

Metadata description - GlobalFungi, Release 5, file "GlobalFungi_5_sample_metadata.txt"

Fieldname	Field content and format	Values from vocabulary	Note
sample_ID	unique identifier of the sample	NO	
paper_ID	unique identifier of the source paper	NO	
paper_title	title of paper	NO	
paper_year	year when paper was published	NO	
paper_authors	list of authors of the paper	NO	
paper_journal	journal where paper was published	NO	
paper_doi	DOI of the paper	NO	
paper_sample_name	sample name used in the paper	NO	
latitude	latitude as decimal number	NO	Using the WGS 84 World Geodetic System
longitude	longitude as decimal number	NO	Using the WGS 84 World Geodetic System
elevation_study	elevation of the sample location (m)	NO	Relative to sea level, negative values = below sea level
continent	continent or ocean	YES (Africa; Antarctica; Asia; Australia; Europe; North America; South America; Atlantic Ocean; Arctic Ocean; Indian Ocean; Pacific Ocean; Southern Ocean)	
country	country	NO	In marine samples, country closest to sampling location
location	geographic description of sample location	NO	Example: "Tajen Forest Station" or "Deschutes National Forest, Oregon"
year_of_sampling_from	start of the year range when sampling was performed	NO	
year_of_sampling_to	end of the year range when sampling was performed	NO	
month_of_sampling	month when the sample was collected	NO	
day_of_sampling	day when the sample was collected	NO	
sample_type	one of defined sample types	YES (air; algae; coral; deadwood; dust; fungal sporocarp; glacial ice debris; lichen; litter; moss; rhizosphere soil; root; root + rhizosphere soil; sediment; shoot; soil; topsoil; water)	"soil" includes the samples belonging to any soil layer excluding litter, such as organic horizon or humus, mineral horizon, subsoil, substratum or parent rock, bedrock; "shoot" includes any part of plant that are aboveground, such as leaf, fruit, branch, stem; "topsoil" includes combined sample of litter and soil or litter, soil and roots; "rhizosphere soil" is soil under immediate influence of root, typically soil attached to root when plant is taken out of soil and gently shaken
sample_type_specification	specification of sample type	NO	Additional information specifying sample type. For "soil", e.g., "organic horizon=humus", "mineral horizon", "subsoil", "substratum=parent rock", "bedrock"; for "water" / "sediment" samples, e.g., "water from a filter", "bedrock from lower oceanic crust", "sediment from seafloor"; for "root" / "shoot", e.g., "deadwood, heartwood", "lichens, whole plant"; "leaf, several randomly selected needles", "fruit, several grape berries", "petiole", "root endosphere", "seed", "roots, ectomycorrhizae (ECM) root tips", etc.
environment_type	coarse level environment classification	YES (aquatic; anthropogenic; cropland; desert; grassland; forest; mangrove; shrubland; tundra; wetland; woodland)	
ecosystem_classification	ecosystem classification based on ENVO hierarchy	YES (all ENVO terms)	Best possible specification of sampled ecosystem from the hierarchical classification at http://www.ontobee.org/ontology/catalog/ENVO?iri=http://purl.obolibrary.org/obo/ENVO_01001110 and http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO_00000446 that is applicable to the sample
dominant_plant_species	scientific name(s) of dominant plant species in the sampled location or species from which the sample was collected	NO	For "soil" or "litter" samples, dominant primary producer(s) at the sampling site, for "root", "shoot", or "rhizosphere soil", plant species sampled, for "fungal sporocarp", "lichen", "moss", "algae" or "coral", the organism(s) sampled
other_plant_species	additional information on vegetation of the sample or sampled location	NO	Any relevant information on vegetation, such as list of observed taxa or taxa that contribute to root or litter pool
manipulated	information if the sample was experimentally manipulated or no	YES (YES; NO)	
experimental_manipulation_type	the nature of manipulation applied to the sampled location	YES (temperature; precipitation; nitrogen; temperature x precipitation; temperature x nitrogen; precipitation x nitrogen; temperature x precipitation x nitrogen)	E.g., increase, decrease, exclusion, addition
experimental_manipulation_direction	the direction of manipulation	NO	E.g., natural, synthetic
experimental_manipulation_vegetation	vegetation of manipulated site	NO	
experimental_manipulation_duration	duration of manipulation before sampling	NO	
experimental_manipulation_frequency	frequency of manipulation applied	NO	
experimental_manipulation_application_detail	detail on how experimental manipulation was achieved	NO	
experimental_manipulation_intensity	detail on how large the manipulation was	NO	
pH	pH in sample material	NO	Estimated by the authors of the paper (not retrieved from external sources)
pH_method	method used to measure pH	YES ("H2O"; "KOH"; "KCl"; "CaCl2"; "other extract")	
organic_matter_content	organic matter content in sample (%)	NO	
organic_C_content	organic C content in sample (%)	NO	Estimated by the authors of the paper (not retrieved from external sources)
total_N_content	N content in sample (%)	NO	Estimated by the authors of the paper (not retrieved from external sources)
total_Ca	Ca content in sample (ppm)	NO	Estimated by the authors of the paper (not retrieved from external sources)
total_P	P content in sample (ppm)	NO	Estimated by the authors of the paper (not retrieved from external sources)
total_K	K content in sample (ppm)	NO	Estimated by the authors of the paper (not retrieved from external sources)
MAT_study	mean annual temperature of the sample location (°C)	NO	
MAP_study	mean annual precipitation of the sample location (mm)	NO	
area_GPS	precision of GPS localization (m2)	NO	If GPS location is not exact (e.g., multiple forest stands sampled with the same GPS location information), estimate of the area in square meters across which the sample may be localized; area covered by imprecise GPS value (52.5N, 12.3E))
area_sampled	area covered by sampling (m2)	NO	Area of sampled location, in square meters, covered by sampling when multiple subsamples were collected and pooled
number_of_subsamples	number of subsamples combined to represent the sample	NO	
sampling_info	additional information specifying sampling design	NO	Additional information specifying sampling design; e.g., for soil cores, core diameter in cm; for roots, sampled length and/or mass; for deadwood, sampled volume.
sample_depth	depth from sample surface or range of depths from sample surface (cm)	NO	For soil samples, depth from the soil surface (excluding litter) up to the sampled depth or depth range sampled, e.g., "5", "5", "20 to 40"
sample_info	additional information on the sample	NO	Additional information available for the sample including additional metadata (such as other chemical, physical or biological properties of the sample/location,).
DNA_extraction_sample_mass	total mass of the sample used for DNA extraction (g)	NO	
DNA_extraction_size	size of sample used for DNA extraction if mass is not available	NO	E.g., number and description of filters for air and water samples or volumes of material
DNA_extraction_method	identification of DNA extraction method plus reference	NO	
barcoding_region	barcoding region targeted by sequencing	YES (ITS1; ITS2; ITSboth)	
PCR_primers	identification of primers used for PCR amplification of the barcoding region	NO	
PCR_primers_sequence	sequence of forward and reverse primer without barcodes	NO	
sequencing_platform	sequencing platform used for HTS data generation	YES (454Roche; DNBSQ-G400; Illumina; IonTorrent; Oxford Nanopore; PacBio; SOLID)	
ITS1_extracted	number of ITS1 sequences extracted from sequencing data	NO	
ITS2_extracted	number of ITS2 sequences extracted from sequencing data	NO	
ITS_total	sum of ITS1 and ITS2 extracted sequences	NO	
date_added	date of addition to the GlobalFungi database	NO	
submitted_by	identification of the submitting person	NO	

"NA" in the table indicates that data are not available