

Readme file for the data set for *Freshwater mollusc community screening – classical and eDNA monitoring to detect rare, indicator and invasive species*

Author: Sonja Leidenberger, University of Skövde, Höskolevägen, Box 408, 541 28 Skövde, sonja.leidenberger@his.se

Supplementary Figure S1A_C_NCBI_BOLD_240625.docx

The figure shows

A) The detected target mollusc species (y = number of species, presence – absence, in %; 100 = one species, 50 means = two species etc.) per sample site (X = site ID 1-5) and season (spring vs. autumn) and sample method (Plankton net, sediment Kick-net, Sediment M42) on NCBI (National Center for Biotechnology Information; <https://www.ncbi.nlm.nih.gov/>). In total, six different species were detected.

B) The detected target mollusc species (y = number of species, presence – absence, in %; 100 = one species, 50 means = two species etc.) per sample site (X = site ID 1-5) and season (spring vs. autumn) and sample method (Plankton net, sediment Kick-net, Sediment M42) on BOLD (Barcode of life data system – v.4; <https://www.boldsystems.org>). In total, 11 different species were detected.

C) All detected freshwater phyla searched in the databases NCBI and BOLD (y = number of species, presence – absence, in %; 100 = one species, 50 means = two species etc.) per sample site (X = site ID 1-5) and season (spring vs. autumn) and sample method (Plankton net, sediment Kick-net, Sediment M42). In total, 13 different phyla were detected.

Supplementary Figure S2_InvasiveSpecies.png

The figure represents the detected invasive species (number of species) per site (Sample ID 1-5) and season (spring vs. autumn) per sampling method (Plankton net, Sediment Kick-net, Sediment M42).

Supplementary Table 1_SampleSites.docx

Sample site information for the localities investigated in this freshwater mollusc study in Sweden. The Site_ID, Area, Locality_name, Type_of_Habitat, Sampling_position in the WGS84 coordinate system, and sampling_date are given in the table.

Supplementary Table 2_SequencesBarcodeTree.docx

List of studied bivalve specimens in the barcode tree including specimen code from cytochrome c oxidase subunit I (COI) sequences in GenBank/NCIB, BOLD (^) and sequences produced by ourselves (PP204098-PP204157). Additionally, information is given for the locality and voucher details (material stored at GNM: Gothenburg Museum of Natural History, Sweden; MNHN: Museum National d'Histoire Naturelle Paris, France; UGSB:

Systematics and Biodiversity Collection of the University of Giessen, Germany) as well as the available reference.

Supplementary Table 3_Morphological_Records_FM_240829b.xlsx

Morphological records for freshwater molluscs in the study. The table contains the actual species name, synonym, locality, municipality, region and coordinates (both WGS84 and SWEREF 99 coordinate system), accuracy of the coordinates, habitat type (lake vs. river), the water_ID, the substrate, the height above the sea level (AMSL), sampling date, sampling time (start and stop), season (white background: spring, light-grey background: autumn), age-status (adult vs. juvenile), the number of collected individuals, determination date, identification method, uncertainty when determined, site ID, picture available/uploaded on Artportalen (yes vs. no), Barcode project running numbers, identifier for barcode sequence on GenBank, link to GenBank, identifier at the Gothenburg Natural History Museum (GNM), official collection where the specimen/voucher/barcode is stored and comments. All occurrence records were also uploaded on www.artportalen.se under the project title “*End biodiversity loss through improved tracking of threatened invertebrates*”.
[cf.: confer = uncertainty in determination; NA: not available]

Supplementary Table 4A_List_InvasiveSpecies.docx

Invasive species list after the search criteria ‘invasive species in Swedish waters’ and ‘freshwater’ on the database accessible from the Swedish Agency for Marine and Water Management (<https://www.havochvatten.se/arter-och-livsmiljoer/invasiva-frammande-arter/sok-frammande-arter.html>). The criteria ‘fish’ and ‘plants’ were excluded. All listed species had a COI barcode available in either GenBank/NCBI or BOLD (accessed 12/03/2024).

Supplementary Table 4B_Results_InvasiveSpecies_Metabarcoding.xlsx

Results of the detected invasive species with eDNA metabarcoding method. The table consist of the site_ID, site name, method used for sampling, coordinates (WGS84), season, habitat type (lake or river), class after BOLD, class after NCBI, order, phylum, genus, family and species name, similarity with GenBank/BOLD (%), similarity with BOLD (%) and running sample name in the project.

Supplementary Table 5A_IndicatorList_.xlsx

Indicator list that shows all existing species in Sweden that are listed under the family names in *Bottenfauna i vattendrag*, 2018, Table 3 (<https://www.havochvatten.se/download/18.717d870b18831fdb013666f/1708680145727/rapport-2018-35-bottenfauna-vattendrag-vagledning-for-statusklassificering.pdf>).

The table gives the Taxon.id, the global.sorteringsorder, the scientific name, the Swedish name (when available), the category (species, subspecies, morfotype) and the redlist category (LC: least concern, NT: near threatened, VU: vulnerable, EN: endangered, CR: critically endangered, NE: not evaluated).

Supplementary Table 5B_Indicator_NCBI_BOLD.DataFootprint.xlsx

Output-file when the indicator list (Table 5A) was run through the data workflow published by Bourlat et al. 2023 (https://gitlab.leibniz-lib.de/smartin/collect_data_footprint). The table list the species name (taxon), taxon ID, NCBI order, NCBI family and available barcodes (COI and 16S gene) on GenBank/NCBI and BOLD. Additionally, the results for the search on GBIF (Global Biodiversity Information Facility; <https://www.gbif.org>) are shown for Sweden: GBIF_taxon name, GBIF_comment (accuracy of the match), GBIF_species key, the number of records found in Sweden between the years 2003-2021 and the records found in Sweden between the years 2023-2022. The workflow was run 12/03/2024.

Supplementary Table 5C_Results_IndicatorSpecies_Metabarcoding.xlsx

Results of the detected indicator species with eDNA metabarcoding method. The table shows the site_ID, site name, sample method, coordinates of sample site (WGS84 coordinate system), season (spring vs. autumn), habitat type (lake vs. river), the indicator value (1-10) for the species, the class, the order, the phylum, the genus, the family, the species name, the similarity to NCBI/GenBank in %, the similarity to BOLD in %, the red list category (LC: least concern, NT: near threatened, VU: vulnerable, EN: endangered, CR: critically endangered, NE: not evaluated) and the running sample name in the project.