

Deliverable D.T1.2.3

Eco-AlpsWater
Innovative Ecological Assessment and Water Management Strategy
for the Protection of Ecosystem Services in Alpine Lakes and Rivers

Priority 3: Liveable Alpine Space. SO3.2 - Enhance the protection, the conservation and the ecological connectivity of Alpine Space

Project Eco-AlpsWater
Work Package WPT1
Activity A.T
Deliverable D.T1.2.3
Version 1.0
Date October 2021
Coordination: Isabelle Domaizon

Deliverable D.T1.2.3

Database for the classification and distribution of aquatic species accessible by stakeholders

Ute Mischke¹, Nico Salmaso², Rainer Kurmayer³, Tina Elersek⁴, Giulia Riccioni², Hans Rund³, Isabelle Domaizon⁵

¹ Bayerisches Landesamt für Umwelt, Ref. 83, Wielenbach, Germany

² Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

³ University of Innsbruck, Research Department for Limnology, Mondsee, Austria

⁴ National Institute of Biology, Slovenia

⁵ French National Institute for Agriculture, Food and Environment, Thonon les Bains, France



Deliverable D.T1.2.3

1. Introduction

This delivery provides the description of the so-called “EAW taxa analysis tool” which is based on the stakeholder interests listed in guidelines for digital accessing of data obtained during the project by stakeholders and citizens (D.T4.1.1).

Following these guidelines, the Eco-AlpsWater project has created a new aquatic species classification and distribution database accessible to stakeholders (D.T1.2.3. This database contains the site classification by national metrics of all samples or sites and indicator values used for water assessment in the Alpine Space. Additionally, this “EAW taxa analysis tool” allows stakeholders to easily compare traditional and metabarcoding taxa inventories.

The “EAW taxa analysis tool” was successively filled with the project data until the end of the project, and intermediate versions were already in use for regional events.

Stakeholders have raised three fields of interest to the project results:

- I) Degree of direct matches when comparing taxa inventories gained by traditional methods (EU-WFD, WHO-CH) and by metabarcoding approaches
- II) Rating the applicability of metabarcoding approaches in terms of cost, practical handling and processing and in terms of assessing the ecological quality of a water body
- III) Which additive and supporting information the metabarcoding approaches can provide (biodiversity of all freshwater organisms; bio-geographic distribution)

Each field of interest on metabarcoding data need an adopted format of digital access for stakeholders. The following chapters describe the database solution developed to support field of interest I and partly II.

When checking the matching level of taxa, users want to check the proportion of taxa matches by comparing taxa list for each sample or for each water body (level “site name”). These kind of data tables are selectable “on demand” for sampling sites. Using a set of filters, the target organisms of the WFD biocomponents can be extracted from the entire and huge Eco-AlpsWater dataset.

Regarding diversity aspects, all genotypes should be listed, while for comparing taxa inventories by genetic results with those by traditional methods, the target organisms of the WFD biocomponents are of main interest. In any case, the data must be reliable and checked by scientists.

The project data produced are very complex and a large number of samples were analysed.

37 lakes and 53 river sites were assessed using traditional methods (counting lists) and additionally with the Eco-AlpsWater metabarcoding approach.

For each of the applied primers the bioinformatic pipelines automatically produced large HTS (high throughput sequencing) record tables called MOTU-by-sample matrix. The user receives the taxonomy of a sequence if there is a match with a genebank (taxonomic assignment).

Deliverable “Eco-AlpsWater Tool-box for the implementation of innovative monitoring approaches” (D.T3.5.1) describes the bioinformatic analyses in detail in addition to the primer specific protocols.

For database needs, the MOTU-by-sample matrix produced by the metabarcoding approach were converted into simple lists in the EAW taxa analysis tool. Sample track codes connect them to counting results and to environmental data.

In Figure 1 the various kinds of project data are described and listed in groups. In the EAW taxa analysis tool level 1 and level 2 data are stored.



Deliverable D.T1.2.3

Level 1: taxa inventories <i>Prerequisite: harmonisation and data check</i> <ul style="list-style-type: none"> all records of one taxon are named/coded the same Agreed confidence level Contamination from connected habitats? proof for synonyms check by BLAST analysis etc. 	Two origins of taxa lists: a) Traditional method: Standardized national bio-component methods with operational names and codes for taxa per sample restricted on WFD bio-components (diatom, fish etc) a) <u>Metabarcoding</u> : Taxonomic assignment according curated reference data bases differs for primer and target group (12, 16, 18s, <u>rbcl</u>) using actual nomenclature and systematics
Level 2: environmental data	Harmonized data according units and outlier check a) Trophic scores: Nutrients, chlorophyll for pressure-taxa-relationships b) Identification of extreme habitats within EAW sites
Level 3: Geographic information and sample origin	Distribution maps of sites and taxa using the coordinates Categories of sampling type (eg. plankton, biofilm etc.)
Level 4: Quality check and documentation data	Metadata to understand sampling, extraction and bioinformatics Storage of DNA sequences, sharing with international data bases
Level 5: Aggregated data: Index results for biodiversity and assessment a) Diversity indices b) Application of common biological metrics on eDNA data c) Suggesting new metrics based on other biological groups than EU-WFD	

(co) Eco-AlpsWater: WP4 Deliverable -4-1-1 graph 1

Fig. 1 – Demonstration of the complexity of data

2. Short description of the storage data base

A data storage facility managed and hosted by the Edmund Mach Institute at S. Michele all'Adige was used to upload and store all the raw data collected during the project. This facility was intended to be used exclusively by the project partners to upload environmental and HTS data in a common digital and controlled environment. This choice was dictated by the need to use a common input system to avoid the use of different digital platforms and formats. In this way, after downloading the data, a unique database was created that can be used for downstream analyses of environmental data (excel files) and bioinformatic analyses of FASTQ files.

The data storage facility was intended to:

- List all samples with EAW sample codes and bar codes
- Store metadata of samples including environmental data
- Store all raw data (FASTQ) obtained by the Illumina MiSeq sequencing of the 12S rRNA, 16S rRNA, 18S rRNA, and rbcL genes with sample codes (linked to environmental samples)
- Store taxa inventory lists gained by traditional methods (counts and taxa by light microscopy).

The data storage facility, containing raw data that are not annotated and non-quality checked, cannot be used for final data distribution, but only for long-term storage. The raw data have been used in downstream analyses, following these steps:

- Environmental data have been analysed carefully, checking the consistency of measurement units, outliers, and completeness of inputs, i.e. checking the actual absence of raw data;
- FASTQ files (raw reads data) were processed by using specific state-of-the-art bioinformatic pipelines (described in detail in WPT1); the analyses included both, a quality control step, which provided a final list of sequences in FASTA format, and a taxonomic step, which allowed to



Deliverable D.T1.2.3

associate taxonomy names to the DNA sequences analysed; the taxonomic attributions were further checked and analysed by using BLAST/BLASTN queries for selected taxa and groups (e.g. cyanobacteria);

- The final classified sequences, the environmental data, as well as the microalgal and selected fish data obtained using traditional methods have been integrated in the “EAW taxa analysis tool”.

The downstream analyses described in the previous steps allowed also to check for potential incongruences in the raw data, and to obtain final quality checked data. This will allow the huge amount of information to be used for many purposes, including public access and preparation of scientific papers. More specifically, the raw FASTQ sequences will be shared publicly and uploaded into the European Nucleotide Archive (ENA, <https://www.ebi.ac.uk/ena/browser/home>). Besides long-term storage, this will allow global access to the original reads through the International Nucleotide Sequence Database Collaboration (INSDC) (<https://www.insdc.org/>), which include, besides ENA, also NCBI-GenBank and DDBJ.

Full descriptions of the data storage facility is provided in Deliverable D.T1.2.2 Protocol data storage, sharing, retrieval and Deliverable D.T1.2.1 Data storage in cloud storage services.

3. Comparing taxa inventories

The “EAW taxa analysis tool” is a program to execute analysis steps automatically, which are necessary to compare the taxa inventories after their listing preparation in common tables. The tool was also developed in response to user questions identified during the preliminary regional data analysis by the project partners fed into the EAW teleconference in December 2020.

The “EAW taxa analysis tool” is specified Access database with data tables and prepared costumer queries. The pre-defined costumer navigation area (left side in Fig. 2) helps the user to find the proper queries.

With tool version 8, query results for selected site(s) are separated into groups "Cyanobacteria", "diatoms" and "18S_eukayotes" and "MiFish-U fish".

A short user instruction for EAW taxa analysis tool is in **Appendix I** (also available in the ZIP-file of tool).

A full list of all prepared queries with legends are in **Appendix II** ((also available in the ZIP-file of tool).

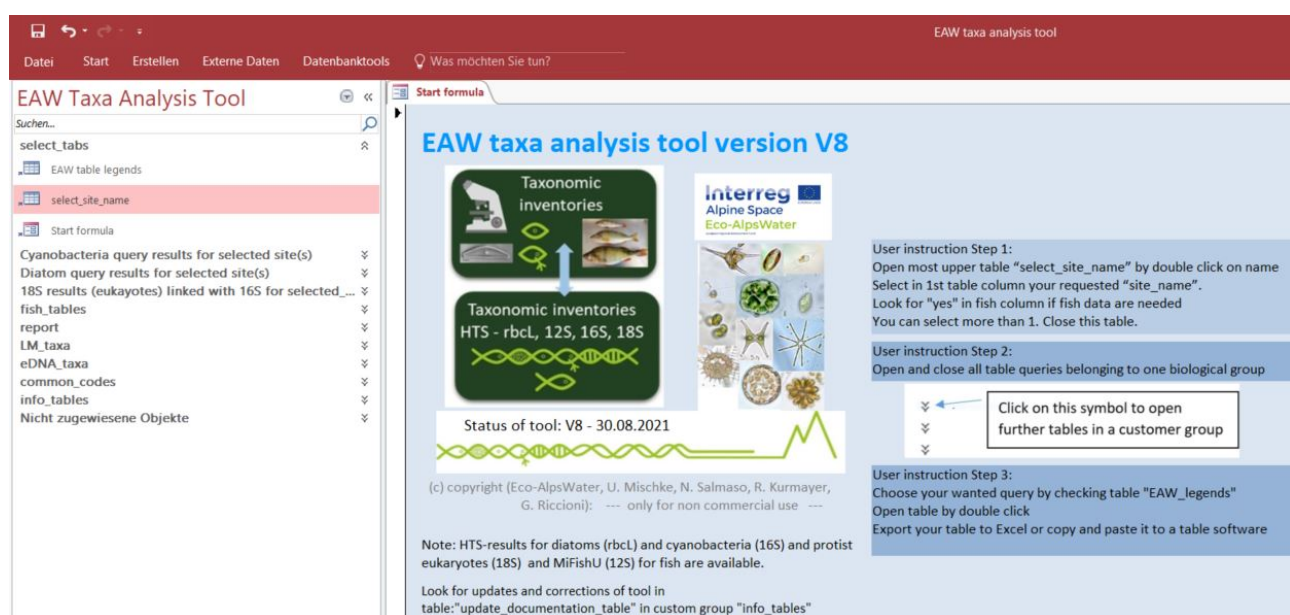


Figure 2: Start view on the Access tool with predefined costumer navigation area at left side.

Deliverable D.T1.2.3

The EAW taxa analysis tool is an output of the project, and linked to the project webpage.
<https://www.alpine-space.eu/projects/eco-alpswater/en/project-results/taxa-analyses-tool>

There are various possibilities to compare and visualise traditional and eDNA results for taxa inventories in the investigated lakes and rivers. To explore the data and analyses, the stakeholder should contact project partners from their countries for demonstration. Version 8 is available and some further functionalities are added until the end of the project (Oct 2021).

In the tool, all results of metabarcoding (HTS) and light microscopy (LM) are listed in tables, which can be linked to each other. Each sample and each taxon of phytoplankton and benthic diatoms had a unique code. Samples are linked to their specific site name and the user select them in a first step. Secondly, prepared queries help to screen specific results (=taxa lists) of all samples belonging to the selected site. Thirdly, the “GAP”-queries compare taxa results by HTS to LM or visa wise.

In the following part we describe functions of the tool and provide background information.

3.1 Tool tables with taxa inventories detected by metabarcoding

The obtained raw data are very complex especially for the metagenomics results and are huge in terms of samples and analysis steps.

When focusing on microorganisms (phytoplankton, diatoms, bacteria and fungi) 153 plankton and 177 biofilm samples from 37 lakes and 53 river sites were assessed with multiplex primers.

In the MiFish-U investigation in total 86 water samples with fish eDNA were analysed.

For each of the primers large HTS (high throughput sequencing) record tables were automatically produced by the bioinformatic pipelines.

These large tables combine all biological EAW samples with all extracted DNA sequences and give the detected signal (rarefied read counts) in the cross field.

11.468 ASV sequences of 18S by Silva 138 (all eukaryotes including phytoplankton)

37.530 ASV sequences of 16S by Protist Ribosomal Reference database-PR2 (all bacteria, fungi)

1.285 ASV sequences of “chloroplast 16S” (some eukaryotes)

1.602 ASV sequences of rbcL selected for diatoms by library data base Diat.barcode

246 ASV sequences of 12S selected for fishes by in-house Fish based on EMBL (vertebrates), with MiFish-U primers (Miya et al., 2015).

3.1.1 The output of one DNA sequence with different taxonomic levels

The HTS record tables provide the taxonomy of a sequence in case there was a match to a genebank (see bioinformatic deliveries D-T1.1.3 1-4). This match can be to different taxonomic levels such as order, family, genus and species. It is notable, that for the more universal primers such as 16S (bacteria, fungi) or 18S (eukaryotes) only a small part of all sequences may belong to one specific genus or species. If the taxa name does not directly match between HTS and traditional method (light microscopy (LM)) 1:1, there can still be a match on the second level (genus level): both methods may have detected the same genus, but different species.

Therefore, the “EAW taxa analysis tool” delivers match tables on genus or on species separately for cyanobacteria and eukaryotes.

This is relevant for water assessment since indicator lists in biological WFD metrics mainly contain taxa on genus or species level (see D.T2.1.1- List of species). Therefore, the national methods for phytoplankton



Deliverable D.T1.2.3

and benthic diatoms require analysis on species or at least genus level. In result, it is impossible to link an ASV to an indicator species found by counting with light microscopy, if its taxonomy match identify only family or order level. This strongly reduce the share of direct matches between the both method approaches.

3.1.2 One species is frequently listed under several DNA sequences in the metabarcoding output – realisation of the intra-specific genetic diversity!

It is an important information for users and stakeholders that one taxon can be detected with several ASVs. Since the DNA sequence of each ASV differs, this diversity presents “genotypes” of a single species. Still, a user of the data just want to know, if a species or genus is present or not. The tool therefore brings all sequences together, which belong to the same taxon and aggregate the result in one “present” record (additional information of interest: maximal signal, number of ASVs and first ASV found in one sample). The counted number of ASVs, which belong to one taxon, is the intra-specific genetic diversity found in the selected site.

3.1.3 Selection of target taxa

With the aim to finally compare the taxa inventories gained by traditional methods (EU-WFD, WTO-CH) and by the metabarcoding approaches the focus is on the biological target groups, the so-called bio-components of the WFD, phytoplankton including cyanobacteria and benthic diatoms.

The user needs support to filter all taxa classes relevant for its target group from all the other organisms groups, which the genetic marker might also detect (e.g. trees, mammalians etc.).

While for the HTS output lists for fishes (12S MiFish-U) and diatoms (rbcl) such a filter was already applied before the delivery to the users, all sequences (ASVs) were delivered for the more unspecific primers 16S or 18S including heterotrophic bacteria, fungi, mammalia, higher plants and ciliates.

Out of all ASVs found with 16S and 18S, only a small part belong to these target groups and their classes. Therefore, a “phytoplankton” filter helps to select phytoplankton taxa within the 18S taxonomy and a “cyanobacteria filter” to select within the 16S taxonomy. These filters are part of the “EAW taxa analysis tool” and a product of the data preparation in WP4.

Applying all the selection and aggregation steps mentioned above to look at the WFD target groups finally

- the 16S ASVs extract belong to 88 cyanobacteria species or genera,
- the 18S ASVs extract to 882 phytoplankton species or genera.
- the rbcl-ASVs extract present 226 benthic diatom species or genera.

These three lists are the “detectable” target taxa, which are part of the “EAW taxa analysis tool”.

The taxonomic assignment at the species level is based on exact matching (no base pair differences) between ASVs and reference strains. The preparation of library followed a standard procedure in use at the FEM, IT sequencing facility directed by M. Pindo, as described e.g. in Salmaso et al. (2018) and for diatoms according the diatom-barcode V9 taxonomy at INRA, FR.

3.2 Data with taxa inventories detected by counting

The produced project data comprise the results by microscopical counting and by fishing.

All project partners were requested to use a common table and coding template to facilitate subsequent data analysis.

Each country had applied their national methods according the EU-WFD or WTO-CH, still there were a common understanding of the depths of biological analysis.

For phytoplankton, all project partners used the Utermöhl-counting technique at light microscope (cells /ml), common determination keys and geoforms to calculate and estimate the biovolume of each taxon (mm³/l).



Deliverable D.T1.2.3

Since the operational taxon names may include various synonyms, which hindered the comparison, all records were additionally assigned to a common code, the Rebecca-code (see freshwaterecology.com). For benthic diatoms all project partners prepared diatom slides and counted at highest magnification at light microscope (about 400 valves), used common determination keys and calculated the proportion of a taxon to total (% valves). Since the operational taxon names may include various synonyms, all records were additionally assigned to a common code, the VALID-code (see freshwaterecology.com).

3.3 Common code for taxa inventories and actuality of species names and systematics

While the harmonisation of names in fish lists is less problematic, the systematic and nomenclature of microorganisms are subject to constant change in a very short period of time. Common codes and taxa names for phytoplankton and benthic diatoms are necessary for comparing lists.

In fact, the operational taxa lists for biomonitoring are frequently more conservative, as determination keys and assessment tools are not adapted very quickly to new results. In biomonitoring, the taxa names are frequently synonyms of actual taxa names and they are attached to higher taxonomic categories of a former, even old-fashion systematic.

On the other side, the HTS taxonomy (NCBI) is modern in the metabarcoding list, and thus, many biological names maybe new for the user, because they are commonly not in use for the biological check lists of WFD/WTO monitoring. In detail, also the NCBI code differs from most actual names in few cases and includes errors and synonyms, so it needs to be checked.

Common codes for phytoplankton and benthic diatoms were originally redrawn from the REBECCA-code for phytoplankton and the VALID-code for benthic diatoms (see freshwaterecology.com; Notification: meanwhile both taxa lists have been updated).

There are historical and practical reasons why the “diatom taxonomy” was kept separate from the more common phytoplankton list (includes planktonic diatoms): There are two tradition approaches for taxa determination of diatoms, in which analysis with diatom slides providing much deeper determination than those with invers light microscopy, and this level is required for diatom indicator lists in water assessment. The basic taxa lists were extended for additional species and genera found in freshwaters of the Alpine Space in the Eco-AlpsWater project.

The two taxa coding lists are in table “EAW_diatom_code” e.g. “EAW_PP_code” in customer group “common codes” of the “EAW taxa analysis tool”.

Using the internet platform “algaebase” , the common coding lists were checked and linked to the most actual taxa names. Such a translation table from common code to the modern names is part of the final “EAW taxa analysis tool” to enable updated naming of the taxa inventory in output tables.

4. Redraw information about diversity and other bio-components

It is innovative that the whole microorganism community is recorded by one sample.

Using the taxa inventories of organisms groups, which are no target groups in the frame of the traditional water assessment according WFD, there are multiple chances to gather further information about the ecological function of freshwater systems.

For example, protist taxa such as ciliates provide supporting information to access freshwaters.

18S marker in the Eco-AlpsWater project detected 652 genotypes (ESVs) of ciliates in EAW plankton samples. These genotypes belong to 54 different ciliate taxa, which were otherwise overlooked (see APPENDIX 3 in D.T4.1.1).

The “EAW taxa analysis tool” provides “eDNA_taxa the queries” “Ciliaten_18S_Plankton” and “Ciliaten_18S_Plankton_all_seqs” in customer groups. Advanced Access-tool users can easily modify and rename these queries for searching other groups by changing the query criteria (example “fungi”) in output column “Division_18S”.



Deliverable D.T1.2.3

5. References

- Kurmayer, R., Sivonen, K., Wilmotte, A., and Salmaso, N. (2017). *Molecular Tools for the Detection and Quantification of Toxigenic Cyanobacteria*. Wiley, Chichester.
- Salmaso, N. (2019). Effects of habitat partitioning on the distribution of bacterioplankton in deep lakes. *Front. Microbiol.* 10, 2257. doi:10.3389/fmicb.2019.02257.
- Salmaso, N., Albanese, D., Capelli, C., Boscaini, A., Pindo, M., and Donati, C. (2018). Diversity and Cyclical Seasonal Transitions in the Bacterial Community in a Large and Deep Perialpine Lake. *Microb. Ecol.* 76, 125–143. doi:10.1007/s00248-017-1120-x.
- Water Framework Directive (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Off. J. Eur. Parliam.* doi:10.1039/ap9842100196.



Deliverable D.T1.2.3

APPENDIX 1

Short user instruction for EAW taxa analysis tool

Unzip the ZIP-file “EAW taxa analysis tool V8” to a local drive.

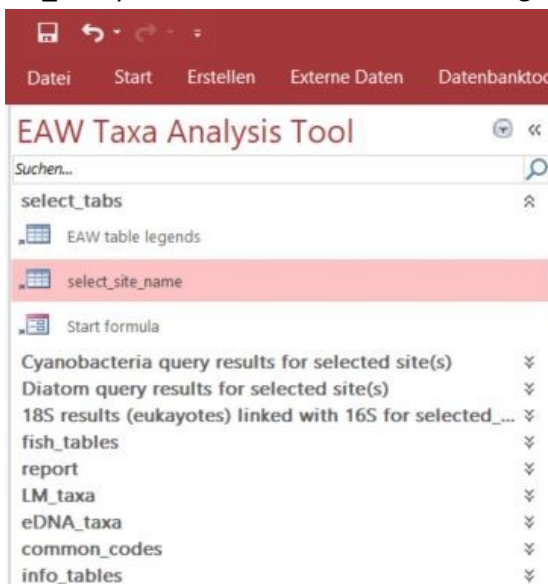
Open Access-file by double click.

Located along the left side of the Access start view, the navigation pane provides a quick way to access the objects in Microsoft Access. You can hide or show the navigation pane at any time press: F11.

The “navigation pane” on left side of the Access window is in special view of “EAW Taxa Analysis Tool”

All user tables and queries are linked into predefined custom groups.

With tool version 8, query results for selected site(s) are separated into groups cyanobacteria, diatoms, 18S_eukayotes and fish tables. The custom groups are listed in the figure below.




Open table “select_site_name” by double click

Select in first column your requested “site_name” by click. You can select more than one site. You can move up and down in the Access table. Waterbody type “L” for lakes and “R” for rivers.

Close the table by clicking on the “x” at most upper right corner of the Access table

(Please note that only pilot sites are with full data records and most miss records for biofilm (column 9))

Start formula		select_site_name						
select	country	waterbody	site_name	N EAW	FISH sa	WB_type	pilot site	count of other algae in BFM than diato
<input type="checkbox"/>	SI	L	Tivoli	1 no	lake	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
<input type="checkbox"/>	IT	L	Varese	1 no	lake	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	CH	R	Vedeggio	3 no	River	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	SI	L	Vogrscek	4 no	lake	<input type="checkbox"/>	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	DE	R	Wertach	13 no	River	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	DE_fish	R	Wertach	2 yes	River	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

Click on symbol “double arrow downwards”  to view a full list of table names of one custom group, for example “Cyanobacteria query_results_for_selected_site(s)”

Choose your query according the legend table (Appendix II) or just explore and produce result table.

Export your taxa analysis table! : Mark the table name in the navigation panel, use the right mouse click, chose “Export” “export to Excel” in the appearing Access menu.

Deliverable D.T1.2.3

All table/query names starting with

- "HTS..." provide the taxa inventory revealed by eDNA.
- "LM..." provide the taxa inventory revealed by light microscope counts.
- "GAP..." provide comparisons of taxa inventory revealed by both methods.

LM with quantitative field "biovolumes mm3/l"

HTS with quantitative field "signals" (rarefied)

All taxa are assigned weather to EAW_diatom code (benthic diatoms) or to EAW_PP_code (phytoplankton)

APPENDIX 2

Legend tables for each custom group in APPENDIX 2 – 2.1 -2.4

Table legends "EAW taxa analysis tool" (Access Database) in project Eco-AlpsWater Version 8 (V8)

With tool version 8 query results for selected site(s) are separated into groups cyanobacteria, diatoms and 18S for eukayotes and fish tables detected by MiFish-U.

APPENDIX 2 – 2.1

Legends for tables in Group "select tabs" and "Cyanobacteria" and "diatoms query results"

User group in Navigation Pane	name of prepared query table in customer group	legends to tables of EAW taxa analysis tool V8
select_tabs	select_tabs	
select_tabs	Start formula	Double click on Start formula to see infos about the EAW tool
select_tabs	select_site_name	First step: Double click on table "select_site_name" for opening the Access table // Select per click one or more EAW sites and save and close Acces table //
select_tabs	EAW table legends	Second step: Double click on one name of the prepared queries listed in the Access navigation panel

Cyanobacteria query results for selected site(s)		bio-component	habitat	tax. LEVEL
16S_cyano_biofilm_in_selected_site	Cyanobacteria in biofilms with 16S; Relevant only in case others than diatoms were analyzed in biofilm samples	cyanobacteria	biofilm	taxa
16S_cyano_genera_in_site_crosstable	16S cyanobacteria results grouped to genus level	cyanobacteria	plankton	genus
GAP_HTS_to_LM_Cyano_in_PL_genera	All genera of 16S cyanobacteria in plankton samples marked with those found in light microscope	cyanobacteria	plankton	genus
GAP_LM_to_HTS_Cyano_in_PL_genera	All genera of cyanobacteria in plankton samples counted by light microscope marked with those found in 16S cyanobacteria and those detectable with 16S	cyanobacteria	plankton	genus
HTS_chloroplast_16S_taxa_in_selected_site	phytoplankton taxa found by chloroplast 16S seq's in selected sites	phytoplankton	plankton	taxa to add for 18S
HTS_cyano_species_PL_in_selected_site	planctonic cyanobacteria detected by 16S - ESV seq-numbers and HTS signal linked to common-taxon names	cyanobacteria	plankton	ESV to taxa
LM_cyano_species_site_crosstable		cyanobacteria	plankton	species

Diatom query results for selected site(s)		bio-component	habitat	tax. LEVEL
GAP_HTS_to_LM_diatoms_BFM_in_selected_site	All diatom taxa by rcbl marked with those found in light microscope	diatom	biofilm	taxa
GAP_LM_to_HTS_diatom_genera_BFM_in_select_site	All diatom genera counted by light microscope marked with those found in rcbl and those detectable with rcbl	diatom	biofilm	genus
GAP_LM_to_HTS_diatom_species_BFM	All diatom species counted by light microscope marked with those found in rcbl and those detectable with rcbl detectable by rcbl_V9	diatom	biofilm	species
HTS_ESV_diatom_taxa_in_selected_site		diatom	biofilm	ESV to taxa
HTS_BFM_diatom_rcbl_identified_in_selected_site	All diatom taxa indentified by rcbl and linked to EAW_diatom_code (incl. VALIDCODES)	diatom	biofilm	taxa
HTS_BFM_diatom_rcbl_identified_in_site_crosstable	benthic diatom signals per sampling station in selected site	diatom	biofilm	taxa
LM_BFM_diatoms_in_select_site	See list of diatom taxa original provided at each sampling point found by light microscopy linked to common code	diatom	biofilm	taxa
LM_BFM_diatoms_in_select_site_crosstable	See % of diatom taxa (common code) at each sampling point found by light microscopy	diatom	biofilm	taxa
Venn_diagramm_diatom_taxa_number_in_site	number of taxa by LM, HTS and shared to prepare Veen diagrams	diatom	biofilm	N taxa

Deliverable D.T1.2.3

APPENDIX 2 – 2.2

Legends for tables in Group “18S results” and “fish tables” and “report”

18S results (eukayotes) linked with 16S for selected_site(s)		bio-component	habitat	tax. LEVEL
HTS_18S_eukaryotes_incl_phytoplankton_selected_site	See list of all eukaryote taxa with all detected sequences in plankton samples of selected sites	eukaryotes	plankton	ESV to taxa
HTS_18S_phytoplankton_selected_site	See extract for phytoplankton taxa in plankton samples of selected sites (sequences belong to one taxon are merged)	phytoplankton	plankton	taxa
HTS_18S_biofilm_all_eukaryotes_at_selected_site	See list of all eukaryote taxa with all detected sequences in biofilm samples of selected sites	eukaryotes	biofilm	ESV to taxa
HTS_ciliats_BFM_at_site	See extract for ciliates taxa in plankton samples of selected sites (sequences belong to one taxon are merged)	protists	plankton	taxa
LM_phytoplankton_PL_in_selected_sites	See list of phytoplankton taxa original provided for each sampling date found by light microscopy in selected site and linked to common PP code	phytoplankton	plankton	taxa
LM_phytoplankton_PL_in_selected_sites_crosstab	See cross-table list of phytoplankton taxa original provided for each sampling date found by light microscopy (LM)	phytoplankton	plankton	taxa
GAP_18S_to_LM_biofilm	eukaryotes found by 18S in biofilm samples at selected site with matching taxa found by light microscopy (LM)			
GAP_18S_to_LM_phytoplankton_genera_level	Genera found by 18S in plankton with link to all taxa by light microscopy which belong to this genus in selected sites	phytoplankton	plankton	taxa
GAP_18S_to_LM_phytoplankton_taxa_PL	Taxa and speices found by 18S in plankton with link to all species by light microscopy (species match!) in selected sites	phytoplankton	plankton	taxa
GAP_LM_to_18S_16S_Ch16S_PP_genera	Merging all HTS results (18S, 16S, chloroplast16S) on genus level with link to light microscopy taxa in selected site samples (Plankton samples)	cyano & phyto	plankton	taxa

fish_tables		bio-component	habitat	tax. LEVEL
HTS_12S_fish_taxa_all_seqs_samples_selected_site	All fish taxa identified by 12S MiFishU by any sequences - Please note: Separate runs (parallels) and stations have their own record here	fish	pelagic or littoral	species or genus
HTS_12S_signal_sumtaxon_selected_site_stations	Taxalist fishes identified by 12S MiFishU - sum of all counts (=reads;=signals) of serveral sequences, which belong to one taxon, each sampling station at water body site			
HTS_12S_fish_taxa_list_in_selected_site	Taxalist fishes identified by 12S MiFishU summarized for one water body site (lake or river)	fish	pelagic or littoral	species or genus
HTS_12S_count_seqs_per_taxon_selected_site_stations	Taxalist fishes identified by 12S MiFishU - count of sequences, which detect one taxon for each sampling station at water body site	fish	pelagic or littoral	species or genus
MiFishU_samples	list of all MiFish samples sampled by VIGI or by Sterivex	fish	pelagic or littoral	
N genotypes 12S seq to fish taxon	Numer of fish genotypes found in all EAW MiFish sample	fish	pelagic or littoral	
taxonomy_seqs_12S_all	246 DNA-sequences found by MiFishU for 12S fish taxonomy			ESV to taxa

	report	queries supporting REPORT D.T3.2.1/3.2.2_Report on 6 key lakes and rivers - select your pilot site before		
report	LM_BFM_diatoms_in_select_site_crosstable	table requested by users - easy overview to diatom results per sampling station at pilot site		
report	report 2 plankton taxon list LM	check table of your pilot site before importing to report	biofilm	
report	report_algal_class_biovolt_cross_table	check table of your pilot site before importing to report	plankton	
report	report_biofilm_18S_taxa list	check table of your pilot site before importing to report	plankton	
report	report_biofilm HTS_cyanotaxa	check table of your pilot site before importing to report	biofilm	
report	report_correspond_diatoms	check table of your pilot site before importing to report	biofilm	
report	report_GAP HTS_to_LM_diatom list	check table of your pilot site before importing to report	biofilm	
report	report_GAP LM_to HTS_diatom list	check table of your pilot site before importing to report	biofilm	
report	report_suppl_table 3	check table of your pilot site before importing to report	plankton	
report	report_suppl_table_1	check table of your pilot site before importing to report	plankton	

Deliverable D.T1.2.3

APPENDIX 2 – 2.3

Legends for tables in Group “LM-taxa” and “eDNA_taxa”

	results with traditional methods e.g. LM = light microscopy tables not specified for selected sites	
LM_taxa	all_biofilm_countings	See full list of all diatom and phytoplankton taxa original provided for each biofilm sampling site found by light
LM_taxa	all_PL_counts_to_HTS16S_Cyano_species	See full list of planctonic cyanobacteria taxa found by light microscopy and linked to 16S results when taxon name is exactly the same
LM_taxa	all_plankton_lake_countings	See full list of all phytoplankton taxa original provided for each sampling date found by light microscopy and linked to common PP code
LM_taxa	LM_BFM_lake_diatoms	See benthic diatoms found in lakes which can be linked to VALIDCode in each biofilm sampling site detected by light microscopy
LM_taxa	LM_BFM_lakes_no_diatom_taxa	See benthic PP taxa (no diatoms) found by 16S which can be linked to Rebecca-Code in each lake sampling site found by light microscopy
LM_taxa	LM_BFM_river_diatoms	See benthic diatoms found in rivers which can be linked to VALIDCode in each biofilm sampling site detected by light microscopy
LM_taxa	LM_BFM_rivers_no_diatom_taxa	See benthic PP taxa (no diatoms) found by 16S which can be linked to Rebecca-Code in each river sampling site found by light microscopy
LM_taxa	LM_diatom_taxa_found_in_biofilm_L_R	list of diatom taxa found in lakes and rivers with counts of records (no station information)
LM_taxa	LM_taxalist_found_in_BFM_all	list of diatoms and other taxa found in biofilm (no station information)
LM_taxa	LM_taxalist_found_in_PL_lakes	list of phytoplankton taxa found in plankton samples lakes (no station information)

Deliverable D.T1.2.3

	results with HTS metabarcoding method - tables not specified for selected sites	
eDNA_taxa	<i>Eukaryotes detectable by 18S</i>	full taxalist found by 18S in all EAW samples - detectable microbial and other eukaryotes in water and biofilm by 18S
eDNA_taxa	<i>Cyanotaxa detectable with 16S</i>	selected taxalist of cyanobacteria (target group) found by 16S in all EAW samples - detectable cyanobacteria in water and biofilm by 16S
eDNA_taxa	<i>diatom taxa detectable with rcbL V9</i>	taxalist of benthic diatoms (target group) found by rcbL in all EAW samples - detectable diatoms in water and biofilm by rcbL
eDNA_taxa	HTS_16S_PL_cyano_genera	cyanobacteria genera found in each plankton samples
eDNA_taxa	HTS_16S_PL_taxa	plankton samples, signals (=read;=count) 16S with link to matching seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_16S_seqs_all	16S taxonomy-seq list: with link of matching seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_chloroplast_seq_16S_toRebecca	chloroplast16S taxonomy-seq list: with link of matching seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_rcbL_seqs_V9_to_VALIDCODE	rcbL taxonomy -seq list: link of rcbLseq's to EAW_diatom_code (including VALID-Code)
eDNA_taxa	HTS_rcbL_signal_list	
eDNA_taxa	HTS_chloroplast_taxa_PL_samples_all	link of chloroplast 16S seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	Ciliaten_18S_Plankton	taxalist of all ciliats found by 18S in EAW plankton samples
eDNA_taxa	HTS_16S_BFM_taxa	biofilm samples, signals (=read;=count) per all 16S-DNA sequences with link to matching seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_18S_all_ASV_to_EAW_PP_Code	18S taxonomy-seq list: with link of matching seq's to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_chloroplast_16S_taxa_in_samples	phytoplankton taxa detected additionally by chloroplast16S per sample and only, when linked to EAW_PP_code (including Rebecca-Code)
eDNA_taxa	HTS_DIATOM_IN_SAMPLES_TO_VALIDCODE_all	biofilm samples, signals per all diatom- rcbL-DNA sequence ASV's with link to matching seq's to EAW_diatom_code (including VALID-Code)

APPENDIX 2 – 2.4

Legends for tables in Group “common codes” and “info tables”

common_codes		
common_codes	EAW_diatom_code	Fixed diatom code table - basic VALID code downloaded from freshwaterecology.info extended for Eco-AlpsWater new taxa records
common_codes	EAW_PP_code	Fixed phytoplankton code table - basic Rebecca- code downloaded from www.freshwater ecology.info (Dec 2019) extended for Eco-AlpsWater new taxa records
common_codes	EAW_sample_tracks	
common_codes	runs_rcbL_with_EAW_sample_code	Info, which samples have one or more rcbL runs
info_tables		
info_tables	info_16S_taxonomy_Silva	
info_tables	HTS_chloroplast_16S_taxa_list	
info_tables	HTS_Cyanobacteria_88_taxa_total	
info_tables	ES_national_method_biofilm_sites_or_samples	Ecological status of biofilm samples according national method
info_tables	ES_plankton_national_methods_partly_per_sample	
unassigned objects	several queries to test and prepare EAW data- no legends, not checked for public use	