Molecular Ecology (2005) **14**, 2445–2457 doi: 10.1111/j.1365-294X.2005.02590.x

A mitogenic view on the evolutionary history of the Holarctic freshwater gadoid, burbot (*Lota lota*)

J. K. J. VAN HOUDT,* L. DE CLEYN,* A. PERRETTI*† and F. A. M. VOLCKAERT*
**Laboratory of Aquatic Ecology, Katholieke Universiteit Leuven, Ch. de Bériotstraat 32, B-3000 Leuven, Belgium, †University of Padova, Dipartimento di Biologia, Via Ugo Bassi 58-B, I-35121 Padova, Italy*

Abstract

Climatic oscillations during the Pleistocene epoch had a dramatic impact on the distribution of biota in the northern hemisphere. In order to trace glacial refugia and postglacial colonization routes on a global scale, we studied mitochondrial DNA sequence variation in a freshwater fish (burbot, *Lota lota*; Teleostei, Gadidae) with a circumpolar distribution. The subdivision of burbot in the subspecies *Lota lota lota* (Eurasia and Alaska) and *Lota lota maculosa* (North America, south of the Great Slave Lake) was reflected in two distinct mitochondrial lineages (average genetic distance is 2.08%). The *lota* form was characterized by 30 closely related haplotypes and a large part of its range (from Central Europe to Beringia) was characterized by two widespread ancestral haplotypes, implying that transcontinental exchange/migration was possible for cold-adapted freshwater taxa in recent evolutionary time. However, the derived mitochondrial variants observed in peripheral populations point to a recent separation from the core group and postglacial recolonization from distinct refugia. Beringia served as refuge from where *L. l. lota* dispersed southward into North America after the last glacial maximum. Genetic variation in the *maculosa* form consisted of three mitochondrial clades, which were linked to at least three southern refugia in North America. Two mitochondrial clades east of the Continental Divide (Mississippian and Missourian clades) had a distinct geographical distribution in the southern refuge zones but intergraded in the previously glaciated area. The third clade (Pacific) was exclusively found west of the Continental Divide.

Keywords: control region, fish, Gadiformes, mitochondrial DNA, phylogeography

Received 23 December 2004; revision received 01 April 2005; accepted 01 April 2005

Introduction

Since the beginning of the Quaternary (2.4 Myr ago), climatic oscillations have driven the periodic advance and recession of ice sheets in the northern hemisphere (Hewitt 1996; Hewitt 2000; Hewitt 2004). This process strongly determined the evolutionary history of northern temperate freshwater fishes (Bernatchez & Wilson 1998). The Pleistocene glaciations destroyed and altered hydrological systems across Eurasia and North America and shaped the Holocene drainage basins (Gibbard 1988; Dyke *et al*. 2002; Grosswald & Hughes 2002). The chance that aquatic organisms survived these dramatic habitat shifts was determined by hydrological continuity to escape the advancing ice sheet and by their capacity to face a changing environment. During glacial advances, the displaced populations had to survive along the glacial margins in a compressed biome, inevitably leading to reduced genetic diversity (Hewitt 1996, 2000). Subsequent warming melted the large ice mass, and the melt waterblocked by the retreating ice sheet, accumulated in large proglacial lakes, which interconnected previously isolated regions and facilitated dispersal over vast geographical ranges (Bernatchez & Wilson 1998; Krinner *et al*. 2004).

Since aquatic species inhabited the same hydrological systems when reacting to the changing environment, congruent phylogeographical patterns are expected (Taberlet *et al*. 1998; Hewitt 1999). Comparative phylogeographical analysis revealed that northern freshwater fish species are

Correspondence: Dr Jeroen Van Houdt, Katholieke Universiteit Leuven, Laboratory of Aquatic Ecology, Ch. de Bériotstraat 32, B-3000 Leuven, Belgium. Fax: +32 16 32 45 75; E-mail: jeroen.vanhoudt@bio.kuleuven.ac.be

characterized by relatively young clades (averaging 0.5–2% sequence divergence over the entire mitochondrial genome) and are dominated by fewer but more widely dispersed haplotypes (Bernatchez & Wilson 1998). However, there is a less trivial concordance among them in regards to clade distribution, which might be attributed to a number of factors. The survival and sorting of lineages depends on stochastic demographic processes, which might result in different phylogeographical patterns regardless of common biogeographical events. But also differences in ecologically relevant traits and the competitive ability during dispersal and colonization might influence the observed patterns (Turgeon & Bernatchez 2001). Furthermore, the comparison of ancient and contemporary DNA sequences revealed that late Pleistocene histories may be more complex than those inferred from either the fossil record or contemporary DNA sequences alone (Leonard *et al*. 2000; Barnes *et al*. 2002). Therefore, information on the phylogenetic structure of wide-ranging fishes from cold temperate regions are required to identify which factors, ecological and/or demographic, are truly predictive of phylogeographical structure of young but distinct evolutionary lineages (Turgeon & Bernatchez 2001).

The burbot (*Lota lota*; Teleostei: Gadidae) is a cold-adapted freshwater fish that lives as a benthic carnivore in slow-running rivers and oligotrophic lakes in Eurasia and North America, southward to about 40°N (Nelson 1994). Along with northern pike (*Esox lucius*), burbot has the widest longitudinal range of any freshwater fish. Fossil evidence suggests that the genus already inhabited European rivers in the early Pliocene (*Lota hulai*; Pietschmann 1934), indicating that this genus is at least 5 Myr old. However, a molecular survey of cytochrome *b* polymorphism revealed that the present-day genetic variation is merely 1 Myr old (Van Houdt *et al*. 2003). The two major evolutionary lineages (Eurasian-Beringian: EB and North American lineage: NA) differed on average 1.25% and their geographical distribution was congruent with the historically described subspecies, *L. lota lota* and *L. lota maculosa*. The latter taxon lives in North America, south of the Great Slave Lake, and the former in Eurasia and the remainder of the Nearctic region. Thus, burbot is a widespread genetically young species that is able to disperse over large distances (Sorokin 1971) and live in close association with glaciated regions. Therefore, the species is eminently suited to study the impact of Pleistocene glaciations on global phylogeographical patterns and genetic diversity.

This study traces back how burbot evolved in space and time, which is essential to understand population divergence and speciation processes. It provides an evolutionary framework for local genetic and ecological studies. In order to elucidate the evolutionary history of burbot, we studied the genetic variation at the fast-evolving mitochondrial control region locus. Genealogical analysis provided information on the chronology of past evolutionary events while a spatial analysis of molecular variance revealed the geographical structuring of the genetic diversity. Additionally, the demographic history was evaluated by means of mismatch analyses and neutrality tests.

Materials and methods

Samples

Burbot was sampled across the Holarctic. Fin clips were collected and preserved in 100% ethanol. A total of 255 individuals were collected from 45 different locations, representing the major river basins across its distribution range. Sampling information is listed in Appendix 1 and sampling locations are pictured in Fig. 1.

DNA extraction, amplification and sequencing

DNA was extracted using the QIAGEN Tissue Kit (Westburg) according to the manufacturers instructions. The complete mitochondrial control region (CR) was amplified using the L19ProGm (5′-CCACTAGCTCCCAAAGCTAGA-3′) and H17Phe primer (5′-ACTTTCTAGGGTCCATCCTA-3′), respectively, adjusted from the L19 and H17 primers of Bernatchez *et al*. (1992). This resulted in a fragment of approximately 900 bp. A smaller fragment of 450 bp at the 5′ end of the mitochondrial CR was amplified with the L19ProGm primer and the HDL400Ll primer (5′-GATTTAGGATTTATGTACTCC-3′). The polymerase chain reaction (PCR) mix (25 μL) consisted of 2 mM $MgCl_2$, 0.8 μM of each primer, 200 μM of each dNTP, 0.5 units *Taq* polymerase (Silverstar, Eurogentec), and 1× reaction buffer (Eurogentec). The reaction product was purified with the GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences), eluted in 25 μL of 10 mM Tris-Cl (pH 8.0) and sequenced with SequiTherm Excell II Kit (Epicentres Technologies) according to the manufacturers' protocol. The reactions were performed bidirectionally and both sequencing primers were used in the same reaction. The reaction products were analysed on an automatic sequencer LI-COR GeneReadIR DNA System (LI-COR). Control region sequences determined in this study have been deposited in the GenBank/EMBL database under Accession nos AY656840–AY656915.

Genealogical and genetic diversity analysis

The CR sequences were aligned using CLUSTAL_X with the default options (Thompson *et al*. 1997) and checked by eye. TCS version 1.3 (Clement *et al*. 2000) was used to build a statistical parsimony network (SP: Templeton *et al*. 1992) and NETWORK version 4.0.0.1 (http://www.fluxus-engineering.com) was used to construct a reduced median

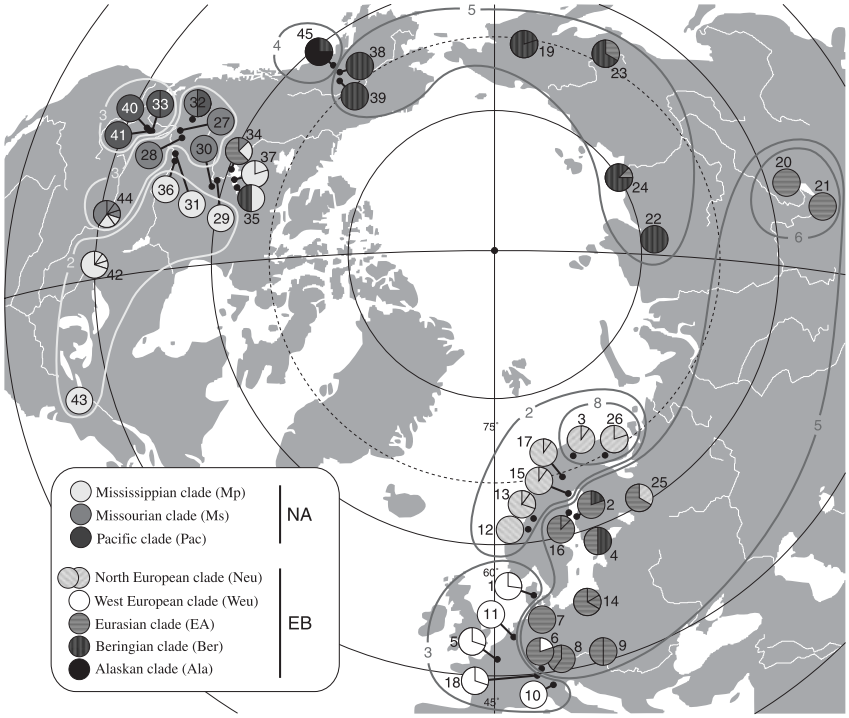

Fig. 1 The 45 sampling locations of burbot (*Lota lota*) across the Holarctic region and the distribution of the two major mitochondrial lineages. The Eurasian-Beringian lineage (EB) and the North American lineage (NA) correspond to the subspecies *Lota lota lota* and *Lota lota maculosa* respectively (Van Houdt *et al*. 2003). Additional information on the locations and haplotype frequencies are listed in Appendix 1. The pies depict the geographical distribution of the various mitochondrial clades, and the group structure defined by the SAMOVA is delineated by light grey lines for the NA phylogroup en dark grey lines for the EB phylogroup; the associated number refers to the grouping level (for more details see text).

network (RM: Bandelt *et al*. 2000) and a median-joining network (MJ: Bandelt *et al*. 1999). Information on allele frequencies and geographical subdivision was used to infer the most plausible minimum spanning tree (MST) according to the maximum variance parsimony criterion (Excoffier & Smouse 1994). These calculations were performed in ARLEQUIN version 2.0 (Schneider *et al*. 2000) using the geographical structure that was determined by SAMOVA. Mitochondrial DNA polymorphism within groups was estimated as haplotypic (nucleon) diversity (h) and nucleotide diversity (π; Nei & Kumar 2000). A spatial analysis of molecular variance (SAMOVA version 1.0; Dupanloup *et al*. 2002) was performed to define groups of populations that were geographically homogenous and maximally differentiated from each other. Hereto the method combines geographical information with an analysis of molecular variance (AMOVA; Excoffier *et al*. 1992). A simulated annealing procedure is used for maximizing the proportion of total genetic variance due to differences between groups of geographically homogenous populations. The SAMOVA was done based on pairwise differences and for a varying number of groups (K); for each grouping the simulated annealing process was repeated 100 times, each time starting from a different initial position.

Mismatch analysis

Historic demographic/spatial expansions were investigated by Tajima's (1989) test of neutrality (Bertorelle & Slatkin 1995; Aris-Brosou & Excoffier 1996) and by examining the frequency distribution of pairwise differences between sequences (mismatch distribution; Slatkin & Hudson 1991; Rogers & Harpending 1992). Because population structure has a limited effect on the mismatch distribution (Rogers 1995; Ray *et al*. 2003; Excoffier 2004), the sudden-expansion model was tested for different assemblage levels using ARLEQUIN version 2.0 (Schneider *et al*. 2000). Various expansion parameters (τ, θ_0, and θ_1) were estimated by a general nonlinear least-squares approach, and confidence intervals were obtained by parametric bootstrapping. Here, τ, θ_0 and θ_1 represent, respectively, the mutational time since expansion, the mutation parameter before expansion and the mutation parameter following expansion for the theoretical mismatch distribution that best describes the observed one (Schneider & Excoffier 1999). The concordance of the observed with the expected distribution under the sudden-expansion model was tested using a parametric bootstrap approach with the sum of squared deviations (SSD) between observed and expected mismatch as test statistic.

Results

Initially, we sequenced a fragment of approximately 830 bp for 141 burbot originating from various locations across their natural range. The fragment corresponds to the region between nucleotide 15 701 and 16 527 of the complete mitochondrial genome of *Lota lota* (AP004412;

Miya *et al*. 2003). The first 23 bp were part of the tRNA-Proline gene. The length of the remaining fragment varied from 803 to 805 bp and represented the complete control region (CR) of *L. lota*. After aligning the 141 burbot sequences, we observed 26 polymorphic sites: five singleton polymorphic sites and 21 parsimony informative sites. Interestingly, more than 90% (24 polymorphic sites) of the observed variation was located in the first 400 bp of the CR. Consequently, further sequencing efforts focused on the 5′ part of the CR and the total number of analysed fish was increased to 255. For a fragment of 425 bp, we found 44 different haplotypes determined by 28 polymorphic sites; seven singletons and 21 parsimony informative sites.

Genealogical analysis

The statistical parsimony, median-joining, and reduced median-joining analyses revealed two distinct mitochondrial lineages (Fig. 2), previously identified as the Eurasian-Beringian (EB) and the North American (NA) lineage by means of cytochrome *b* sequences (Van Houdt *et al*. 2003). It took at least five mutations to link haplotypes from both assemblages, whereas the most divergent haplotypes

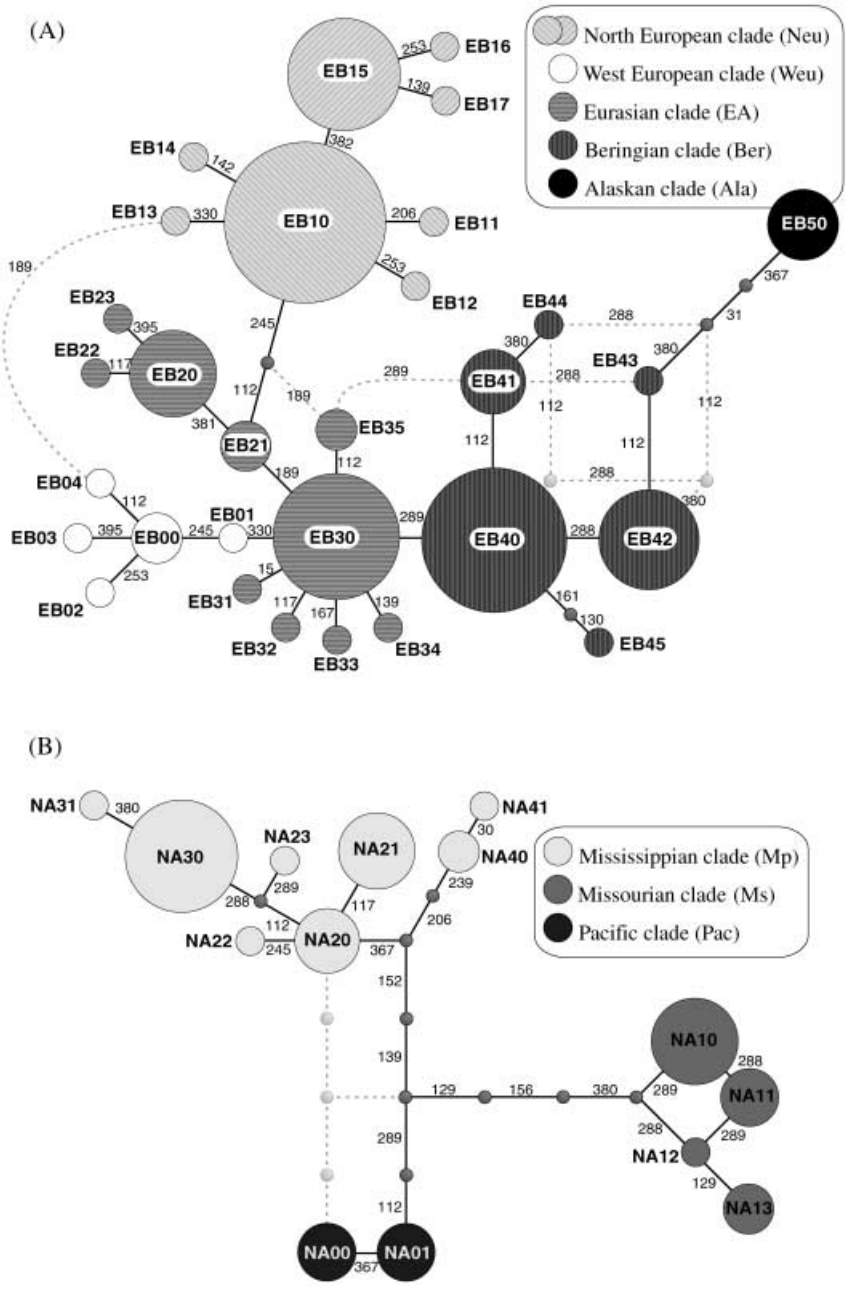

Fig. 2 Reduced median network representing the genealogical relationships between the 44 mitochondrial control region haplotypes of burbot (*Lota lota*): (A) the Eurasian-Beringian lineage (EB) and (B) the North American lineage (NA). The surface of the circles is proportional to their abundance. The geographical distribution of the mitochondrial variants is pictured in Fig. 1; detailed information is provided in Tables 1 and 2.

differed by 14 mutations. On average, both groups differed by 8.8 mutations or 2.08%. Despite these small differences, the precise relationship between EB and NA assemblage could not be clearly determined. Homoplasy at some hypervariable positions in the CR fragment resulted in multiple connections through equally parsimonious pathways.

The different genealogical approaches (SP, RM, MJ and MST) produced a similar output for the EB-lineage. A consensus genealogy with alternative connections (broken lines) is pictured in Fig. 2. Initially, haplotype EB04 and EB13 were disregarded from the network analysis as they enforced the addition of multiple dimensions in the RM network. Furthermore, they connected EB10 to E00 and were as such considered as internal instead of derived haplotypes. Given that both haplotypes were encountered only once in a single population, a tip position seems more likely. The RM network evoked a third dimension (broken line), depicting the alternative pathways to connect EB50 to the EB network. Theoretical predictions from the coalescent theory were used to resolve the cycles in the network (Posada & Crandall 2001) and preserve the most plausible pathways (solid lines). EB50 differed in three mutations from the EB lineage, whereas it required at least five mutations for connection to the NA lineage.

Within the NA assemblage, three clades could be distinguished, namely the Mississippian (NA20-NA23, NA30, NA31, NA40, NA41), the Missourian (NA10-NA13) and the Pacific (NA00, NA01). In the statistical parsimony and median-joining analyses, the latter clade served as an intermediate connection between the two former ones, and the three clusters could be linked unambiguously. However, in the reduced median network, a more complex relation was revealed. The three clades were linked through multiple, equally parsimonious evolutionary pathways (Fig. 2), suggesting homoplasy at some positions. The indel at position 380 and substitutions at positions 112, 129, 288 and 289 occurred at least twice at different positions in the NA network. Under the maximum parsimony criterion, the solid pathway connecting the three Na clades was most plausible. But it did not resolve the reticulations within the Missourian clade.

Geographical distribution of mitochondrial diversity

The Eurasian-Beringian (EB) phylogroup was found in Europe, throughout Asia, Siberia, Alaska and in Canada as far south as the Great Slave Lake. The North American phylogroup (NA) was observed in North America from the southern margin of the distribution range (approximately 40° NB) up to the Great Slave Lake. A SAMOVA was performed separately for both regions because of their distinct geographical association. The samples from the Great Slave Lake, including Pine Point (34), Redcliff Island (35) and Simpson Island (37; Table 2; GSL-group), were disregarded in these analyses since both evolutionary lineages (EB and NA) were present.

The between-group variance for the EB lineage was maximized for a varying number of groups (K = 2–8), and changed from 0.48 to 0.74 with increasing K, but no maximum was reached for a given number of groups. For all levels, the variance components were highly significant. Each time K was increased, an additional group of geographical homogeneous populations was defined. These groups were always characterized by a distinct mitochondrial clade or variant. In Fig. 1 the geographical delineation for the various K-values is pictured, except for K = 7 because this did not result in a clear geographical pattern. In the latter case, the Central European populations were subdivided in a group characterized by haplotypes EB20-22 (2, 6, 9, 25) and a group mainly constituted of EB30-33 (4, 7, 8, 14, 16).

For the North American assemblage, F_{CT} increased from 0.50 to 0.81 with an increasing number of assumed groups (K = 2–6) and in each case the variance components were highly significant. In Fig. 1, only the structure up to K = 3 is pictured for graphical simplicity. The assumption of four groups (K = 4) separated the Leech Lake sample (42) from the Mississippian group (K = 2), because it was composed of distinct Mp haplotypes (mainly NA21), compared to the other populations that shared the NA30 haplotype. In a five-group structure (K = 5), populations 32 and 44 were joined into an additional group, because of the shared NA13 mitochondrial variant. When K was increased to 6, both populations were considered as a separate group.

We calculated the diversity within the EB and NA phylogroup and for the various groups defined by the SAMOVA (Table 3). The haplotype diversity (h) was approximately the same for the EB and NA assemblage, but the nucleotide diversity was considerably higher in the latter region.

Demographic analysis

A plot of pairwise differences among all CR sequence data was clearly bimodal (Fig. 3). One mode corresponded to the number of differences among EB and NA, and the other to differences among individuals within both lineages. When only EB haplotypes were considered, the mismatch distribution (Fig. 3) fitted well the predictions under a model of sudden expansion ($P > 0.05$). The associated D statistic was negative (−0.31) but did not differ significantly from zero. The observed value of the age expansion parameter (τ) was 3.57 units of mutational time. Although the sudden expansion model could not be rejected for the NA lineage (τ = 7.71), it will not be used for further inference because of the positive D statistic (Table 3).

Table 1 Distribution of the control region haplotypes of burbot for the Eurasian-Beringian lineage. The populations are grouped according to the SAMOVA analysis (see text for details) and the haplotypes according to the associated clades. For sample abbreviations see Appendix 1

	WEu group					NEu group						EA group										Ber group						Ala
	Lin	Ver	Gar	Meu	Luc	Sel	Sna	Lul	Tor	Mud	Umb	Con	One	Ipo	Kyr	Isa	Vis	Por	Elb	Sav	Fro	Kha	Kht	Ana	Kol	Len	Fie	Hea	Tol
EB00	8	7			7																								
EB01			1		3																								
EB02		3																											
EB03	3																												
EB04				1								1																	
EB10						7	6	9	9																				
EB11						2																							
EB12									1																				
EB13								1																					
EB14						1																							
EB15										9	4		2																
EB16										1																			
EB17											1																		
EB20													4	1	4														
EB21																2	1												
EB22																				1									
EB23												4																	
EB30																3	4	2	9							1			
EB31														1															
EB32																				7									
EB33																	1												
EB34																					3	9							
EB35																									2				
EB40																								4	4	6	7		
EB41															1			2					2						
EB42																												1	2
EB43																								1					
EB44																													
EB45																										1			
EB50																													6

Table 2 Distribution of the control region haplotypes of burbot from the North American lineage. The populations are grouped according to the results of the SAMOVA (see text for details) and the haplotypes according to the associated clades. For sample abbreviations see Appendix 1

	Pac group			Ms group					GSL group			Mp group				
	Koc	Bou	Kou	Goa	Sak	At1	At2	Col	Pin	Red	Sim	Mis	Bay	Dev	Sas	Nei
EB35									5							
EB44										1						
NA00	2	1		1												
NA01	2		1	1												
NA10					1	2	3	2	1							
NA11					4											
NA12					1											
NA13				2	1											
NA20											4	1				
NA21												7				
NA22												1				
NA23												1				
NA30									2	1	1			1	1	8
NA31													1			
NA40					2											
NA41					1											

Table 3 Diversity indices for the defined phylogeographical units in *Lota lota*

Phylogeographical unit	*n*	Number of haplotypes	Haplotype diversity (h)	Nucleotide diversity (× 100) (π)	Tajima's D
North America	**53**	**14**	**0.90**	**1.13**	**1.69***
Mississippi	26	6	0.74	0.40	0.63
Missouri	21	8	0.82	0.87	0.54
Pacific	6	2	0.60	0.14	1.45*
Great Slave Lake	10	4	0.71	1.20	0.31
Eurasia-Beringia	**192**	**29**	**0.92**	**0.74**	**−0.34**
Northern Europe	51	8	0.57	0.16	−1.21
Western Europe	33	5	0.54	0.15	−0.93
Eurasia	62	12	0.85	0.47	−0.64
Beringia	38	7	0.63	0.22	−0.60
Alaska	8	2	0.43	0.40	0.41

n, number of individuals per region, number of haplotypes observed, haplotype diversity (*H*), nucleotide diversity × 100 (π) and Tajima's *D* statistic (values significant are marked with *) the two major phylogeographic units are displayed in bold.

Major demographic changes might be a consequence of dramatic changes in global climate. Assuming that each termination of a glacial period provides an opportunity for a population expansion, we calculated the corresponding divergence rate for subsequent glacial terminations. An expansion across the Eurasian continent after termination of the last glacial period (14 kyr) implies a divergence rate of 60.00% per Myr to account for the τ-value observed in the EB lineage. Expansions in response to the earlier glacial terminations (TermII, TermIII, TermIV, and TermV) approximately 130 kyr, 240 kyr, 325 kyr, and 420 kyr ago (Petit *et al*. 1999) would result in the observed mismatch distribution if the divergence rate for the first variable part of the control region would be 6.46%, 3.50%, 2.58%, or 2.00% per Myr, respectively.

Discussion

The first part of the mitochondrial control region of burbot was obviously the most variable, confirming the findings of Lee *et al*. (1995) that this is the most suitable part for population genetic studies in most fishes. Various molecular clock rates have been assumed, ranging from 0.8% (Englbrecht *et al*. 2000; Susnik *et al*. 2001) to 3.6% divergence

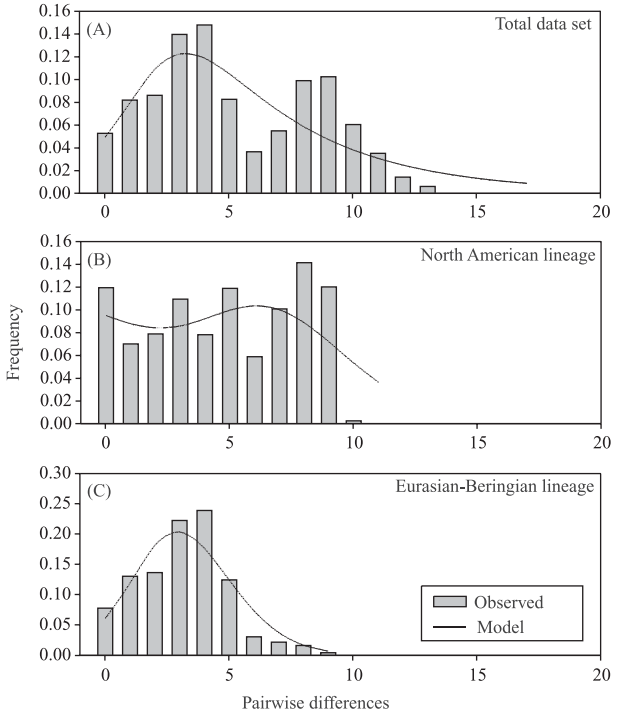

Fig. 3 Frequency distribution of the pairwise number of mutational differences observed (grey bars) in the first part of the mitochondrial control region of burbot (*Lota lota*): (i) all individuals (ii) the North American lineage (ii) the Eurasian-Beringian lineage. The solid line represents the expected mismatch distribution for the sudden-expansion model.

per Myr (Donaldson & Wilson 1999) and even a much higher rate in sturgeon ($1.1–1.3 \times 10^{-7}$ nucleotide per site per year; Brown *et al*. 1993). For the cytochrome *b* gene of burbot, a maximum and a minimum divergence rate were estimated at, respectively, 3.0% and 0.9% divergence per Myr (Van Houdt *et al*. 2003). The average divergence of the control region between the NA and EB lineage was 2.0%, whereas this amounted only to 1.25% for the cytochrome *b* gene. As such, the divergence rate from the control region is most likely within the interval of 2–6%. The latter values are concordant with the divergence values derived from the possible expansion events.

Historical demography

The genus *Lota* probably separated 10 Ma from the other Lotinae (Van Houdt *et al*. 2003) and fossil evidence suggests that some 5 Ma, a freshwater form existed in Central Europe (Pietschmann 1934). However, the genetic divergence between the two previously identified lineages for cytochrome *b* (EB and NA lineage) was not older than 1 Myr, indicating that a major component of the genetic legacy of the genus *Lota* has been lost. The distribution of the two phylogroups, EB and NA, was congruent with the subspecies, *L. lota lota* (EB lineage) and *L. lota maculosa* (NA lineage) respectively (Van Houdt *et al*. 2003). From variation in the mitochondrial control region, we clearly identified both lineages; the average divergence (2.08%) suggests that their separation dates back to the late Pleistocene. Solely, the samples from the Great Slave Lake (GSL group) harboured the *L. l. maculosa* and *L. l. lota* mitochondrial genomes.

The mitochodrial diversity of the Eurasian burbot carries signatures of a historic expansion. Its gene genealogy consists of closely related haplotypes, with widely distributed central haplotypes (EB30 and EB40) and derived clades that were found at the periphery of the subspecies' range (Fig. 2). Furthermore, the expansion model could not be rejected and the mitochondrial diversity was characterized by a negative Tajima's *D* statistic (not significant) and a highly significant negative F_S. The latter is sensitive to demographic expansion, which generally leads to large negative F_S values (Fu 1997). Additionally, simulations demonstrated that range expansions may lead to a molecular signature quite similar to that observed after a pure demographic expansion (Ray *et al*. 2003). As such, a past geographical range expansion could explain the observed mitochondrial diversity in *L. l. lota*. If we assume that such expansions coincide with the termination of a glacial period, we can derive the mutation rate necessary to generate the observed mitochondrial diversity pattern since the initiation of the expansion. This illustrates that it is unlikely that the observed mitochondrial diversity of *L. l. lota* is the result of a transcontinental expansion after the last glacial maximum (LGM), since this would have required a mutation rate of 60% per Myr. Any expansion associated with the later glacial terminations (II to V) results in acceptable mitochondrial mutation rates. However, since we used the most variable part of the control region, it is more likely that Eurasian burbot spread across the Palearctic after the second (130 kyr) and/or third (240 kyr) glacial termination.

Within the NA lineage, three distinct mitochondrial clades were observed, separated by several mutations. The sudden expansion model could not be rejected for the mismatch distribution of *L. l. maculosa*, but the significantly positive Tajima's *D* statistic and the three clearly separated mitochondrial clades are not consistent with a historic expansion. The shape of the haplotype network (Fig. 2) suggests that the three NA assemblages diverged approximately simultaneously from a common ancestor. Under the assumptions of a molecular clock, the average genetic distance of 1.32% would be the result of an allopatric divergence that initiated less than 700 kyr ago.

European traces of the Ice Ages

Burbot became re-established in Europe relatively recently and has only been subjected to three or four glacial cycles.

Nevertheless, three different burbot assemblages with a distinct mitochondrial composition and geographical distribution could be discriminated: (i) the WEu group includes populations in various Atlantic drainages from France to Denmark; (ii) the NEu group inhabits Scandinavia; and (iii) the EA group covers the Danube system up to Finland in the north, including populations from Lake Baikal (Fig. 1). In general, the southern peninsulas of Iberia, Italy and the Balkans, along with the Caspian/Caucasus region are considered important refugia for many species (Hewitt 2004). Several phylogeographical (Bernatchez 2001; Kotlík & Berrebi 2001; Volckaert *et al*. 2002) and zoogeographical (Bănărescu 1992) studies illustrate the importance of the Danubian refuge for survival of the European freshwater fauna during the climatic oscillations. Also, the most diverse southern populations of burbot were found in this region. The higher diversity observed in the Baltic samples results from a mix of multiple glacial races. However, it will take complementary samples from the Balkans and Caspian/Caucasus region to reconstruct for burbot the colonization routes in Central Europe.

In Western Europe (Atlantic basin), most freshwater fishes (Durand *et al*. 1999; Nesbø *et al*. 1999; Bernatchez 2001; Kotlík & Berrebi 2001) including burbot display a limited mitochondrial diversity of closely related variants, which points to a recent colonization (Hewitt 2004). The distinct mitochondrial variants are probably a consequence of geographical isolation and subsequent rapid sorting of new haplotypes during the most recent expansion/contraction events. In Northern Europe, we observed mitochondrial genomes that could not be assigned to a particular southern refuge. Likely, they originated from a multispecies refugium located east of Scandinavia (Nesbø *et al*. 1999; Koskinen *et al*. 2000).

Trans-Siberian exchange of freshwater taxa

Phylogeographical studies on Arctic charr (Brunner *et al*. 2001), whitefish (Bernatchez & Dodson 1994) and *Daphnia* (Weider *et al*. 1999) revealed closely related genomes across Eurasia. Similarly, a few common closely related mitochondrial variants of burbot were distributed from central Europe to eastern Siberia. This suggests that recent transcontinental dispersal of cold-adapted freshwater species must have taken place. During an early phase of the last Ice Age, about 90 kyr ago, an ice sheet covered the shallow Barents and Kara Seas and large ice-dammed lakes were formed (Krinner *et al*. 2004). According to Grosswald & Hughes (2002), a continuous Arctic ice sheet could impound and divert all the north-flowing rivers and result in the formation of a Trans-Siberian drainage system that joins hydrographic basins from the Lena River in the east to the Caspian and Black Seas in the west. Such a system possibly facilitated exchange of cold-adapted species between Eastern Europe and Siberia. It provides a plausible explanation for the shared haplotypes in geographically distant populations (Table 1).

Continental exchange across the Beringian 'land' bridge

Two mitochondrial EB assemblages were found in Beringia (northeastern Siberia and Alaska) (Fig. 1). The Alaskan clade (Ala, represented by a single haplotype, EB50) was only found in the Tolsona population, whereas the Beringian clade (Ber) was found from eastern Siberia to the Great Slave Lake. This confirms that Beringia served as a refugium for burbot as suggested by McPhail (1997) and by Cumbaa *et al*. (1981). The Beringian gate between the Nearctic and Palearctic has opened and closed several times over the past million years (Sher 1999). Zoogeographical data suggest that the exchange of freshwater fishes was asymmetric and occurred mostly from east to west (Lindsey & McPhail 1986). Hence the Beringian and Alaskan clade most likely evolved independently at each side of the Bering Strait followed by a dispersal of Siberian burbot into the Yukon system. Interestingly, a similar pattern was observed in *Thymalus arcticus* (Stamford & Taylor 2004). It was explained as the result of an allopatric differentiation in a northern and southern refuge in Beringia.

The contribution of the Beringian refuge to the colonization of North America is apparently variable. Beringian lake whitefish (*Coregonus clupeaformis*) did not reach the Mackenzie Basin and the glaciated region was almost exclusively repopulated with Mississippian fishes (Bernatchez & Dodson 1991). Whereas arctic grayling (*T. arcticus*; Stamford & Taylor 2004) and lake trout (*Salvelinus namaycush*) dispersed extensively southward after the LGM. Almost two-thirds of the present range of lake trout was recolonized by Beringian fish and displayed an extensive secondary intergradation among refugial groups (Wilson & Hebert 1996). The Siberian burbot (Ber clade) was apparently either the most successful or had better access to dispersal routes than the Alaskan burbot (Ala clade). The latter mitochondrial variants were restricted to Alaska, whereas the Beringian burbot were found as far south as the Great Slave Lakes.

Crossing the Continental Divide

The current North American range of burbot was almost completely covered with ice 10 000 years ago. There are no indications that *L. l. maculosa* was present in the Beringian refuge during the LGM. As such, this phylogroup must have survived the Late Wisconsinan glacial advance in refugia south of the ice sheet. Crossman & McAllister (1986) suggested that burbot survived in a Missourian and

a Mississippian refuge, whereas McPhail (1997) assumed an additional Pacific refuge based on the typical morphology of the burbot from the Columbia-Kootenay system. The latter is confirmed by the presence of distinct Pacific mitochondrial variants (Pac clade: Fig. 1) in the Kootenay River. East of the Continental Divide, the Missourian and Mississippian refuges are supported by the observation of two distinct mitochondrial clades (Mp and Ms: Fig. 2). Both clades have a distinct distribution in the southern populations but intergrade substantially in the previously glaciated regions (Fig. 1). The postglacial dispersal of the Missourian and Mississippian clade north to the Great Slave Lake was possible via the extensive water connections that were formed after deglaciation (Lindsey & McPhail 1986; Dyke *et al*. 2002). The present distribution of the Missourian and Mississippian clades suggests that the Athabasca River was recolonized from the Missouri refugium via the Saskatchewan-Peace dispersal corridor, whereas burbot from the Mississippian refuge apparently used the Clearwater dispersal corridor (Rempel & Smith 1998).

The Missourian group was characterized by a relatively high nucleotide diversity (Table 3) because it contained, in addition to closely related Missourian haplotypes, mitochondrial variants more similar to the Mississippian clade (NA40-41: Fig. 2), that were not found elsewhere. As such, the Sakakawea Lake population was probably recolonized after the LGM by burbot that survived in the upper Missouri River (Missouri refugium with Ms clade) and burbot from the lower Missouri river (Mississippian refuge with NA40-41: Mp clade). In this case, more than one gene pool must have survived the LGM in the Mississippian refuge. Alternatively, the haplotype composition was the result of a historical exchange and subsequent differentiation of the Mississippian variants in the upper Missouri.

The establishment of burbot at both sides of the Continental Divide probably dates back to the initial colonization from Eastern Siberia into North America (less than 700 kyr ago). Nevertheless, we observed in the Goat River (tributary of the Fraser River) both Pacific and Missourian haplotypes (Fig. 1, 32). The former originated from the Pacific refugium, whereas the latter suggests a recent exchange across the Continental Divide. Zoogeographical and geological data indicate that, approximately 10 000 years ago, faunal exchange was possible via a Fraser-Peace connection (Lindsey & McPhail 1986). Alternatively, this might be the result of human-mediated transfer, since burbot from Alberta have been introduced in British Columbia (Crossman 1991).

Holarctic evolutionary history of burbot

Divergence estimates suggest that the differentiation within the Nearctic phylogroups was initiated prior to the divergence of the observed Palearctic lineages. The former process started approximately between 220 and 660 kyr ago, whereas the latter process was initiated between 130 and 420 kyr ago. Fossil records suggest that the genus arose in the Palearctic (Van Houdt *et al*. 2003), but the genetic signature points to a Nearctic origin of the present-day burbot populations. In this case, the ancestral Palearctic burbot must have been extirpated, where after the Old World was reinvaded from a Nearctic source population. Alternatively, *L. l. lota* was reduced to a single ancestral gene pool in the Palearctic region after the successful colonization of North America. Subsequently, it expanded again across Eurasia and invaded the Nearctic continent for a second time. In our opinion, this is a more likely scenario to explain the central position of the common European EB30 haplotype and the peripheral position of Alaskan EB50 haplotype. Due to expansion/contraction events, derived mitochondrial variants became fixed in isolated regions at the periphery of the range. On the other hand, the Trans-Siberian drainage system homogenized the genetic composition by allowing exchange across a vast geographical range.

The ancestral population of *L. l. maculosa* probably entered the North American continent via the Bering Strait and continued its southern dispersal. *L. l. maculosa* survived subsequent glacial oscillations in the established southern refugia (Pacific, Missourian and Mississippian), whereas it was lost, or replaced, by *L. l. lota* in the Beringian region at a later stage. Apparently, the Pacific group was isolated since the formation of the Continental Divide and only after the LGM Missourian burbot were able to invade the Fraser River system. East of the Continental Divide, Mississippian and Missourian burbot remained isolated in their respective refugia, only to mix across a vast geographical area after each glacial retreat.

Conclusions

It is thought that the geographical features of the Palearctic and Nearctic regions affect the structuring of genomes by climatic oscillations differently (Hewitt 2000). However, one might expect that the Nearctic burbot, whose present-day distribution range was almost completely covered with ice during the LGM, should have been most severely impacted by the past glaciations. Nevertheless, the species managed to preserve even more genetic diversity than in Eurasia where the glacial advance was less extensive. Apparently, the latitudinally orientated landscape in North America provided better opportunities to cope with the climatic oscillations, whereas the east–west orientated mountain ranges in southern Europe limited the possibilities to escape the advancing ice sheets. The shallow and homogenous genetic structure in the remainder of the Palearctic region suggests that even the limited ice

coverage resulted in a dramatic perturbation of the Eurasian aquatic habitats by blocking the northern outlets of Eurasian drainage systems. These massive inundation basins provided opportunities to the freshwater fauna for transcontinental exchange or colonization.

Acknowledgements

JVH received a scholarship from the Institute for Scientific and Technological research (IWT) for which we are grateful. We would like to thank R. Zardoya and an anonymous reviewer for their comments on the manuscript and we are very grateful to M. Mina for coordinating the sampling in Russia. We thank D. De Charleroy, I. Chereshnev, T. Eroes, P. Fischer, S. Fisher, W. Franzin, M. Hansen, W. Hughson, E. Ieshko, C. Johnson, A. Kirillov, J. Kjellman, D. Kucharczyk, S. Kusmin, G. Low, M. Luczynski, R. Müller, V. Paragamian, F. Parker, L. Perry-Plake, A. Raat, A. Schreiber, M. Steinhilber, S. Stridsman, T. Taube, J. Trautner and T. Van De Valk for providing samples. Research was funded by the Ministry of the Flemish Community (TWOL-AMINAL project AMINAL/BG/V 97.2).

References

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution*, **13**, 494–504.

Bănărescu P (1992) *Distribution and dispersal of freshwater animals in North America and Eurasia*. Aula-Verlag, Wiesbaden.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Bandelt HJ, Macaulay V, Richards M (2000) Median networks: speedy construction and greedy reduction, one simulation and two case studies from human mtDNA. *Molecular Phylogenetics and Evolution*, **16**, 8–28.

Barnes I, Matheus P, Shapiro B, Jensen D, Cooper A (2002) Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science*, **295**, 2267–2270.

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, Dodson JJ (1991) Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation 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 mitochondrial DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 240–251.

Bernatchez L, Guyomard R, Bonhomme F (1992) DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Molecular Ecology*, **1**, 161–173.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Bertorelle G, Slatkin M (1995) The number of segregating sites in expanding human populations, with implications for estimates of demographic parameters. *Molecular Biology and Evolution*, **12**, 887–892.

Brown JR, Beckenbach AT, Smith MJ (1993) Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Molecular Biology and Evolution*, **10**, 326–341.

Brunner PC, Douglas MR, Osinov A, Wilson CC, Bernatchez L (2001) Holarctic phylogeography of arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. *Evolution*, **55**, 573–586.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Crossman EJ (1991) Introduced freshwater fishes — a review of the North American perspective with emphasis on Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 46–57.

Crossman EJ, McAllister DE (1986) Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungava Bay and the Arctic Archipelago. In: *The Zoogeography of North American Freshwater Fishes* (eds Hocutt CH, Wiley EO), pp. 53–103. John Wiley and Sons, New York.

Cumbaa SL, McAllister DE, Morlan RE (1981) Late Pleistocene fish fossils of *Coregonus*, *Stenodas*, *Thymallus*, *Catostomus*, *Lota* and *Cottus* from the Old Crow Basin, northern Yukon, Canada. *Canadian Journal of Earth Sciences*, **18**, 1740–1754.

Donaldson KA, Wilson RR (1999) Amphi-Panamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution*, **13**, 208–213.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Durand JD, Persat H, Bouvet Y (1999) Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, **8**, 989–997.

Dyke AS, Andrews JT, Clark PU *et al.* (2002) The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. *Quaternary Science Reviewss*, **21**, 9–31.

Englbrecht CC, Freyhof J, Nolte A, Rassmann K, Schliewen U, Tautz D (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.

Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene trees within a species — molecular variance parsimony. *Genetics*, **136**, 343–359.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes — application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitch-hiking and background selection. *Genetics*, **147**, 915–925.

Gibbard PL (1988) The history of the great northwest European rivers during the past 3 million years. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **318**, 559–602.

Grosswald MG, Hughes TJ (2002) The Russian component of an Arctic ice sheet during the Last Glacial Maximum. *Quaternary Science Reviews*, **21**, 121–146.

Hewitt GM (1996) Some genetic consequences of Ice Ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Postglacial recolonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the Quaternary Ice Ages. *Nature*, **405**, 907–913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183–195.

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.

Kotlík P, Berrebi P (2001) Phylogeography of the barbel (*Barbus barbus*) assessed by mitochondrial DNA variation. *Molecular Ecology*, **10**, 2177–2185.

Krinner G, Mangerud J, Jakobsson M, Crucifix M, Ritz C, Svendsen JI (2004) Enhanced ice sheet growth in Eurasia owing to adjacent ice-dammed lakes. *Nature*, **427**, 429–432.

Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of Teleost mitochondrial control regions. *Journal of Molecular Evolution*, **41**, 54–66.

Leonard JA, Wayne RK, Cooper A (2000) Population genetics of ice age brown bears. *Proceedings of the National Academy of Sciences, USA*, **97**, 1651–1654.

Lindsey CC, McPhail JD (1986) Zoogeography of fishes of the Yukon and Mackenzie Basins. In: *The Zoogeography of North American Freshwater Fishes* (eds. Hocutt CH Wiley EO), pp. 639–674. John Wiley & Sons, New York.

McPhail JD (1997) *A Review of Burbot (*Lota lota*) Life History and Habitat Use in Relation to Compensation and Improvement Opportunities*. Department of Fisheries and Oceans, Canada.

Miya M, Takeshima H, Endo H *et al.* (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **26**, 121–138.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Nelson JS (1994) *Fishes of the World*. John Wiley and Sons, New York.

Nesbø CL, Fossheim T, Vollestad LA, Jakobsen KS (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, **8**, 1387–1404.

Petit JR, Jouzel J, Raynaud D *et al.* (1999) Climate and atmospheric history of the past 420 000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429–436.

Pietschmann V (1934) *Lota hulai*, eine neue Fischart aus dem Wiener Becken. *Palaeontologische Zeitschriftm*, **16**, 48–52.

Posada D, Crandall KA (2001) Intraspecific genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, **16**, 37–45.

Ray N, Currat M, Excoffier L (2003) Intrademe molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, **20**, 76–86.

Rempel LL, Smith DG (1998) Postglacial fish dispersal from the Mississippi refuge to the Mackenzie River basin. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 893–899.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN version 2.0: a software for population genetics data analysis. University of Geneva, Geneva.

Sher A (1999) Traffic lights at the Beringian crossroads. *Nature*, **397**, 103–104.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Sorokin VN (1971) The spawning and spawning grounds of the burbot (*Lota lota* (L.). *Journal of Ichthyology*, **11**, 907–915.

Stamford MD, Taylor EB (2004) Phylogeographical lineages of Arctic grayling (*Thymallus arcticus*) in North America: divergence, origins and affinities with Eurasian *Thymallus*. *Molecular Ecology*, **13**, 1533–1549.

Susnik S, Snoj A, Dovc P (2001) Evolutionary distinctness of grayling (*Thymallus thymallus*) inhabiting the Adriatic River system, as based on mtDNA variation. *Biological Journal of the Linnean Society*, **74**, 375–385.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

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.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876–4882.

Turgeon J, Bernatchez L (2001) Mitochondrial DNA phylogeography of Lake Cisco (*Coregonus artedi*): evidence supporting extensive secondary contacts between two glacial races. *Molecular Ecology*, **10**, 987–1001.

Van Houdt JK, Hellemans B, Volckaert FAM (2003) Phylogenetic relationships among Palearctic and Nearctic burbot (*Lota lota*): Pleistocene extinctions and recolonization. *Molecular Phylogenetics and Evolution*, **29**, 599–612.

Volckaert FAM, Hanfling 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.

Weider LJ, Hobaek A, Colbourne JK, Crease TJ, Dufresne F, Hebert PDN (1999) Holarctic phylogeography of an asexual species complex I. Mitochondrial DNA variation in Arctic *Daphnia*. *Evolution*, **53**, 777–792.

Wilson CC, Hebert PDN (1996) Phylogeographic origins of lake trout (*Salvelinus namaycush*) in eastern North America. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 2764–2775.

This paper is part of the PhD project of Jeroen Van Houdt on the phylogeography and conservation genetics of burbot. He is currently studying the effects of habitat fragmentation on the genetic structure of freshwater fish and involved in the genetic identification and phylogeny of Antarctic fishes. Lotje De Cleyn and Alessandra Perretti (EU Socrates exchange student) contributed to the genotyping and interpretation as MSc thesis students. Filip Volckaert has a special interest in the ecological genetics and genomics of fish (http://www.kuleuven.be/bio/eco).

Appendix 1

Sampling data of burbot (*Lota lota*) across the northern hemisphere including site (country and location), abbreviation (ID), longitude and latitude and number of fish collected (*n*)

nr	Country	ID	Location	Coordinates	*n*
1	Denmark	Lin	Linaa River	056°09′ N – 009°41′ E	11
2	Finland	Kyr	Kyronjoki River	063°12′ N – 022°03′ E	5
3	Finland	Mud	Muddusjarvi Lake	069°00′ N – 026°50′ E	10
4	Finland	Por	Porkkala Bay	060°00′ N – 024°25′ E	4
5	France	Ver	Automne River	049°24′ N – 002°49′ E	10
6	Germany	Con	Lake Constance (Birnau)	047°44′ N – 009°12′ E	5
7	Germany	Elb	Elbe River, Geesthacht	053°26′ N – 010°22′ E	9
8	Germany	Isa	Isar River	048°30′ N – 012°00′ E	5
9	Hungary	Ipo	Ipoly River	048°01′ N – 018°53′ E	2
10	Italy	Gar	Garda Lake	045°35′ N – 010°35′ E	1
11	Netherlands	Meu	Meuse River	051°46′ N – 005°20′ E	1
12	Norway	Sel	Selbu Lake	063°16′ N – 010°45′ E	10
13	Norway	Sna	Snasa Lake	064°15′ N – 012°20′ E	6
14	Poland	Vis	Vistula River	053°57′ N – 018°53′ E	6
15	Sweden	Lul	Luleälven, Lake Langas	064°53′ N – 021°19′ E	10
16	Sweden	Sav	Sävarån River	063°48′ N – 020°34′ E	8
17	Sweden	Tor	Torneälven River	067°45′ N – 020°30′ E	10
18	Switzerland	Luc	Lucerne Lake	047°00′ N – 008°30′ E	10
19	Russia	Ana	Anadyr River	064°40′ N – 170°25′ E	5
20	Russia	Fro	Froliha Lake	055°36′ N – 109°34′ E	3
21	Russia	Kha	Khakusi Bay (Lake Baikal)	051°52′ N – 104°48′ E	9
22	Russia	Kht	Katanga River	071°57′ N – 102°26′ E	2
23	Russia	Kol	Kolyma River	062°26′ N – 152°19′ E	6
24	Russia	Len	Lena River	072°38′ N – 126°35′ E	8
25	Russia	One	Onega Lake	061°51′ N – 035°49′ E	6
26	Russia	Umb	Umbozero Lake	067°31′ N – 034°31′ E	5
27	Canada	At1	Athabasca River, Lambert Creek	053°30′ N – 117°00′ W	2
28	Canada	At2	Athabasca River, Lac Beauvert	052°55′ N – 118°00′ W	3
29	Canada	Bay	Bayonet Lake	059°56′ N – 110°17′ W	1
30	Canada	Col	Colin Lake	059°33′ N – 110°07′ W	2
31	Canada	Dev	Devils Lake	053°43′ N – 114°06′ W	1
32	Canada	Goa	Goat River	053°31′ N – 120°35′ W	4
33	Canada	Koc	Kootenai River	049°00′ N – 116°30′ W	4
34	Canada	Pin	Great Slave Lake, Pine Point	060°58′ N – 114°23′ W	8
35	Canada	Red	Great Slave Lake, Redclif Island	062°26′ N – 111°06′ W	2
36	Canada	Sas	Saskatchewan River	054°03′ N – 112°46′ W	1
37	Canada	Sim	Great Slave Lake, Simpson Island	061°50′ N – 012°30′ W	5
38	USA	Fie	Fielding Lake	063°10′ N – 145°40′ W	7
39	USA	Hea	Healy Lake outlet,Tanana River	064°00′ N – 144°47′ W	10
40	USA	Bou	Boundary Creek	048°55′ N – 116°36′ W	1
41	USA	Kou	Kootenai River	049°00′ N – 116°30′ W	2
42	USA	Mis	Mississippi, Leech Lake	047°09′ N – 094°25′ W	10
43	USA	Nei	Oneida Lake	043°10′ N – 075°55′ W	8
44	USA	Sak	Missouri River, Sakakawea Lake	047°37′ N – 102°14′ W	10
45	USA	Tol	Tolsona Lake	062°06′ N – 146°04′ W	8