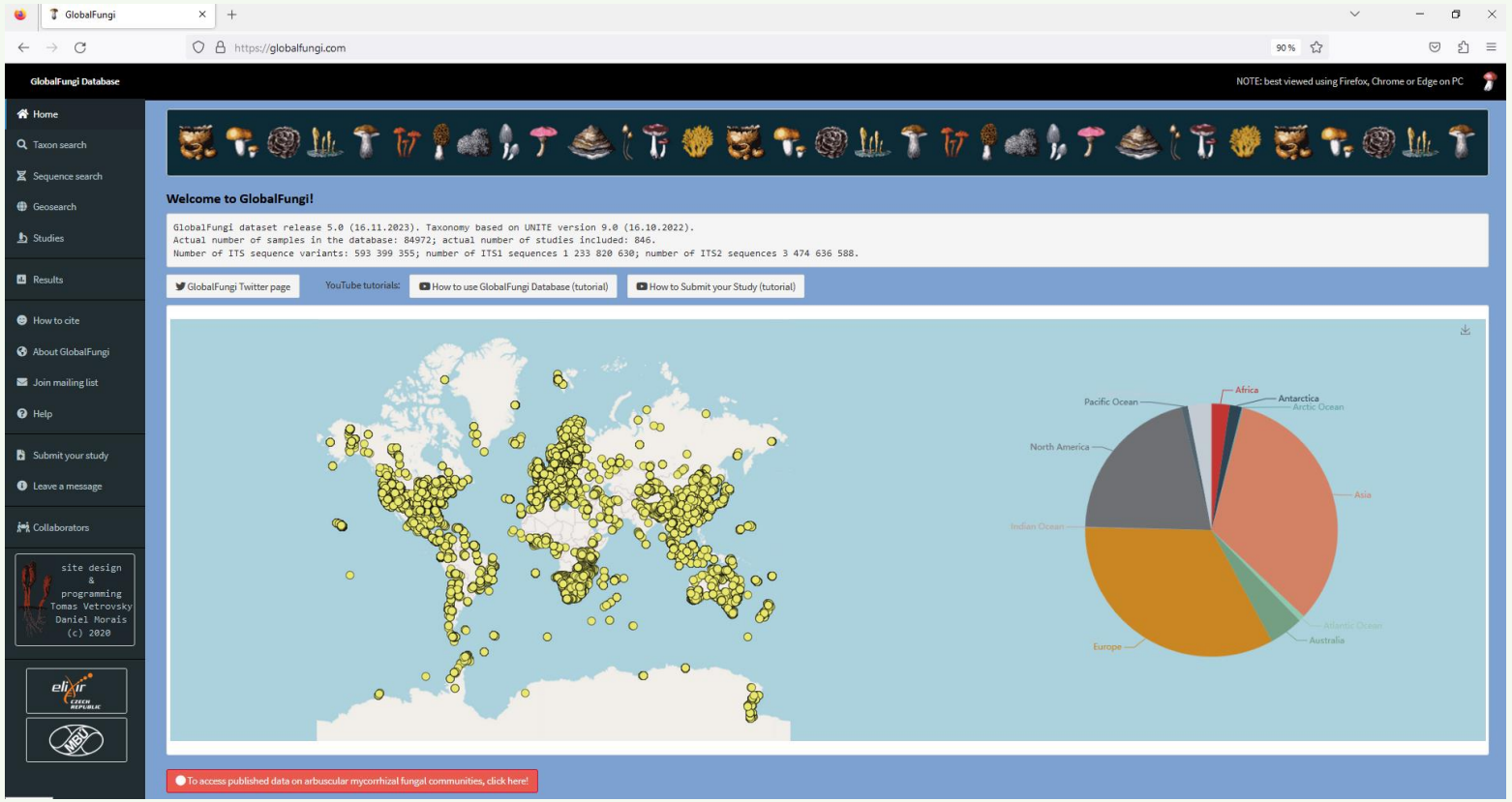
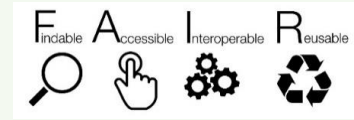
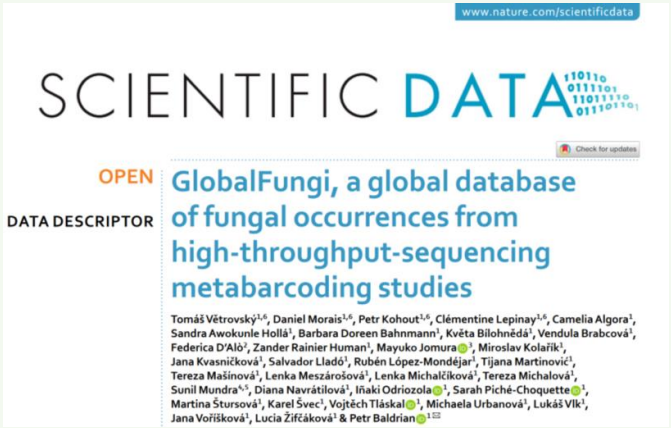
The GlobalFungi database



https://globalfungi.com/





The Czech Academy of Sciences

Initiative of the GlobalFungi database

► Initiative from the Laboratory of Environmental Microbiology from the Institute of Microbiology at the Czech Academy of Sciences

Petr Baldrian

> project leader

Tomáš Vetrovsky

> perform bioinformatic, design online database

Petr Kohout

> valorization of the database

> coordinate data acquisition

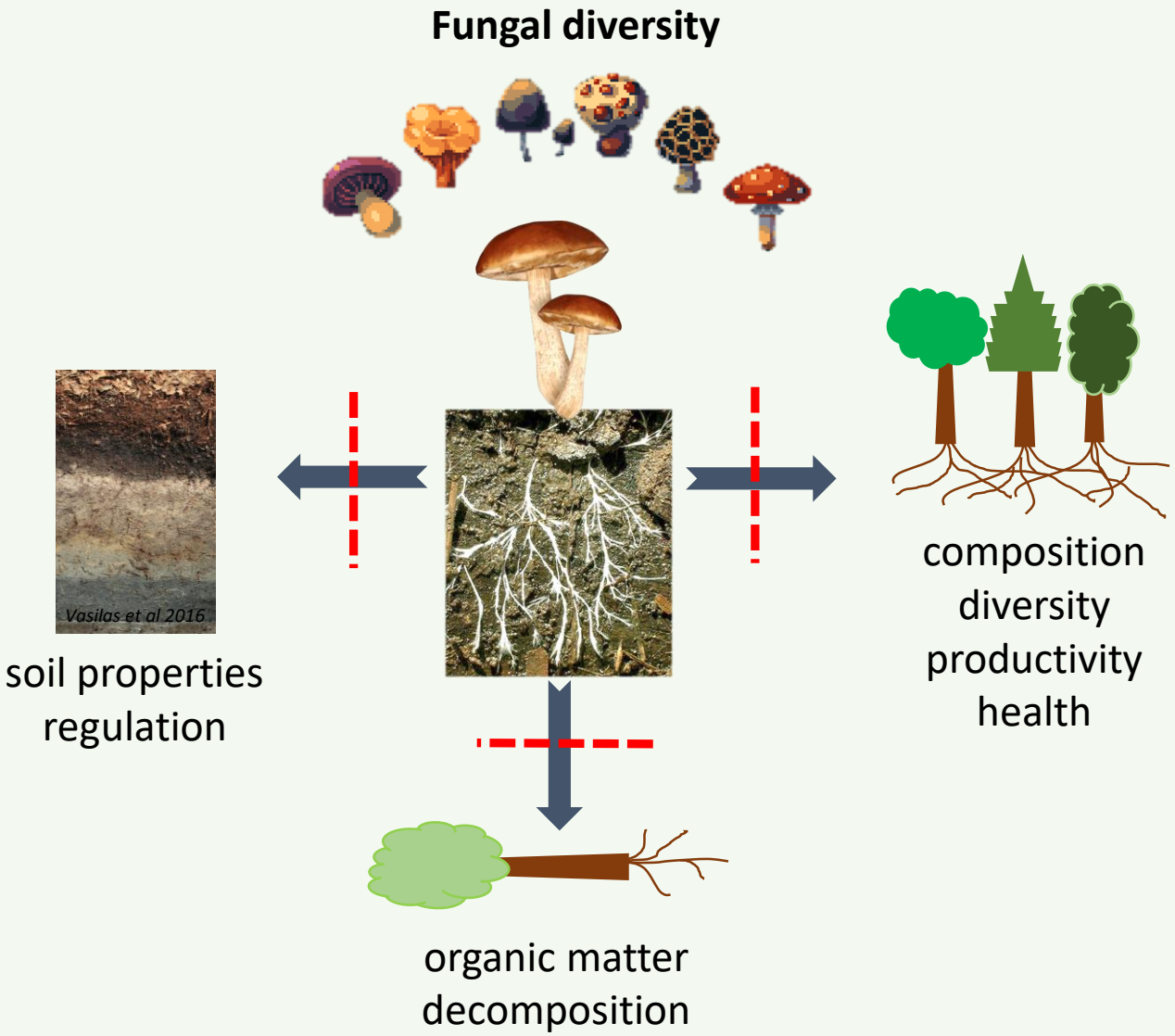
Clémentine Lepinay

A **COLLABORATIVE** project on a **national** and **international** scale

Project funded by grants from the Czech Science Foundation. ELIXIR CZ research infrastructure project by the Ministry of Education, Youth and Sports of the Czech Republic is hosting the database and provides access to computing and storage facilities.

Scientific context

► Major ecological and economic roles of fungi



Ecosystem services

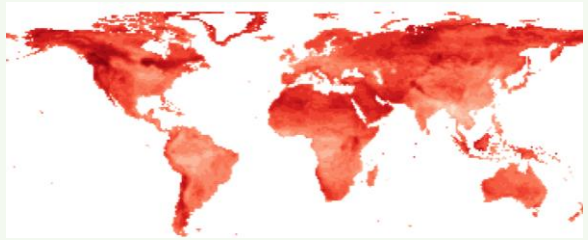
Environmental filters determining fungal diversity and distribution

- Climate
- Soil characteristics
- Plant community composition

At global scale…
What is where and why ?

Objectives of the GlobalFungi database

- Characterize the distribution of fungal diversity worldwide

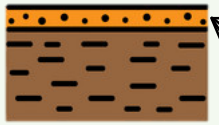


- Determine which environmental factors could best explain the distribution of fungal diversity

Climate

Temperature
Precipitation

Soil characteristics



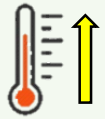
Bulk density
Soil pH
Cation exchange capacity
Organic carbon content

Plant community

Enhanced vegetation index
Gross primary productivity

- Assess the impact of global change on fungal diversity

Warming

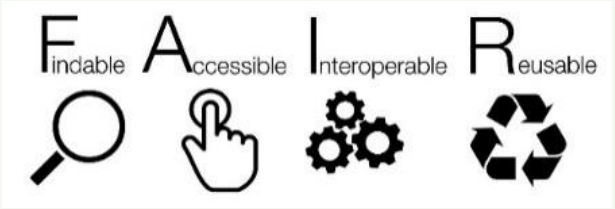


Drought

Nitrogen deposition

Elevated CO_2

- Create an open database fitting with the FAIR principles and continuously enriched with recent data



Methods (1/2) – Source of data

► **Join the efforts** of the scientific community to collect **fungal diversity data** worldwide by compiling **high-throughput sequencing results** from papers reporting fungal community composition in **environmental samples**

- **Environmental samples** (soil, air, water, coral, fungal sporocarp, lichen, mosses, algae, plant material)
- **Exclude manipulated** environments/studies (microcosms, greenhouse experiments…) → *specific category for manipulated samples (T°C, water, CO_2, N) in field conditions*
- **Include** samples from environments following **usual agricultural practices**
- Sequencing primers should be **general for fungi** and target the **ITS region of rDNA** → *specific category for AMF fungi*
- Sequencing data must result from **high-throughput sequencing** methods
- **raw sequencing data** (FASTQ format) available into the paper or requested to the authors
- **GPS coordinates** must be provided

Published papers

- Internal collaboration
- International collaboration
- Annual paper search with keywords

Check for eligibility criteria

Extraction of a set of data and metadata

- GPS coordinates
- Raw sequencing data
- Set of metadata (x16 mandatory and x32 optional; *geographic, sampling, DNA extraction, soil characteristics, sequencing, description of manipulation for specific category of manipulated samples*)

Checking/homogenization

- Ensure data quality
- Communication with authors

Retrieval of metadata from external databases

- Climate
- Soil characteristics
- Vegetation

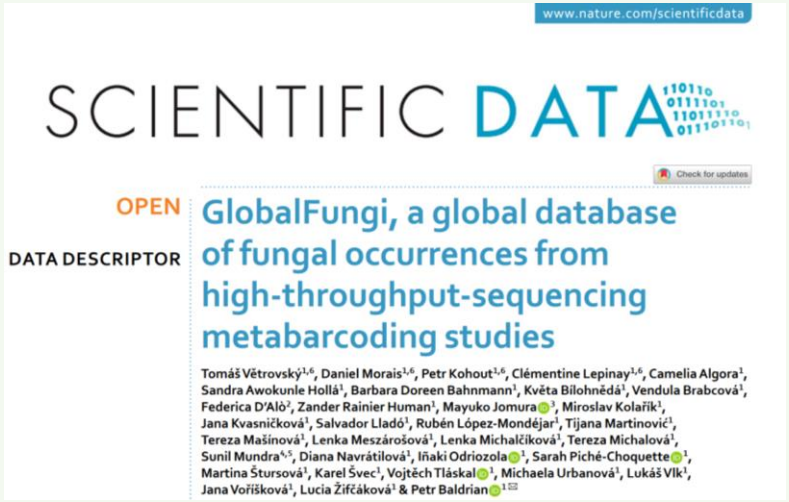
CHELSA

ISRIC World Soil Information SoilGrids

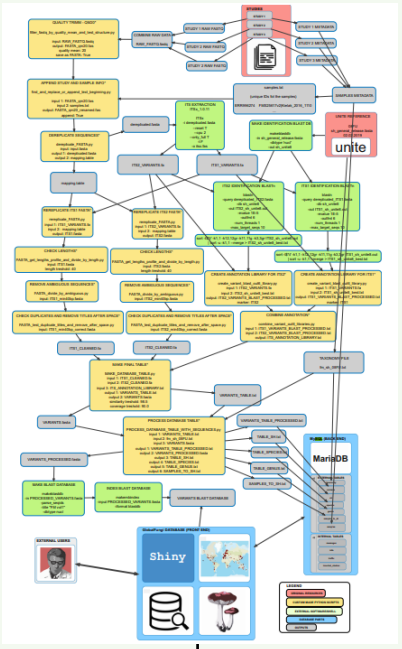
NASA MODIS satellite data

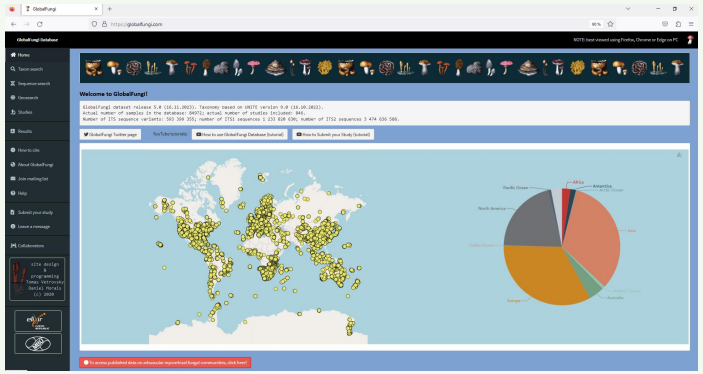
Ontobee

Methods (1/2) – Data processing



Data processing



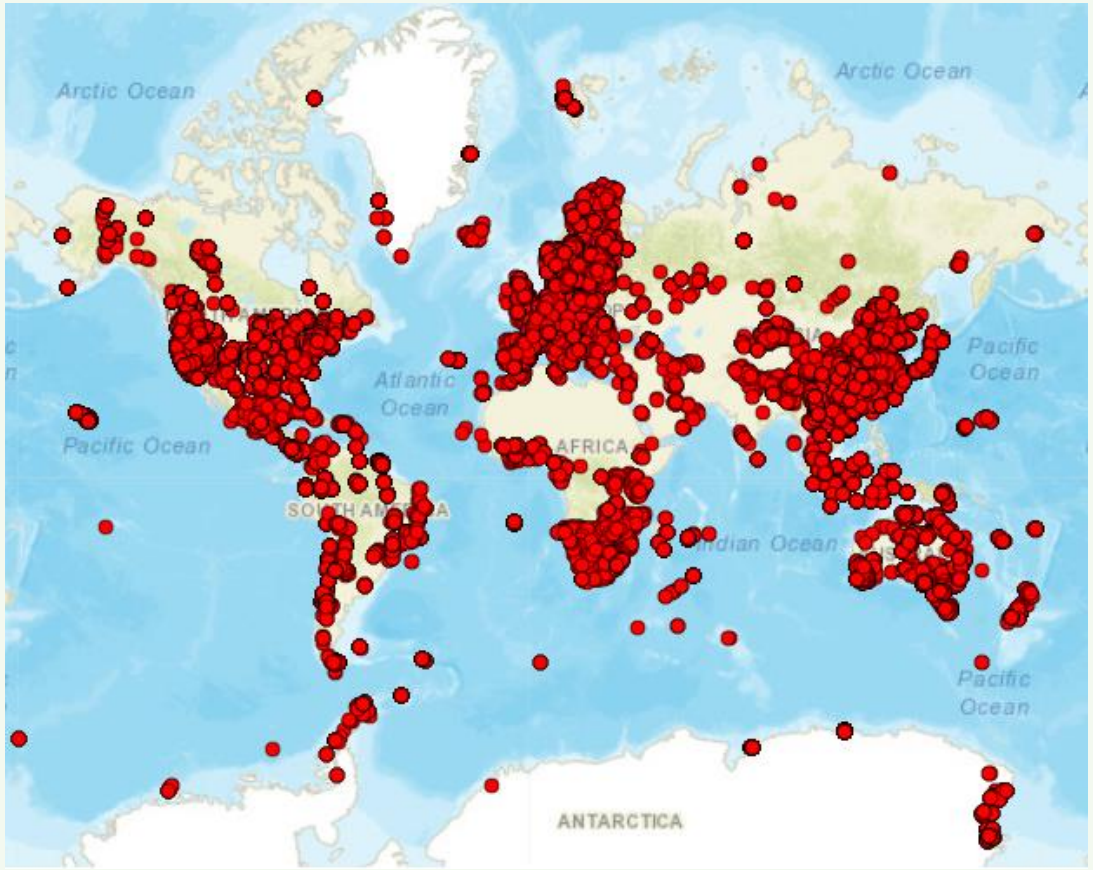


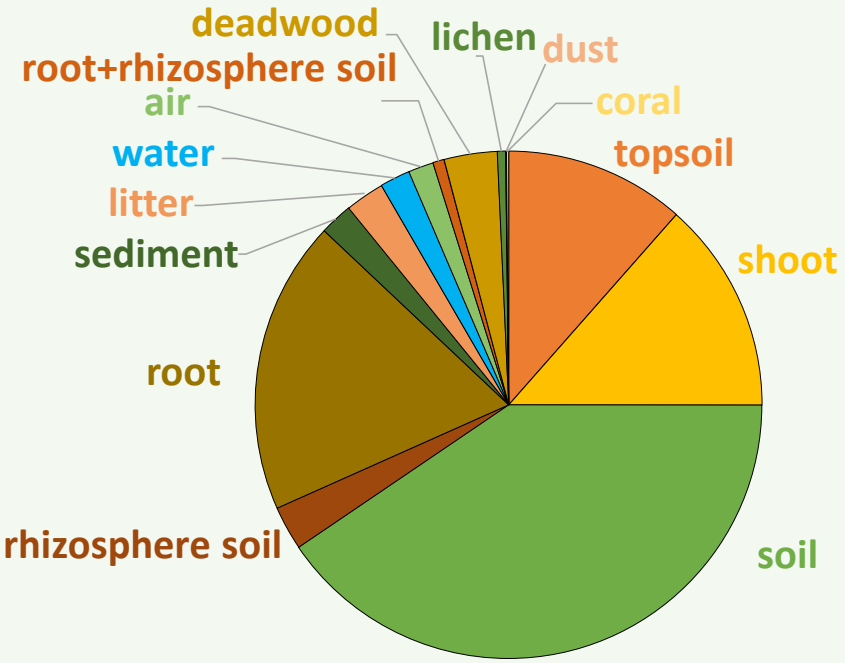
SEED pipeline
(Větrovský *et al*. 2018)

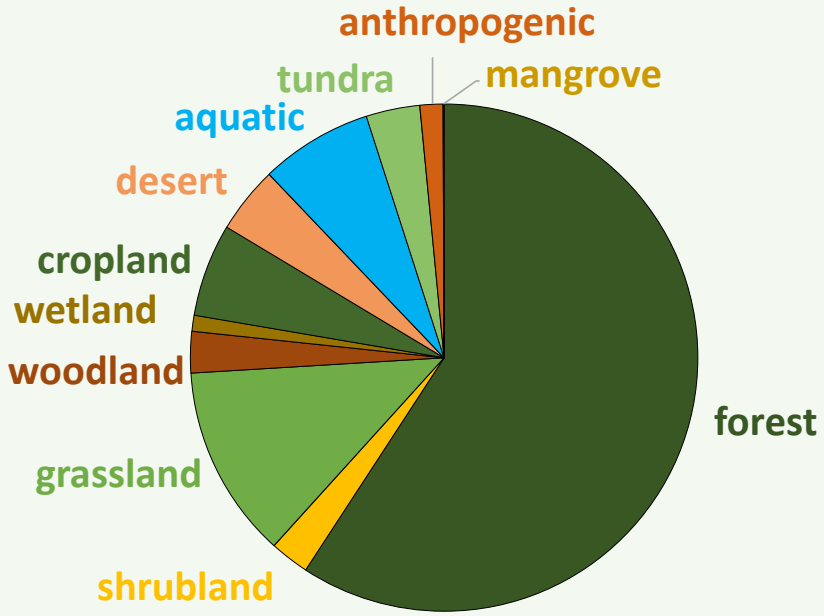
- Raw sequencing data quality filtered
- Extraction ITS1 and/or ITS2 fungal region
- ITS sequences classified according to the representative sequence of the closest UNITE species hypothesis (SH) using BLASTn (98.5% similarity threshold)
- All representative sequences unclassified were used to build database library of unique nucleotide sequences (sequence variants)

Current status of the GlobalFungi database

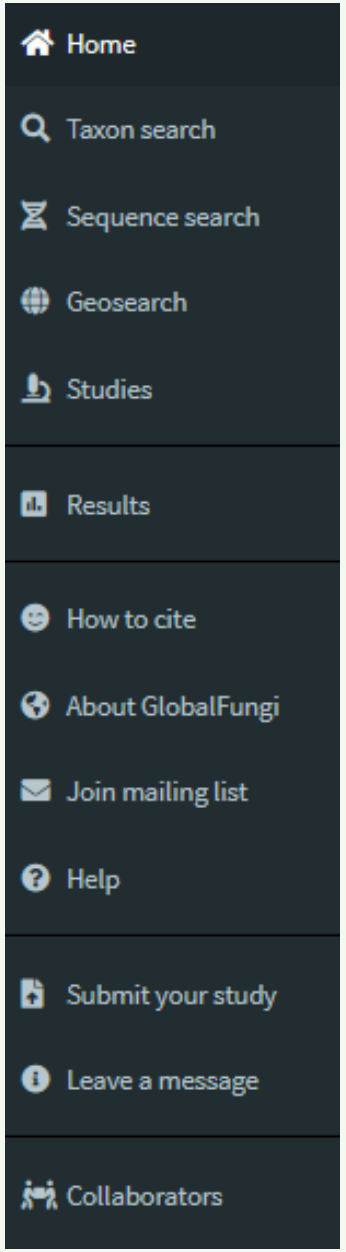
→ Actual number of samples in the database: 84972
→ Actual number of studies included: 846



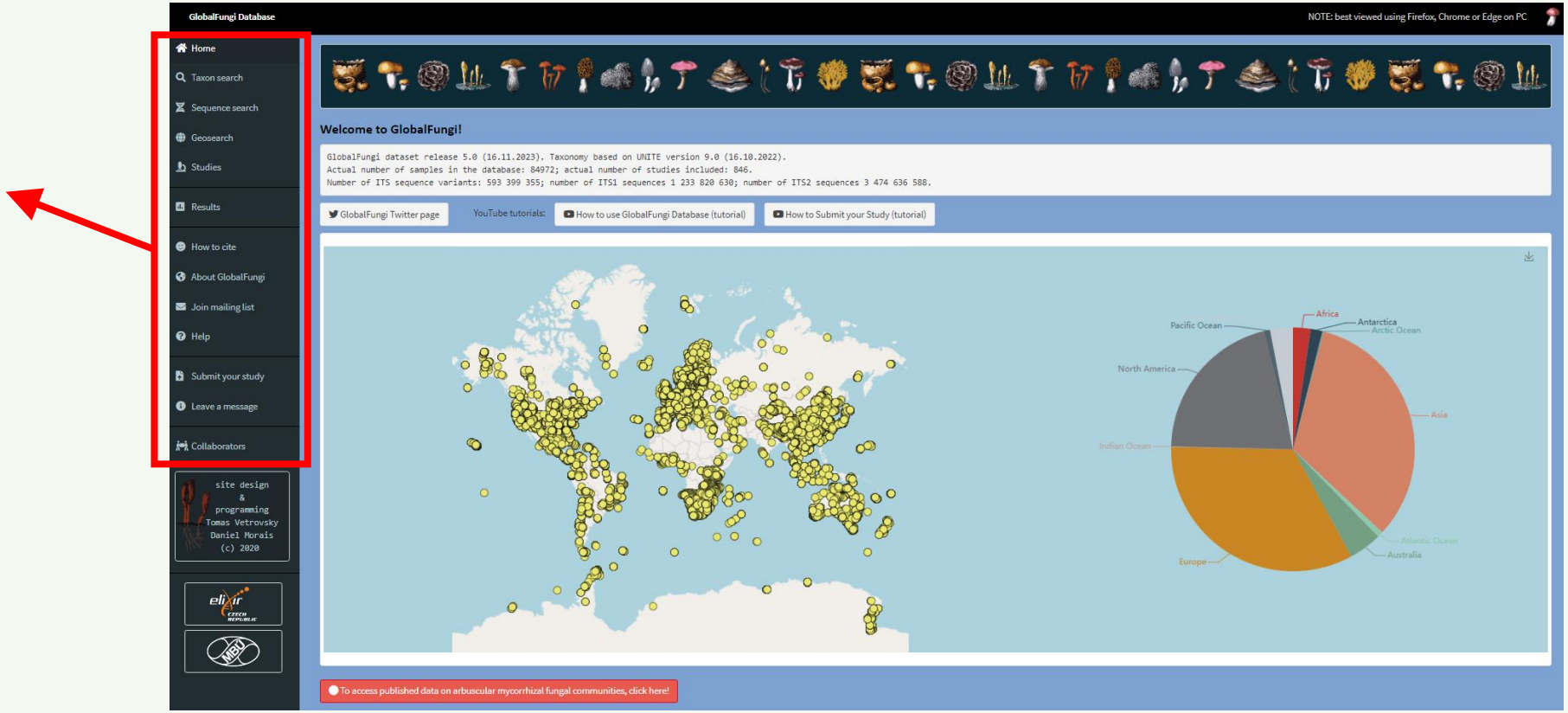




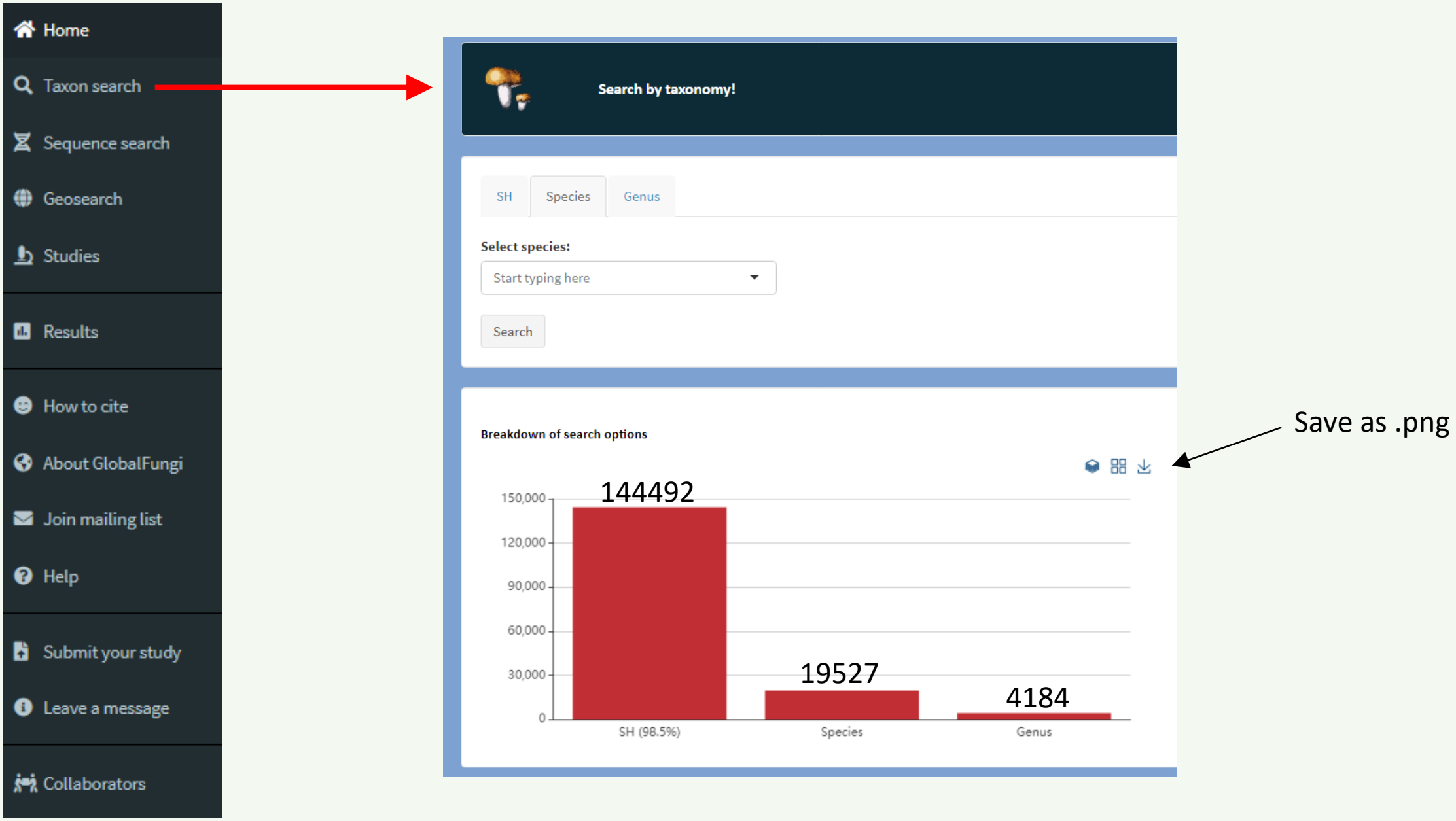
How to use the GlobalFungi database



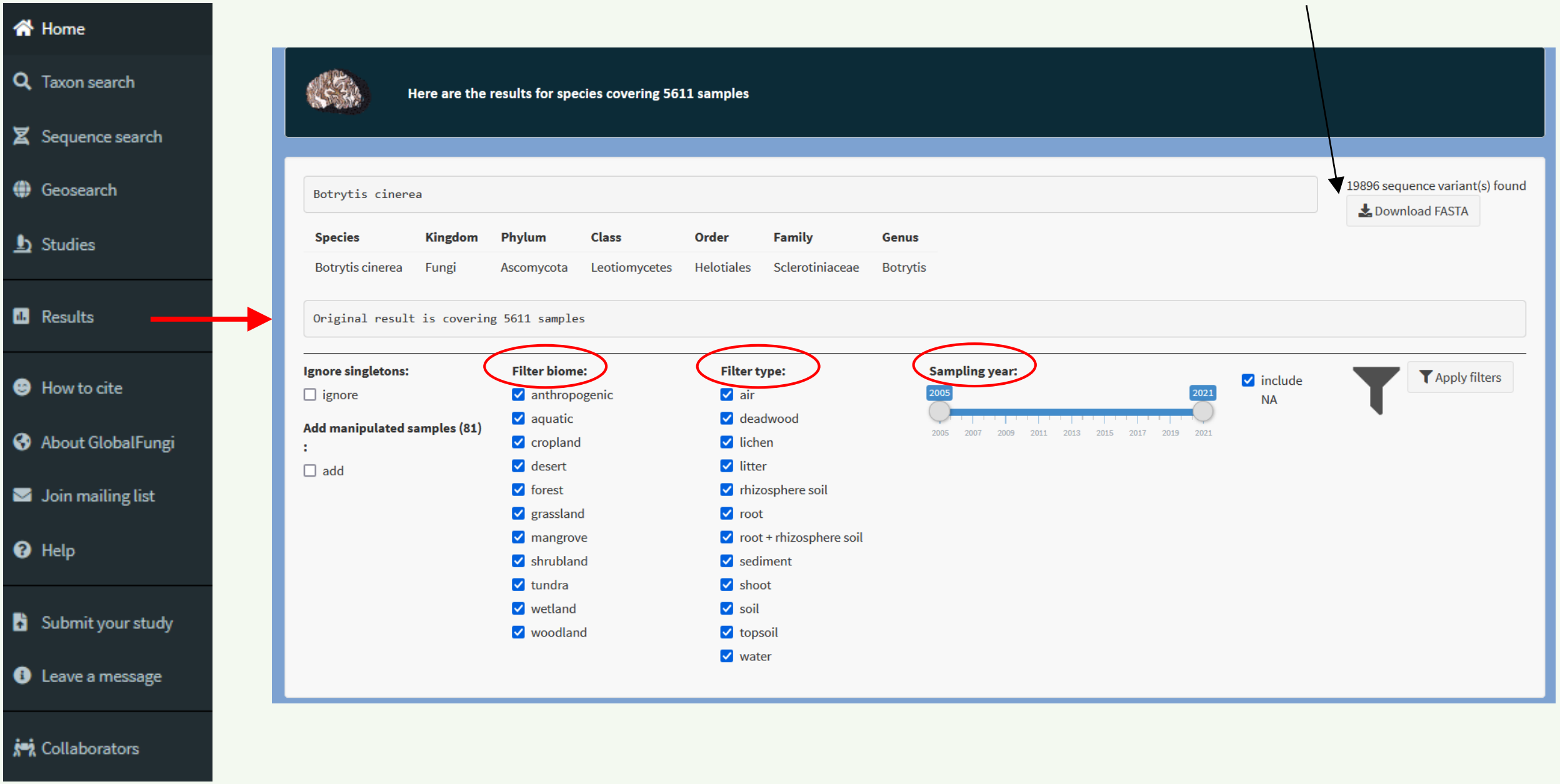
https://globalfungi.com/



How to use the GlobalFungi database



How to use the GlobalFungi database



How to use the GlobalFungi database

- Home
- Taxon search
- Sequence search
- Geosearch
- Studies
- Results
- How to cite
- About GlobalFungi
- Join mailing list
- Help
- Submit your study
- Leave a message
- Collaborators

Filtered result is covering 5611 samples (NO FILTERS APPLIED)

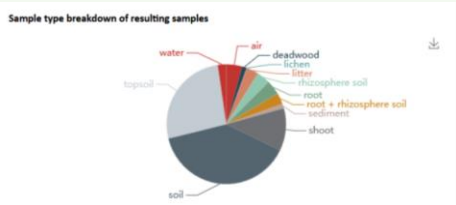
SHs | Sample type & Biome | MAT & MAP | pH | Geography | Map | Metadata

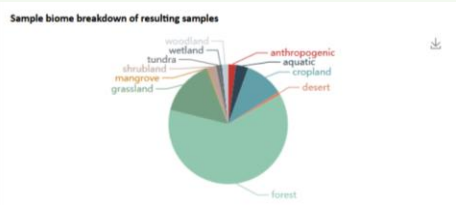
Download SH list ← Save as .txt

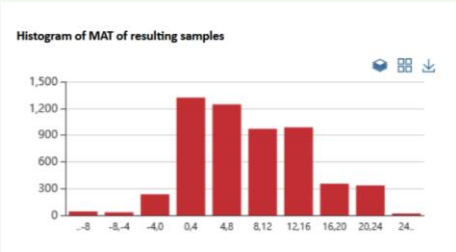
Show 10 entries Search:

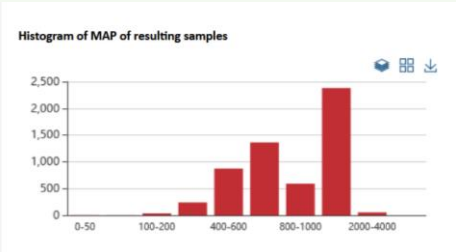
	SH	Kingdom	Phylum	Class	Order	Family	Genus	Species
1	SH1152451.08FU	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis cinerea
2	SH1152452.08FU	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis cinerea
3	SH1152462.08FU	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis cinerea
4	SH1152477.08FU	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis cinerea

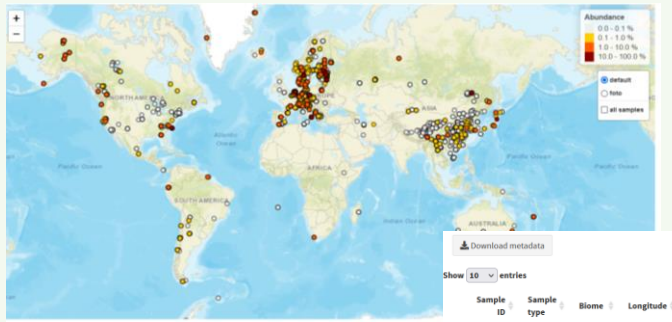
Showing 1 to 4 of 4 entries Previous 1 Next

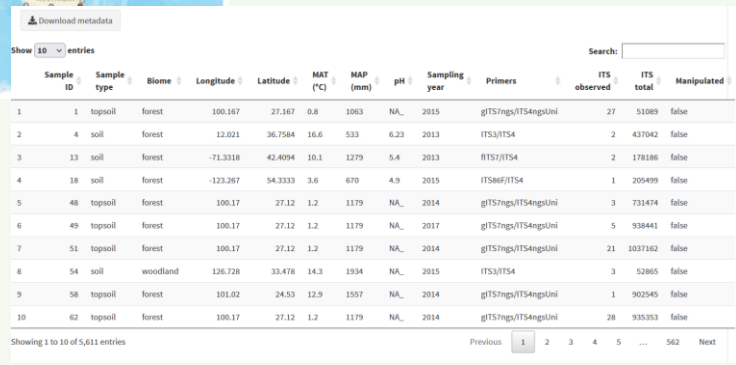




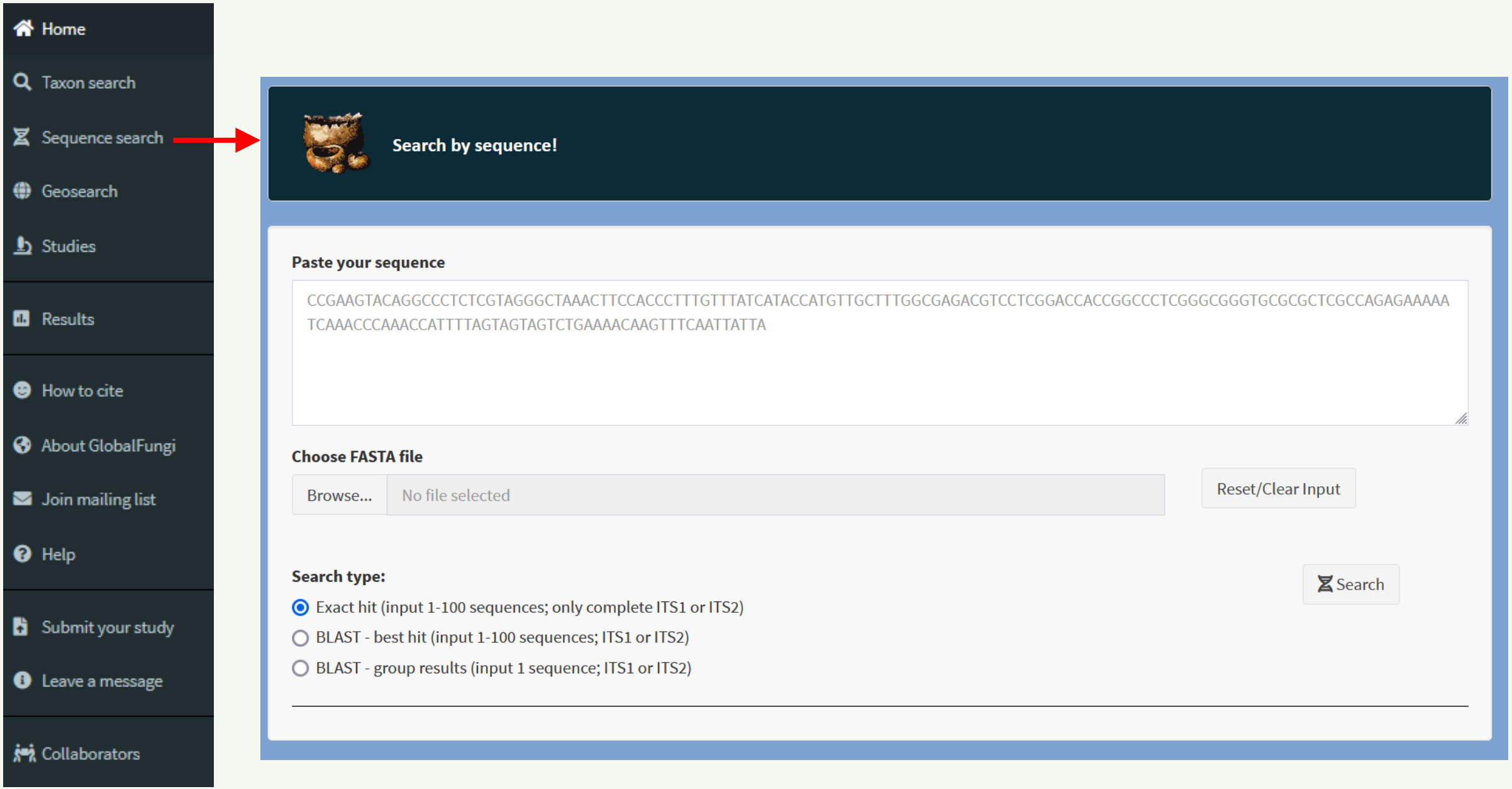




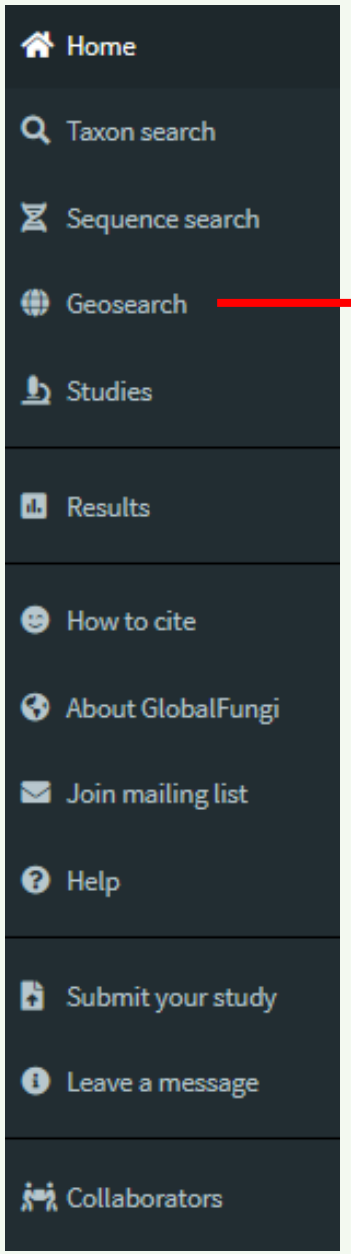


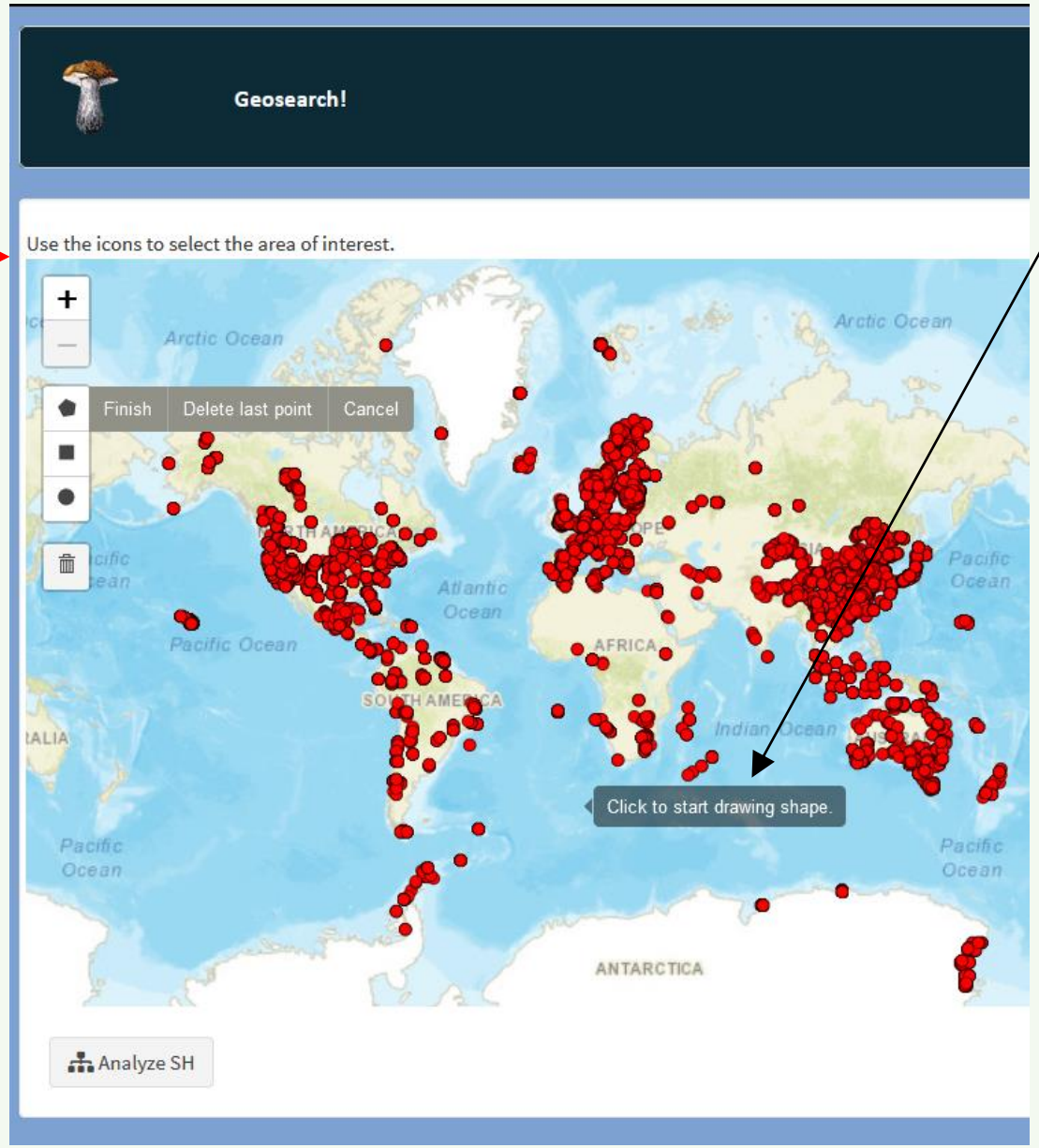


How to use the GlobalFungi database

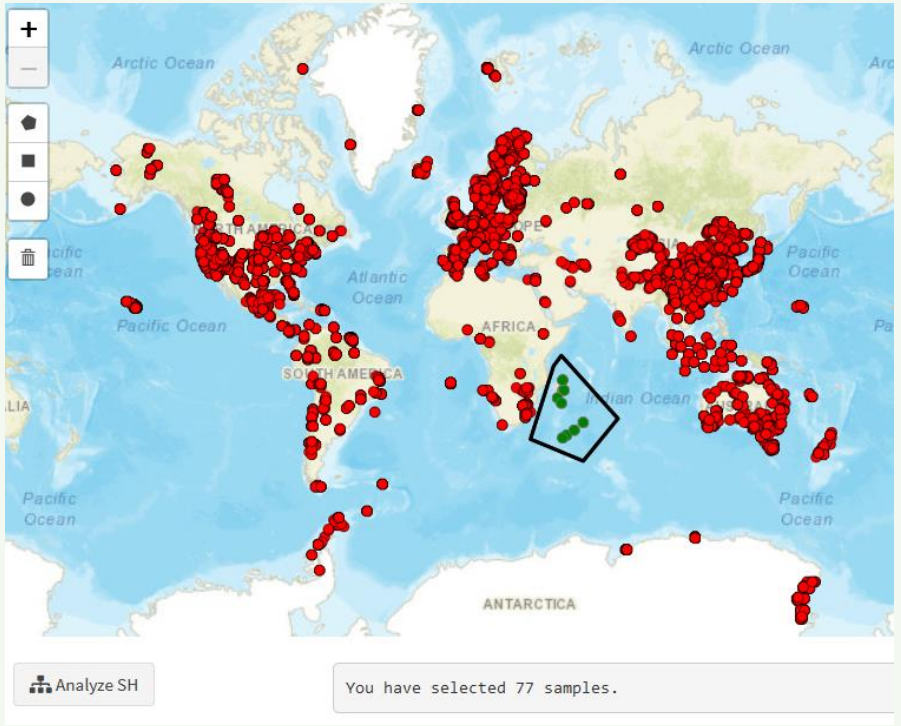


How to use the GlobalFungi database



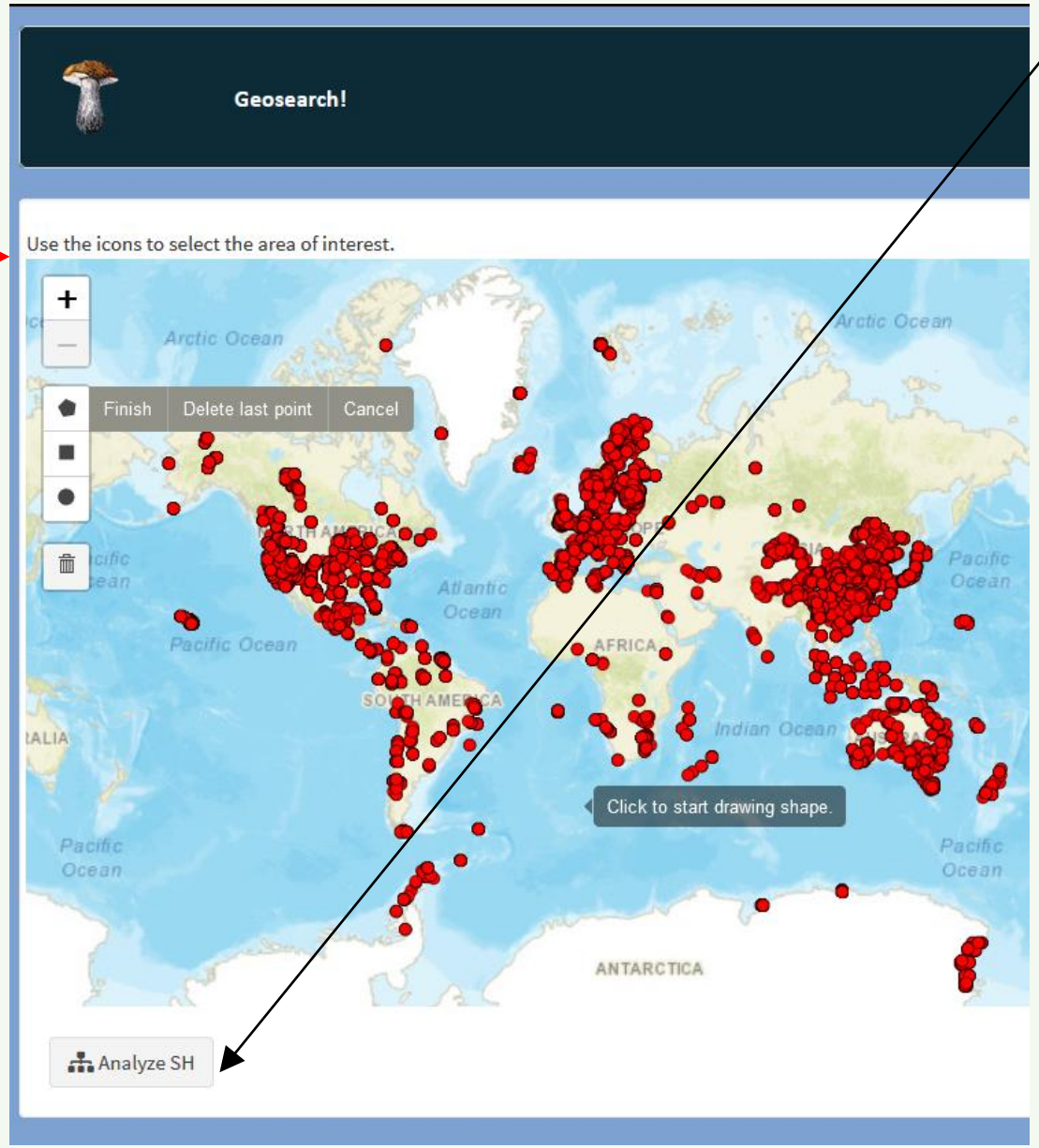


1) Delimit an area



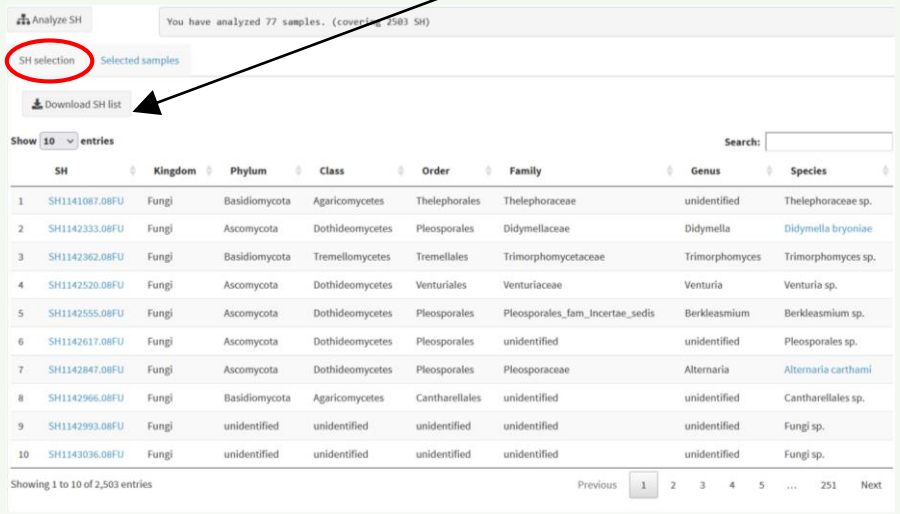
How to use the GlobalFungi database

- Home
- Taxon search
- Sequence search
- Geosearch
- Studies
- Results
- How to cite
- About GlobalFungi
- Join mailing list
- Help
- Submit your study
- Leave a message
- Collaborators

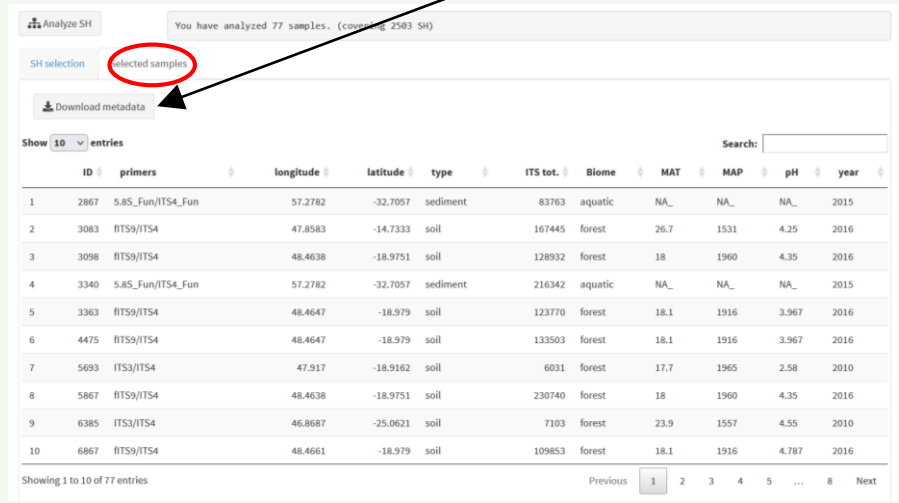


2) Analyse the results

Save as .txt

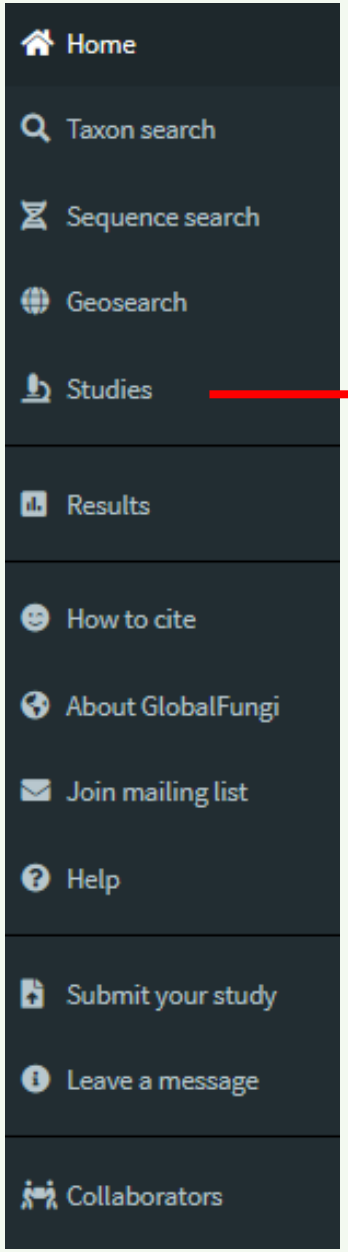


Save as .txt



How to use the GlobalFungi database

Information about all the samples from the study



Studies

Show 10 entries Search:

	Date	Title	Authors	Journal	Year	DOI	manipulated	Actions
1	21.07.2022	Revealing cues for fungal interplay in the plant-air interface in vineyards.	Abdelfattah, A., Sanzani, S.M., Wisniewski, M., Berg, G., Cacciola, S.O. and Schena, L.	Frontiers in plant science	2019	10.3389/fpls.2019.00922	false	Show
2	21.07.2022	Reindeer grazing history determines the responses of subarctic soil fungal communities to warming and fertilization.	Ahonen, S.H., Ylänne, H., Väisänen, M., Ruotsalainen, A.L., Männistö, M.K., Markkola, A. and Stark, S.	New Phytologist	2021	10.1111/nph.17623	true	Show
3	21.07.2022	Ectomycorrhizal community composition and function in a spruce forest transitioning between nitrogen and phosphorus limitation.	Almeida, J. P., Rosenstock, N. P., Forsmark, B., Bergh, J., & Wallander, H.	Fungal Ecology	2019	10.1016/j.funeco.2018.05.008	true	Show
4	21.07.2022	Plant invasion impacts on fungal community structure and function depend on soil warming and nitrogen enrichment.	Anthony, M.A., Stinson, K.A., Moore, J.A. and Frey, S.D.	Oecologia	2020	10.1007/s00442-020-04797-4	true	Show
5	21.07.2022	Fungal Diversity and Composition of the Continental Solar Saltern in Añana Salt Valley (Spain).	Azpiazu-Muniozguren, M., Perez, A., Rementeria, A., Martinez-Malaxetxebarria, I., Alonso, R., Laorden, L., Gamboa, J., Bikandi, J., Garaizar, J. and Martinez-Ballesteros, I.	Journal of Fungi	2021	10.3390/jof7121074	false	Show
6	21.07.2022	Next-generation sequencing of root fungal communities in continuous cropping soybean.	Bai, L., Sun, H., Zhang, X. and Cai, B.	Chilean journal of agricultural research	2018	10.4067/S0718-58392018000400528	false	Show
7	21.07.2022	Analyzing Ash Leaf-Colonizing Fungal Communities for Their Biological Control of Hymenoscyphus fraxineus.	Becker, R., Ulrich, K., Behrendt, U., Kube, M. and Ulrich, A.	Frontiers in Microbiology	2020	10.3389/fmicb.2020.590944	false	Show
8	21.07.2022	Evidence for Co-evolutionary History of Early Diverging Lycopodiaceae Plants With Fungi.	Benucci, G. M. N., Burnard, D., Shepherd, L. D., Bonito, G. and Munkacsi, A. B.	Frontiers in Microbiology	2020	10.3389/fmicb.2019.02944	false	Show
9	21.07.2022	Partner turnover and changes in ectomycorrhizal fungal communities during the early life stages of European beech (Fagus sylvatica L.).	Boeraeve, M., Everts, T., Vandekerkhove, K., De Keersmaeker, L., Van de Kerckhove, P. and Jacquemyn, H.	Mycorrhiza	2021	10.1007/s00572-020-00998-0	false	Show
10	21.07.2022	Forest edge effects on the mycorrhizal communities of the dual-mycorrhizal tree species Alnus glutinosa (L.) Gaertn.	Boeraeve, M., Honnay, O. and Jacquemyn, H.	Science of the Total Environment	2019	10.1016/j.scitotenv.2019.02.290	false	Show

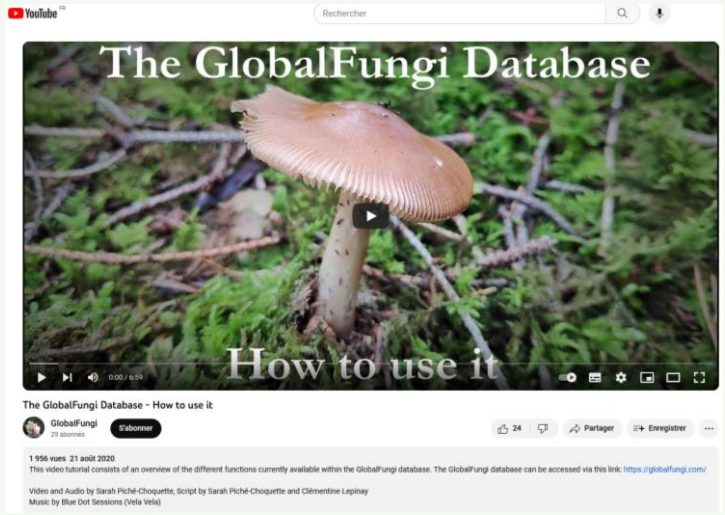
Showing 1 to 10 of 515 entries Previous 1 2 3 4 5 … 52 Next

How to use the GlobalFungi database

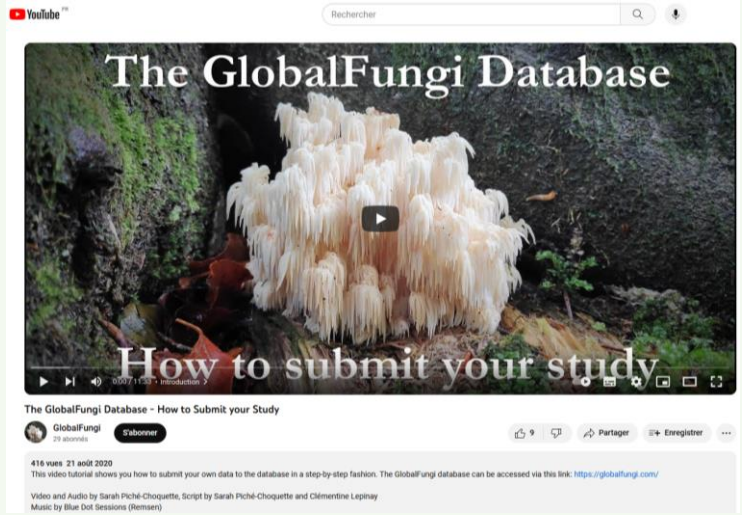
- Home
- Taxon search
- Sequence search
- Geosearch
- Studies
- Results
- How to cite
- About GlobalFungi
- Join mailing list
- Help
- Submit your study
- Leave a message
- Collaborators

Větrovský T., Morais D., Kohout P., Lepinay C., *et al.* GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Scientific Data 7, 228 (2020). https://doi.org/10.1038/s41597-020-0567-7

Brief presentation, record of each release, links to tutorial videos …



https://www.youtube.com/watch?v=0_opE1hOXwY



https://www.youtube.com/watch?v=HmGyr26Hhso

Mailing list (new releases)

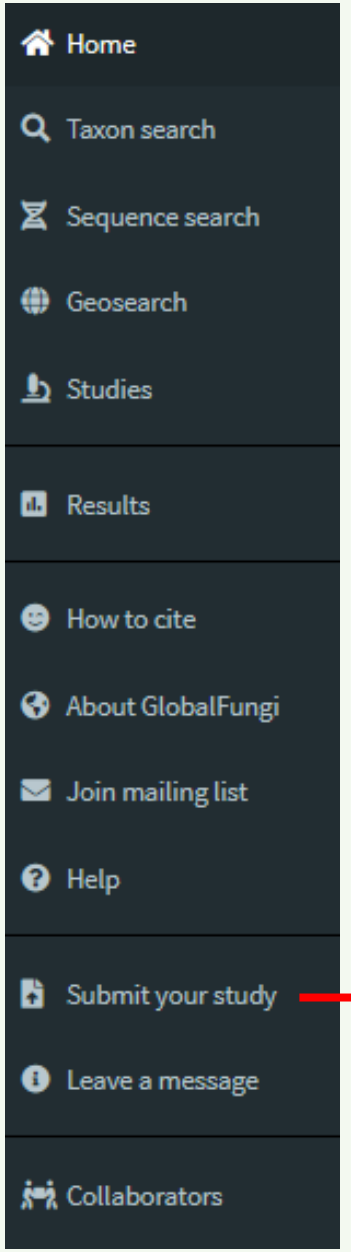
https://twitter.com/globalfungi

Presentation of each section

info@globalfungi.com

List of all the collaborators

Call for collaborations !



- ► Follow the written instructions + video tutorial + request help if necessary (Clémentine or info@globalfungi.com)
- ► Easy process: one Excel file to fill (including the accession numbers allowing to download the Fastq/Fasta files of the raw sequencing data from public repositories)

	A	B	C
2	Please use dot . as floating point!		
3	Empty fields fill with "NA_"		
4			
5	**VARIABLE**	**OPTIONS**	**HINTS**
6	***Latitude***	***TEXT - Mandatory!***	Geographical latitude where the sample has been collected, in decimal degrees and using the WGS 84 geodetic system (i.e., following the format: "12.345" if North or "-12.345" if South. If necessary, you can convert 'degrees, minutes, seconds' in 'decimal degrees' with https://www.gps-coordinates.net/).
7	***Longitude***	***TEXT - Mandatory!***	Geographical longitude where the sample has been collected, in decimal degrees and using the WGS 84 geodetic system (i.e., following the format: "56.789" if East or "-56.789" if West. If necessary, you can convert 'degrees, minutes, seconds' in 'decimal degrees' with https://www.gps-coordinates.net/).
8	Elevation study (m)	TEXT - optional	Elevation in meter above sea level, , or meter below sea level for aquatic samples, with a minus sign (example: "-3300").
9	***Continent/Ocean***	***One of the following: Africa/Antarctica/Asia/ Australia/Europe/North America/South America/Atlantic Ocean/Arctic Ocean/Indian Ocean/Pacific Ocean/Southern Ocean***	***Specify the continent/ocean where the sample has been collected (select from the proposed list).***
10	Country	TEXT - optional	Country where the sample has been collected. For aquatic samples, please specify the country to which the water area is reattached or leave empty if you are not sure.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	***TEXT - Mandatory!***	***TEXT - Mandatory!***	TEXT - optional	***One of the fc***	TEXT - option	TEXT - option	***One of the fc***	***One of the follo***	TEXT - optional	TEXT - option	TEXT - option	***TEXT - Mand***	TEXT - option
2	***Latitude***	***Longitude***	Elevation study (m)	***Continent/ Ocean***	Country	Location	***Sample type***	***Environment classification***	Environment classification (ENVO)	MAT (°C) study	MAP (mm) study	***Sample ID in study***	Area covered by sampling (m2)
3													
4													
5													
6													

Why to submit your study?

+ your data will be accessible in an **easily accessible form**
+ your work **gets visibility** and the international community gets the access to **additional resources**
+ **data sharing** is important
+ your name and affiliation can be added into the **list of collaborators**
+ you can have your name added to the **GlobalFungi Authorship Group** to be mentioned in future publications

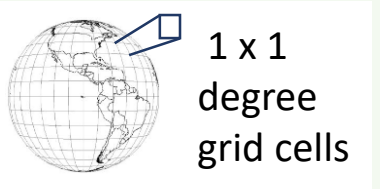
What about our initial research questions?

Detailed analysis of 3084 soil samples from 36 studies (Větrovský, T., *et al.* 2019 Nature Communications)

Fungal diversity data → Uniform sequence processing → OTU number / Chao's index / ecological guilds

\+

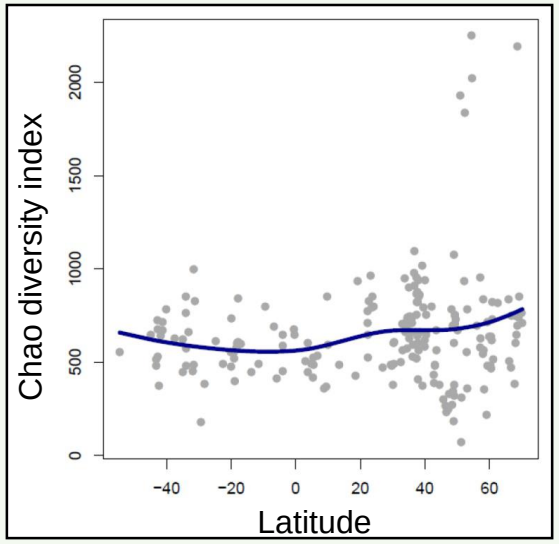
GPS coordinates → Retrieval of environmental data → Environmental variables (climate, soil, vegetation)

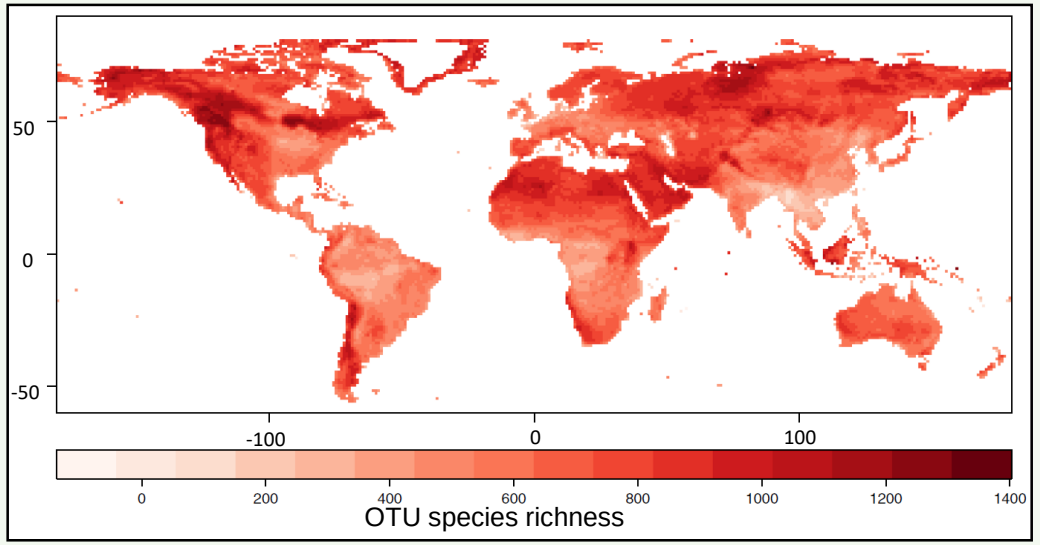


► <u>Characterize the distribution of fungal diversity worldwide</u>:

→ Non-parametric smoothing (Generalized linear models (GLMs) with a second-degree polynomial function)

→ Climate-based generalized linear models (18 million GLMs with combinations of bioclimatic variables (max. 10 variables/model; R package *leaps,* Lumley 2017). Best model based on adjusted R^2, AIC, and BIC)



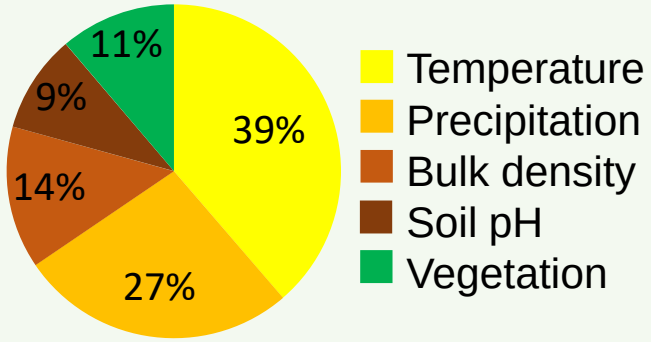


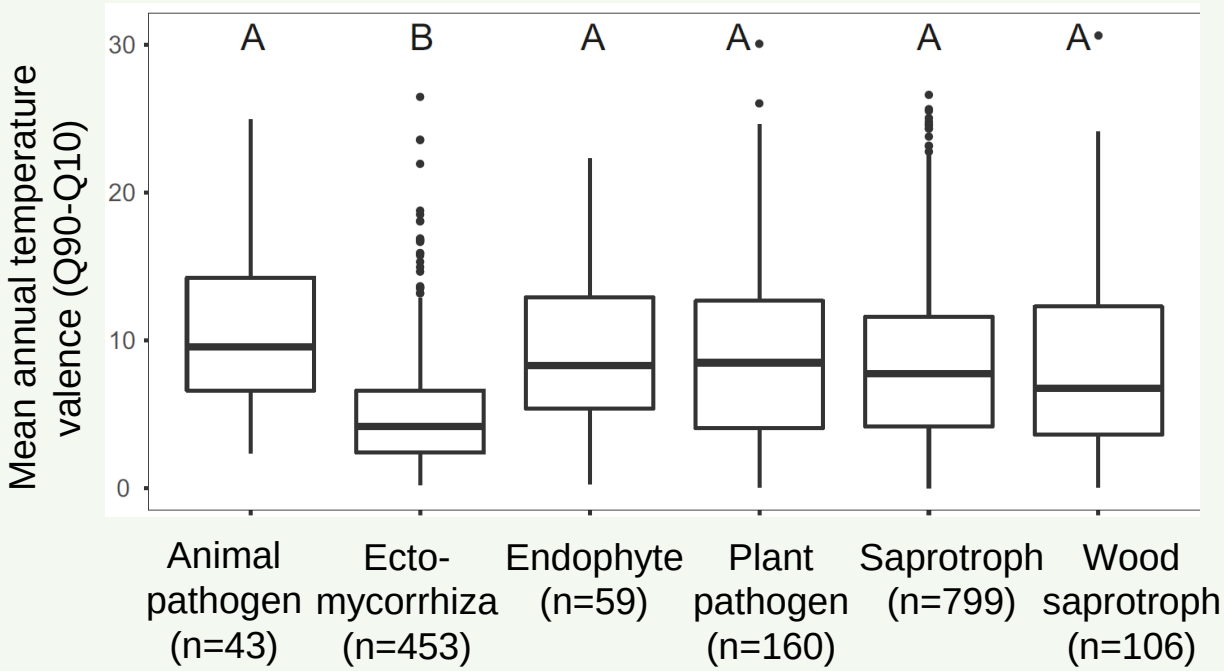
- High latitudes harbour the most diverse fungal communities
- Significant variation in diversity in temperate regions (peaks in the boreal forests of Eurasia and North America)
- Low diversity in the Southern Hemisphere

What about our initial research questions?

► <u>Determine which environmental factors could best explain the distribution of fungal diversity</u>:

→ Random forest (Comparison between trees containing randomly sampled 2/3 of the observations and 1/3 of the predictors and the remaining observations (Breiman 2001; Liaw & Wiener 2002 for R package *randomForest*))





- Climate variables contribute to the explanation of biogeographic distribution for 97% of the 469 top taxa (Species Hypothesis in >5% of samples)
- Ectomycorrhizal fungi have a narrower climate niche than other guilds

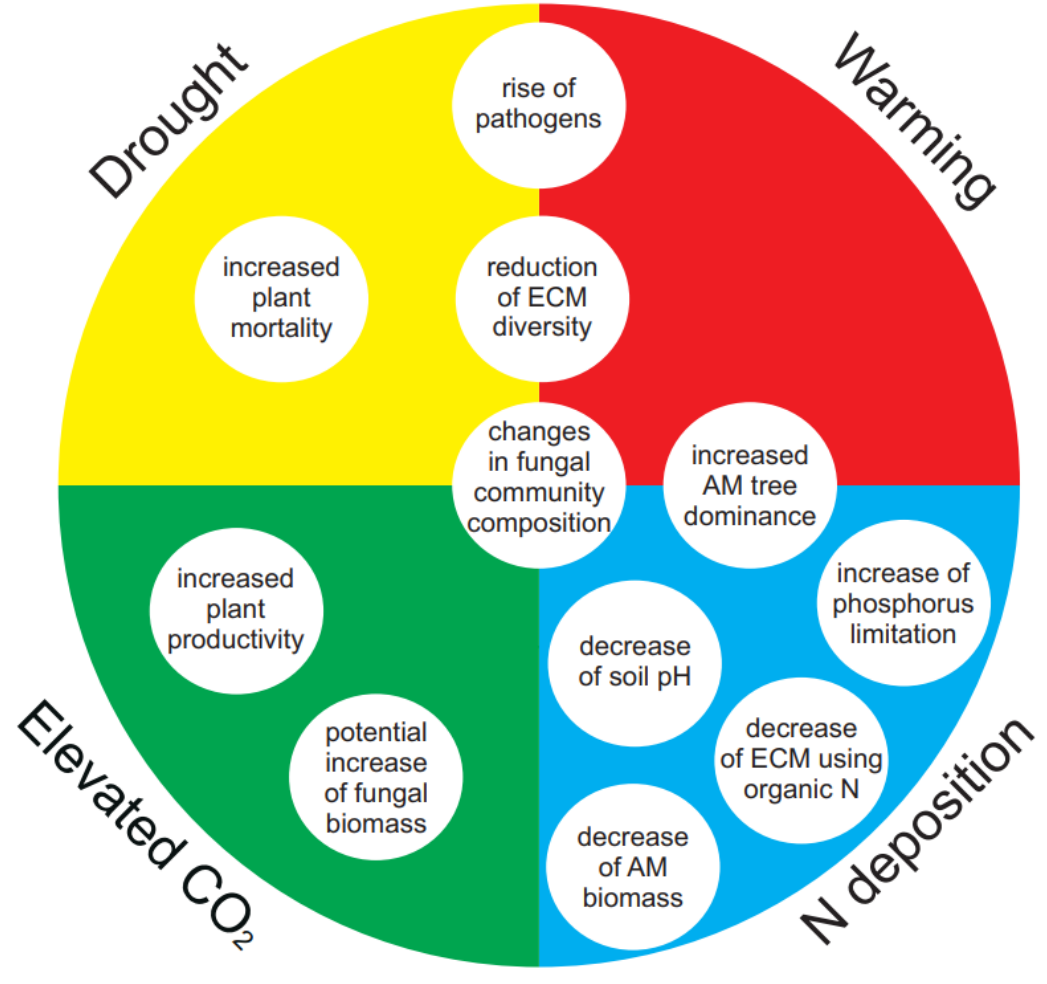
What about our initial research questions?

Available online at www.studiesinmycology.org
STUDIES IN MYCOLOGY 103: 1–24 (2022).

Fungal communities in soils under global change

P. Baldrian*, L. Bell-Dereske, C. Lepinay, T. Větrovský, P. Kohout

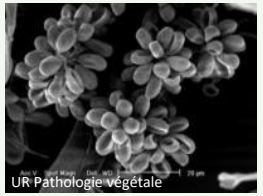
► Assess the impact of global change on the distribution of fungal diversity



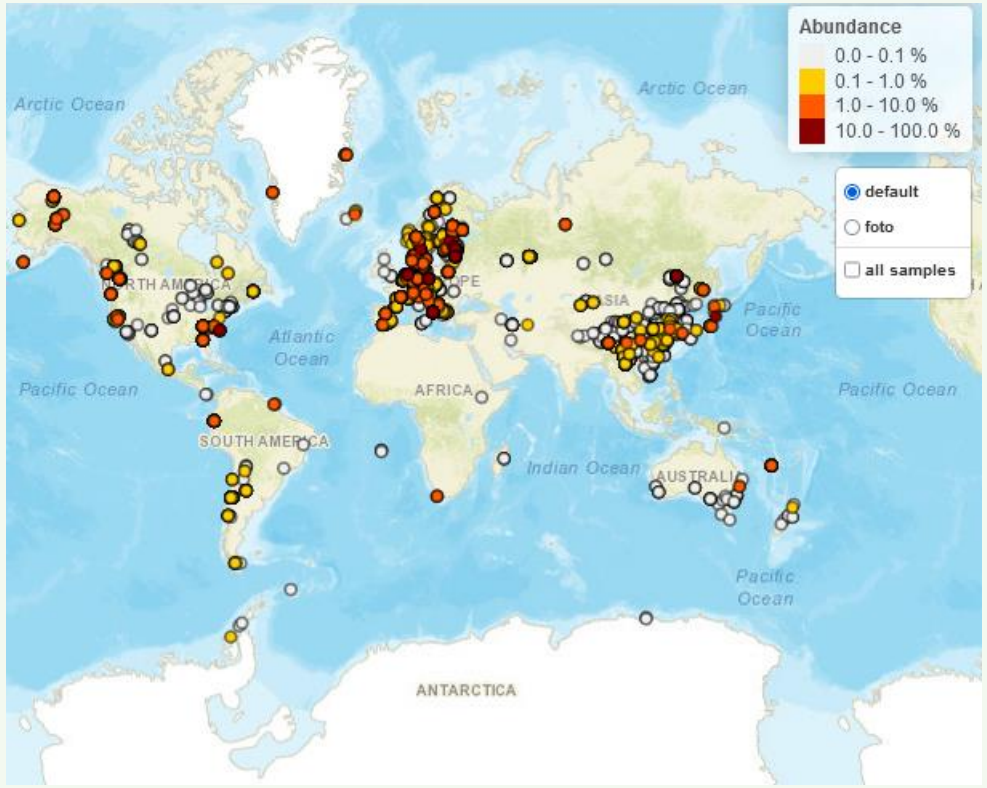
- Strong effect of N deposition on fungal communities
- Increase in abundance and dispersal of plant pathogenic fungi
- plant-mutualistic fungal guilds (ectomycorrhizal fungi and arbuscular mycorrhizal fungi) strongly and negatively affected by N deposition and warming
- Effect of size and duration of change + local conditions

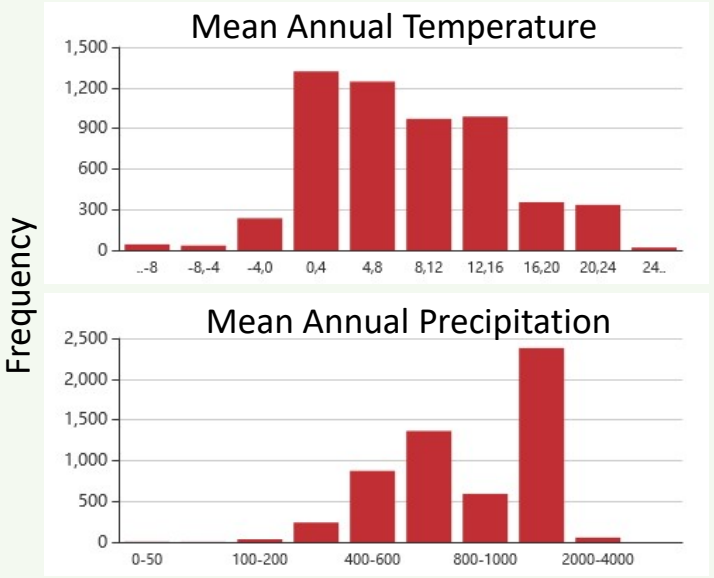
How to use the database in plant pathology?

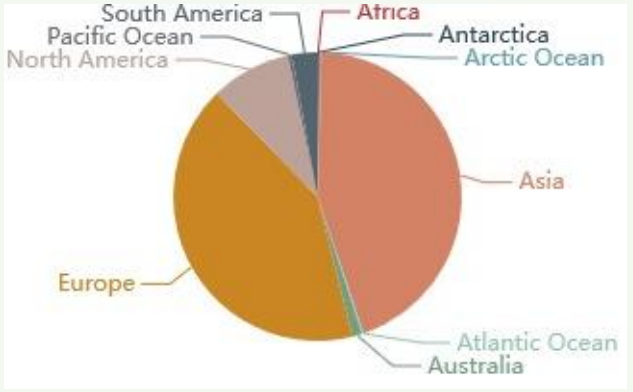
► Example of the airborne pathogenic fungi *Botrytis cinerea*

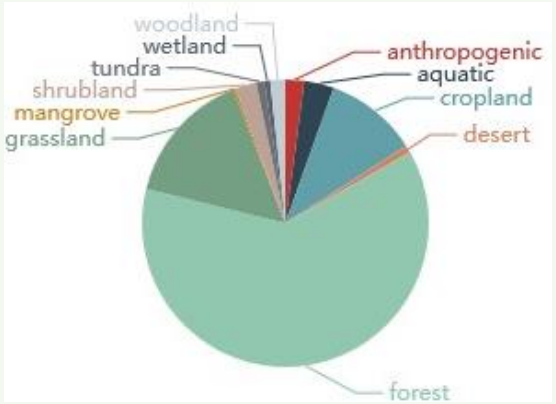


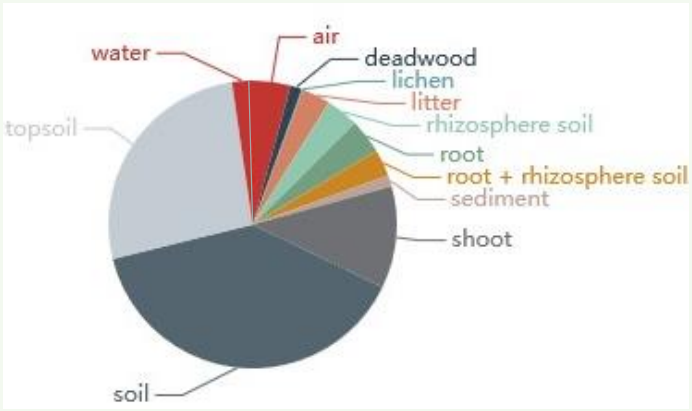
Botrytis cinerea

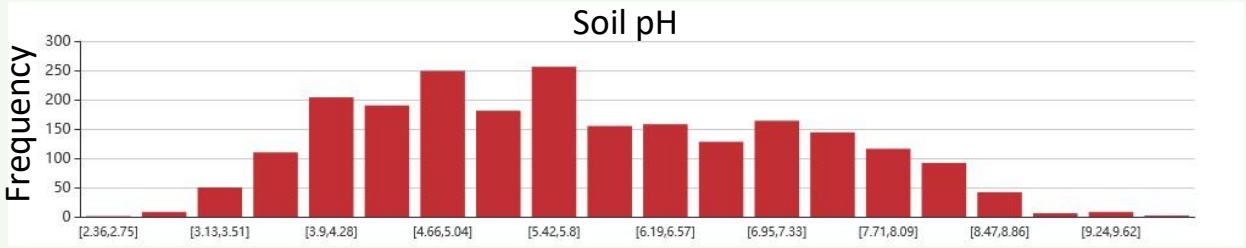












Perspectives of the GlobalFungi database

- Open to many other scientific questions …

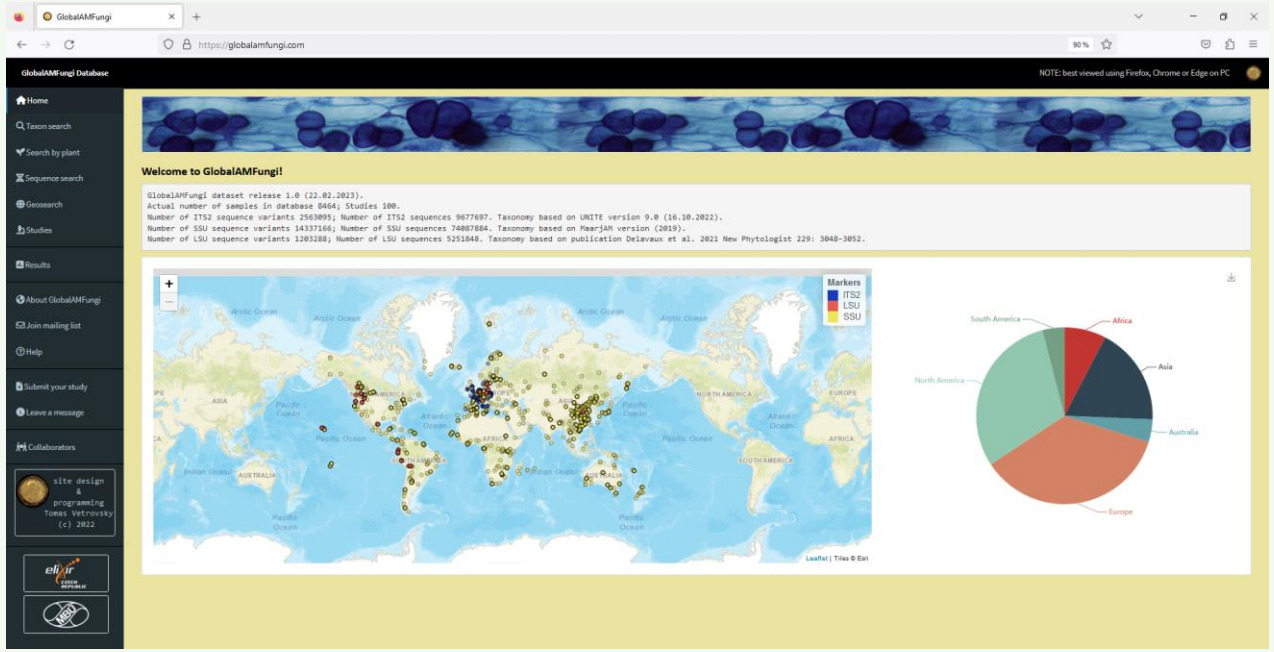
Fungal Diversity
https://doi.org/10.1007/s13225-021-00472-y

Check for updates

High-throughput sequencing view on the magnitude of global fungal diversity

Petr Baldrian[1] · Tomáš Větrovský[1] · Clémentine Lepinay[1] · Petr Kohout[1]

- Similar project developed for arbuscular mycorrhizal fungi (AMF): https://globalamfungi.com/



New Phytologist
Research

Methods

GlobalAMFungi: a global database of arbuscular mycorrhizal fungal occurrences from high-throughput sequencing metabarcoding studies

Tomáš Větrovský[1], Zuzana Kolaříková[2], Clémentine Lepinay[1], Sandra Awokunle Hollá[1], John Davison[3], Anna Fleyberková[1], Anastasiia Gromyko[1], Barbora Jelínková[1], Miroslav Kolařík[1], Manuela Krüger[2], Renata Lejsková[1], Lenka Michalčíková[1], Tereza Michalová[1], Mari Moora[3], Andrea Moravcová[1,4], Štěpánka Moulíková[1], Iñaki Odriozola[1], Maarja Öpik[3], Monika Pappová[1], Sarah Piché-Choquette[1], Jakub Skřivánek[1,4], Lukáš Vlk[1], Martin Zobel[3], Petr Baldrian[1] and Petr Kohout[1,4]

Accepted: 4 September 2023

Thank you for your attention









Clémentine Lepinay – INRAE – Unité de Pathologie Végétale