

RESEARCH

Open Access



Phylogenetic analysis of museum specimens of houting *Coregonus oxyrinchus* shows the need for a revision of its extinct status

R. Kroes^{1*}, Y. Winkel¹, J. A. J. Breeuwer¹, E. E. van Loon¹, S. P. Loader², J. S. Maclaine², P. F. M. Verdonschot^{1,3} and H. G. van der Geest¹

Abstract

According to the IUCN Red List the anadromous houting *Coregonus oxyrinchus* is categorized as 'extinct'. However, this extinct status might be incorrect because taxonomic difference between *C. oxyrinchus* and the closely related *C. lavaretus* is based on a disputable morphological comparison. Also, phylogenetic studies on mtDNA only focused on recent obtained coregonids. We are the first to perform a mtDNA analysis on both historic and recent specimens, including the syntype specimen which was used for species description by Linnaeus originally. Two primer pairs for mitochondrial *CytB* and *ND3* were used to extract sequences for phylogenetic analysis. Sequences from 14 out of 21 *C. oxyrinchus* museum specimens were successfully obtained and compared with sequences from recent obtained *C. lavaretus*. The sequences were combined with GenBank data from a previous phylogenetic study on houting to create a phylogenetic tree and two minimum spanning haplotype networks. Results show that *C. oxyrinchus* and *C. lavaretus* form a clade with limited genetic variation. Low bootstrap values also show weak support for geographical patterns in distribution of mitochondrial haplotypes. Statistical analysis of the haplotype networks also shows that historic and recent specimens are similar species. Our results suggest that *C. oxyrinchus* is a junior synonym of *C. lavaretus*. A definitive taxonomic revision could not be made because only *CytB* sequencing was successful for the syntype specimen. We discuss taxonomic consequences and the species-specific focus in nature conservation. We propose a shift in nature conservation to a more functional approach based on traits rather than species.

Keywords Phylogeny, Coregonids, Houting, Ancient DNA, IUCN, Linnaeus, Nature conservation

Background

Containing nearly 80 species the whitefish genus *Coregonus* is a diverse salmonid group inhabiting lakes, rivers and open fresh and sea water habitats [1]. The genus is also of high conservation interest given the high number of extinct and endangered species. However, the complex evolution of phenotypic traits has caused taxonomic confusion resulting in a number of nomenclatural issues [2–4].

One of the representatives is the anadromous houting or 'North Sea houting' *Coregonus oxyrinchus*, once geographically widespread in the Wadden Sea area and the coastal zones of the southern North Sea [5]. This species

*Correspondence:

R. Kroes
robkroes@uva.nl

¹ Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94240, 1090 GE Amsterdam, The Netherlands

² Natural History Museum, Cromwell Rd, South Kensington, London SW7 5BD, UK

³ Wageningen Environmental Research, Wageningen University & Research, Droevendaalsesteeg 3-3 A, Wageningen 6708 PB, The Netherlands



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

was first described by the Dutch ichthyologist Gronovius in 1754–56, although not in binominal nomenclature and with genus ‘*Salmo*’ rather than ‘*Coregonus*’ [6]. Linnaeus used the work of Gronovius to describe the species in the 10th and 12th edition of *Systema naturae*. Dried fish skins from Gronovius’ collection used for the species description were later incorporated in the collection of the Natural History Museum of London, including the type specimen for *C. oxyrinchus* [6].

Present-day, according to the IUCN Red List *C. oxyrinchus* is considered extinct [7]. This claim is based on a comparison of two morphological traits, snout length and number of gill rakers, taken from museum specimens (non-type material) and recent collected coregonids by Freyhof & Schoter [8]. However, phylogenetic studies from the last two decades challenge the taxonomic status of *C. oxyrinchus* and its IUCN listing as being extinct [5, 9–15]. Morphological traits might be plastic and change during the lifetime of a fish, for example when they migrate back and forth between fresh and salt water or change diet during growth. Also, homoplasy of morphological traits is common in coregonids because of rapid adaptation and differentiation to trophic ecology, reproductive behavior and diet acquisition [16–19]. This homoplasy hypothesis is supported by the absence of a relationship between the mitochondrial DNA (mtDNA) phylogeny and the number of gill-rakers within the *C. lavaretus* complex [15, 20]. Furthermore, mitochondrial genome sequences of European *Coregonus* spp. are less than 2% different (313 out of 16,600 nucleotides) suggesting that genetic diversification is recent and perhaps the result of isolation by distance [20]. Besides questionable morphologic traits, also geographic origin seem to be used for identification of *C. oxyrinchus*. Specimens originating from the Wadden Sea area and the coastal zones of the southern North Sea are usually identified as *C. oxyrinchus*. Specimens from the Baltic and its tributary rivers are usually identified as *C. lavaretus* [20, 21].

Because of the morphological and phylogenetic discussions, it remains unclear if *C. oxyrinchus* is a distinct

species or an ecotype of *C. lavaretus* to date [22]. Until detailed comparisons are made between type, historic and current material it cannot be ruled out that the *C. oxyrinchus* is a separate lineage. Therefore, in this study our aim is to examine molecular phylogenetic patterns of European whitefish to better understand the groups taxonomy. Specifically, we will compare various parts of older and recent *Coregonus* genomes, including never previously sampled historic material, to reveal the taxonomic status of *C. oxyrinchus*. We hypothesize that ancient and recent *Coregonus* spp. form a clade, and that this will have taxonomic and conservation implications for the species *C. oxyrinchus*.

Results

Phylogenetic analysis

Thirty-eight successful PCR samples were sent for sequencing of which $n=23$ sequences remained with consistent peak patterning of which $n=9$ recent obtained and $n=14$ museum specimens (see S.1 for sequence success overview and GenBank accession numbers). Together with the dataset from Østbye et al. [15], the resulting dataset comprised 8 specimens with 187 aligned bp for *CytB* and 187 bp for *ND3*. The percentage of phylogenetically informative (PI) sites in the aligned dataset was 4.5% for *CytB* and 7.8% for *ND3* (Table 1).

This dataset was used to construct a phylogeny using the maximum likelihood method and Kimura 2-parameter substitution model with a discrete Gamma distribution (5 categories (+G, parameter=0.1463)) to generate bootstrap support for clades. The topology of this tree resembled the phylogeny of Østbye et al. [15] and the same clades appeared in our extended phylogeny (Fig. 1). Only small differences in the splitting order of some clades were observed. These were probably due to the small number of informative sites and inconsistencies between informative sites. The latter is evidenced by the low bootstrap values, less than 70%, for these clades. In addition, most of the informative sites separated *C.*

Table 1 General information on alignment of the markers and substitution models used for creating Maximum Likelihood trees. #sequences gives the number of sequences available per marker in this study, whereas the ones in brackets are from Østbye et al. [15]. Positions is the number of nucleotides per alignment, V stands for the number of variable sites in each alignment, Pi is the number of phylogenetically informative sites per alignment, %Pi explains the percentage of the sequence that is phylogenetically informative. Model is the substitution model test (log likelihood-ratio test) carried out in MEGA7 to find the best model for constructing a Maximum Likelihood (ML) tree and base composition explains the nucleotide composition of the alignment

Genetic marker	#sequences	Positions	V	Pi	%Pi	Model	Base composition			
							T	C	A	G
Mitochondrial Cytochrome b (<i>CytB</i>) & NADH dehydrogenase 3 (<i>ND3</i>)	22+(62)	460	51	27	5.9	K2+G	0.301	0.342	0.199	0.158

lavaretus from *C. clupeaformis* and *C. albula*, which were used as outgroups.

The museum specimens, collected at the turn of the twentieth century from Ireland, the United Kingdom, the Netherlands, and Belgium, clustered together with recent specimens from Norway, Sweden, Finland, Russia and the Baltic Sea area [15] and two recent obtained specimens in this study from freshwater bodies in the Netherlands. However, bootstrap support for this clade was low (= 62%).

All museum specimens belonged to a larger weakly supported group of specimens indicated by Østbye et al. [15] that have been collected from the Barents Sea (part of the Arctic Ocean), the Baltic Sea and its Scandinavian freshwater tributaries, and Danish tributaries to the North Sea (clade A-K). Sequence differences between specimens in this group were very small ($d=1.51$ bp, 0.4% sequence divergence) and did not exceed more than 5 base pair differences, all in *ND3*, indicating that the museum specimens are not different from recent obtained specimens. Sequencing of the syntype and neotype specimens was only successful for *CytB* and were therefore not included in the concatenated phylogeny of *CytB* and *ND3*. Recent obtained specimens from the Netherlands were spread out across various other clades (P, R, C, A and F) that were indicated by Østbye et al. [15]. The recent obtained specimens in the Netherlands and Germany do not form a monophyletic clade nor do they belong to a specific clade and therefore are not a distinctive species.

The minimum spanning haplotype network for both *CytB* and *ND3* supports the above results as specimens from the ingroup (*C. lavaretus*, museum and recent obtained specimens) cluster together and are separated from the outgroup (*C. albula* and *C. clupeaformis*) (Fig. 2). P-distance within groups was 0.01 (ingroup) and 0.02 (outgroup). Between ingroup and outgroup, p-distance was 0.03. AMOVA shows that variation between groups was significant (AMOVA: Fixation index $\Phi_{ST}=0.818$, $\Phi_{SC}=0.517$, $\Phi_{CT}=0.622$; Significance of Φ_{ST} , Φ_{SC} and $\Phi_{CT}<0.001$). In the minimum spanning haplotype network for only *CytB*, the neotype and syntype were additionally added to the ingroup. The *CytB* network shows that both the syntype and the

neotype specimens cluster with the other included *C. lavaretus* and *C. oxyrinchus* specimens from the ingroup (Fig. 3). P-distance within groups was 0.01 (ingroup) and 0.02 (outgroup). Between ingroup and outgroup, p-distance was 0.02. AMOVA shows that variation between groups was not significant (AMOVA: Fixation index $\Phi_{ST}=0.818$, $\Phi_{SC}=0.517$, $\Phi_{CT}=0.622$; Significance of $\Phi_{ST}<0.001$, $\Phi_{SC}=0.063$ and $\Phi_{CT}=0.09$). However, *CytB* sequences from both type specimens were equal to both museum and recent obtained specimens in the clades C and E from the phylogenetic tree and differed 1 basepair from recent obtained specimens in clades R and P. Thus, based on this sequence divergence between ingroup and outgroup there is no strong genetic signal to further divide the studied European *Coregonus* specimens in different (phylogenetic) species.

Discussion

Phylogenetics

In 2008, *C. oxyrinchus* was categorized as 'extinct' on the IUCN Red List of Species. This claim is based on the morphological analysis of Freyhof and Schöter [8] on gill raker numbers. We sampled both specimens they used from the Natural History Museum in London (MH17 and MH20, original identification BMNH1844.11.11.15 and BMNH1862.11.20.1) as well as the syntype (MH21, original identification BMNH 1853.11.12.160). MH17 was included in our phylogenetic tree and the *CytB*+*ND3* haplotype network. The syntype and neotype were also included in the *CytB* haplotype network. Our mtDNA sequence analysis shows that all older and recent *Coregonus* spp., formerly classified as *C. oxyrinchus* and *C. lavaretus* respectively, clustered together and did not form separate clades or lineages in the phylogenetic tree. This conclusion is supported by the low bootstrap values on branches. Also, the older *C. oxyrinchus* show no monophyletic geographic distribution in comparison to the recent obtained *C. lavaretus*. Statistical analysis on the *CytB* and *ND3* haplotype network further supports our conclusion. The mean p-distance within ingroup and outgroup was 2–3 times lower than p-distance between the groups. Variance between ingroup and outgroup for *CytB* and *ND3* was also significant as shown by AMOVA. Observed significant variance between determined

(See figure on next page.)

Fig. 1 The Maximum Likelihood phylogenetic tree of the mtDNA (*CytB*+*ND3*) sequences of *Coregonus* spp. with node support from 500 bootstraps on branches. The geographical distribution of major branches is given. The 1-step nested clades are marked in colour. Representatives from *Coregonus clupeaformis* and *Coregonus albula* are included as the genetic outgroup. The recent obtained specimens of *Coregonus* spp. are indicated by FH and a black fish behind the name for the ones from the Netherlands and a blue fish for the one from the fish farm in Germany. Museum specimens of *Coregonus* spp. are indicated by MH with a red fish. The clades within the geographical distribution are given in brackets. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site

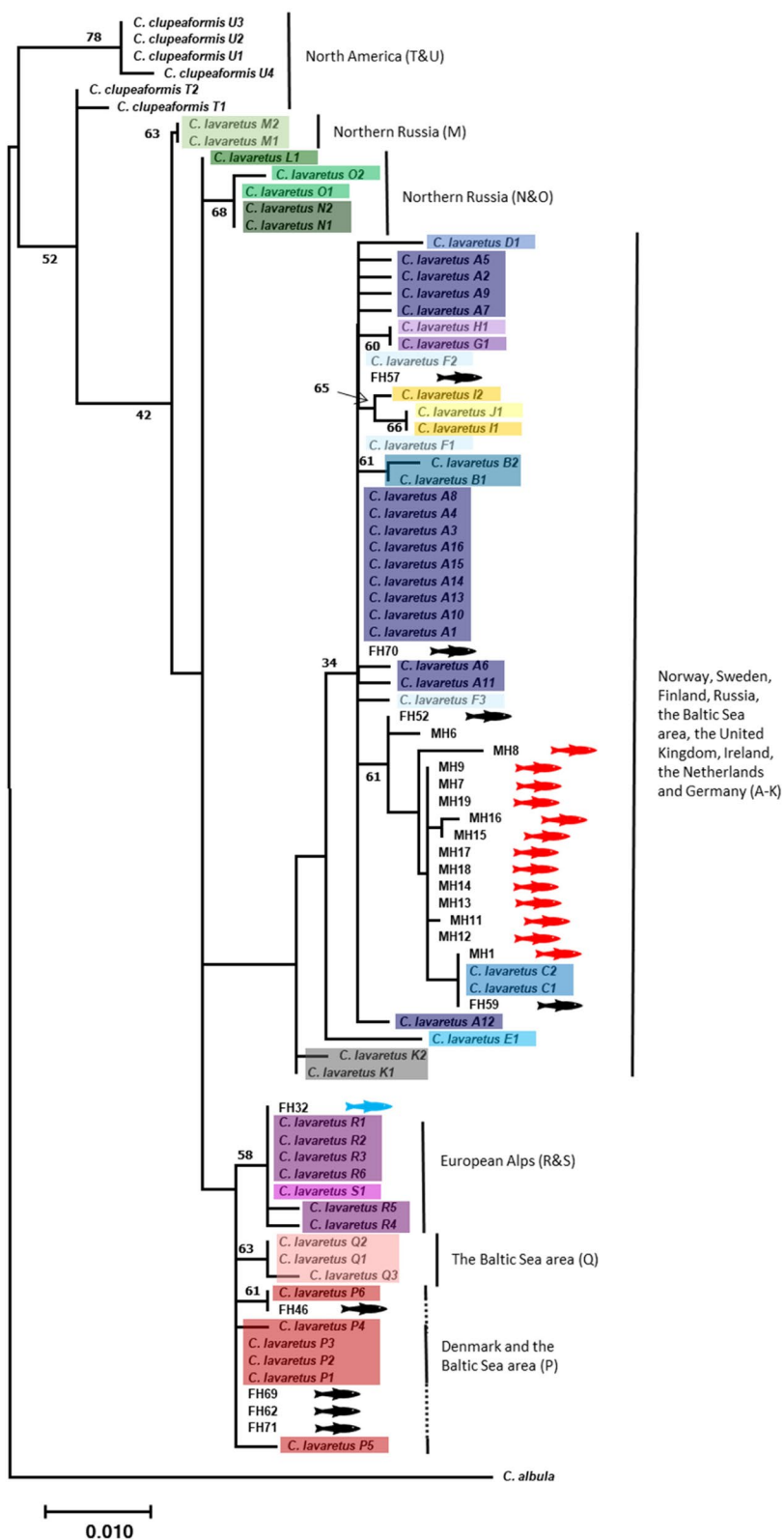


Fig. 1 (See legend on previous page.)

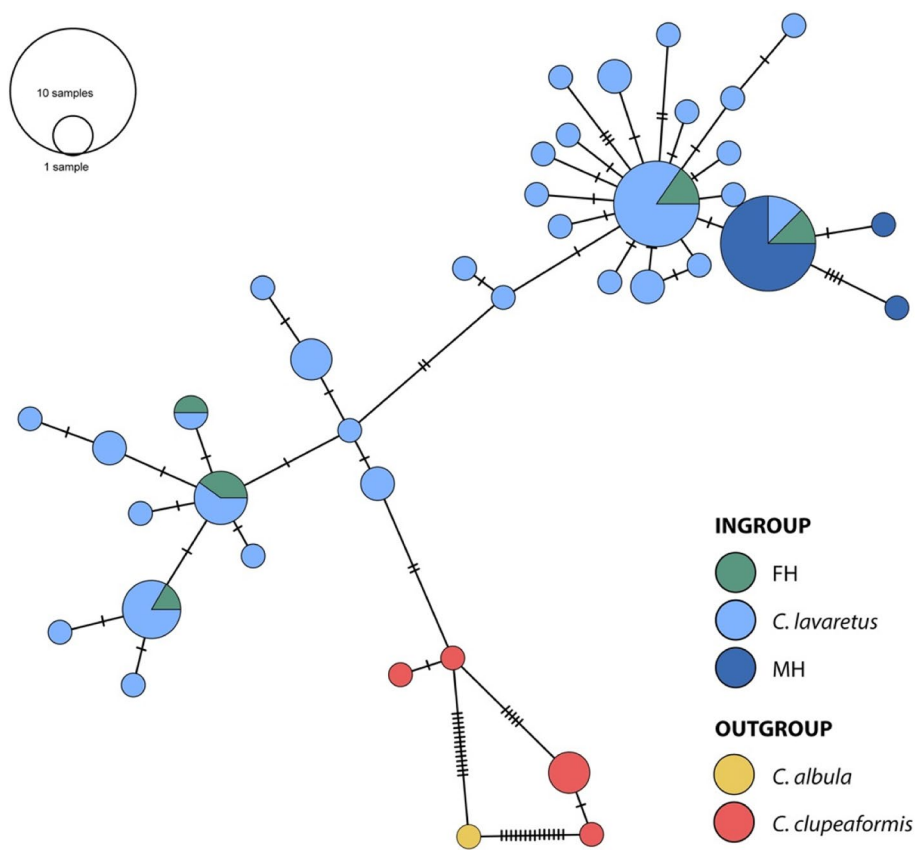


Fig. 2 Minimum spanning haplotype network of the mtDNA (*CytB* + *ND3*) sequences of *Coregonus* spp. The relative contribution of the different taxonomic groups per node are marked in colour. Representatives from *C. clupeaformis* and *C. albula* are selected as the genetic outgroups. Dashes indicate number of substitutions; the length of the edges is not informative

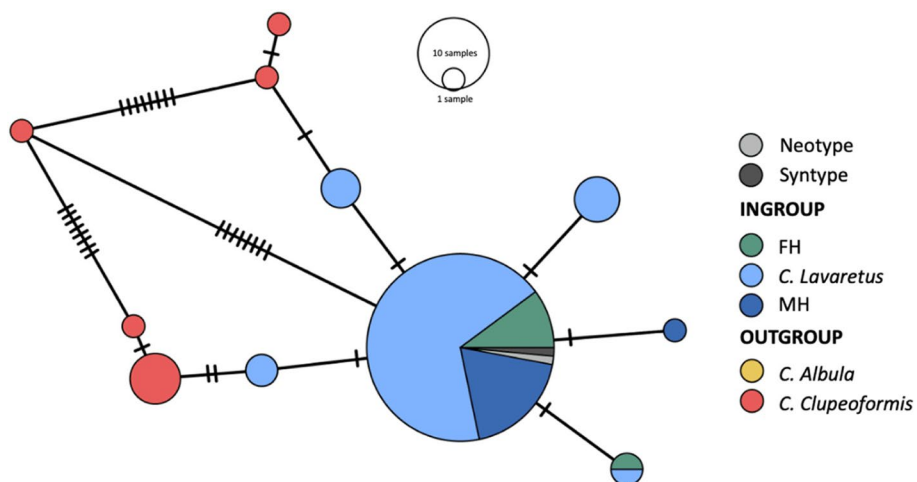


Fig. 3 Minimum spanning haplotype network of the mtDNA (*CytB*) sequences of *Coregonus* spp. The relative contribution of the different taxonomic groups per node are marked in colour. Representatives from *C. clupeaformis* and *C. albula* are selected as the genetic outgroups. The neotype and syntype were added to this network. Dashes indicate number of substitutions; the length of the edges is not informative

species is caused by the diverse outgroups used in this study. Excluding the outgroup showed that there is no significant variance between taxonomic groups from the ingroup (significance of $\text{Phy_ST}=0,481$). In the *CytB* haplotype network, the syntype and neotype also cluster with other taxonomic groups from the ingroup. Although no significant variance was found between ingroup and outgroup for *CytB*, this can be explained by the limited numbers of mutations between taxonomic groups and therefore less informative site of *CytB*. Thus, our hypothesis that ancient and recent *Coregonus* spp. form a single clade is accepted. This result has taxonomic and conservation implications towards recognition of the taxon *C. oxyrinchus*.

We argue for a revision of the IUCN conservation status 'extinct' of *C. oxyrinchus*. The currently used phenotypic traits weakly support the species status of *C. oxyrinchus*. Although not analyzed in the present study because of destructiveness of museum material, there is ample evidence from literature that gill raker numbers show high intraspecific variation, mainly because of rapid and recent evolution [12, 15, 17, 23]. This conclusion is supported by Etheridge et al. (2012) who showed that gill raker number is an inappropriate trait for taxonomic subdivision in UK-specific coregonids [24].

Previous phylogenetic studies on the *Coregonus* genus showed the presence of multiple distinctive lineages, mainly geographically clustered [5, 15]. However, bootstrap values are low to moderate and only weakly supported the geographic clusters indicated by Østbye et al. [15]. In addition, Hansen et al. [5] only used a small geographic area for their samples. Identification of geographical clusters of *Coregonus* spp. is further confused by introgression and admixture which are attributed to both natural and anthropogenic events, such as restocking [9, 20, 25–27]. More detailed studies with microsatellites more strongly supported the distinction of different *Coregonus* populations [16, 26]. A complete mtDNA genome analysis of Jacobsen et al. [20] showed that the anadromous "North Sea houting" from the Vidå river differs from non-anadromous or freshwater inhabiting coregonids. However, these morphological, geographical and genetic criteria are used to discriminate within species level, not between.

Based on plasticity of morphological traits used for identification, weak support on geographical patterns in distribution of mitochondrial haplotypes and our analysis of older material, we conclude that North Sea Houting *C. oxyrinchus* is not and has not been a biological species. Our results suggest that *C. oxyrinchus* is a junior synonym of *C. lavaretus*. A taxonomic revision awaits further analysis of the syntype from the Gronovius collection since alternative relatively shorter markers should also be

sequenced to obtain sufficient phylogenetically informative sites. However, complications for further analysis are caused by limited amplicon length that can be obtained from older museum specimens. More importantly, morphological identification of specimens that are used for sequencing is uncertain.

Based on our findings we hypothesize that syntype and neotype will cluster together in future analyses as well. Since the syntype is extant in the Natural History Museum, the *C. oxyrinchus* neotype BMNH1862.11.20.1 designated by Freyhof and Schöter [8] is probably invalid. The absence of gill rakers for identification of the syntype specimen should not have been used as an argument to designate a neotype. This morphological trait is plastic and more important, the original type material for the first binominal species description from Linnaeus can be used for phylogenetic analysis.

Management implications

In 2000, the European Union launched the Water Framework Directive (WFD) to restore the ecological status of European freshwater ecosystems. One of the measures to reach this goal was to reconnect water systems and hence mitigate migratory fish species [28]. Effects of these measures are often monitored by taxonomic inventories [29]. The IUCN uses these inventories as a primary source of distribution data. Because of the difficulties of fish taxonomy in general and coregonids in particular, legal problems consequently arise for taking the right actions to protect migrating fish if taxonomy fails. For example, in Europe houting *C. oxyrinchus* is both extinct (IUCN) and a present protected species (Bern convention, Habitats Directive, OSPAR). Without taxonomic agreement, it is unknown to which species the alive and swimming fish in Europe belong.

For ecosystem functioning, taxonomic inventories are less important. Recent studies show that ecosystem functioning and stability depends on functional diversity and its underlying genetics rather than taxonomic diversity [30–32]. Villéger et al. [32] also show that anthropogenic impacts have a six-fold higher effect on functional homogenization than taxonomic homogenization, with stronger homogenization when fish species are translocated. The latter happened with coregonids that originated from a Danish population and were reintroduced to the Rhine area in the 1990's [33]. A study from Borcharding et al. [9] showed facultative migration of *Coregonus* spp. (classified as *C. oxyrinchus* in their study) in Lake IJsselmeer, but populations in other lakes and rivers in The Rhine delta have not been studied yet. It is therefore unknown to what extent the reintroduced fish express their original anadromous way of life and how they affect ecosystem functioning.

Conclusions

Functional and phylogenetic diversity in the genus *Coregonus* needs protection. Anadromous populations in the North Sea basin suffered the most from habitat fragmentation and degradation and almost disappeared in the 1980's. A habitat restoration program probably saved the Danish North Sea population from extinction [21]. For taxonomic revision, the syntype from the Gronovian collection requires more detailed research. For the genus *Coregonus*, we propose that policies avoid the species specific focus in nature conservation. Instead management actions should focus on the restoration and functioning of water bodies, including restoring connectivity as stated in the WFD. To do so, more knowledge on the niche conditions of newly established coregonid populations is needed to fully understand its ecological role and conservation status.

Methods

Specimen samples

Pectoral fins from 20 specimens stored in ethanol and muscle tissue from the dried skin of one specimen of *Coregonus* spp. were sampled from the collection of the Natural History Museum in London. Museum specimens

dated from 1844 to 1927 and were classified as *C. oxyrinchus*. The geographical origin of the specimens covered the sea, lakes and rivers in the United Kingdom, Belgium, the Netherlands, Sweden and two unknown origins. From the total of 21 sampled specimens, 19 were non-type material. The neotype designated by Freyhof and Schöter [8] (BMNH 1862.11.20.1) and the syntype of *C. oxyrinchus* (BMNH 1853.11.12.160, dried skin of *Salmo oxyrinchus* Linnaeus, 1758, former name of *C. oxyrinchus*) from the Gronovian collection were also sampled (Fig. 4 for geographic distribution and Supplementary S.1 for all sample data).

In addition, pectoral or tail fins from 23 specimens of *Coregonus lavaretus* from the collection of the Institute of Biodiversity and Ecosystem Dynamics were sampled. These specimens were offered to the Institute between 2014 to 2017 and originated from the Westeinderplassen and the Noordzeekanaal in the Netherlands where they were found dead. All specimens were stored at -80 °C.

Finally, 17 larvae from a *C. lavaretus* hatchery of the Ruhrverband in the Mohnesee (Germany) were obtained in 2017 (exact location included in S1). The tissue used in DNA analysis consisted of the entire individual stored in ethanol.

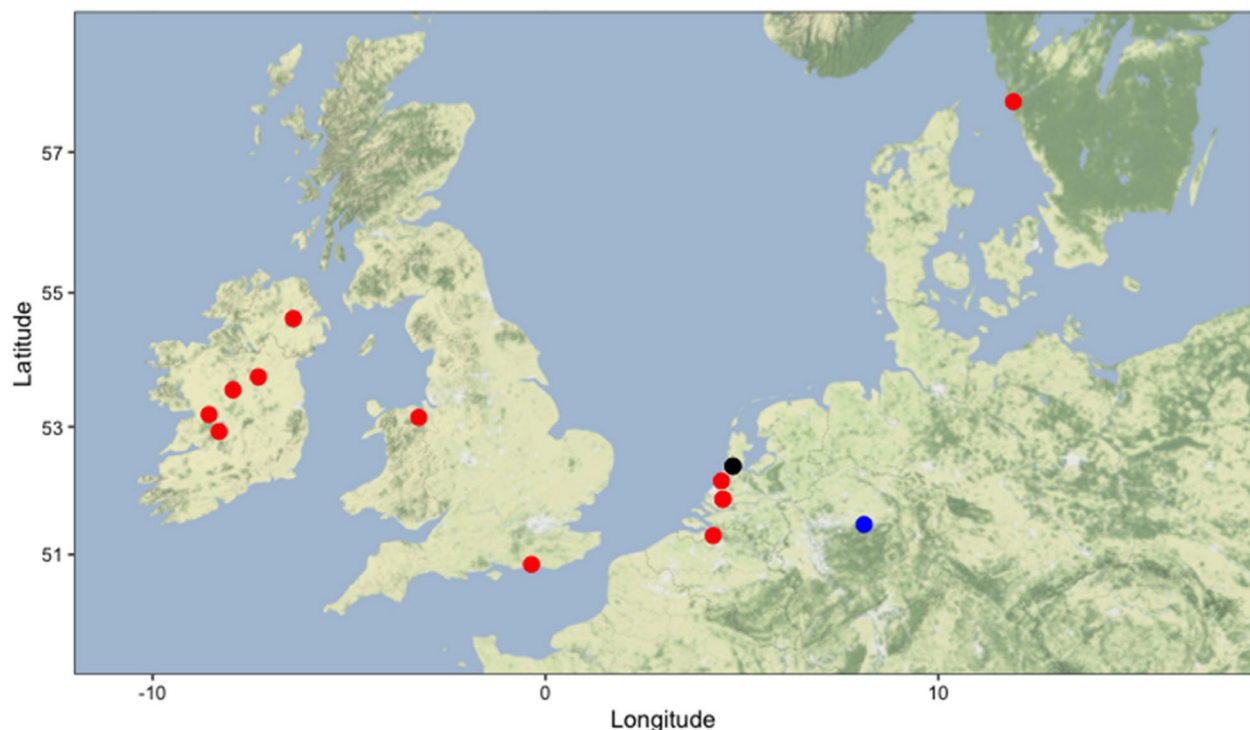


Fig. 4 Origin of *Coregonus oxyrinchus* from the Natural History museum in London and recent obtained whitefish. The red dots indicate the origin of the museum specimens (for two specimens, the location was unknown, and therefore not plotted on the map). The black dot indicates the origin of the recent obtained specimens in the Netherlands. The blue dot indicates the origin of the specimens from the fish farm in Germany (Background source: stamen maps)

DNA extraction

Tissue was cut into small pieces using sterilized scissors and dried on filter paper. Next, tissue was placed in 0.97 mL lysis buffer (200 mM Tris-HCl (pH 8.0), 100 mM EDTA and 250 mM NaCl), 5 μ L Proteinase K (20 mg/mL) and 60 μ L 10% SDS and incubated for a minimum of 2 h at 48°C. After incubation, DNA was extracted with phenol:chloroform:isoamyl alcohol and isopropanol precipitated with ammonium acetate [34]. The DNA pellet was dissolved in 50 μ L water. In addition, 1 μ L RNase (10 mg/mL) was added to the DNA samples of the recent obtained specimens to digest RNA according to Younas [35]. Isolated DNA was stored at -20°C until further use.

To prevent cross-contamination between historic and recent obtained DNA samples, we extracted DNA from the museum samples (likely to have low DNA concentration) before the recent obtained samples. In addition, all extractions were carried out in a fume hood that was decontaminated by UV light for a minimum of 30 min before starting the actual extraction. After every extraction the pipettes were cleaned using 96% ethanol and during the extraction only filter pipet tips were used to avoid transfer of aerosols.

PCR and sequencing

Two primer pairs for mitochondrial *CytB* and *ND3* previously used in phylogeographic research on European *Coregonus* species were used in this study [15, 36] (Table 2). Each PCR was done in a volume of 10 μ L, containing 3.5 μ L sterile milliQ water, 2 μ L Phire hotstart Buffer (5x), 2 μ L dNTP's (1mM), 0.2 μ L of each primer (10 μ M), 0.1 μ L Phire hotstart polymerase (ThermoFisher) and 2 μ L of template DNA (20–50 ng). As with DNA extraction, we also used filter pipet tips for preparing PCR's. Each PCR reaction contained one negative and one positive sample. For the negative sample sterile milliQ water was used and for the positive sample a good DNA sample from a recent obtained specimen was used. PCR was performed 3 times until negatives were blank. PCR conditions for the mitochondrial primers were 30s at 98°C; 35 cycles of 10s at 98°C, 10s at 58°C and 20s at 72°C; 1 min at 72°C and hold at 10°C. After PCR, products were run on 1.5% agarose gel and checked for DNA

quality, *i.e.* brightness and length of amplicon and presence of unspecific amplicons. Subsequently, 1 μ L of each amplicons was Sanger sequenced by Macrogen Europe Laboratories in a final volume of 10 μ L, 1 μ L of the forward primer (10 μ M) and 8 μ L sterile milliQ water.

Phylogenetic analysis

Low quality ends of sequences were trimmed using standard settings CLC Main Workbench 20.0.4 (Qiagen). After trimming, sequences were aligned per marker and observed differences were checked in the chromatograms by eye once to confirm that differences were true. Specimens with both *CytB* and *ND3* sequences available were concatenated by 100 N's into a single sequence. The concatenated sequences were combined with the GenBank PopSet 1015647666 data from Østbye et al. [15].

To find the best nucleotide substitution model for each alignment a model test was implemented in MEGA version 7.0.26 [37]. The evolutionary history was inferred by using the Maximum Likelihood method based and the corresponding substitution model ss determined by Modeltest in MEGA. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances using the Maximum Composite Likelihood (MCL). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. The dataset was bootstrapped 500 times to assess support for clades in the phylogeny with the highest log likelihood.

In addition to the phylogenetic tree, the same dataset was used to construct a minimum spanning haplotype network for *CytB*+*ND3* using PopArt [38]. In the network 5 groups were labeled: Crête-lafrenière et al. (2012) was used to select *C. albula* and *C. clupeaformis* for the outgroup, the ingroup comprised *C. lavaretus*, museum and recent obtained specimens [39]. An additional minimum spanning haplotype network was constructed for *CytB* only, in which the syntype and neotype were additionally added to the ingroup. The networks were used to analyze whether the outgroups from both networks were significantly different from

Table 2 *ND3* and *CytB* markers used for *Coregonus oxyrinchus*. The name stands for the part of the amplified gene. The forward and reverse primer give the sequence of the primer, length is the length of the sequence that this primer amplifies. Reference stands for the study where the primer is derived from

	Name	Forward primer (5'→3')	Reverse primer (5'→3')	Length	Reference
Mitochondrial	Cytochrome b (<i>CytB</i>)	CTTCGCTACGCAATCCTAC	GGCTCATTGAGGGCTTTAT	282 bp	Østbye et al. 2005 [15]
	NADH dehydrogenase 3 (<i>ND3</i>)	CATCACCATCGCACTATCCA	CCTCCTTGGGTTCCTCGTA	241 bp	Østbye et al. 2005, Schulz et al. 2006 [15, 36]

the ingroups. For this, MEGA version 7.0.26 [37] was used to determine p-distance within and between groups. In addition, Phy-statistics were calculated within and between groups with AMOVA as implemented in PopArt [38].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-023-02161-7>.

Additional file 1: S.1. Sample origin and analysis results. For each sample, code used in this study, original reference code, type status, GenBank accession numbers, species characteristics, origin and analysis results are given, if available. 'NA' refers to data that is not available.

Additional file 2: S.2. NCBI data for ND3. Given are sample data and ND3 analysis results for samples included in the phylogenetic tree and haplotype networks, as submitted to GenBank. See S.1 for accession numbers for corresponding samples.

Additional file 3: S.3. NCBI data for ND3. Given are sample data and ND3 analysis results for samples included in the phylogenetic tree and haplotype networks, as submitted to GenBank. See S.1 for accession numbers for corresponding samples.

Acknowledgements

We thank Reptile, Amphibian & Fish Conservation Netherlands (RAVON), Sportvisserij Nederland, Theo Rekelhof and Geert Timmermans for offering recent specimens to the collection of the Institute of Biodiversity and Ecosystem Dynamics.

Permission statement

The authors confirm that permission was obtained from the Natural History Museum London to include their samples in this study.

Authors' contributions

RK was the initiator of this study and main writer of all parts of the manuscript. YW and HB analyzed and interpreted the results from the phylogenetic analysis and were contributors in writing and revising the manuscript. SL and JM were contributors to the experimental design, writing and revising the manuscript. EL, PV and HG helped interpreting the data and were contributors in writing and revising the manuscript. All authors read and approved the final manuscript.

Funding

This work is part of the research program 'Doctoral Grant for Teachers' with project number 023.008.037, financed by the Dutch Research Council (NWO). The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The datasets generated and/or analyzed during the current study are deposited under embargo at GenBank under accession numbers OP723928—OP723955 for ND3 and OP723956—OP723986 for CytB. Accession numbers per sample are listed in table S.1 (supplementary data).

Declarations

Ethics approval and consent to participate

Recent obtained specimens included in this study were obtained from commercially available fish or dead fish from the wild. According to the Dutch law on animal testing "Wet op de diersoepen" and European DIRECTIVE 2010/63/EU on the protection of animals used for scientific purposes, ethical approval was not applicable for this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 7 October 2022 Accepted: 23 August 2023

Published online: 27 September 2023

References

- Alonso JDG, Häberle S, Plogmann HH, Schibler J, Schlumbaum A. Millennial-long co-existence of two major European whitefish (*Coregonus* spp.) lineages in Switzerland inferred from ancient mitochondrial DNA. *Divers*. 2017;9(3).
- Himberg MKJ, Lehtonen H. Systematics and nomenclature of coregonid fishes, particularly in Northwest Europe. *Adv Limnol*. 1995;46:39–47.
- Kottelat M, Freyhof J. Handbook of European freshwater fishes. Berlin: Kottelat, Cornol and Freyhof; 2007. p. 1–646.
- Svårdson G. Postglacial dispersal and reticulate evolution of Nordic coregonids. *Nord J Freshw Res*. 1998;74:3–22.
- Hansen MM, Fraser DJ, Als TD, Mensberg K-LD. Reproductive isolation, evolutionary distinctiveness and setting conservation priorities: the case of European lake whitefish and the endangered North Sea houting (*Coregonus* spp.). *BMC Evol Biol*. 2008;8(May 2016):137.
- Wheeler AC. The Gronovius fish collection: a catalogue and historical account. *Bull Br Museum (Natural Hist)*. 1958;1(5):185–249.
- Freyhof J, Kottelat M. *Coregonus oxyrinchus*. The IUCN Red List of Threatened Species 2008. e.T5380A11126034. 2018. <https://doi.org/10.2305/IUCN.UK.2008.RLTS.T5380A11126034.en>.
- Freyhof J, Schöter C. The houting *Coregonus oxyrinchus* (L.) (Salmoniformes: Coregonidae), a globally extinct species from the North Sea basin. *J Fish Biol*. 2005;67(3):713–29.
- Borcherding J, Pickhardt C, Winter H V., Becker JS. Migration history of North Sea houting (*Coregonus oxyrinchus* L.) caught in Lake IJsselmeer (The Netherlands) inferred from scale transects of 88Sr:44Ca ratios. *Aquat Sci*. 2008;70(1):47–56.
- De Leeuw JJ, Buijse AD, Griff RE, Winter HV. Management and monitoring of the return of riverine fish species following rehabilitation of Dutch rivers. *Large Rivers*. 2005;15(1–4):391–411.
- Douglas MR, Brunner PC, Bernatchez L. Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the Central Alpine region of Europe represent species flocks? *Mol Ecol*. 1999;8(4):589–603.
- Häkli K, Østbye K, Kahilainen KK, Amundsen PA, Præbel K. Diversifying selection drives parallel evolution of gill raker number and body size along the speciation continuum of European whitefish. *Ecol Evol*. 2018;8(5):2617–31.
- Kranenborg J, Winter H V., Backx JJGM. Recent increase of North Sea houting and prospects for recolonization in the Netherlands. *J Fish Biol*. 2002;61(SUPPL. A):251–3.
- Lelek A, Buhse G. *Fische des Rheins - früher und heute*. Springer-Verlag. (No. C115/17). Wageningen Marine Research; 2013.
- Østbye K, Bernatchez L, Næsje TF, Himberg KJM, Hindar K. Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Mol Ecol*. 2005;14(14):4371–87.
- Østbye K, Amundsen PA, Bernatchez L, Klemetsen A, Knudsen R, Kristoffersen R, et al. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. *Mol Ecol*. 2006;15(13):3983–4001.
- Kahilainen, K and Ostbye K. Morphological differentiation and resource polymorphism in three sympatric whitefish *Coregonus lavaretus* (L.) forms in a subarctic lake. *J Fish Biol*. 2006;68:63–79.
- Douglas MR, Brunner PC, Douglas ME. Evolutionary homoplasy among species flocks of central alpine *Coregonus* (Teleostei: Salmoniformes). *Copeia*. 2005;2005(2):347–58.
- Bernatchez L, Vuorinen JA, Bodaly RA, Dodson JJ. Genetic evidence for reproductive isolation and multiple origins of sympatric trophic ecotypes of whitefish (*Coregonus*). *Evolution* (N Y). 1996;50(2):624–35.
- Jacobsen MW, Hansen MM, Orlando L, Bekkevold D, Bernatchez L, Willerslev E, et al. Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and

- ecologically distinct European whitefish (*Coregonus* spp.). *Mol Ecol*. 2012;21(11):2727–42.
21. Jepsen N, Deacon M, Koed A. Decline of the north sea houting: Protective measures for an endangered anadromous fish. *Endanger Species Res*. 2012;16(1):77–84.
 22. Jacobsen MW, da Fonseca RR, Bernatchez L, Hansen MM. Comparative analysis of complete mitochondrial genomes suggests that relaxed purifying selection is driving high nonsynonymous evolutionary rate of the NADH2 gene in whitefish (*Coregonus* spp.). *Mol Phylogenet Evol*. 2016;95:161–70. <https://doi.org/10.1016/j.ympev.2015.11.008>. Elsevier Inc.
 23. Ozerov MY, Himberg M, Aykanat T, Sendek DS, Hägerstrand H, Verliin A, et al. Generation of a neutral FST baseline for testing local adaptation on gill raker number within and between European whitefish ecotypes in the Baltic Sea basin. *J Evol Biol*. 2015;28(5):1170–83.
 24. Etheridge EC, Adams CE, Bean CW, Durie NC, Gowans ARD, Harrod C, et al. Are phenotypic traits useful for differentiating among a priori *Coregonus* taxa? *J Fish Biol*. 2012;80(2):387–407.
 25. Rougeux C, Gagnaire PA, Bernatchez L. Model-based demographic inference of introgression history in European whitefish species pairs. *J Evol Biol*. 2019;32(8):806–17.
 26. Dierking J, Phelps L, Präbel K, Ramm G, Prigge E, Borcherding J, et al. Anthropogenic hybridization between endangered migratory and commercially harvested stationary whitefish taxa (*Coregonus* spp.). *Evol Appl*. 2014;7(9):1068–83.
 27. Borcherding J, Breukelaar AW, Winter HV, König U. Spawning migration and larval drift of anadromous North Sea houting (*Coregonus oxyrinchus*) in the River IJssel, the Netherlands. *Ecol Freshw Fish*. 2014;23(2):161–70.
 28. Kallis G, Butler D. The EU water framework directive: Measures and implications. *Water Policy*. 2001;3(2):125–42.
 29. van der Lee GH, de Baat ML, Wieringa N, Kraak MHS, Verdonschot RCM, Verdonschot PFM. Structural and functional assessment of multi-stressed lowland waters. *Freshw Sci*. 2020;39(4):621–34.
 30. Su G, Logez M, Xu J, Tao S, Villéger S, Brosse S. Human impacts on global freshwater fish biodiversity. *Science* (80-). 2021;371(6531):835–8.
 31. Craven D, Eisenhauer N, Pearse WD, Hautier Y, Isbell F, Roscher C, et al. Multiple facets of biodiversity drive the diversity–stability relationship. *Nat Ecol Evol*. 2018;2(10):1579–87. <https://doi.org/10.1038/s41559-018-0647-7>. Springer US.
 32. Villéger S, Grenouillet G, Brosse S. Functional homogenization exceeds taxonomic homogenization among European fish assemblages. *Glob Ecol Biogeogr*. 2014;23(12):1450–60.
 33. Borcherding J, Heynen M, Jäger-Kleinicke T, Winter HV, Eckmann R. Re-establishment of the North Sea houting in the River Rhine. *Fish Manag Ecol*. 2010;17(3):291–3.
 34. Ausubel FM. *Current protocols in molecular biology*. Wiley; 1995. ISBN 047150338X 9780471503385 0471503371 9780471503378.
 35. Younas, T. An efficient method of DNA isolation from fish fin. *Pak J Agric Sci*. 2016;53:843–850. <https://doi.org/10.21162/PAKJAS/16.3998>.
 36. Schulz M, Freyhof J, Saint-Laurent R, Østbye K, Mehner T, Bernatchez L. Evidence for independent origin of two spring-spawning ciscoes (*Salmoniformes: Coregonidae*) in Germany. *J Fish Biol*. 2006;68(SUPPL. 1):119–35.
 37. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33(7):1870–4.
 38. Leigh, JW, Bryant. PopART: full-feature software for haplotype network construction. *Methods Ecol Evol*. 2015;6(9):1110–1116.
 39. Crête-Lafrenière A, Weir LK, Bernatchez L. Framing the Salmonidae Family Phylogenetic Portrait: a more complete picture from increased taxon sampling. *PLoS ONE* 2012;7(10):e46662. <https://doi.org/10.1371/journal.pone.0046662>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

