

Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers

K. ØSTBYE,*† L. BERNATCHEZ,‡ T. F. NÆSJE,* K.-J. M. HIMBERG§ and K. HINDAR*

*Norwegian Institute for Nature Research (NINA), Tungasletta 2, NO-7485 Trondheim, Norway, †Norwegian University of Science and Technology (NTNU), Department of Biology, NO-7491 Trondheim, Norway, ‡Département de Biologie, Québec-Océan, Université Laval, Québec G1K 7P4, Canada, §Laboratory of Aquatic Pathobiology, Åbo Akademi University, BioCity, Artillerigatan 6, FI-20520 Åbo, Finland

Abstract

We compared mitochondrial DNA and gill-raker number variation in populations of the European whitefish *Coregonus lavaretus* (L.) species complex to illuminate their evolutionary history, and discuss mechanisms behind diversification. Using single-strand conformation polymorphism (SSCP) and sequencing 528 bp of combined parts of the cytochrome oxidase *b* (cyt *b*) and NADH dehydrogenase subunit 3 (ND3) mitochondrial DNA (mtDNA) regions, we documented phylogeographic relationships among populations and phylogeny of mtDNA haplotypes. Demographic events behind geographical distribution of haplotypes were inferred using nested clade analysis (NCA) and mismatch distribution. Concordance between operational taxonomical groups, based on gill-raker numbers, and mtDNA patterns was tested. Three major mtDNA clades were resolved in Europe: a North European clade from northwest Russia to Denmark, a Siberian clade from the Arctic Sea to southwest Norway, and a South European clade from Denmark to the European Alps, reflecting occupation in different glacial refugia. Demographic events inferred from NCA were isolation by distance, range expansion, and fragmentation. Mismatch analysis suggested that clades which colonized Fennoscandia and the Alps expanded in population size 24 500–5800 years before present, with minute female effective population sizes, implying small founder populations during colonization. Gill-raker counts did not commensurate with hierarchical mtDNA clades, and poorly with haplotypes, suggesting recent origin of gill-raker variation. Whitefish designations based on gill-raker numbers were not associated with ancient clades. Lack of congruence in morphology and evolutionary lineages implies that the taxonomy of this species complex should be reconsidered.

Keywords: adaptive radiation, glacial refugia, ice age, mismatch distribution, natural selection, nested clade analysis

Received 27 April 2005; revision accepted 10 August 2005

Introduction

The evolutionary significance of taxonomic designations needs to be validated in several settings such as before prioritizing among biological units for conservation (see

Fraser & Bernatchez 2001). Comparing patterns of molecular and phenotypic trait variation can be a powerful means towards this goal (e.g. Pinceel *et al.* 2004). Yet, despite criteria for species delimitation (Avice & Ball 1990; Mallet 1995), rather few studies have attempted to study species limits, or testing traditional taxonomy, using a combination of morphological and molecular analyses in a phylogeographic context (but see Puerto *et al.* 2001; Wiens & Penkrot 2002). One reason for this is the contention regarding what constitutes a theoretical and operational species under

Correspondence: K. Østbye, Present address: University of Oslo, Department of Biology, PO Box 1066 Blindern, NO-0316 Oslo, Norway. Fax: +4773801401; E-mail: kjartan.ostbye@nina.no

different species concepts. de Queiroz (1998) states that most species concepts agree that species are evolutionary lineages united by gene flow in sexual organisms. Moreover, as different species concepts may illuminate different phenomena associated with separation of lineages, data need to be tested against predictions from several criteria (Templeton 2001; Sites & Marshall 2003). A first step towards this objective is to test whether the variation in diagnostic phenotypical traits of suggested taxa is associated with evolutionary lineages defined by molecular markers.

Northern freshwater fishes have for decades intrigued scientists due to the occurrence of polymorphism in many genera (Behnke 1972; Bodaly 1979; Svärdsön 1979; Robinson & Wilson 1994; Taylor 1999). The repeated occurrence of intralacustrine morphs suggests that evolutionary processes such as adaptive phenotypic plasticity, divergent selection and/or colonization of numerous lineages have been involved in generating the tremendous phenotypic diversity among these fishes (Schluter 2000; Bernatchez 2004; Dieckmann *et al.* 2004). Indeed, distinct colonization pathways, hybridization events, and new selective regimes in new niches following deglaciation may be inferred from the phylogeographic structure of many organisms (e.g. Nesbø *et al.* 1999; Hewitt 2000; Seehausen 2004). Since intralacustrine divergence between populations may reflect early stages of species formation, or secondary contact of allopatric lineages, intralacustrine fish morphs constitute pertinent targets in evolutionary ecology (Bernatchez 2004; Gavrillets 2004).

The European whitefish *Coregonus lavaretus* (L.) species complex is ideal for contrasting the role of historical contingency and evolutionary events, as it displays levels of variation in morphology, life history, ecology, and spawning preferences that surpasses that documented in most northern freshwater fish species (Svärdsön 1957; Amundsen *et al.* 2004; Kahilainen *et al.* 2004; Østbye *et al.* 2005). The complex pattern of phenotypic diversity has generated taxonomic splitting or lumping under different authorships. Many taxonomic designations have been put forward (Linnaeus 1758; Järvi 1928; Berg 1962; Svärdsön 1979; Kottelat 1997), with more than 200 intraspecific forms described (Reshetnikov 2004). Traditional whitefish taxonomy has relied on the number of gill-rakers (Himberg 1970) for assessing systematic relationships. This trait is assumed to be a reliable marker with high heritability (Svärdsön 1970; Hermida *et al.* 2002), and temporal stability, even when transferred to new environments (Svärdsön 1950, 1979; Sandlund 1992). However, its association with foraging (e.g. Bergstrand 1982; Amundsen *et al.* 2004) suggests that this trait is also exposed to selection. Moreover, the parallel occurrence of allopatric and sympatric populations suggests that morphological species designations may not commensurate with evolutionary history (Bernatchez 1995). Indeed,

a previous study by Bernatchez & Dodson (1994) of the lake whitefish (*Coregonus clupeaformis*) × European whitefish (*C. lavaretus*) complex in their natural range in the Northern Hemisphere revealed five glacial refugia, and a mismatch between taxonomic designations and genetic groups. Moreover, *C. lavaretus* comprised a southern and a northern glacial refugia in Europe.

In this study, we illuminate the evolutionary history of the European whitefish species complex by documenting patterns of variation in mitochondrial DNA (mtDNA) and gill-raker numbers. We first establish phylogenetic relationships among mtDNA haplotypes mainly from central and northern Europe. Second, we use a nested clade analysis (NCA) (Templeton *et al.* 1987, 1992) and mismatch statistics (Rogers & Harpending 1992) to resolve the demographic patterns resulting from past isolation and colonization. Third, we test potential concordance between gill-raker variation and genetic structure at various levels of temporal separation. Finally, we test whether operational taxonomical groups based mainly on gill-raker numbers commensurate with phylogeographic patterns in the European whitefish.

Materials and methods

Sample collection

We sampled 129 populations of *Coregonus lavaretus* (L.) from 94 geographical sites comprising lakes, rivers and brackish water, in the natural distribution area in Europe, including two localities in Siberia, north Russia (Table S1, Supplementary material). These sampling sites cover a large part of the natural distributional range of *C. lavaretus* in Europe (Bernatchez & Dodson 1994). Each population was characterized by geographical coordinates, main river, migratory habits (stationary or anadromous), and spawning habitat (lakes, rivers or brackish water), and mean (\pm SD) numbers of gill-rakers (see Table S1). Most counts were done by ourselves, but for many populations, we relied on counts from literature, which may introduce some noise into the data set (Table S1). However, the temporal stability of gill-rakers justifies this procedure (Svärdsön 1979). We chose to use this single trait as it is heritable, frequently reported in scientific papers, and influenced by natural selection in whitefish (Bernatchez 2004). A recent operational taxonomical grouping of *C. lavaretus* into several ecological forms (e.g. CLF = the fera group in Table S1) by Kaukoranta *et al.* (2000), based on Svärdsön's (1979) study of gill-raker variation, was tested on our populations (Table S1). In using this definition, we weighted gill-rakers more than the spawning habitats when in conflict.

Genetic variation was screened in two mtDNA regions for 4–30 fish per population (Table S1). We also included 29 lake whitefish *Coregonus clupeaformis* (Mitchill) from North

America, representing the Mississippian, Atlantic, Acadian and Beringian glacial refugia for comparison with earlier studies (Bernatchez & Dodson 1991, 1994). A vendace *Coregonus albula* L. from Lake Orrevann, Norway (#107; Table S1), was used as an outgroup for rooting the internal relationships in the phylogenetic analysis of whitefish. The mtDNA material comprised 1341 fish, with 1311 European whitefish.

Single-strand conformation polymorphism analysis and mtDNA sequencing

Tissue samples were stored in ethanol, and DNA was extracted using phenol–chloroform (Sambrook *et al.* 1989). We designed primers amplifying a 282-bp region of cytochrome oxidase *b* (cyt *b*), and a 246-bp region of NADH dehydrogenase subunit 3 (ND3) (Schulz *et al.* in press); cyt *b*_F (CTTCGCCTACGCAATCCTAC)/ cyt *b*_R (GGCTCATTCGAGGGCTTTAT), ND3_F (CATCACCATCGCACTATCCA)/ ND3_R (CCTCCTTGGGTTCACTCGTA). Polymerase chain reaction (PCR) was performed in 12.5- μ L reaction volumes with 25–50 ng genomic DNA, 800 nM primer, 75 μ M CTP, GTP and TTP, 5 μ M ATP, 0.2 μ Ci ³⁵S-ATP, 1.5 mM MgCl₂, and 0.25 unit *Taq* polymerase. The PCR cycles were 1 \times (3 min at 95 °C), 32 \times (1 min at 95 °C, 45 s at 60 °C, 1 min at 72 °C), 1 \times (10 min at 72 °C).

Prior to the single-strand conformation polymorphism (SSCP) runs, 5 μ L formamide blue was added to the amplified products and denatured at 95 °C for 5 min and put on ice. Products migrated for 16 h at 8 °C for cyt *b* (at 2000 V, 300 mA, 12 W, 0.5 \times TBE buffer), and 18 h at 8 °C for ND3 (at 2000 V, 300 mA, 20 W, 1.0 \times TBE buffer) on polyacrylamide gels [7.0% acrylamide (1 : 49 bisacrylamide : acrylamide) +7.2% glycerol]. SSCP conformers were visualized by autoradiography (Sambrook *et al.* 1989). Reference SSCP patterns were used in each run, and representatives of all SSCP conformers were run on comparative gels. In order to insure reliability of the mtDNA-SSCP approach, and to resolve genetic relationships between different SSCP conformations, we sequenced two or more representatives of all conformers from two or more geographically distant sites. Here, only one SSCP-conformant harboured more than one sequence, in which case all individuals with this conformant were sequenced. In total, 156 fish were sequenced for cyt *b* and 199 fish were sequenced for ND3. Sequences were run on ABI automated sequencers after PCR amplification in 25- μ L reaction volume (conditions as above), and purification of PCR products using the QIAGEN purification kit. All sequences were aligned in SEQAPP 1.9a157 (D. Gilbert, University of Indiana). The sequences were aligned to a complete mtDNA sequence of *C. lavaretus* from Miya & Nishida (2000; GenBank Accession no. AB034824). Sequences (with composite coding) are deposited in GenBank, Accession no: DQ173305–DQ173428.

Phylogenetic analysis

The best nucleotide substitution model was found in MODELTEST 3.06 (Posada & Crandall 1998) after initial analysis in PAUP 4.0b10 (Swofford 2001). The TrN + G model (Tamura & Nei 1993) fitted the data best using the hierarchical likelihood-ratio test (hLRT). A total of 34 characters were phylogenetically informative, base frequencies were A = 0.2049, C = 0.3424, G = 0.1647, T = 0.2880, and estimate of the gamma shape parameter was 0.1885. We performed a maximum-likelihood (ML) analysis under heuristic search in PAUP 4.0b10 (Swofford 2001) with 10 random stepwise addition replicates and tree-bisection–reconnection swapping (TBR). One hundred bootstrap replicates were run to assess the statistical support in the tree nodes.

Nested clade analysis

We first estimated a haplotype network based on mutational differences for the European whitefish using the program tcs 1.13 (Clement *et al.* 2000). Ambiguous loops between haplotypes were resolved as suggested by Crandall & Templeton (1993). All haplotypes with less than 12 substitutions could unambiguously be connected into one single network at 95% confidence level (Templeton *et al.* 1992). Second, an NCA was performed using the GEODIS version 2.0 program (Posada *et al.* 2000). Failure to reject the null hypothesis of no geographical association of clades, can be due to inadequate sampling (e.g. gaps) or panmixia (i.e. no population structure). As these two cannot be differentiated (Hammer *et al.* 1998), biological inference was confined only to clades with rejected null hypothesis, applying the inference key of Templeton (1998).

Genetic divergence among European whitefish populations

The net sequence divergence (D_A ; Nei 1987) and the TrN + G model in MEGA 2.1 (Kumar *et al.* 2001) were used for estimating genetic divergence of populations. We used evolutionary clocks of 1% and 2% sequence divergence per million years for timing mtDNA divergence (e.g. Bermingham *et al.* 1997; Bernatchez 2001).

Demographic analysis

We performed a mismatch analysis (Rogers & Harpending 1992; Rogers 1995) on clades defined by NCA. Under the sudden expansion model of Rogers & Harpending (1992), a population at equilibrium with initial effective female population size N_0 changes its size to N_1 at τ units of mutational time in the past, and the process generates a mismatch distribution described by $\theta_0 = 2N_0\mu$ (before population expansion), $\theta_1 = 2N_1\mu$ (after expansion) and $\tau =$

μt (being the age of expansion). Here, μ is the mutation rate per site per generation, and t is the time in number of generations. For time estimates, we used a mean generation length of 6 years. The analysis was performed in DNASP 4.0 (Rozas *et al.* 2003). Assumption of neutrality was tested by Tajima's D -test (Tajima 1989a, b). As this test has low statistical power (Ramos-Onsins & Rozas 2002), we additionally used the F value of Fu (1997), and R^2 value of Ramos-Onsins & Rozas (2002), with higher power (depending on N). Since significant Tajima's tests may indicate population expansion, selective sweep or purifying selection against weakly deleterious variants (Rogers 1995; Excoffier 2004), and as variances in estimates of timing of the population expansion are large, results are suggestive only.

Concordance between mtDNA, gill-rakers and operational taxonomical groups

To test associations between gill-raker numbers and genetic structure, we first used a nested ANOVA for partitioning variation in gill-raker numbers of individual fish ($N = 535$) between different mtDNA clade levels. Second, concordance between evolutionary lineages (defined by 3-step clades), and operational taxonomical groups was tested by a chi-squared homogeneity test with significance levels determined by 10 000 simulations of the exact distribution of the test statistic under the null hypothesis (computer program; Steinar Engen, Norwegian University of Science and Technology, Trondheim). This test was applied to 74 populations belonging to either one of two pure clades (3-1 and 3-3), and to the total data set (118 populations) including populations with mixed clades.

To infer mtDNA and gill-raker relationships among sympatric populations, we constructed a dendrogram using UPGMA in MEGA 2.1 (Kumar *et al.* 2001) from genetic distances (D_A ; Nei 1987) between populations. Populations on this dendrogram were marked for their 3-step clade levels (3-1, 3-3 and a mix of 3-1 and 3-3, see Results), reflecting the oldest event in the genetic hierarchy. From visual contrast we evaluated intralacustrine polymorphism with regard to clade assignment.

Results

Sequence variation

SSCP and sequencing revealed a total of 61 mtDNA haplotypes in the joint sequence of *cyt b* (282 bp) and ND3 (246 bp) among 1340 whitefish from North America and Europe (excluding the outgroup) (Tables S1 and S2, Supplementary material). Of these, 55 haplotypes were found in *Coregonus lavaretus* in Europe and Asia (the Russian Republic). Haplotype A1 was found in 68% of all fish, and only four additional haplotypes had a frequency higher

than 2% ($R1 = 10.8\%$, $P1 = 5.2\%$, $C1 = 3.5\%$, $P2 = 2.0\%$). Of the remaining haplotypes, 27 were singletons.

Phylogenetic analysis

The intraspecific relationship within the European whitefish revealed several haplotype groups supported by moderate bootstrap values (Fig. 1). One group (termed the 'North European lineage') comprised haplotypes (K1 to I1), which were mainly distributed from the Kola Peninsula through the Baltic Sea area, but also entering into Denmark and Alpine lakes. This lineage displayed a starlike phylogeny, harbouring several closely related singletons, but also more divergent haplotypes. The A1 haplotype had an extensive distribution from Yukon (in a fish defined as *Coregonus clupeaformis* belonging to the Beringean lineage) to the European Alps, while other haplotypes, such as C1, were geographically confined to the Baltic Sea area. Another group (termed the 'Siberian lineage') comprised divergent haplotypes (L1 to O2) observed only in the two Siberian populations. The third group (termed the 'South European lineage'), which harboured three geographical haplotype clusters (P1 to R5), was mainly distributed from South Scandinavia to the central Alpine region. One haplotype cluster was mainly found in Denmark (P1 to P2), another in the Baltic Sea area (Q1 to Q3), and a third cluster in the Alps (R1 to R5). The maximum-likelihood tree (Fig. 1) also revealed that the European whitefish were separated from the Beringean and Acadian + Atlantic + Mississippian mitochondrial lineage in the North American lake whitefish (Bernatchez & Dodson 1991, 1994).

Nested clade analysis

The minimum-spanning network for European whitefish gave three 3-step clades, eight 2-step clades, 20 1-step clades and 55 0-step clades (haplotypes) (Fig. 2). Overall, this network closely matched the haplotype clustering resolved in the ML tree (Fig. 1), the main difference being the clustering of K1–K2 with the Siberian lineage in NCA. Thus, clade 3-1 was mainly distributed as the 'North European lineage' (Figs 1 and 3a). Fish belonging to this clade were also found throughout the study area from Siberia (e.g. population #2), to the Alpine lakes (#127), but at low frequencies outside northern Europe. Although characterized by a high haplotype diversity, clade 3-2 was seen only in a few sites along a line connecting southwest Norway (#108) with Siberia (#2) through the Baltic Sea area (e.g. #35) (Figs 1 and 3a). Clade 3-3 was distributed similarly as the 'South European lineage' (Figs 1 and 3a). The Siberian clade harboured high genetic diversity compared to the North and South European clades. Contact zones were found between the North and South European clade (3-1 and 3-3; #84, 93, 105), and between the North European and Siberian clades (3-1 and 3-2; #1).

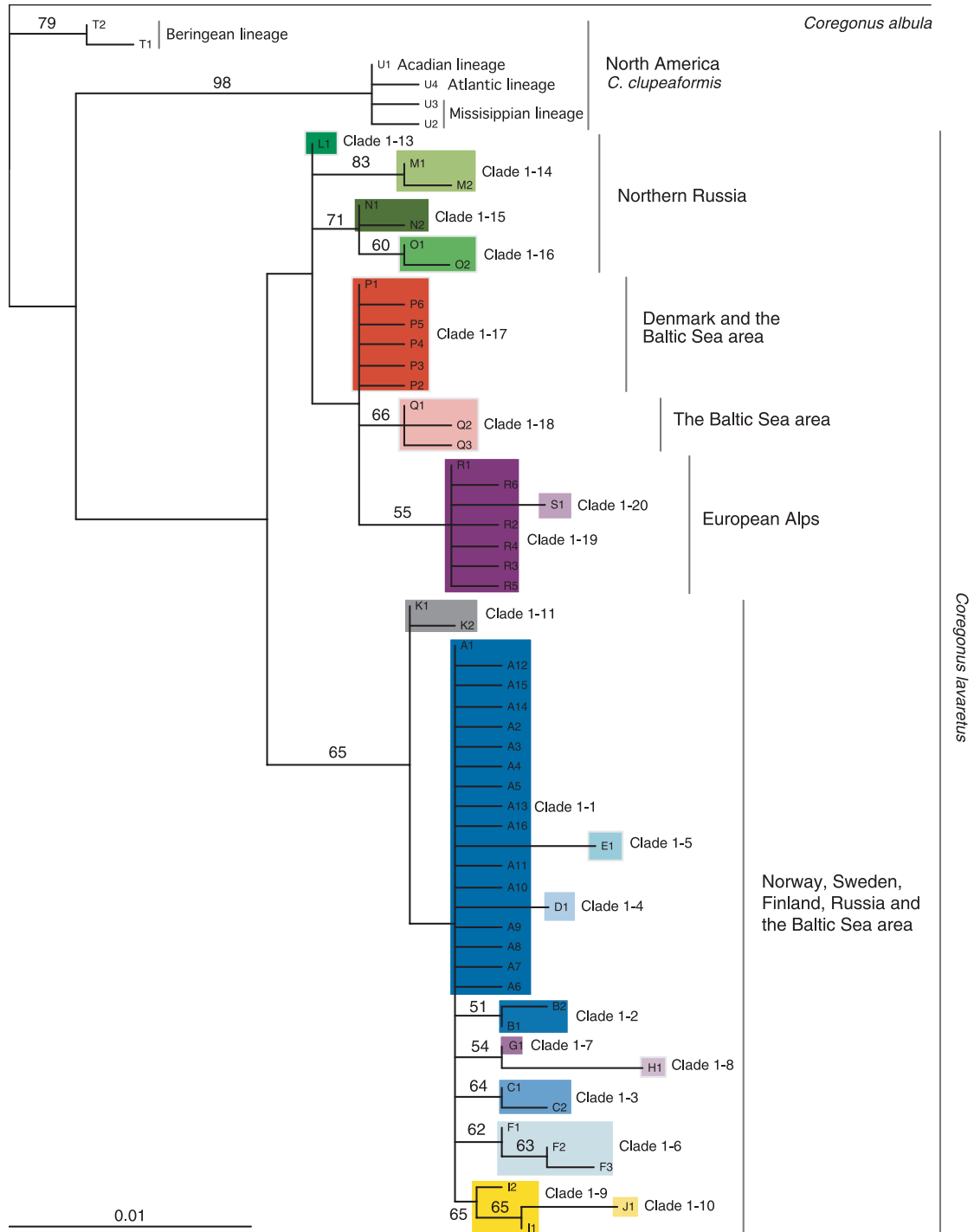


Fig. 1 The maximum-likelihood (ML) tree of the two composite mtDNA sequences in *Coregonus lavaretus* with node support from 100 bootstraps on branches. Geographical distribution of major branches is given. The 1-step nested clades are marked in colour. Representatives from *Coregonus clupeaformis* is given with *Coregonus albula* as the genetic outgroup.

At the 2-step clade level, clade 3-1 comprised clade 2-1 which was the major clade with a distributional range largely matching clade 3-1 (Fig. 3b), clade 2-3 with haplotypes (I1 and I2) northeast and northwest of the Baltic

Sea, respectively, and the small clade 2-2 confined to the Baltic Sea. Clade 3-2 harboured clade 2-5 and 2-6 seen only in Siberia, and clade 2-4 from southwest Norway to Siberia. Clade 3-3 comprised clade 2-7 from Denmark

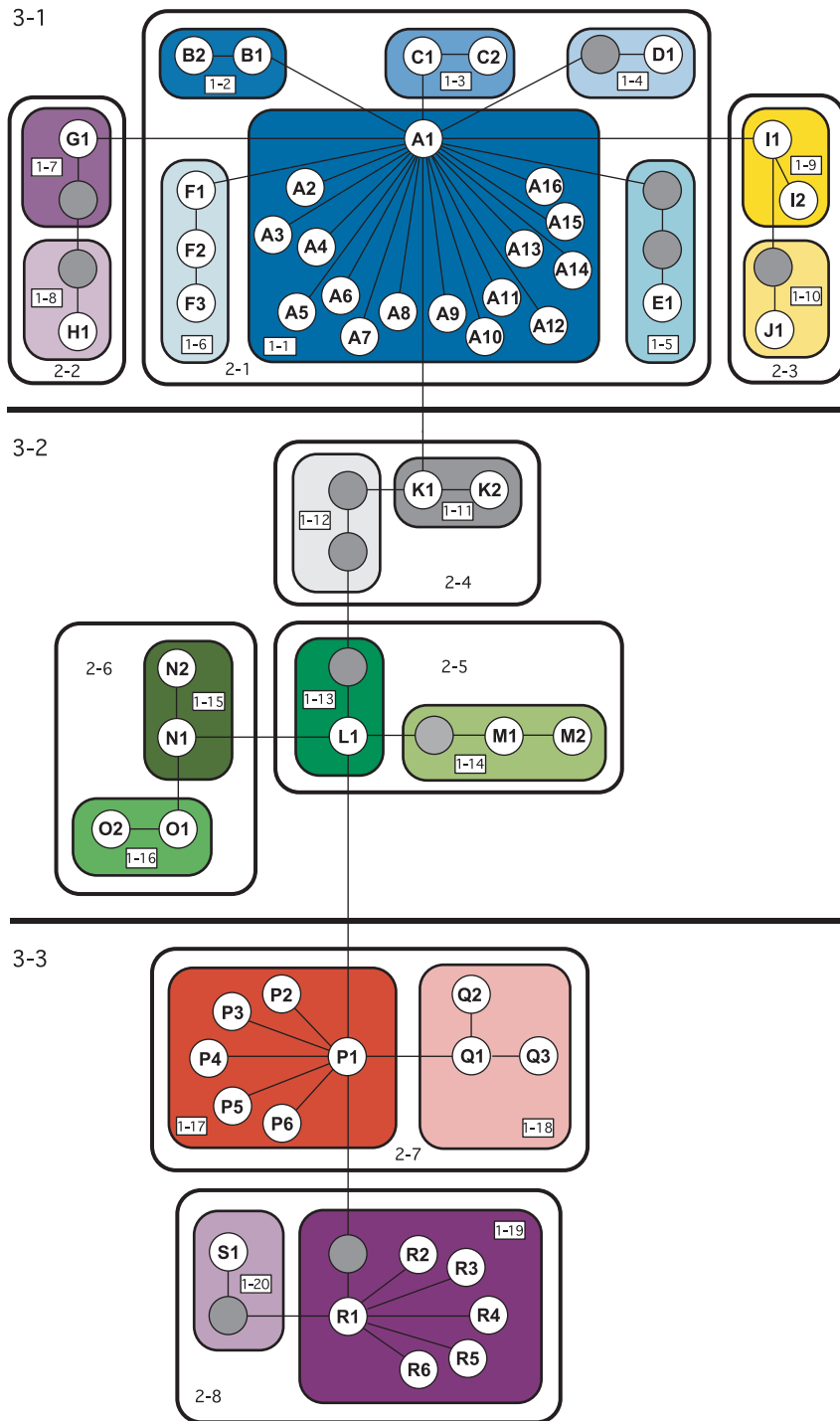


Fig. 2 The most-parsimonious haplotype network based on mutational differences among 55 *Coregonus lavaretus* composite mtDNA sequences found in Europe and the Russian Republic. One mutation links haplotypes, open circles denote missing haplotypes in the estimated network.

to Baltic Sea, and clade 2-8 being only seen in the Alp lakes.

At the 1-step clade level (Fig. 3c), clade 1-1 matched the distribution of its higher nested clades (2-1 and 3-1). Clades 1-2, 1-3 and 1-6 were only seen in Baltic Sea area. Clade 1-9 had same distribution as its higher-level clade 3-2. Clades 1-13 and 1-16 were restricted to Siberia. Clade 1-17 was

found in Denmark and Baltic Sea area, while clade 1-18 was seen only in Baltic Sea. Clades 1-19 and 1-20 were found in European Alp lakes only.

A total of 13 clades significantly differed from a random geographical distribution, and this group was thus considered for demographical-spatial inference (Table S3). The whole nesting hierarchy (4-1) was significant with inference

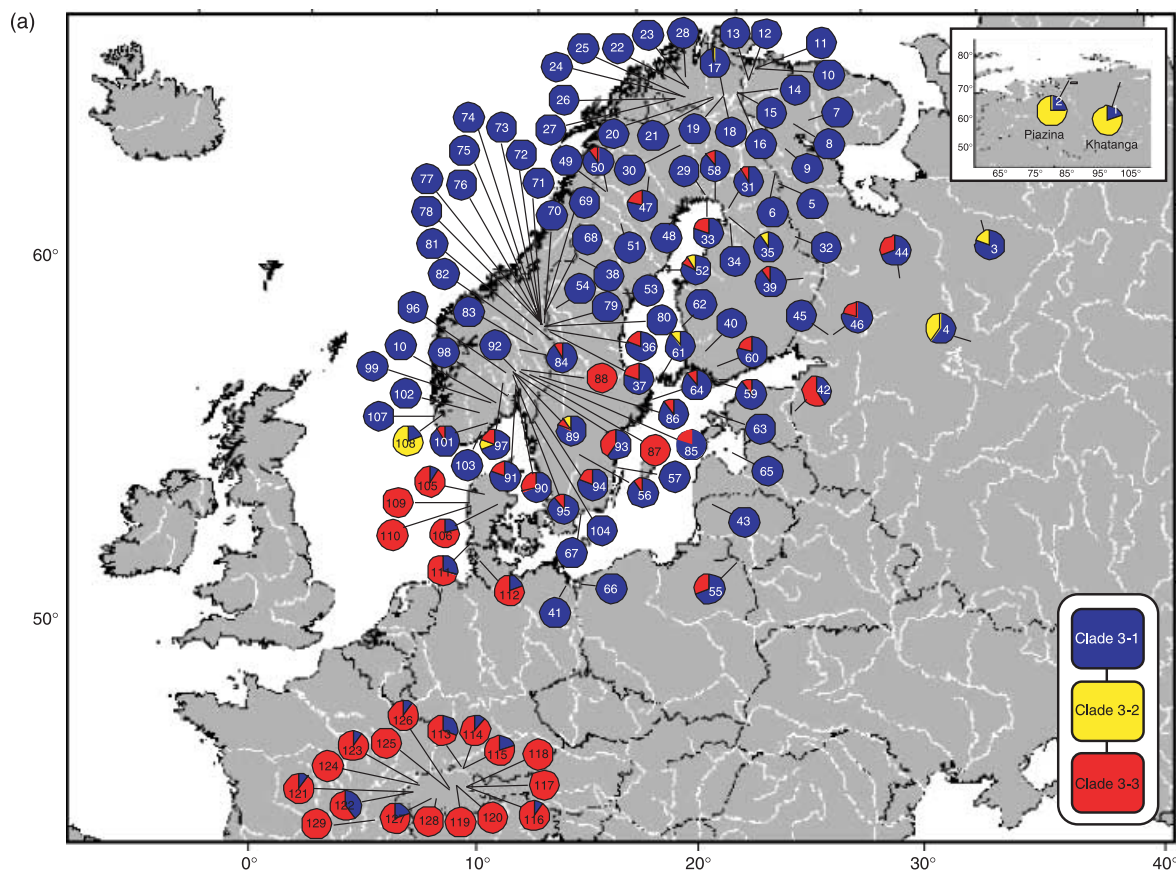


Fig. 3 Geographical distribution of 129 populations of *Coregonus lavaretus* partitioned into (a) 3-step clades, (b) 2-step clades, and (c) 1-step clades.

of past fragmentation (and/or long-distance colonization). Clade 3-1 had inference of restricted gene flow with isolation by distance, clade 3-2 of contiguous range expansion, and clade 3-3 allopatric fragmentation. Within clade 3-1, clade 2-1 had inference of restricted gene flow/dispersal but with some long-distance dispersal. Within clade 3-3, clade 2-7 had contiguous range expansion, and clade 2-8 had restricted gene flow with isolation by distance. Within clade 2-1, clade 1-1 showed restricted gene flow/dispersal but with some long-distance dispersal, clade 1-3 contiguous range expansion, and clade 1-6 long-distance colonization. Within clade 2-3, clade 1-9 had allopatric fragmentation. Within clade 2-7, clade 1-17 and 1-18 had an inference of restricted gene flow with isolation by distance. Only clade 1-19 had inconclusive demographic-spatial inference.

Mismatch analysis

Mismatch analysis of the nested clades showed that five clades had a significantly negative Tajima's D value (Table 1). Given the redundancy between clade levels, we only consider details of 1-step clades. Clade 1-1 ($D = -2.145$,

$P < 0.01$), being the main clade in north Fennoscandia, had an estimated time for demographic expansion roughly 11 700–5800 years before present (BP). Moreover, effective population size at the onset of expansion was extremely small (i.e. $N_e \sim 0$). Clade 1-19 ($D = -1.703$, Fu's F_S with $P = 0.007$), the main clade found in Alpine lakes, had time of expansion roughly 24 500–12 300 BP, also with $N_e \sim 0$. Clade 1-11 ($D = -0.592$, R^2 -test with $P = 0.095$), a clade within Siberian clade 3-2, had demographic expansion 35 600–17 800 BP. Clade 1-18 ($D = -1.120$, R^2 -test with $P = 0.087$), confined to the Baltic Sea area, had demographic expansion 57 400–28 700 BP, with $N_e \sim 0$.

Concordance between mtDNA, gill-rakers and operational taxonomical groups

Nested ANOVA of fish with individually scored gill-raker numbers onto the nested genetical hierarchy gave a significant association (whole model; $R^2 = 0.07$, $N = 535$, $P = 0.0097$). However, only the haplotype (0-step) level was significant [3-step (SSQ = 10.4, $F = 0.16$, $d.f. = 2$, $P = 0.848$), 2-step (SSQ = 5.3, $F = 0.08$, $d.f. = 5$, $P = 0.919$), 1-step (SSQ = 68.1,

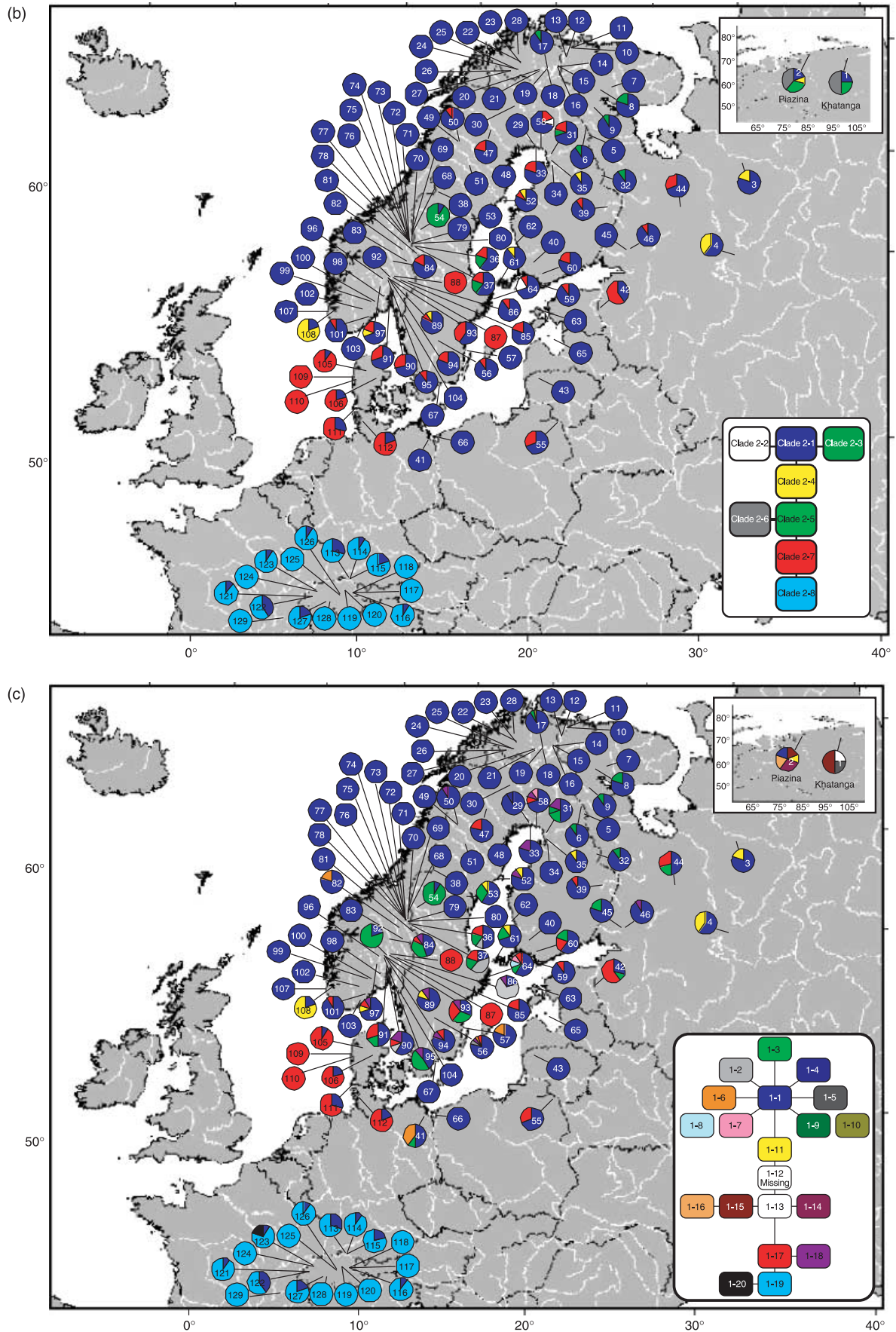


Fig. 3 Continued

Table 1 Pairwise sequence divergence statistics (mismatch analysis) based on frequency of haplotypes in *Coregonus lavaretus* clades identified by the nested clade analysis. The number of sequences (N), polymorphic sites (S), mean number of pairwise differences (k), mutational time since demographic expansion (τ), and the mutation parameter (θ) as observed and prior to (θ_{t_0}) expansion. R_i , Harpending's raggedness index. Only clades with a significant (or marginally significant $P < 0.10$) negative value under Tajima's D -test (or R^2 or F_u 's test) are given estimates for time and effective population (N_e) size at expansions by 1–2% sequence divergence per million years. * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$

Clade	N	S	k	Tau (τ)	Theta obs.	Theta _{t₀}	R_i	Tajima's D	Time since expansion	N_e
4-1	1300	56	2.857	0	7.229	3.023	0.153	-1.595 **	0–0	95 063–47 531
3-1	989	33	0.271	0	4.416	0.280	0.384	-2.373 ***	0–0	8805–4403
3-2	31	12	2.994	0.735	3.004	2.258	0.120	-0.011 NS	—	—
3-3	280	17	1.502	1.255	2.737	0.247	0.103	-1.136 NS	—	—
2-1	965	27	0.219	0	3.625	0.240	0.466	-2.311 **	0–0	7547–3774
2-2	3	3	2.000	1.000	2.000	1.000	1.000	—	—	—
2-3	22	3	0.667	0.667	0.823	0	0.164	-0.491 NS	—	—
2-5	5	3	1.500	1.500	1.636	0	0.083	-0.754 NS	—	—
2-6	6	2	0.933	0.933	0.876	0	0.222	0.311 NS	—	—
2-7	126	8	0.834	0.834	1.479	0	0.144	-1.022 NS	—	—
2-8	154	7	0.180	0.001	1.248	0.180	0.542	-1.883 *	189–94	5660–2830
1-1	901	15	0.062	0.062	2.032	0	0.775	-2.145 **	11 698–5849	0–0
1-2	5	1	0.400	0.400	0.480	0	0.200	-0.817 NS	—	—
1-3	48	1	0.042	0.042	0.225	0	0.842	-1.107 NS	—	—
1-6	9	2	0.944	0.944	0.736	0	0.176	0.975 NS	—	—
1-9	20	1	0.442	0.442	0.282	0	0.209	1.026 NS	—	—
1-11	20	1	0.189	0.189	0.282	0	0.422	-0.592 NS ¹	35 660–17 830	0–0
1-15	4	1	0.667	0.667	0.545	0	0.556	1.633 NS	—	—
1-17	107	5	0.603	0.603	0.953	0	0.164	-0.773 NS	—	—
1-18	19	2	0.304	0.304	0.572	0	0.255	-1.120 NS ²	57 358–28 679	0–0
1-19	152	5	0.130	0.130	0.893	0	0.576	-1.703 NS ³	24 528–12 264	0–0

(1) The more powerful R^2 test of Ramos-Onsins & Rozas (2002) gave a marginally significant value $P = 0.095$. (2) The R^2 test gave a marginally significant value $P = 0.087$. (3) The more powerful test using Fu (1997) F_S statistics gave a significant value $P = 0.007$.

$F = 0.43_{d.f.=5}$, $P = 0.826$), 0-step (SSQ = 700.4, $F = 1.92_{d.f.=11}$, $P = 0.025$).

A frequency distribution of gill-raker numbers shows that most whitefish in the North European clade (3-1), belong to a single haplotype (A), and encompass all of the variation in gill-raker numbers seen among the two other major clades (Fig. 4).

In those populations that could be assigned to the operational taxonomical groups of Kaukoranta *et al.* (2000) (Table S1), no significant association was found between taxonomical groups and pure 3-step clades ($\chi^2 = 4.75$; d.f. = 4; $P = 0.275$) or between these groups and a combination of pure and mixed 3-step clades ($\chi^2 = 16.74$; d.f. = 16; $P = 0.354$).

Small genetic distances were seen between sympatric whitefish populations, in north and south European lakes (Fig. 5). No distinct genetic clusters of populations with similar gill-raker means were evident. Sympatric populations occurred in lakes with one clade only (either 3-1 or 3-3), and in lakes having admixtures of these clades. The largest difference in gill-raker mean values between sympatric populations was found in lakes having a single 3-step clade (e.g. #69 vs. #73 in Lake Femund, and #17 vs. #19 in Lake Muddusjärvi).

Discussion

Glacial refugia and early dispersal

The European whitefish species complex comprises three major mtDNA clades, where the total NCA structure suggested a history of past fragmentation, with restricted gene flow with isolation by distance, contiguous range expansion, and allopatric fragmentation being explanations for its 3-step clades. This corroborates that the current distribution of the north and south European clade was largely covered with ice prior to c. 20 000 BP, while huge ice-free areas existed east of the Scandinavian ice sheet and surrounding the glaciated European Alps (Andersen & Borns 1994; Donner 1995). Our results are largely in agreement with the hypothesis of two main glacial refugia in the European whitefish based on an mtDNA-RFLP (mtDNA-restriction fragment length polymorphism) analysis (Bernatchez & Dodson 1994). The few samples from the Russian Republic imply that the whitefish phylogeography from this area cannot be well resolved.

We suggest that the North European clade (3-1) had a glacial refugium to the northeast of the current distribution

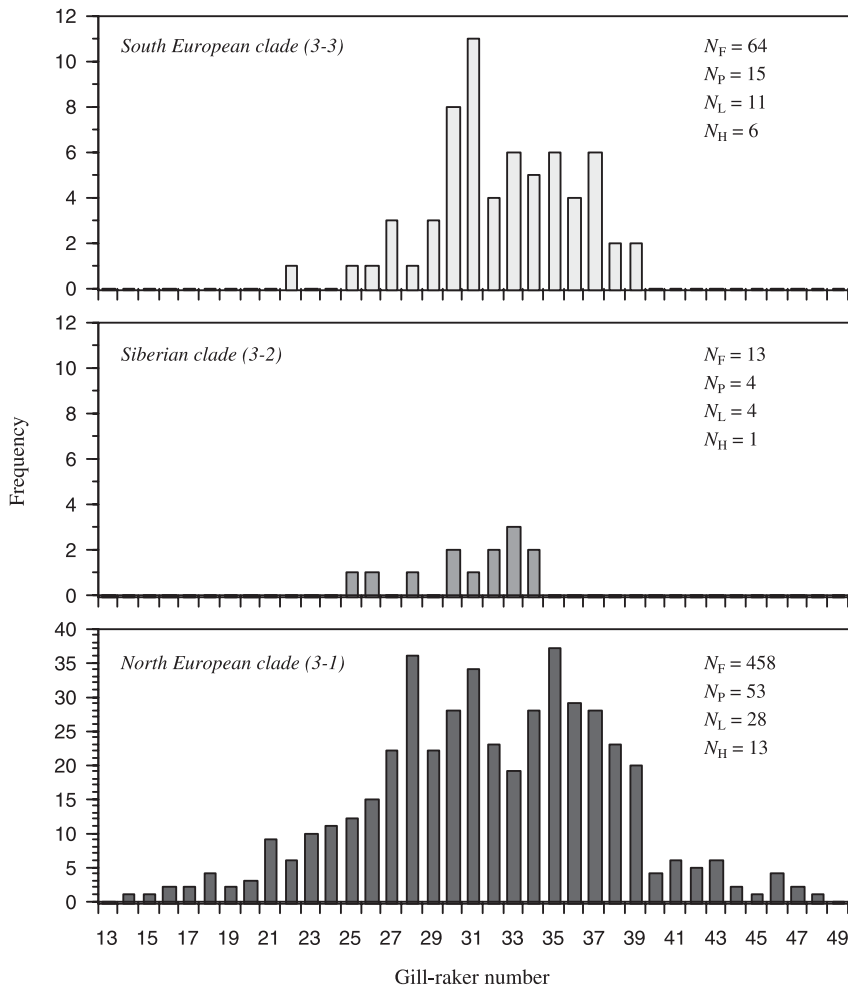


Fig. 4 The distribution of gill-raker numbers in *Coregonus lavaretus* grouped into the three observed major nested clades (N is number of: F, fish; P, populations; L, lakes; H, haplotypes). Following haplotypes were applicable; clade 3-3 (P1, P2, Q1, R1, R4, R6), clade 3-2 (K1), clade 3-1 (A1, A3, A6, A7, A9, A10, B1, B2, C1, F1, I1, I2, J1).

area in ice lakes on the western side of the Ural mountain ridge. Here, reversion of the water flow at c. 90 000 BP, created the Komi Ice Lake (2400 km²) and the connected White Sea Ice Lake (15 000 km²) (Svendsen *et al.* 1999; Mangerud *et al.* 2001a, b). Presence of whitefish in Pechora (Pymva Shor) is documented by retrieval of *Coregonus* spp. remains dated 13 000–11 000 BP (A.K. Hufthammer, Museum of Zoology, University of Bergen, Norway, personal communication). Early dispersal likely took place through the outlets of these ice lakes for long time periods towards Lake Onega and to the dislocated freshwater Baltic Sea (Maslenikova & Mangerud 2001; Mangerud *et al.* 2004). The demographic inference of the North European clade is in agreement with such a refugial occupation and dispersal pattern.

The glacial refugia for the Siberian clade (3-2) was likely on the eastern side of the Ural, in the huge West Siberian ice lake (15 000 km²) formed at the same time as the Lake Komi and White Sea ice lakes (Svendsen *et al.* 1999; Mangerud *et al.* 2001a, b). The high genetic diversity, and large number of endemic haplotypes, largely lacking in Europe, suggests a large historical effective population

size, and low gene flow with the North European (3-1) and the South European clades (3-3). The ice-free areas in Siberia, and eastwards, are likely candidates for speciation hotspots in coregonids (Smith 1957; Politov *et al.* 2004; Sukhanova *et al.* 2004). A further connection westwards was possible, with a likely transgression north of the Ural, then via the Tsilma Pass in the Timan Ridge (Mangerud *et al.* 2004), following the above-mentioned dispersal route. Demographic inference is concordant with these scenarios.

The South European clade (3-3) probably had its glacial refugia south of 53°N, which matches the maximum southward extension of Saalian glaciation (180 000–130 000 BP). Fish remains (salmonids, percids and cyprinids) from central European rivers (Oder, Elbe, Weser and Rhine) document long-term occupation from the Eemian interglacial (125 000 BP), and older, to late Pleistocene (Böhme 1997; Torke 1998). The Rhine and Elbe River were likely dispersal routes towards the Alps and Denmark after deglaciation 20 000 BP (e.g. Hansen *et al.* 1999). The demographic inference is concordant with glacial refugia in this area.

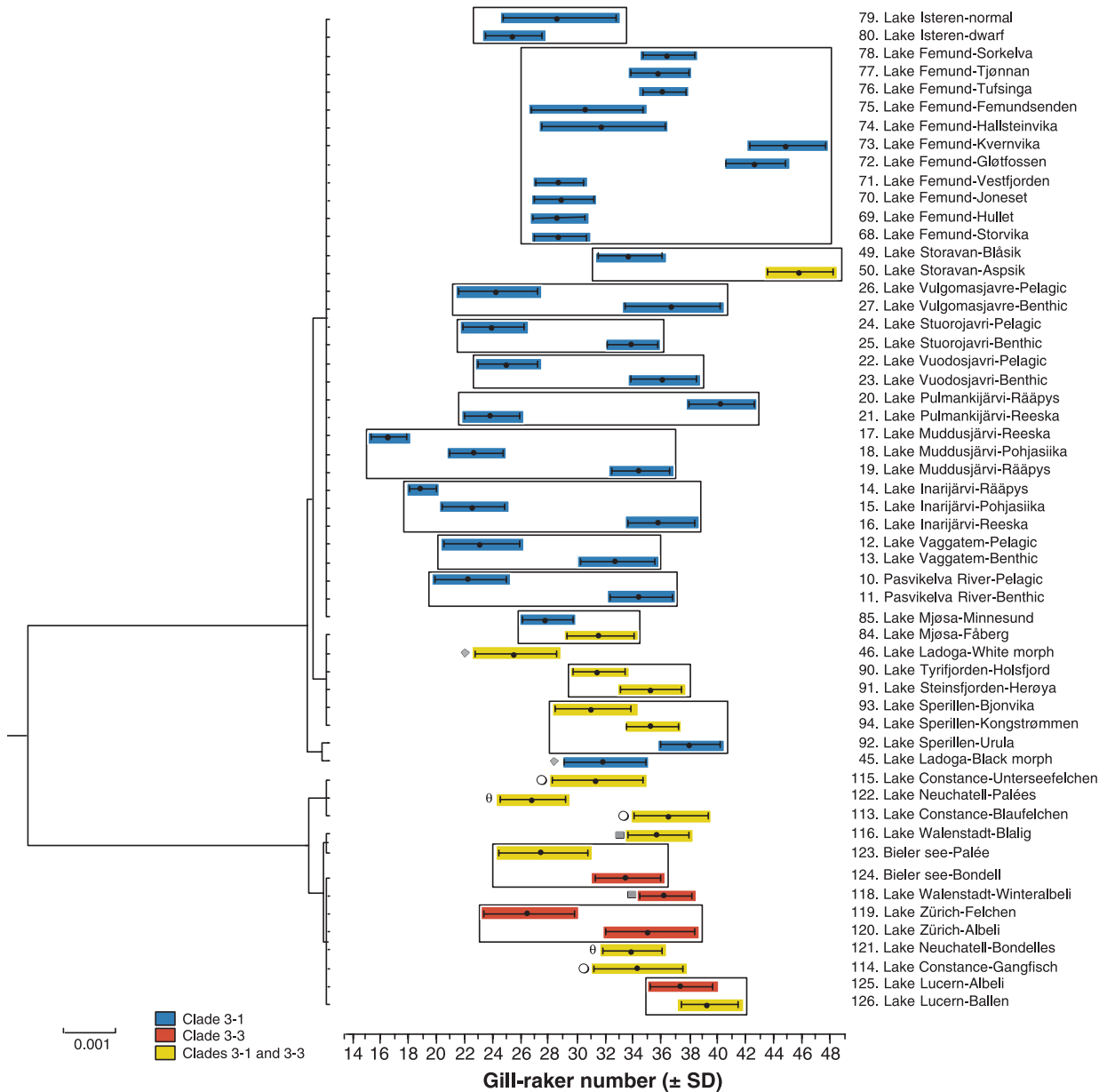


Fig. 5 The UPGMA tree from genetic distances (D_A) between sympatric populations of *Coregonus lavaretus*. Fifty-five populations (20 lakes) are compared using mean gill-raker numbers (\pm SD). Colours signify 3-step clades. Sympatric populations are enclosed within boxes, or marked using symbols.

Colonization of the Baltic Sea area from the northeast

The understanding of how European whitefish was subsequently distributed in Fennoscandia is closely related to the dynamics of the Baltic Sea (Björck 1995; Donner 1995). Remains of *C. lavaretus* in the southern Baltic Sea have been dated to 10 500 BP (K. A. Sørensen, Zoological Department, University of Copenhagen, Denmark, personal communication). Whitefish could have colonized the Baltic Sea area through the early drainages (prior to 40 000 BP) of the huge

ice lakes (Mangerud *et al.* 2004) to the dislocated pre-Baltic ice lake 100 000–25 000 BP (Houmark-Nielsen 1989). A later colonization route likely commenced through a temporal-stepwise connection from the White Sea via Lake Onega-Ladoga to the Baltic Sea (Saarnisto *et al.* 1995). However, Sendek (2004) suggests that whitefish in Lake Ladoga did not originate from the White Sea. Haplotypes (A1, A2, K1, K2) in the River Dvina (#3) and Lake Kubenskoye (#4) could potentially represent an ancestral population from the Komi Ice lake, and therefore a contact zone of the North

European and Siberian clades. The high frequency of the K1 haplotype (but also the A1 haplotype) in southwest Norway (#108), but a low frequency in the Baltic Sea, could suggest an early occurrence in, e.g. the Dogger Ice Lake 25 000–20 000 BP (Houmark-Nielsen 1989; Andersen & Borns 1994; Verhart 1995). Whitefish may even have used the early meltwater drainages after 17 000–15 000 BP for dispersal through the southern Peribaltic area toward central Poland and via 'Urstromtaler' to the Elbe River (Mangerud *et al.* 2004). Also, introgressed *Coregonus sardinella* mtDNA sequences in *Coregonus albula* in northern Germany, supports an early and long range dispersal event from Siberia (Schulz *et al.* in press). Demographic inference of the North European clade (clade 1-1 being its major component) is in line with such a scenario, with demographic expansion 11 700–5800 BP, corresponding with the development of the Baltic ice lake around 13 000–10 000 BP, and also the later stages of the Baltic Sea.

Colonization of the Baltic Sea area from the southwest

The contact zone between the north and south European clade matches remarkably well with the maximum extent of the Weichselian ice sheet 20 000 BP (Andersen & Borns 1994; Donner 1995), suggesting the latter clade colonized the Baltic Sea later. Several early to late Weichselian freshwater lakes have been discovered in Denmark west of the glacial maximum (Houmark-Nielsen 1995), where *C. lavaretus* was dated in a lake 13 000 BP (Aaris-Sørensen 1995). The North Sea Continent (Verhart 1995) could also hold whitefish in, e.g. Lough Hibernia (Maitland 1970), Dogger Lake (Svärdson 1998), and Kattogat Ice Lake (17 000–10 000 BP) (Houmark-Nielsen 1989). At 10 300 BP, the Baltic ice lake burst into the Kattogat-Skagerrak area (Björck 1995) with a massive outflow north- and westwards along the Norwegian coast, where narrow passages in the North Sea Continent concentrated freshwater in the surface layer (Stabell & Thiede 1986; Jiang *et al.* 1998). At this stage, whitefish could colonize coastal areas in Denmark, Sweden, and Norway (may be reaching Lake Liavann #108). During the marine influx 10 300–9500 BP in the Yoldia Sea stage, or in the Ancylus Lake stage 9500–8000 BP, whitefish in Denmark, or from the ice lakes, could have colonized northwards along the ice-margin (Lagerlund & Houmark-Nielsen 1993). An earlier possible connection between the Baltic ice lake and the Kattogat could have occurred 12 000 BP (Björck 1995), likely used by few immigrants. The south European clade transgressed widely into the Baltic Sea, with colonization of the adjacent lowland reaching Lake Onega and Ladoga (but see also Sendek 2004). Demographic inference suggests that clade 1-18 which originated from the Danish clade 1-17, had a demographic expansion 57 500–28 000 BP. This suggests an expansion in the North Sea Continent, or more south (likely in an ice lake), and thus,

geographical segregation of clades 1-17 and 1-18 prior to the Baltic Sea colonizations.

Colonization of the central European Alpine lakes

During or prior to the collapse of the Baltic ice lake, whitefish from the North European clade was possibly flushed out and/or could migrate to Denmark, meeting the south European clade either through the coastal areas or through early meltwater drainages via the southern Peribaltic area (see Björck 1995; Mangerud *et al.* 2004). Whitefish from these clades could then have used the Elbe, which drained to the west coast of Denmark, and the Rhine River further south to enter Alpine lakes. Hansen *et al.* (1999), based on an mtDNA survey, suggested that whitefish colonized Denmark through the large Elbe River system. We suggest that the southern clade was the first to colonize the Alps due to its high frequency and distribution in all lakes of this area, which are also closer to a southern refugium. A demographic expansion was estimated at 24 500–12 300 BP for clade 1-19, coinciding well with the retreat of the late Weichselian Ice in the Alps from 20 000 BP (Anderson & Borns 1994). Böhme (1997) refers to *Coregonus* sp. dating 11 000–10 000 BP in the River Rhine area close to Köln, which demonstrates the early presence of whitefish in that watercourse.

Phylogeographic inference compared with other freshwater fishes

Phylogeographic studies of European freshwater fishes are rapidly accumulating, with species showing similarities and differences with the patterns seen in European whitefish. In the Eurasian perch *Perca fluviatilis* L., Nesbø *et al.* (1999) found three mtDNA clades in northeast, southwest, and southeastern Europe. The first two clades match our North and South European whitefish clades. These are replicated in mtDNA lineages in the European grayling *Thymallus thymallus* L. (Koskinen *et al.* 2000). More genetic groups and clades in central European grayling have been observed (Weiss *et al.* 2002; Gum *et al.* 2005). Van Houdt *et al.* (2003) and Kontula & Väinölä (2001) studying burbot *Lota lota* L. and bullhead *Cottus gobio* L. found remarkably similar mtDNA lineages with those in perch and grayling (Koskinen *et al.* 2000), including the North and South European clades as seen in whitefish. In a global survey of burbot, Van Houdt *et al.* (2005) also found an Eurasian clade extending from central and north Europe to Lake Baikal, a pattern which could bear resemblance to our Siberian whitefish clade. More genetic groups and clades in bullhead were resolved in central Europe (Engelbrecht *et al.* 2000; Volckaert *et al.* 2002). North of the European Alps, Bernatchez (2001) mainly observed two mtDNA clades in the brown trout *Salmo trutta* L., again one clade in the southwest and

another in the northeast. Even in anadromous Atlantic salmon *Salmo salar* L., a northeast and southwest dichotomy are inferred in allozymes (Koljonen *et al.* 1999), but with a less consistent pattern in mtDNA (Nilsson *et al.* 2001).

With regards to NCA and historical demography, few studies exist in north European fish. In perch, past fragmentation into glacial refugia was inferred, while contiguous range expansion was the main force behind mtDNA haplotype distribution at lower clades, but also with restricted gene flow due to isolation by distance (Nesbø *et al.* 1999). The starlike phylogeny at lower clades in perch was also seen in lower clades in whitefish. The brown trout had past fragmentation as main historical factor, but with contiguous range expansion, and restricted gene flow at lower levels (Bernatchez 2001). The grayling had past fragmentation and range expansion as main historical events, and restricted gene flow (Weiss *et al.* 2002; Gum *et al.* 2005). Van Houdt *et al.* (2005) suggested that a range expansion could explain mtDNA diversity in burbot in Eurasia. In whitefish, all these demographic inferences are suggested along with long-distance colonization. The Atlantic lineage in brown trout had demographic expansion at 26 800–13 400 BP (Bernatchez 2001), being similar to demographic expansions of whitefish in this area.

The repeated phylogeographic pattern of these species is clearly associated with the maximum ice coverage, and the occupation of similar glacial refugia. Subsequent dispersal and secondary contact zones of genetic groups are probably species-specific due to differences in migration ability and environmental tolerance. Future attempts to localize multiple-species glacial refugia should take into consideration the dynamics of glacial refugia in time and space.

Operational taxonomical groups within the C. lavaretus species complex

The most ancient mtDNA clades (3-step) were not associated with operational taxa based on gill-rakers. This lack of congruence in genetic and phenotypic patterns questions the validity of gill-raker numbers for taxonomic designations in the *C. lavaretus* species complex. The use of gill-raker numbers may be valuable for identifying ecological distinctiveness and differentiation for practical and local management (Kaukoranta *et al.* 2000). The naming of such groups by trinomial Latin names, should however, be abandoned.

Gill-rakers, were significantly related to haplotype variation within lower-level clades, albeit with a very low degree of explained variation (R^2) in the nested model. This nesting of individual gill-raker phenotypes upon the NCA-hierarchy revealed a lack of associated demographic fragmentation events. This implies that the observed phenotypic diversification in formerly glaciated areas is of recent origin. Indeed, the observation that the most wide-

spread single haplotype (A1) encompasses the variation in gill-raker numbers seen among all the other haplotypes (Fig. 5) suggests it may reflect adaptive radiation during colonization into unoccupied habitats.

Our results support Bernatchez & Dodson (1994) who found no clear association between population genetic clustering and traditional taxonomic designations in a global phylogeny of the joint *C. clupeiformis* × *C. lavaretus* species complex using mtDNA-RFLP. Lack of concordance was also suggested for the North Sea Houting *Coregonus oxyrhynchus* compared to other European whitefish in Denmark (Hansen *et al.* 1999). Also, allozyme patterns are not highly concordant with traditional taxonomy in Eurasian coregonid species, albeit with exceptions (Ferguson *et al.* 1978; Politov *et al.* 1999; Sendek 1999, 2004).

Svärdson (1957, 1979, 1998), hypothesized the occurrence of five whitefish species in Fennoscandia based on morphology and distribution. Four species were suggested to have colonized Fennoscandia from the northeast via pro-glacial lakes, while a fifth anadromous species came from southwest of the Baltic Sea. Our results corroborate Svärdson's suggestion for colonization from northeast and southwest of the Baltic Sea area, but give no support for five species. Svärdson (1998) further suggested that these five species evolved in allopatry, but retained capacity to produce fertile offspring via hybridization. He also stressed the importance of reticulate evolution, and viewed introgressed populations as incipient species where new adaptive gene variants could generate species flocks. Such processes should produce intermediate phenotypes (e.g. Voloshenko 1973; Pethon 1974), and would blur their evolutionary history as assessed with mtDNA and morphology, because maternally inherited genomes would become decoupled from the nuclear genome and from polygenic traits such as gill-raker numbers (e.g. Turgeon & Bernatchez 2003; Seehausen 2004).

The four morphs in Lake Femund (# 68–78, Fig. 5), were considered allopatrically evolved species by Svärdson (1979, 1998; personal communication). However, a microsatellite comparison of these morphs with populations along potential routes of immigration into Lake Femund, revealed an intralacustrine origin of the morphotypic diversification (Østbye *et al.* 2005), which is supported by their minimal mtDNA diversity. Also, in other Norwegian lakes (#10–13, 22–27, Fig. 5), morphologically similar pairs of whitefish seem to have originated independently based on microsatellite data, mtDNA and life history (Østbye *et al.*, unpublished). Similar results were found by Douglas *et al.* (1999), who suggested that sympatric whitefish in European Alps likely were intralacustrine species flocks.

Our results suggest that sympatric populations may result from adaptive radiation within a single clade, and occasionally from secondary contact between clades, as suggested for lake whitefish (Bernatchez 2004). Repeated

evolution of similar phenotypes is most likely driven by divergent selection, which could be a general phenomenon throughout the species' range. In such a case, gill-raker diversity may reflect an ancient liability towards euryphagy in opportunistic whitefish colonizing areas with few competitors and predators. Coupled with the diverse spawning sites and timing, a postglacial adaptive diversification could be rather fast. A polyphyletic origin of morphs has been shown in lake whitefish (Bodaly *et al.* 1992; Lu *et al.* 2001), in ciscoes *Coregonus artedii* complex (Turgeon & Bernatchez 2003), and other members of the subgenus *Leucichthys* (Bernatchez & Dodson 1991; Politov *et al.* 2004). Thus, parallel events of diversification by natural selection into ecological niches appears to be a general mechanism for explaining diversity in European whitefish and other freshwater fishes in northern environments (Robinson & Wilson 1994; Amundsen *et al.* 2004; Campbell & Bernatchez 2004; Kahilainen *et al.* 2004; Rogers & Bernatchez 2005).

Acknowledgements

We thank all the very nice people who contributed whitefish samples (Table S1, Supplementary material), without whom a phylogeographic study would be impossible. Laboratory assistance, and discussions, were kindly provided by R. Saint-Laurent, C. Potvin, C. Landry, J. Turgeon, J. Blais, G. Lu, M. Schulz, V. Castric, S. Rogers, D. Fraser, M. Valcourt, L. Papillon, D. Garant, T. Balstad and K. Kvaløy. Thanks to C. L. Nesbø for help with phylogenetic analysis, and E. Østbye, A. Vøllestad, A. Nesje, and three anonymous referees for comments that improved the paper. The Research Council of Norway and NorFA provided funding.

Supplementary material

The supplementary material is available from <http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC2737/MEC2737sm.htm>

Table S1 *Coregonus lavaretus* population information of geographical coordinates (map code refers to Fig. 3, latitude and longitude in north and east), ecological characteristics (living habitat in lake, river or brackish water, and stationary or anadromous migration, where spawning habitat are denoted as: * river spawning; # brackish water spawning; and § known lake spawning), genetic data (haplotype distribution), and meristic data (gill-raker mean \pm SD). §, long nose. N_{mtDNA} , N_{h} and N_{GR} denote number of analysed fish in mtDNA-SSCP, observed haplotypes, and numbers used to calculate gill-raker means. Parenthesis under SD is gill-raker range. In the last column operational taxonomical group designations of *C. lavaretus* are given, based on Kaukoranta *et al.* (2000). The North American *Coregonus clupeaformis* material are summarized in the first footnote. The samples collector and sources for gill-raker counts and tissue are given in footnote using map code #. '-' denotes missing data.

Table S2 The composite mitochondrial DNA sequences (left panel: 282 bp of cyt *b*, and right panel: 246 bp of ND3) in North American (*Coregonus clupeaformis*) and European (*Coregonus lavaretus*)

whitefish. The vendace, *Coregonus albula* is used as outgroup. The most frequent haplotype (A1) is used as a base for aligning sequences, where the number given over variable nucleotide sites is comparable to the sequence of *C. lavaretus* in Myia & Nishida (2000). Haplotype positions in increasing nested clade levels are given in the first four columns as follows; 0-step (i.e. haplotypes), 1-step, 2-step and 3-step clades.

Table S3 Nested clade inference based on 55 mtDNA sequences in *Coregonus lavaretus*. Sequences are ordered in an increasing nested hierarchical level from left to right (0-step to 3-step clades). The statistically significant patterns were interpreted using the inference key at <http://darwin.uvigo.es/>. D_c , clade distance; D_{nr} , nested clade distance, I-T, interior-tip distance, Nc, not calculated (singletons), L (S) denotes significantly large (small) D_c or D_{nr} . Grey boxes are internal haplotypes or clades. Inference chains given under significant clades (in bold) are RGF-LDD, restricted gene flow/dispersal but with some long distance dispersal; CRE, contiguous range expansion; LDC, long-distance colonization (perhaps associated with recent fragmentation); AFR, allopatric fragmentation; RFG-IBD, restricted gene flow with isolation by distance; INO, inconclusive outcome; and PFR, past fragmentation (and/or long-distance colonization).

References

- Aaris-Sørensen K (1995) Paleoeecology of a Late Weichselian vertebrate fauna from Nørre Lyngby, Denmark. *Boreas*, **24**, 355–365.
- Amundsen P-A, Knudsen R, Klemetsen A, Kristoffersen R (2004) Resource competition and interactive segregation between sympatric whitefish morphs. *Annales Zoologici Fennici*, **41**, 301–307.
- Andersen BG, Borns HW Jr (1994) *The Ice Age World*, 1st edn. Scandinavian University Press, Oslo, Norway.
- Avice JC, Ball RM (1990) Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology*, **7**, 43–67.
- Behnke RJ (1972) The systematics of salmonid fishes of recently glaciated lakes. *Journal of the Fisheries Research Board of Canada*, **29**, 639–671.
- Berg LS (1962) *Freshwater Fishes of the USSR and Adjacent Countries*, 4th edn. Israel Program for Scientific Translations, Jerusalem.
- Bergstrand E (1982) The diet of four sympatric whitefish species in Lake Parkijaure. *Report of the Institute of Freshwater Research, Drottningholm*, **60**, 5–14.
- Bermingham E, McCafferty S, Martin A (1997) Fish biogeography and molecular clocks, perspectives from the Panamanian Isthmus. In: *Molecular Systematics of Fishes* (eds Kocher T, Stepien C), pp. 113–128. Academic Press, San Diego, California.
- Bernatchez L (1995) A role for molecular systematics in defining evolutionary significant units in fishes. *American Fisheries Society Symposium*, **17**, 114–132.
- Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351–379.
- Bernatchez L (2004) Ecological theory of adaptive radiation. An empirical assessment from *Coregonine* fishes (Salmoniformes). In: *Evolution Illuminated, Salmon and Their Relatives* (eds Hendry AP, Stearns S), pp. 175–207. Oxford University Press, New York.
- Bernatchez L, Dodson JJ (1991) Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) in

- North America and its relationship to Pleistocene glaciations. *Evolution*, **45**, 1016–1035.
- Bernatchez L, Dodson JJ (1994) Phylogenetic relationships among Palearctic and Nearctic whitefish (*Coregonus* sp.) populations as revealed by mtDNA variation. *Canadian Journal of the Fisheries and Aquatic Sciences*, **51**, 240–251.
- Björck S (1995) A review of the history of the Baltic Sea, 13.0–8.0 ka BP. *Quaternary International*, **27**, 19–40.
- Bodaly RA (1979) Morphological and ecological divergence within the lake whitefish (*Coregonus clupeaformis*) species complex in Yukon Territory. *Journal of the Fisheries Research Board of Canada*, **36**, 1214–1222.
- Bodaly RA, Clayton JW, Lindsey CC, Vuorinen J (1992) Evolution of lake whitefish (*Coregonus clupeaformis*) in North America during the Pleistocene: genetic differentiation between sympatric populations. *Canadian Journal of the Fisheries and Aquatic Sciences*, **49**, 769–779.
- Böhme G (1997) Fossile Fischfaunen aus dem jüngeren Känozoikum Deutschlands. *Quartär*, **47/48**, 113–138.
- Campbell D, Bernatchez L (2004) Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, **21**, 945–956.
- Clement M, Posada D, Crandall KA (2000) tcs: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959–969.
- Dieckmann U, Doebeli M, Metz JAJ, Tautz D (2004) *Adaptive Speciation*. Cambridge University Press, Cambridge, UK.
- Donner J (1995) *The Quaternary History of Scandinavia*. Cambridge University Press, Cambridge, UK.
- Douglas MR, Brunner PC, Bernatchez L (1999) Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the central Alpine region of Europe represent species flocks? *Molecular Ecology*, **8**, 589–603.
- Engelbrecht CC, Freyhof J, Nolte A *et al.* (2000) Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-Pleistocene origin of the major central European populations. *Molecular Ecology*, **9**, 709–722.
- Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853–864.
- Ferguson AK, Himberg K-JM, Svärdsön G (1978) Systematics of the Irish pollan (*Coregonus pollan*, Thompson): an electrophoretic comparison with other Holarctic Coregoninae. *Journal of Fish Biology*, **12**, 221–233.
- Fraser D, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for deriving conservation units. *Molecular Ecology*, **19**, 2741–2752.
- Fu Y-X (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics*, **143**, 557–570.
- Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, New Jersey.
- Gum B, Gross R, Kuehn R (2005) Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. *Molecular Ecology*, **14**, 1707–1725.
- Hammer MF, Karafet T, Rasanayagam A *et al.* (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Molecular Biology and Evolution*, **15**, 427–441.
- Hansen MM, Mensberg K-LD, Berg S (1999) Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. *Molecular Ecology*, **8**, 239–252.
- Hermida M, Fernández C, Amaro R *et al.* (2002) Heritability and ‘evolvability’ of meristic characters in a natural population of *Gasterosteus aculeatus*. *Canadian Journal of Zoology*, **80**, 532–541.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Himberg K-JM (1970) A systematic and zoogeographic study of some North European Coregonids. In: *Biology of Coregonid Fishes* (eds Lindsey CC, Woods CS), pp. 219–250. University of Manitoba Press, Winnipeg, Canada.
- Houmark-Nielsen M (1989) The last interglacial-glacial cycle in Denmark. *Quaternary International*, **3/4**, 31–39.
- Houmark-Nielsen M (1995) Late Pleistocene stratigraphy and geomorphology: outline of the excursion area and features. In: *Quaternary Fieldtrips in Central Europe, INQUA XIV International Congress* (ed. Schirmer W), pp. 90–101. Verlag Friderich Pfeil, München.
- Järvi TH (1928) Über die Arten und Formen der Coregonen s. str. in Finland. *Acta Zoologica Fennici*, **5**, 1–259.
- Jiang H, Svensson NO, Björck S (1998) Meltwater discharge to the Skagerrak-Kattegat from the Baltic ice lake during the younger Dryas interval. *Quaternary Research*, **49**, 264–270.
- Kahilainen K, Malinen T, Tuomaala A, Lehtonen H (2004) Diel and seasonal habitat and food segregation of three sympatric *Coregonus lavaretus* forms in a subarctic lake. *Journal of Fish Biology*, **64**, 1–17.
- Kaukoranta M, Koljonen M-L, Koskiniemi J, Pennanen J, Tammi J (2000) *Atlas of Finnish Fishes. English summary. Distribution of lamprey, brook lamprey, salmon, trout, Arctic charr, whitefish, vendace, grayling, asp, vimba, spined loach and bullhead, and status of the stocks*. Research Report, Finnish Game and Fisheries Research Institute, Oy Edita Ab Press, Helsinki, Finland.
- Koljonen M-L, Jansson H, Paaver T *et al.* (1999) Phylogeographic lineages and differentiation pattern of Atlantic salmon (*Salmo salar*) in the Baltic Sea with management implications. *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 1766–1780.
- Kontula T, Väinölä R (2001) Postglacial colonization of Northern Europe by distinct phylogeographical lineages of the bullhead, *Cottus gobio*. *Molecular Ecology*, **10**, 1983–2002.
- Koskinen MT, Ranta E, Piironen J *et al.* (2000) Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Molecular Ecology*, **9**, 1609–1624.
- Kottelat M (1997) European freshwater fishes. An heuristic checklist of the freshwater fishes of Europe (exclusive of former USSR), with an introduction for non-systematists and comments on nomenclature and conservation. *Biologia, Section Zoology*, Bratislava, **52** (Suppl. 5), 1–271.
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.
- Lagerlund E, Houmark-Nielsen M (1993) Timing and pattern of the last deglaciation in the Kattegat region, southwest Scandinavia. *Boreas*, **22**, 337–347.
- Linnaeus C (1758) *Systema Naturae*, 10th edn. Laurentii Salvii, Stockholm, Sweden.
- Lu G, Basley J, Bernatchez L (2001) Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization

- between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. *Molecular Ecology*, **10**, 965–985.
- Maitland PS (1970) The origin and present distribution of *Coregonus* in the British Isles. In: *Biology of Coregonid Fishes* (eds Lindsey CC, Woods CS), pp. 99–114. University of Manitoba Press, Winnipeg, Canada.
- Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology & Evolution*, **10**, 294–299.
- Mangerud J, Astakhov V, Jakobsson M, Svendsen JI (2001a) Huge ice-lakes in Russia. *Journal of Quaternary Science*, **16**, 773–777.
- Mangerud J, Astakhov V, Murray A, Svendsen JI (2001b) The chronology of a large ice-dammed lake and the Barents–Kara ice sheet advances, northern Russia. *Global and Planetary Change*, **31**, 319–334.
- Mangerud J, Jakobsson M, Alexanderson H *et al.* (2004) Ice-dammed lakes and rerouting of the drainage of northern Eurasia during the last glaciation. *Quaternary Science Reviews*, **23**, 1313–1332.
- Maslenikova O, Mangerud J (2001) Where was the outlet of the ice-dammed Lake Komi, northern Russia? *Global and Planetary Change*, **31**, 337–345.
- Miya M, Nishida M (2000) Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. *Molecular Phylogenetics and Evolution*, **17**, 437–455.
- Nei M (1987) *Molecular Evolutionary Genetics*, 1st edn. Columbia University Press, New York.
- Nesbø CL, Fosshem LA, Vøllestad A *et al.* (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, **8**, 1387–1404.
- Nilsson J, Gross R, Asplund T *et al.* (2001) Matrilinial phylogeography of Atlantic salmon (*Salmo salar*) in Europe and postglacial colonization of the Baltic Sea area. *Molecular Ecology*, **10**, 89–102.
- Østbye K, Næsje TF, Bernatchez L, Sandlund OT, Hindar K (2005) Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund, Norway. *Journal of Evolutionary Biology*, **18**, 683–702.
- Pethon P (1974) Naturally occurring hybrids between whitefish (*C. lavaretus*) and cisco (*C. albula*) in Lake Orrevann. *Norwegian Journal of Zoology*, **22**, 287–293.
- Pinceel J, Jordaens K, Van Houtte N, De Vinter AJ, Backejau T (2004) Molecular and morphological data reveal cryptic taxonomic diversity in the terrestrial slug complex *Arion subfuscus/fuscus* (Mollusca, Pulmonata, Arionida) in continental north-west Europe. *Biological Journal of the Linnean Society*, **83**, 23–38.
- Polotov DV, Bickham JW, Patton JC (2004) Molecular phylogeography of Palearctic and Nearctic ciscoes. *Annales Zoologica Fennici*, **41**, 13–23.
- Polotov DV, Yu N, Makhrov G, Makhrov AA (1999) Genetic identification and taxonomic relationships of six Siberian species of *Coregonus*. *Archivum für Hydrobiologia, Special Issues in Advanced Limnology*, **57**, 21–34.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.
- Puerto G, Da Graca Salomao M, Theakston RDG *et al.* (2001) Combining mitochondrial DNA sequences and morphological data to infer species boundaries: phylogeography of lancehead pitvipers in the Brazilian Atlantic forest, and the status of *Bothrops pradoi* (Squamata: Serpentes: Viperidae). *Journal of Evolutionary Biology*, **14**, 527–538.
- de Queiroz K (1998) The general lineage concept of species: species criteria and the process of speciation. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 57–75. Oxford University Press, Oxford, UK.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.
- Reshetnikov YS (2004) Coregonid fishes in arctic waters. *Annales Zoologica Fennici*, **41**, 3–11.
- Robinson BW, Wilson DS (1994) Character release and displacement in fishes: a neglected literature. *American Naturalist*, **144**, 596–627.
- Rogers AR (1995) Genetic evidence for a Pleistocene population expansion. *Evolution*, **49**, 608–615.
- Rogers SM, Bernatchez L (2005) Integrating QTL mapping and genomic scans towards the characterization of candidate loci under parallel directional selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, **14**, 351–361.
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.
- Rozas J, Sánchez-DelBarrio JC, Messegyer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
- Saarnisto M, Grönlund T, Ekman I (1995) Late glacial of Lake Onega — contribution to the history of the eastern Baltic basin. *Quaternary International*, **27**, 111–120.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sandlund OT (1992) Differences in the ecology of two vendace populations separated in 1895. *Nordic Journal of Freshwater Research*, **67**, 52–60.
- Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- Schulz M, Freyhof J, Saint-Laurent R, Østbye K, Mehner T, Bernatchez L (2005) Evidence for independent evolution two spring-spawning ciscoes of the *Coregonus albula* (L.) species group (Salmoniformes: Coregonidae) in Germany. *Journal of Fish Biology*. (in press).
- Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, **19**, 198–207.
- Sendek DS (1999) Electrophoretic studies of Coregonid fishes from across Russia. *Archivum für Hydrobiologia, Special Issues in Advanced Limnology*, **57**, 35–55.
- Sendek DS (2004) The origin of sympatric forms of European whitefish (*Coregonus lavaretus* (L.)) in Lake Ladoga based on comparative genetic analysis of populations in North-west Russia. *Annales Zoologica Fennici*, **41**, 25–39.
- Sites JW, Marshall JC (2003) Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution*, **18**, 462–470.
- Smith SH (1957) Evolution and distribution of the coregonids. *Journal of the Fisheries Research Board of Canada*, **14**, 599–604.
- Stabell B, Thiede J (1986) Paleobathymetry and paleogeography of southern Scandinavia in the late Quaternary. *Meyniana*, **38**, 43–59.

- Sukhanova LV, Smirnov VV, Smirnova-Zalumi NS, Kirilchik SV, Shimuzu I (2004) Grouping of Baikal Omul *Coregonus autumnalis migratorius* Georgi within the *C. lavaretus* complex confirmed by using a nuclear DNA marker. *Annales Zoologica Fennici*, **41**, 41–49.
- Svårdson G (1950) The coregonid problem: II. Morphology of two coregonid species in different environments. *Report of the Institute of Freshwater Research, Drottningholm*, **31**, 151–162.
- Svårdson G (1957) The coregonid problem. VI. The Palearctic species and their intergrades. *Report of the Institute of Freshwater Research, Drottningholm*, **38**, 267–356.
- Svårdson G (1970) Significance of introgression in coregonid evolution. In: *Biology of Coregonid Fishes* (eds Lindsey CC, Woods CS), pp. 33–59. University of Manitoba Press, Winnipeg, Canada.
- Svårdson G (1979) Speciation of Scandinavian *Coregonus*. *Report of the Institute of Freshwater Research, Drottningholm*, **57**, 1–95.
- Svårdson G (1998) Postglacial dispersal and reticulate evolution of Nordic coregonids. *Nordic Journal of Freshwater Research*, **74**, 3–32.
- Svendsen J, Astakhov VI, Bolshiyarov DY *et al.* (1999) Maximum extent of the Eurasian ice sheets in the Barents and Kara Sea region during the Weichselian. *Boreas*, **28**, 234–242.
- Swofford DL (2001) *PAUP* version 4.0b10. Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1989a) Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.
- Taylor EB (1999) Species pairs of north temperate freshwater fishes: taxonomy, evolution and conservation. *Reviews in Fish Biology and Fisheries*, **9**, 299–324.
- Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypothesis about gene flow and population history. *Molecular Ecology*, **7**, 381–397.
- Templeton AR (2001) Using phylogeographic analysis of gene trees to test species status and processes. *Molecular Ecology*, **10**, 779–791.
- Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.
- Torke W (1998) Fische aus jungpleistozänen und holozänen siedlungsplätzen in Baden-Württemberg. *Naturkunde Württemberg*, **154**, 231–259.
- Turgeon J, Bernatchez L (2003) Reticulate evolution and phenotypic diversity in North American ciscoes, *Coregonus* ssp. (Teleostei: Salmonidae): implications for the conservation of an evolutionary legacy. *Conservation Genetics*, **4**, 67–81.
- Van Houdt JK, De Cleyn L, Perretti A, Volckaert AM (2005) A mitogenic view on the evolutionary history of the Holarctic freshwater gadoid, burbot (*Lota lota*). *Molecular Ecology*, **14**, 2445–2457.
- Van Houdt JK, Hellemans B, Volckaert FAM (2003) Phylogenetic relationship among Palearctic and Nearctic burbot (*Lota lota*): Pleistocene extinctions and recolonization. *Molecular Phylogenetics and Evolution*, **29**, 599–612.
- Verhart LBM (1995) Fishing for the Mesolithic. In: *Man and Sea in the Mesolithic. Coastal Settlement Above and Below Present Sea Level* (ed. Fisher A), pp. 291–302. Oxbow Monograph 53, Oxford, UK.
- Volckaert FAM, Hänfling B, Hellemans B, Carvalho GR (2002) Timing of the population dynamics of bullhead *Cottus gobio* (Teleostei: Cottidae) during the Pleistocene. *Journal of Evolutionary Biology*, **15**, 930–944.
- Voloshenko BB (1973) A comparative analysis of the feeding of underyearlings of the Pelyad [*Coregonus peled* (Gmelin)], the Broad whitefish [*Coregonus nasus* (Pallas)] and their hybrids when reared together. *Journal of Ichthyology*, **13**, 569–576.
- Weiss S, Persat H, Eppe R *et al.* (2002) Complex pattern of colonization and refugia revealed for European grayling *Thymallus thymallus*, based on complete sequencing of the mitochondrial DNA control region. *Molecular Ecology*, **11**, 1393–1407.
- Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Systematic Biology*, **51**, 69–91.

Kjartan Østbye is currently studying European whitefish (PhD), but has additional interest in faunal history, ancient DNA, and cave adaptation. L. Bernatchez is interested in understanding patterns and processes of molecular and organismal evolution, and their relevance to conservation. Tor Fredrik Næsje works with population ecology, and effects of human encroachment, resource exploitation and natural environments on fish populations. Mikael Himberg studies taxonomy and evolution of Coregonids (PhD), and variation in biochemical and morphological characters in relation to environment. Kjetil Hindar studies population genetics and ecology of salmonid fishes, and how their genetic structure is affected by humans.
