

Data management and publishing on PlutoF platform

Kessy Abarenkov
University of Tartu
Natural History Museum



Outline of my talk

- Data lifecycle with PlutoF
- UNITE Species Hypotheses workflow & SH matching analysis
- Publishing [DNA-derived occurrence data] in PlutoF

UNITE Community



- UNITE (<https://unite.ut.ee>)
 - Theoretical and practical platform for **calculating, identifying** and **communicating** DNA based species hypotheses that may not have been described as formal species yet
 - Provides **reference datasets** for various analytical pipelines
- PlutoF (<https://plutof.ut.ee>)
 - Data management platform
 - Third-party annotations
 - Analytical services developed by the UNITE community

What is PlutoF?

PlutoF

- An online workbench and computing service provider for biology and related disciplines
- Used for describing biodiversity as well as storing and working with related data (molecular systematics, taxonomy and ecology)
- Features an analysis module by providing analytical services for molecular sequence identification and species discovery from eDNA samples

Data lifecycle with PlutoF

DMP

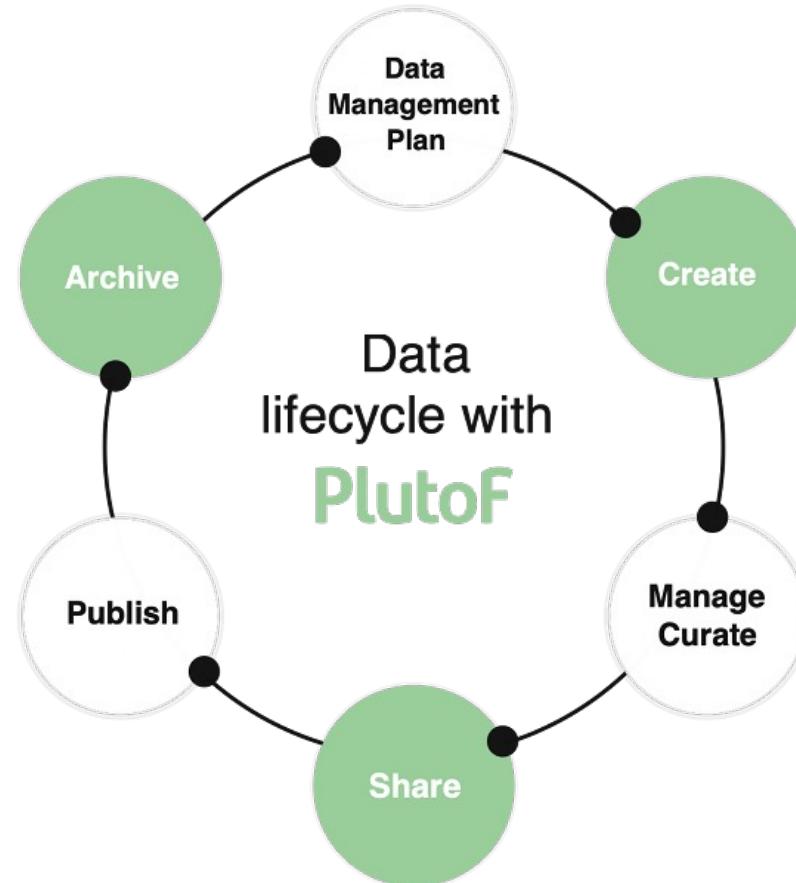
Data Management Plan describes how data are collected, managed and stored during the research project. Compiling DMP with PlutoF is simple because it is a logical part of your project description and associated datasets.

Archive

PlutoF serves as an archive as well because you may store your datasets here for unlimited time. Datasets can be stored as such or they can be published with DOI. Dataset without DOI can be managed and updated through time.

Publish

There are several ways you may publish your data via PlutoF. You may export data in different formats and publish as files. You may ask DOI for your dataset or publish data in GBIF. Then DOI is issued by GBIF. You can also send your data directly to the Pensoft data journal.



Create

Datasets are created in PlutoF either by importing data with template csv files or by uploading data manually via online workbench. Mobile devices can be used for the fast and simple data upload via applications PlutoF GO and Legulus (<https://legulus.tools/#/>)

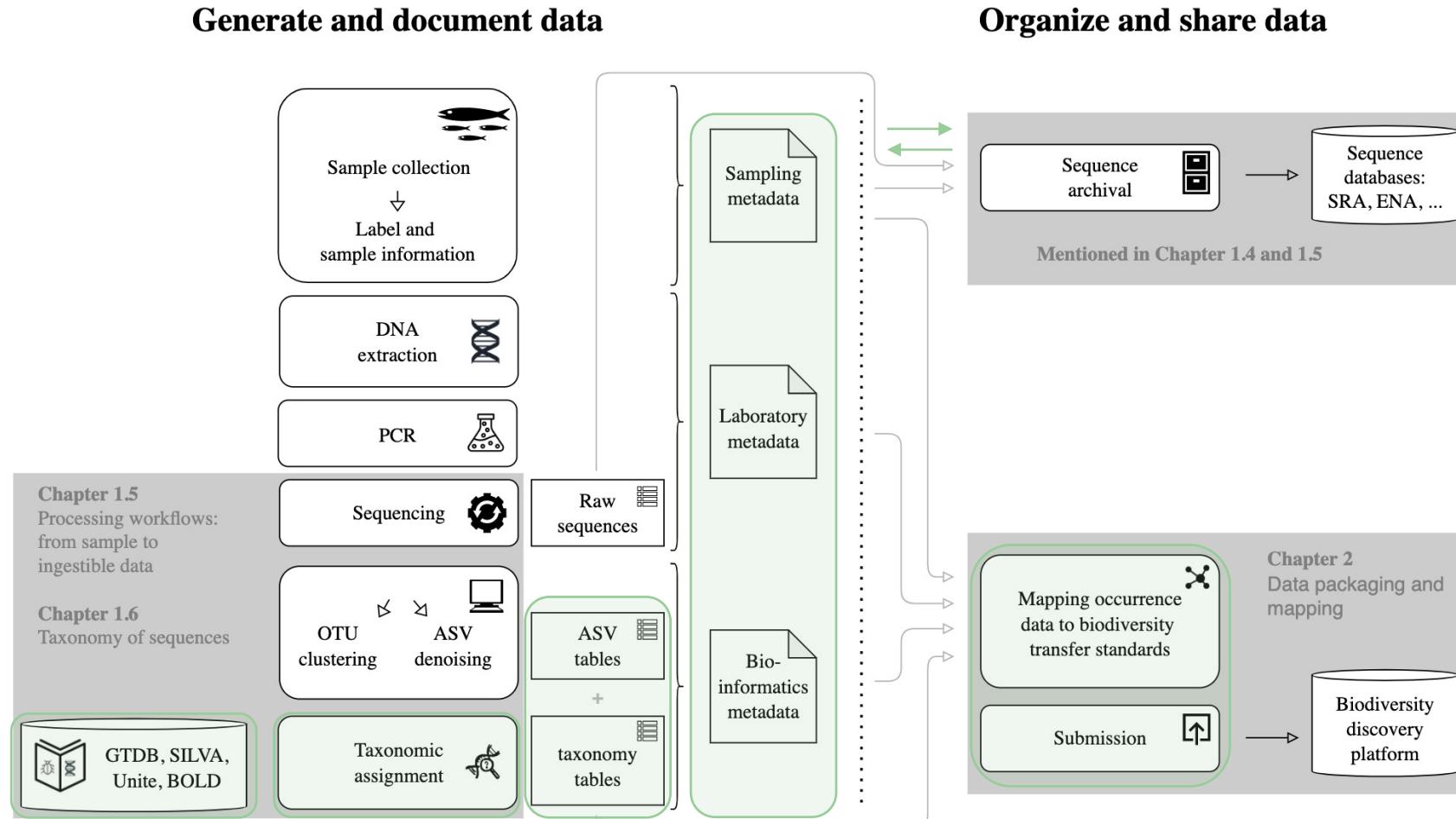
Manage / Curate

Data management and curation is carried out on online workbench.

Share

In PlutoF you can share your datasets with any other user or workgroup. You can give simple access to your data or even right to edit your data. In this way you may collaborate with other scientists, students and citizen scientists.

Guidelines for “Publishing DNA-derived data through biodiversity data platforms”



Technical acronyms:

PCR: Polymerase chain reaction
OTU: Operational taxonomic unit
ASV: Amplicon sequence variant

Databases:

GTDB: <https://gtdb.ecogenomic.org>
SILVA: <https://www.arb-silva.de>
BOLD: <http://www.boldsystems.org>

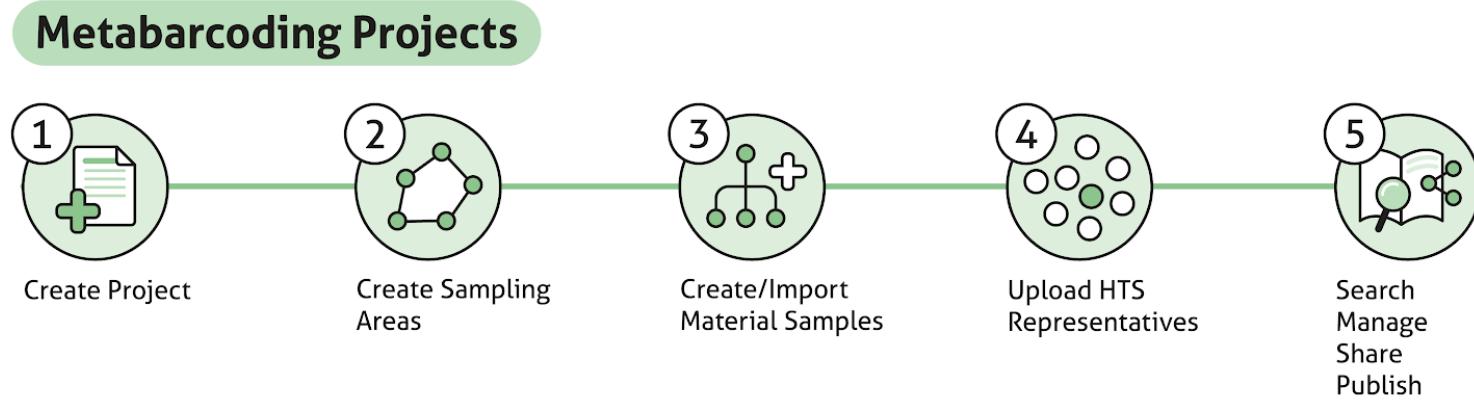
Unite: <https://unite.ut.ee>

SRA: <https://www.ncbi.nlm.nih.gov/sra>
ENA: <https://www.ebi.ac.uk/ena>

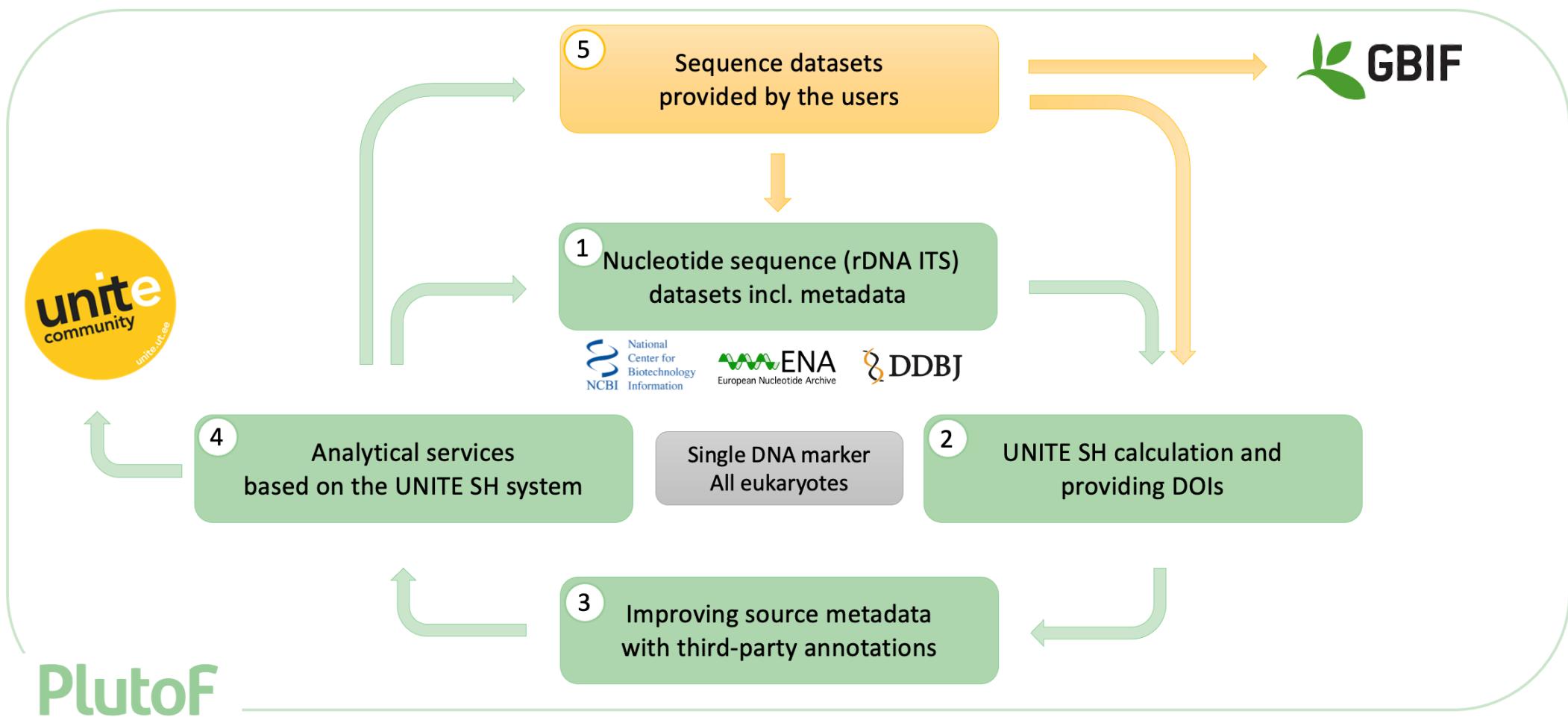
“Figure 1. Overall workflow for DNA sequence-derived biodiversity data” from <https://doi.org/10.3897/mbmg.6.84960>

Metabarcoding / eDNA projects

Metabarcoding is using high-throughput DNA sequencing and DNA based identification for the biodiversity assessment in biological samples like soil, air, water, tissue of organism, dust, etc.



UNITE Species Hypotheses workflow



UNITE SH matching



- Digital service for the global species discovery from eDNA -
 - Distribution patterns, host range, etc.
 - Potentially undescribed new species?
 - Are these SHs found in other studies?
 - Alien or threatened species?
- Output: DOI based stable identifiers for communicating SHs found in your eDNA dataset

UNITE SH matching



source_1

```
>cb583cef8056196748165bf17bba1ae3d5d95a21;size=1024301;
TGAATCATCGAGTTTTGAAACGCAATTGCGCCCGAGGCCCTCGTCGAGGGCACGCTCTGTCTGGGCCTCGCATCTACATACACGCCTCCCTGCAAGGCA
TGGGGGAGGAGCAGGGAGATGGTCGTCGCTCCGTTCCAACCGTGGTCGCGCTGCTGAAATACATGCCATGTTCCGTGTCCTGTGTCGGCGAGCGGTGATATTGTCC
CCTGGTGGGGCGAGTCGGCGTTGGTATGCAAGGGTGTGTCAGCATTGCGGAACCTTGCTTGGCCATCGACTTCCATCTGGTGGAGGTCTTGGCT
CCAAT
>853cafa9df6ff83d9b6635a200d606afc3dc8da6;size=972517;
ATTTGGACCAAAGACCTCACCACAGATGAAAGTCGATGGCAAAGCAATTCCCAGGAAATGCTGCACACACCCCTGCATCAACCAACGCCACTGCCCA
ACAAGGGGACAATATCACCGCTCGCCGACAAGACACCGAACATGCCATGTATTTCAGCCGACCGAACACGTTGGGACGGACGACCATCTCCGCTCT
CCCCCATGCTTCAGGGAGCGCTGTATGAGGATGCAACGCCCAGACAGACGTGCCCTGACCGAGGGCTGGCGCAAATTGCGTTAAAAACTCGAT
GATTCA
>fa4b206430dbcc24e9dca3fd75c7071537df;size=792768;
AGTTGAGGGCGCGCAGTAAACCGCACTTCAGGGTCAAAATGGGCAAACCCGAAACAAACGATGGCATGCACGACGTGACGCCAGGGTTAGCAACCACAC
TTGTCGTATGTCCTGATTGGGACCGGTTAGGGCGACGGATGCGAACGGGGGCCAATATCCGCCACGCTCCGACGTGCCGTTGG
GGCACGGGGAGTGGGTGACGCAATGCGTAGCAGCCAGGACAGCTGCTTCAACCTAATGGCTTCAGGCCAACTTGCGTTCAAAGACTCGATGGTTCA
>1b7a5e60d9482db8004ea85989e3f52a15f26ab;size=769364;
TGAACCATCGAGTTGGCCTGAAGGCAATTAGGTTGAAAGGACCGTCTGCTGGCGTCACGCATTGCGTCACCCACTCCCCCGTGCC
CCAAGCGGGCACGTCGGAGCGTGGCGGATATTGGCCCCCGTTCGCATCCGTGCGCGTGGCTAAAAACGGGCTCCAAATGACGGACATCAGACAA
GTGGTGGTTGCTAACCGTCGGTCACGTCGTGCATGCCATGTTGTCGGGTTGGCCATTGACCTGAAGTGCCTTAACTGCGGCCCTCAACT
>b91cc92d30c9c8de970ade81a77fe48e202b883c;size=433734;
TGAATCATCGAGTTTTGAAACGCAATTGCGCCCGAGGCCCTCGTCGAGGGCACGCTCTGTCTGGGCCTCGCATCCATTCAAACGCCTCCCTGCAATGTG
CTAGGGAGCAGGGAGCTGGTCGTCGCCCCGTGCGGTGCGGTTGGCTGAAATACCTCAAGCGATGTTGCGGGCGCTGGCGAGCGGTGATCTTG
TCCCCCTGGATGGGAGTCGGCGTTAGCCGATGCCGCTTGTGTCGACCTGCTGAACCTGCTCTCTTGTGCTCTGGCCATTGGGTAGGGCGGATT
TAGCTCAAACCTT
>5249c65dc15817780ec057f3a6da39261f9c52c4;size=398454;
AAAGTTGGAGCTAAATCGCCCTACCCATGGGAGGACAAGAGAGAGACAAGCAAGTTGAGCGATGCCACACAAAGCCGCTACGGCTAACGCCACTGC
CCATCCAAGGGGACAAGATCACCCTCGCCGACGCCGCGCACGAAACATCGCTTGGTATTTCAGCCAACCGCACGGGACGGACGACAGCTC
CGCTGCTCCCTAGCACATTGCAAGGGAGCGCTTGAATGGATGCGACGCCAGACAGACGTGCCCTGACCGAGGGCTGGCGCAAATTGCGTTAAAA
ACTCGATGATTCA
>a4ed78ff3ea3fcfd848898d95816c14f60b38cb28;size=252393;
AGTCGAGGTACCGAGTACCTGACGGTACTCGGGGTCGATGAGCCGCAACGACAAGACAGCATGCACGACGCCGTCGGACGGTCTGGCAACCACAC
```

UNITE SH matching tool



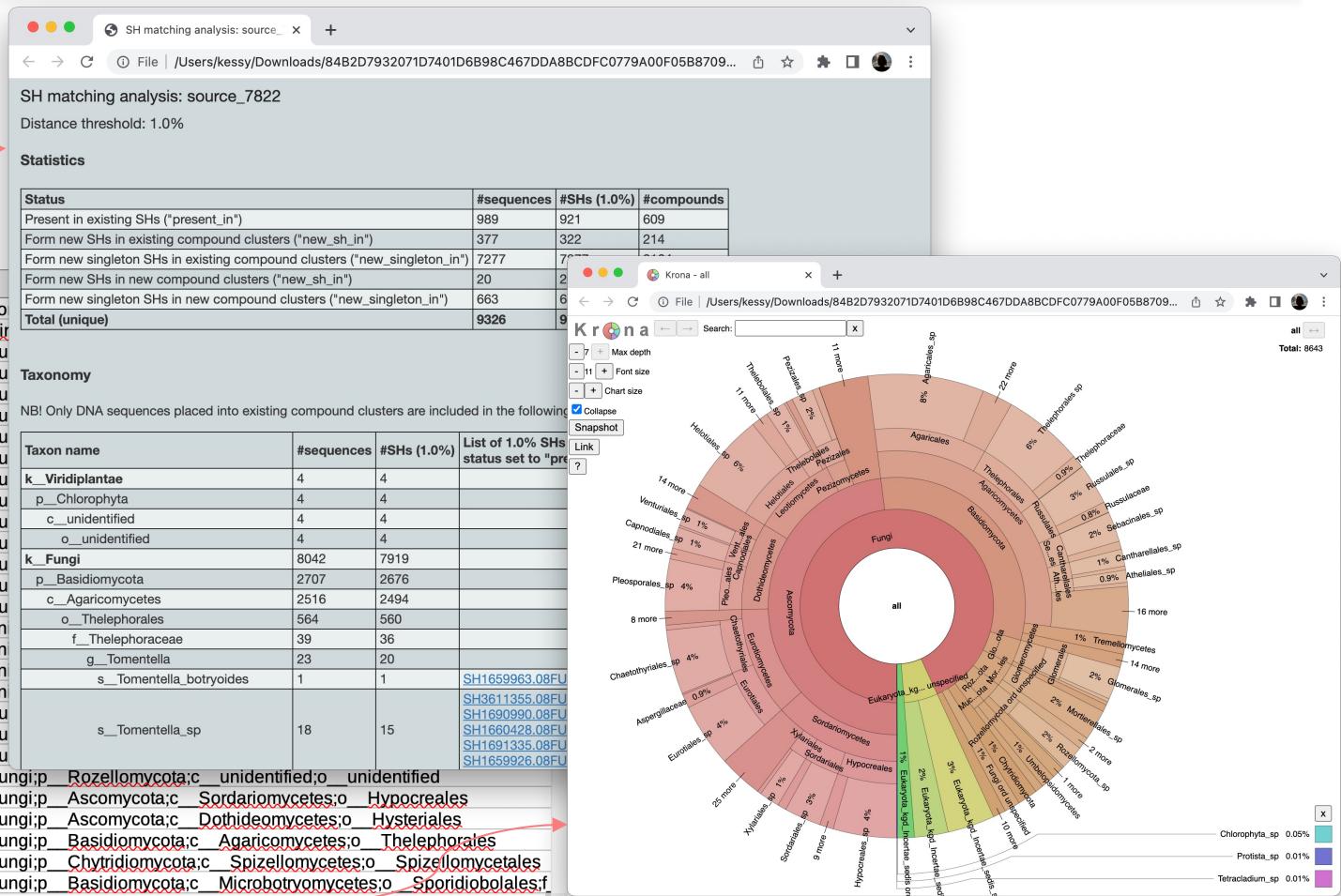
Screenshot of the GitHub release page for **SH MATCHING v1.0.0**. The page shows the release was made by **kessya** 8 days ago, version **v1.0.0**, commit **b811b60**. It includes a note: "First public release of SH MATCHING". Under the "Assets" section, there are links to download **Source code (zip)** and **Source code (tar.gz)**.

Screenshot of the **PlutoF** software interface. The left sidebar shows a navigation tree with categories like Data Management Plans, TAXON OCCURRENCES, Traits and Measurements, LABORATORIES, Analysis Lab (which is selected), Molecular Lab, GIS Lab, Nature Conservation Lab, File Repository, Persons, Organizations, Clipboard & Export, Import, Search, and Annotations. The main panel is titled "New Sequence Analysis" under "General Data". It shows the "Analysis type" set to "SH Matching (DEVELOPMENT VERSION)" and the "Name" field set to "Test analysis". In the "Sequences" section, the "Choose input file in FASTA format" field contains "GSMc_10.fasta". The "Program Parameters" section has "Sequence region" set to "Full ITS". At the bottom are "Start" and "Cancel" buttons.

UNITE SH matching tool



A	B	C	D
1 seq_id tmp	seq_acccn0	status (3.0)	SH code (3.0)
2 7822_2	000dff30d494c3ff0aa877b	new_singleton_in	s1
3 7822_1	000754072de54a4c4c9b8e	present_in	SH1140480.08FU
4 7822_3	0015131d0b361c04ab94	present_in	SH2617547.08FU
5 7822_4	001bbf67397ff073fb7ede6	present_in	SH1155811.08FU
6 7822_5	0022125aab8979027ca80	new_singleton_in	s2
7 7822_8	0036d4f7497485e9dd88a	new_singleton_in	s3
8 7822_10	004491c60c6ef9709c69d	present_in	SH1140446.08FU
9 7822_12	004fed7101687bae98906	new_singleton_in	s4
10 7822_13	00562c1e044d1bb14b53	new_singleton_in	s5
11 7822_14	005c40926c22d1bc867fd	new_singleton_in	s6
12 7822_15	0062ea4d092cebd38e	present_in	SH1180606.08FU
13 7822_16	0069799442170822fdf3561	new_singleton_in	s7
14 7822_17	007014c55a0c62f198ddf	present_in	SH1160354.08FU
15 7822_19	007e7616dec0edc6ba8f6	present_in	SH1140471.08FU
16 7822_21	008ce9f41386319260199	present_in	SH3219083.08FU
17 7822_20	0084f170378293a9830b	new_singleton_in	s8
18 7822_22	009373bb56464d6861120	new_singleton_in	s9
19 7822_23	009a83f0a73c7cc64fc537	new_singleton_in	s10
20 7822_24	00a079d43547c326f83f95	new_singleton_in	s11
21 7822_25	00a660e8f07f08e41ef29d	new_singleton_in	s12
22 7822_26	00ac234f163c5f98881846	new_singleton_in	s13
23 7822_27	00b2be2e3b86cb770bbc7	new_singleton_in	s14
24 7822_28	00b96fab028162a5165e	new_sh_in	1
25 7822_30	00c7a5b296905afab9cf2	new_singleton_in	s15
26 7822_32	00d38bc9a8e56e50e176e	new_sh_in	2
27 7822_33	00d91e0f4c5d6b01a7550	new_singleton_in	s16
28 7822_34	00df349d23d87f0baff74	present_in	SH3565597.08FU



UNITE SH matching tool



<https://unite.ut.ee/sh/SH1502197.08FU>
[doi:10.15156/BIO/SH1502197.08FU](https://doi.org/10.15156/BIO/SH1502197.08FU)

A	B	C	D
seq_id_tmp	seq_accno	status (3.0)	SH code (3.0)
1	000dff30d494c3ff0aa877b	new_singleton_in	s1
2	7822_2	000dff30d494c3ff0aa877b	present_in
3	7822_1	000754072de5a4c4c9b8e	present_in
4	7822_3	0015f131d0b361c04ab94	present_in
5	7822_4	001bbf67397ff073fb7ede6	present_in
6	7822_5	0022125aab8979027ca80	new_singleton_in
7	7822_8	0036d4f7497485e9dd88a	new_singleton_in
8	7822_10	004491c60c6ef9709c69d	present_in
9	7822_12	004fed7101687bae98906	new_singleton_in
10	7822_13	00562c1e044d1bb14b53	new_singleton_in
11	7822_14	005c40926c22d1bc867fd	new_singleton_in
12	7822_15	0062ea4d0992cebd38e	present_in
13	7822_16	006979442170822fdf3561	new_singleton_in
14	7822_17	007014c55a0c621198ddf	present_in
15	7822_19	007e7616dec0edc6ba8f6	present_in
16	7822_21	008ce9f41386319260199	present_in
17	7822_20	0084f170378293a9830b2	new_singleton_in
18	7822_22	009373bb56464d6861120	new_singleton_in
19	7822_23	009a83f0a73c7cc64cf537	new_singleton_in
20	7822_24	00a079d43547c326f83f95	new_singleton_in
21	7822_25	00a660e80f0708e41ef29d	new_singleton_in
22	7822_26	00ac234f163c5f98881846	new_singleton_in
23	7822_27	00b2be2e3b86cb770bbc7	new_singleton_in
24	7822_28	00b96fabb028162a5165e	new_sh_in
25	7822_30	00c7a5b296905afabb9cf2	new_singleton_in
26	7822_32	00d38bc9a8e56e50e176	new_sh_in
27	7822_33	00d91e0f4c5d6b01a7550	new_singleton_in
28	7822_34	00df349d23d87f0baffb74	present_in

UNITE - Species Hypothesis | unite.ut.ee/sh/SH1502197.08FU#fnfdn-panel3

[Run Analysis](#) [Search Pages](#) [Resources](#) [Governance](#) [Statistics](#) [How to cite?](#) [Notes and News](#)

Tomentella fuscocinerea (Pers.) Donk (DOI: TH034569) | SH1502197.08FU (DOI: SH1502197.08FU)

Distance to the closest SH: 1.5
No. of sequences in SH: 435

Placement in the fungal classification
Fungi; Dikarya; Basidiomycota; Agaricomycotina; Agaricomycetes; Agaricomycetes
sbc Incertae sedis; Thelephorales; Thelephoraceae; Tomentella
Index Fungorum: [#439164]

Reference sequence: UDB016492
Chosen by: Urmas Kõljalg
Date: 2019-02-13 15:59

[Statistics](#) [Taxonomy](#) [Ecology](#)

Older version(s) of this SH is/are available

SH code (Count*/Total count**)
SH189360.07FU (32/435); SH189723.07FU (2/435); SH184521.07FU (27/435);

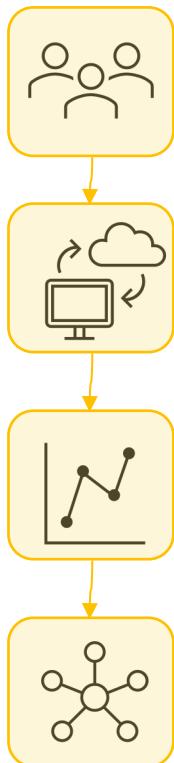
*Number of sequences carried over from previous version
**Total number of sequences composing this SH in current version

Distribution map

*Locations without exact coordinates are displayed as spherical country centroids

Accession number	UNITE taxon name	INSD taxon name	Sequence source	Interacting taxa	Sampling area
**UDB0251842	Thelephoraceae	sample			Latvia
GU220689	Thelephoraceae	Thelephoraceae (uncultured Tomentell...	Orchid mycorrhiza sample	Corallorrhiza striata var.	United States

Publishing your results



Open Access Article

Fungi Recorded on Folivorous Lepidoptera: High Diversity Despite Moderate Prevalence

by Robin Gielen 1,* , Hendrik Meister 1 , Toomas Tammaru 1 and Kadri Pöldmaa 2

¹ Entomology Unit, Department of Zoology, Institute of Ecology and Earth Sciences, Faculty of Science and Technology, University of Tartu, Vanemuise 46, Tartu 51014, Estonia

² Mycology Unit, Department of Botany, Institute of Ecology and Earth Sciences, Faculty of Science and Technology, University of Tartu, Ravila 14a,

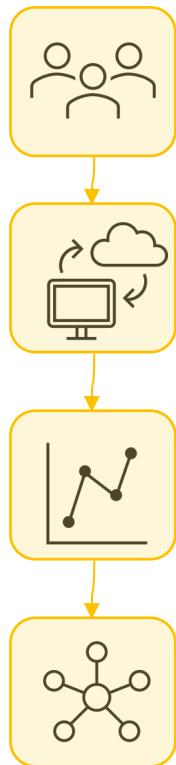
* Author to whom correspondence should be addressed.

J. Fungi 2021, 7(1), 25; <https://doi.org/10.3390/jfungi7010025>

Table 1. Species of fungi isolated from lepidopteran hosts (the field experiment data in brackets). UNITE species hypothesis (SH) codes are presented to facilitate communication on detected fungi [22].

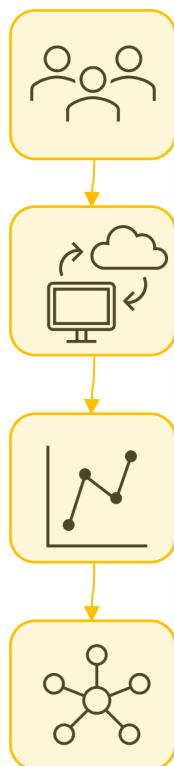
Order/Family	Species	SH DOI *	Hosts **
Hypocreales/Cordycipitaceae	<i>Akanthomyces muscarius</i> (Petch) Spatafora, Kepler & B. Shrestha	SH1886969.08FU	EA 8 CP 4(+25) HP 1
	<i>Cordyceps bifusispora</i> O.E. Erikss.	SH1887323.08FU	CP (2) EA 1
	<i>Cordyceps farinosa</i> (Holmsk.) Kepler, B. Shrestha & Spatafora	SH1524463.08FU	EA 29 CP 6(+10) HP 7 UP 1
	<i>Lecanicillium praecognitum</i> Gorczak & Kislo	SH1524455.08FU	CC 1 EA 4
	<i>Simplicillium aogashimaense</i> Nonaka, Kaifuchi & Masuma	SH1988378.08FU	EA 1

Publishing your results - why publish?



- Share your data with biodiversity researchers, taxonomists, ecologists, etc. (distribution patterns, host ranges, 'dark taxa' fungi, biological conservation)
 - Increase citations to your data (and work) - favoured by funding bodies
 - May invite unexpected collaborations!
 - [Help improving data standards]
- Guidelines for publishing DNA-derived occ. data - <https://doi.org/10.3897/mbmg.6.84960>

Publishing your results - GBIF



GBIF | Global Biodiversity Information Facility

Free and open access to biodiversity data

OCCURRENCES SPECIES DATASETS PUBLISHERS RESOURCES

Search

What is GBIF? About GBIF Estonia

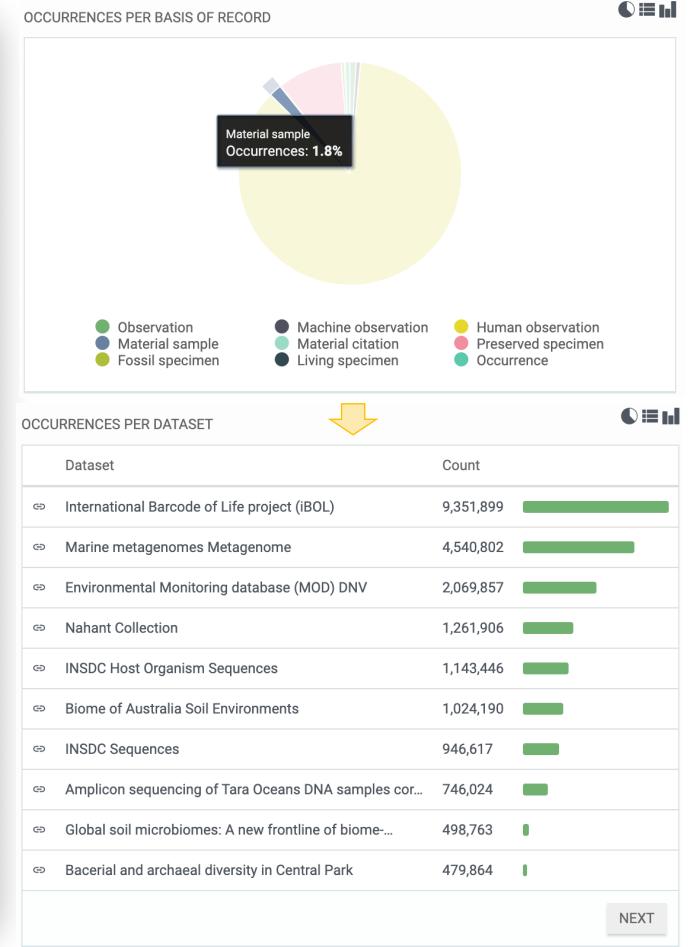
Vanellus lugubris (Lesson, 1826) observed in Garwula, Liberia by mlanguy (CC BY-NC 4.0)

2,208,162,198 Occurrence records

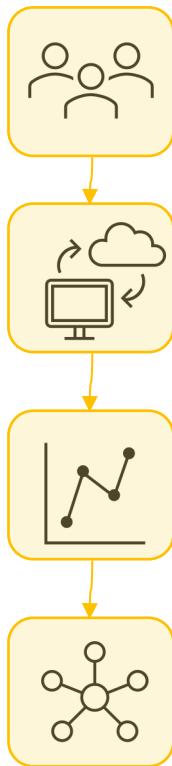
75,354 Datasets

1,874 Publishing institutions

7,513 Peer-reviewed papers using data



Publishing your results



PlutoF

Taxon Occurrence My Menu Settings kess Log out Est

Sequence Analysis Generate report Delete Bookmark Info Back

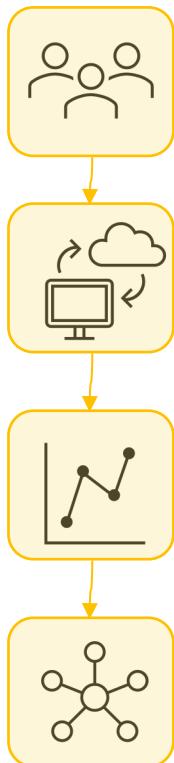
General Data

Analysis type	Name	Status	Estimated finish time
massBLASter SH matching (BLAST+ 2.13.0)	NEFOM SH matching BLAST	Ready	2021-04-22 16:41

Sequences

Accession number	UNITE SH	SH taxon name	Matched sequence	E-Value	Blast score	Percentage	Distance	Identification added
AF122035	SH1516056.08FU	Deconica hartii	KJ832038 (Deconica)	8.59084e-179	629.908	98.19	1.5%	
DQ900978	SH1513434.08FU	Agaricales	HM240521 (Coprinus cordisporus)	0	1392.17	98.35	1.5%	
FJ179475	SH1656359.08FU	Agaricales	FJ179475 (Agaricales)	0	2319.25	100.00	1.5%	
FJ475693	SH1543992.08FU	Basidiomycota	FJ475748 (Basidiomycota)	1.74077e-39	168.087	100.00	1.5%	
FJ475703	SH1577280.08FU	Serendipitaceae	HQ211818 (Serendipita)	0	1156.97	98.85	1.5%	
FJ644697	SH1647242.08FU	Agaricales	FJ644697 (Agaricales)	0	1320.06	100.00	1.5%	
FJ948136	SH1503585.08FU	Leucoagaricus gongylophorus	KF572013 (Leucoagaricus gongylophorus)	0	1124.35	98.81	1.5%	
FJ948137	SH1634762.08FU	Agaricaceae	FJ948145 (Agaricales)	0	1181	100.00	1.5%	
FJ948138	SH1514346.08FU	Agaricales	FJ948138 (Agaricales)	0	1181	100.00	1.5%	
FJ948139	SH1503589.08FU	Agaricaceae	FJ948139 (Agaricales)	0	1249.68	100.00	1.5%	

Publishing your results



PlutoF

UNITE Species Hypotheses

SH code	Taxon name	Threshold	No. of seqs	RefSeq chosen by
SH0008827.09FU	Basidiomycota	3.0 %	152	
SH0271063.09FU	Basidiomycota	2.5 %	140	
SH0564220.09FU	Basidiomycota	2.0 %	140	
SH0916936.09FU	Basidiomycota	1.5 %	139	
SH1434284.09FU	Basidiomycota	1.0 %	130	
SH2433236.09FU	Basidiomycota	0.5 %	28	

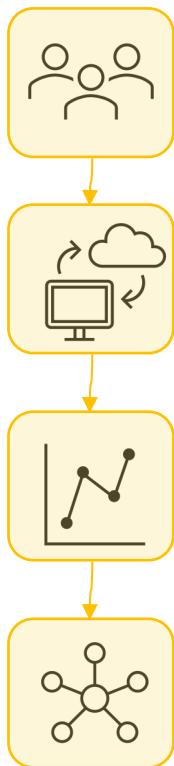
SH identifications ①

UNITE SH SH1655909.08FU	Best sequence match KY322626	Distance (%) 1.5	Identified at 2019-04-28	Identifiers Kai Ilves
E-Value 0	Blast score 861.677	Percentage 99.61		

Identifications ①

Latest Taxon name Fungi R.T. Moore, 1980 (kgd)	Taxon name add.	Date identified	Typification
Assessment	Remarks	Identifiers Kai Ilves	Created 2019-04-26 16:26 (Kai Ilves)

Publishing your results



PlutoF

UNITE Species Hypotheses

SH code	Taxon name
SH0008827.09FU	Basidiomycota
SH0271063.09FU	Basidiomycota
SH0564220.09FU	Basidiomycota
SH0916936.09FU	Basidiomycota
SH1434284.09FU	Basidiomycota
SH2433236.09FU	Basidiomycota

SH identifications 1

UNITE SH	Best sequence match	Distance
SH1655909.08FU	KY322626	1.5

E-Value	Blast score	Percent
0	861.677	99.61

Identifications 1

Latest	Taxon name	Taxon name add.
	Fungi R.T. Moore, 1980 (kgd)	

Assessment

Remarks

unite
community
a non-profit association

Basidiomycota R.T. Moore | SH0916936.09FU

Distance to the closest SH: 1.5
No. of sequences in SH: 139

Placement in the fungal classification
Fungi; Dikarya

Representative sequence:
[UDB02106869|51f6df8f260ba069a2c30765a38d6f628112cba8](#)

Chosen by: automatically by the program
Date: 2013-11-19

Statistics **Taxonomy** **Ecology**

Identifications

Fungi sbk Incertae sedis (84); unspecified (30); Basidiomycota (23); Fungi (2);

Older version(s) of this SH is/are available

SH code (Count/Total count**)
[SH1655909.08FU](#) (56/139);

*Number of sequences carried over from previous version
**Total number of sequences composing this SH in current version

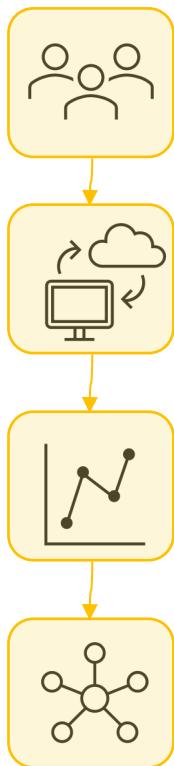
Distribution map

*Locations without exact coordinates are displayed as spherical country centroids

Accession number	UNITE taxon name	INSD taxon name	Sequence source	Interacting taxa	Sampling area	% Alignment
UDB02106869	Fungi sbk Incertae sedis		Soil fungal DNA		Denmark	0.5
UDB02106872	Fungi sbk Incertae sedis		Soil fungal DNA		Denmark	1.0
UDB0666810					Estonia	1.5
UDB0196696(2)					Estonia	2.0
UDB02091851	Fungi sbk Incertae sedis		Soil fungal DNA		United Kingdom	2.5
UDB02104195	Fungi sbk Incertae sedis		Soil fungal DNA		Sweden	3.0
UDB02159519(3)	Fungi sbk Incertae sedis		Soil fungal DNA		Russian Federation	3.5

A sequence alignment visualization showing a green heatmap of aligned bases across multiple sequences. The x-axis represents sequence positions, and the y-axis represents different samples. A color scale at the bottom indicates the percentage of alignment, ranging from 0.5% (light green) to 3.5% (dark green).

Publishing your results



PlutoF

Taxon Occurrence My Menu Settings kess Log out Est

Data Management Plans
Projects
TAXON OCCURRENCES
Traits and Measurements
LABORATORIES
Collection Lab
Publishing Lab
Taxonomy Lab
Analysis Lab
Molecular Lab
GIS Lab
Nature Conservation Lab
File Repository
Persons
Organizations
Clipboard & Export
Import
Search
Annotations

Publishing Laboratory

GBIF Datasets
Manage and publish Global Biodiversity Information Facility datasets

GBIF Datasets

DataCite DOIs
Create and register DataCite Digital Object Identifiers (DOIs) to datasets and other research objects.

DataCite DOIs

References
View, search and manage references.

References

Reference Series
View and create reference series.

Reference Series

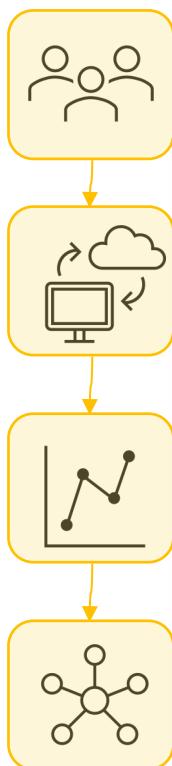
Keywords
View, create and link keywords as thesauri.

Keywords

Keyword Collections
Browse keyword collections and create new ones to catalogue your keywords.

Keyword Collections

Publishing your results



The screenshot shows the PlutoF publishing interface. On the left, a sidebar lists various lab sections like Collection Lab, Publishing Lab, and Analysis Lab. The main area displays 'General Data' for a GBIF Dataset. It includes fields for Title, Record type (Sequence), License (Creative Commons BY 4.0), and a logo. Below this is a 'Taxonomic coverage' section. At the bottom, there's a 'Records' table with two entries:

Sequence	Taxon	Project	District	Commune	Locality text
aSeq115904 UDB0746178	Fungi	Tartu Botanical Garden Endophyte Project	Tartu maakond	Tartu linn	Tartu linn
aSeq127190 UDB0746179	Fungi	Tartu Botanical Garden Endophyte Project	Tartu maakond	Tartu linn	Tartu linn

The screenshot shows the GBIF.org occurrence dataset page for 'Tartu Botanical Garden Root and Leaf Endophytes'. It has a green header and navigation bar. The main content area displays 'OCCURRENCE DATASET | REGISTERED MAY 7, 2019' and 'Tartu Botanical Garden Root and Leaf Endophytes'. It shows 5,649 occurrences and 10 citations. Below this is a 'DATASET' tab showing a summary with a green circle for 5,649 occurrences and 100% with coordinates. A map shows 10 georeferenced records. At the bottom, it says 'APPEARS IN 2 OCCURRENCE DATASETS' with links to 'BIOWIDE eDNA Fungi dataset' and 'Tartu Botanical Garden Root and Leaf Endophytes'.

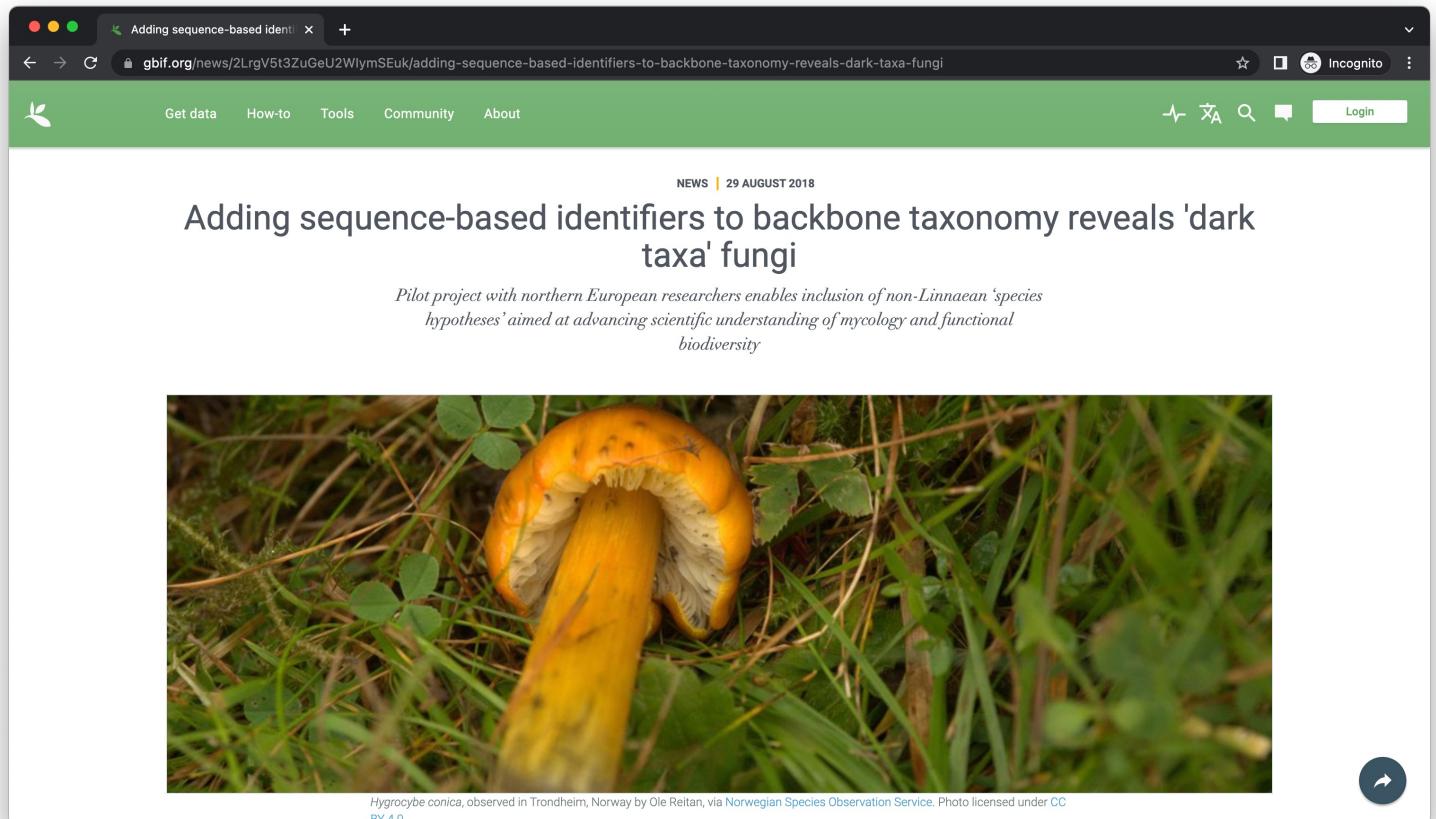
Publishing to GBIF

UNITE reference datasets for -

- QIIME
- mothur
- Crest
- UCHIME/USEARCH/UTAX
- General FASTA

LinkOut from -

- GBIF
- INSD databases (ENA, NCBI)
- GlobalFungi
- etc.



<https://www.gbif.org/news/2LrgV5t3ZuGeU2WlymSEuk/adding-sequence-based-identifiers-to-backbone-taxonomy-reveals-dark-taxa-fungi>

Thank you!



UNIVERSITY OF TARTU
Natural History Museum
and Botanical Garden



PlutoF