

Genetic studies of population history and contemporary microevolution in grayling (*Thymallus*: Salmonidae)

Mikko T. Koskinen

Department of Ecology and Systematics
Division of Population Biology
University of Helsinki
Finland

Academic dissertation

To be presented, with permission of the Faculty of Science of the University of Helsinki, for public criticism in the lecture hall of the Viikki Infocentre, Viikinkaari 11, on November 9th, 2002, at 12 o'clock noon.

Helsinki 2002

- © Mikko T. Koskinen. Pp. 5-33.
- © Blackwell Science. Articles I, II, V.
- © Nature Publishing Group. Articles III, VI.
- © Kluwer Academic Publishers. Article IV.

Cover illustration kindly painted by Meri Liesiö

Author's address:
Department of Ecology and Systematics
Division of Population Biology
P.O. Box 65 (Viikinkaari 1)
FIN-00014 University of Helsinki
Finland

E-mail: mtkoskin@cc.helsinki.fi

ISBN 952-91-5170-5
Oy Edita Ab
Helsinki 2002

The thesis is based on the following articles:

- I Koskinen, M.T., Knitzhin, I., Primmer, C.R., Schlötterer, C. & Weiss, S. 2002. Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of freshwater, Lake Baikal. *Molecular Ecology*, in press.
- II Koskinen, M.T., Ranta, E., Piironen, J., Veselov, A., Titov, S., Haugen, T.O., Nilsson, J., Carlstein, M. & Primmer, C.R. 2000. Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Molecular Ecology*, **9**, 1609-1624.
- III Koskinen, M.T., Nilsson, J., Veselov, A.Je., Potutkin, A.G., Ranta, E. & Primmer, C.R. 2002. Microsatellite data resolve phylogeographic patterns in European grayling. *Heredity*, **88**, 391-401.
- IV Koskinen, M.T., Piironen, J. & Primmer, C.R. 2001. Interpopulation genetic divergence in European grayling (*Thymallus thymallus*, Salmonidae) at a microgeographic scale: implications for conservation. *Conservation Genetics*, **2**, 133-143.
- V Koskinen, M.T., Sundell, P., Piironen, J. & Primmer, C.R. 2002. Genetic assessment of spatiotemporal evolutionary relationships and stocking effects in grayling (*Thymallus thymallus*, Salmonidae). *Ecology Letters*, **5**, 193-205.
- VI Koskinen, M.T., Haugen, T.O. & Primmer, C.R. 2002. Contemporary fisherian life-history evolution in small salmonid populations. *Nature*, **419**, 826-830.

The articles are referred to by their Roman numerals in the text.

Contributions:

Supervised by: Dr. Craig R. Primmer
University of Helsinki
Finland

Reviewed by: Dr. Michael M. Hansen
Danish Institute for Fisheries Research
Denmark

Prof. Eric B. Taylor
University of British Columbia
Canada

Examined by: Prof. Louis Bernatchez
Université Laval
Canada

Contents

Introduction	5
Sources of genetic variation in nature	5
Phylogeography and historical environmental perturbations.....	6
Contemporary evolutionary processes	6
Molecular tools in evolutionary studies	7
Grayling (<i>Thymallus</i> spp.): natural distribution, life-history characteristics and cultural importance	8
Outline of the thesis	10
Summary of the original publications	10
I - Phylogeography of grayling within Lake Baikal and its surrounding watersheds	10
II, III - Phylogeography of European grayling.....	13
IV, V - Microgeographic evolutionary relationships	15
V - Spatiotemporal assessment of stocking effects	17
VI - Contemporary life-history evolution in grayling.....	22
Conclusions	26
Acknowledgements	28
References	29

Introduction

Sources of genetic variation in nature

One of the principal aims of population genetics is to quantify the amount of inheritable variation present in nature. Population genetics thereby attempts to shed light on the fundamental question in evolutionary biology: what determines genetic differences between species, populations and individuals? Among the primary forces that have potential to create and homogenize variation are natural selection, random genetic drift, mutations and gene flow (Fisher 1930; Wright 1931; Simpson 1944; Dobzhansky 1951; Simpson 1953).

It is widely accepted that adaptation of populations to their specific habitats is predominantly the outcome of natural selection due to environment and resource competition (Darwin 1895; Fisher 1930; Simpson 1944; Dobzhansky 1951; Simpson 1953; Schluter 2000). Natural selection operates to select for fitness of individuals, i.e. their contribution of offspring to the next reproducing generation (components of fitness are also commonly known as 'life-history traits'). If fitness differences among individuals are associated with any given gene, it follows that selection will operate on that gene. For genes under selection, parents of different genotypes pass on their genes unequally to the next generation, leading to genetic differences among isolated populations (Falconer & Mackay 1996). The fitness of an individual is the complex total outcome of all of its developmental and physiological processes expressed in nature. The response of fitness to selection is determined by *Fisher's fundamental theorem* (Price 1972; Frank & Slatkin 1992; Edwards 1994). The main conclusion of the theorem is that the increase in fitness (or a life-history trait) at any time is equal to the additive genetic variance of fitness at that time. It follows, that response to selection for a given trait will cease when its additive genetic variance (and, thus, heritability of the trait) is depleted (Price 1972).

By definition, random genetic drift is unpredictable in direction. In the absence of selection, mutations and migration, gene and genotype frequencies remain constant over time in effectively large populations. In small populations, however, the frequencies are always subject to fluctuations between generations due to random sampling of gametes. Independent sampling of gametes in different sub-populations leads to differing gene and genotype frequencies, i.e. random genetic drift (Wright 1931).

Changes that occur in DNA sequence (mutations), can be non-recurrent or recurrent. Non-recurrent mutations are normally considered to have very little evolutionary importance as their chances of surviving in large populations are small (even if only stochastic processes are considered and the potential selective disadvantages of mutations are discarded). For example, when a mutational event gives rise to one representative of a mutated gene in a population, the chance of it being lost in the next generation is one-half. The probability of each surviving copy of the mutant passing on to the third generation is again one-half. As the loss of a mutation is permanent, the probability of indefinite survival of new mutations is very small in large populations (Falconer & Mackay 1996). Recurring mutations are important in evolution, but their frequencies are generally very low (Papadopoulos et al. 1999). Nevertheless, when population divergence times are sufficiently long, i.e. in the order of 10,000 generations, mutations can have influential evolutionary consequences (Remold & Lenski 2001). It is worth noting, however, that mutations can be scarcely detected in species that have long generation intervals (Falconer & Mackay 1996) (note that this is not the case in the microorganisms that were used to reveal the influence of mutations; Papadopoulos et al. 1999; Remold & Lenski 2001).

Migration (gene flow) homogenizes variation among populations. The so-called stepping-stone migration models approximate natural processes by predicting that genetic differentiation (due to drift) will increase with geographical distance, generating a pattern of isolation by distance (IBD; Slatkin 1993). The tendency of migration to change gene frequencies toward an equilibrium becomes stronger as frequency differences among populations increase. Hence, there is a point of balance, known as 'migration-drift equilibrium', in which genetic variation among populations is held in check (Slatkin 1993). Among the principles of the migration-drift equilibrium is that smaller populations require higher levels of gene flow than larger ones to remain at an equal state of differentiation (drift is expected to be stronger in smaller populations). Furthermore, relatively small numbers of immigrants will prevent much differentiation by random drift (Falconer & Mackay 1996). Empirical studies have revealed that many recently founded natural populations (such as those that colonized their present habitats after the Pleistocene glaciations, i.e. after approximately 10,000 years before present; ybp) may have not yet reached a migration-drift

equilibrium (Hutchison & Templeton 1999; Turgeon & Bernatchez 2001).

All of the above-mentioned evolutionary forces act in time and space (Fisher 1930; Wright 1931). In order to understand the complex processes governing the distribution of genetic diversity in nature, it is therefore important to study the history of populations, contemporary evolutionary agents, as well as the combination of historical and ongoing factors (Templeton et al. 1995).

Phylogeography and historical environmental perturbations

Phylogeography is a rapidly expanding field of study concerned with historical processes governing the geographical distribution of genetic variation (Avisé et al. 1987; Avisé 2000). As the demographic and genealogical past of populations is tightly connected to the geological events that provide or limit suitable habitat, phylogeographic inferences are normally made by comparing genetic data with independently generated theories on landscape history (Avisé 2000). For instance, details regarding the formation of the isthmus of Panama or California's Channel Islands are known *a priori* with a high degree of certainty from the geological record. Subsequently, their impact on organismal genealogy or population genetic history can be investigated (Goldstein et al. 1999), used to calibrate rates of molecular evolution (Bermingham & Lessios 1993), or even applied to build support for novel analytical approaches (Goldstein et al. 1999). Alternatively, paleo-landscape scenarios can be made *post-hoc*, in an attempt to explain particular patterns of genetic diversification. For example, this approach was taken to provide a biological perspective on the paleo-hydrological development of Lake Baikal and its surrounding watersheds (I).

Many phylogeographic investigations have demonstrated that past events had a prevailing influence on the distribution of genetic variation we see today (Avisé 2000). In particular, it is gradually becoming evident that environmental perturbations during the Pleistocene ice-ages (often defined as 1,8 million - 8,000 ybp; e.g. Svendsen et al. 1999) were the most significant events to have occurred during the evolutionary lifespan of most extant species (Hewitt 1996; Avisé et al. 1998; Hewitt 1999; 2000). The dramatic influence of glaciations is intuitively easy to accept in many regions of the world. For instance, most of northern Europe (approximately 6,6 million km²) was covered by the up-to 3,000 meter-thick Scandinavian Ice Sheet only 18,000 ybp (Svendsen et al. 1999). The glacial

periods influenced fish species by reducing their habitat sizes, often leading to fragmentation events. Habitat fragmentation resulted in hydrologically isolated and potentially small sub-populations where genetic drift could have been strong (Avisé 2000). Also interglacial periods were probably evolutionarily very important; meltwater from the retreating glaciers formed new lakes and rivers which offered vast dispersal possibilities for populations surviving in the glacial margins (Avisé 2000). Interglacials thereby provided optimal conditions for homogenization of among-population differences through gene flow (Avisé 2000). It is also easy to envision that natural selection during the Pleistocene period was strong owing to constant adaptation of populations to relatively rapidly changing environmental conditions (e.g. Merilä 1997).

Phylogeographic studies of European taxa have outlined three main regions that served as postglacial colonization sources for many species: Iberic Peninsula, Balkans and Italy (Hewitt 1996; Taberlet et al. 1998; Hewitt 1999; 2000). Although each species naturally has its own colonization history, the patterns exemplified by grasshopper (*Chorthippus parallelus*), hedgehog (*Erinaceus* spp.) and bear (*Ursus arctos*) have been suggested to represent the most frequently utilized dispersal routes, and to serve as paradigms across all taxa (Hewitt 1999; 2000). However, some case studies already indicate that fish species may not fall into these major categories. For instance, populations of brown trout (*Salmo trutta*) have been demonstrated to have persisted in five evolutionarily remarkably distinct lineages throughout the Pleistocene and remained allopatric since then (Bernatchez 2001).

By shedding light on the impact of historical events that have influenced the genetic structure of populations, phylogeography has revolutionized modern population genetics (Avisé 2000; Hewitt 2000). In addition to increasing our knowledge on how species and populations have originated, phylogeographic studies are very important for conservation of endangered species: a prerequisite for managing and conserving biodiversity is the recognition of population groups with independent evolutionary histories and phylogeographic studies can provide such information.

Contemporary evolutionary processes

While the Pleistocene glacial epoch can be seen as having a dominant role in sculpting present-day genetic diversity [Hewitt 1996; Avisé et al. 1998; Hewitt 1999; 2000; but see Klicka & Zink (1997) for an opposing view in birds], all evolutionary

forces continuously influence the genetic structure of populations. However, the relative importance of natural selection (Fisher 1930) and random genetic drift (Wright 1931) in contemporary evolution remains poorly understood and has been an enduring subject of discussion since the introduction of the first population genetic models (Kimura 1983; Gillespie 1991; Falconer & Mackay 1996; Schluter 2000).

The classical 'fisherian' evolutionary theory (Fisher 1930), much influenced by the views of Charles Darwin (Darwin 1859), holds that phenotypic differentiation results primarily from positive natural selection (Fisher 1930). In stark contrast, other evolutionary models have emphasized the potential importance of random genetic drift (Wright 1931; Lande 1976; Carson & Templeton 1984; Lynch 1990; Kimura 1995). Some experimental evidence from populations experiencing changing environmental conditions lends support to the importance of natural selection (Gibbs & Grant 1987; Losos et al. 1997; Reznick et al. 1997; Majerus 1998; Rundle et al. 2000; Hendry 2001; Grant & Grant 2002). However, empirical evidence to evaluate the fisherian versus the contrasting evolutionary theories is surprisingly scarce and further study is warranted for several reasons (VI; Schluter 2000).

In addition to naturally occurring dispersal, human-mediated introductions of domestically reared individuals are becoming an increasingly abundant form of gene flow. Hence, introductions can be seen as playing an important role in contemporary microevolution of natural populations (Allendorf & Waples 1996). Introduced individuals can arguably have detrimental genetic effects on the indigenous populations into which they are released, via decreasing the overall inherited diversity, reduction of fitness due to outbreeding depression and loss of local adaptations (Allendorf & Waples 1996). In light of these concerns, it is important to estimate the effect that non-native individuals have had on the genetic composition of wild indigenous populations.

Molecular tools in evolutionary studies

A key to successful application of molecular tools in population genetics is to recognize that different DNA markers vary in their inherent properties. The main issues to consider include sensitivity of the markers and their technical features (the markers of choice must enable collection of an adequate amount of data with the resources available). The sensitivity of a DNA locus is determined by its mutation rate and mode

of inheritance. Mitochondrial DNA and microsatellite DNA were the genetic markers applied in this thesis and a brief description of their advantages and pitfalls follows.

MtDNA analysis is a powerful tool in evolutionary biology (Avice et al. 1987; Moritz et al. 1987). This is due to mtDNA's relatively rapid mutation rate and haploid and dominantly maternal mode of inheritance, which reduce the effective population size of mtDNA to 1/4 of nuclear loci. This increases the diversifying power of genetic drift acting on mtDNA. Empirical studies have revealed that mtDNA can be extremely informative e.g. for inferring population history (Avice et al. 1987), contact zones between genetically divergent population groups (Moritz et al. 1987), and for quantifying intra-population variability for conservation purposes (Hoelzel et al. 1993). However, as is the case with all genetic markers, mtDNA is not without its limitations. At least the following three concerns are noteworthy: (i) mtDNA studies are often based on a small number of genes and always on just one independently segregating locus, potentially leading to erroneous inference at the population level of resolution (Pamilo & Nei 1988); (ii) selection may severely complicate mtDNA patterns (Hey 1997); and (iii) the sensitivity of mtDNA may be limited when populations have diverged recently (Brunner et al. 1998).

Microsatellites are a class of nuclear DNA loci consisting of tandemly repeated sequence motifs of two to six base pairs in length (Litt & Luty 1989; Tautz 1989; Weber & May 1989). Microsatellites would appear to be optimal markers to complement mtDNA studies. First, as microsatellites are inherited biparentally, they can complement mtDNA resolution in revealing complex biological processes taking place in nature. Namely, a combined investigation of maternally inherited mtDNA and nuclear microsatellites can reveal a pattern of sex-biased dispersal (FitzSimmons et al. 1997). Second, due to their relative ease of application, a large number of independently segregating microsatellites can be utilized. Third, microsatellites are neutral (non-coding) fragments of DNA and can thereby be assumed to reflect evolutionary events primarily due to genetic drift. This assumption naturally requires that microsatellites are not physically linked to coding genes driving inter-population differentiation via natural selection, a phenomenon known as 'genetic draft' (Gillespie 2000; Gillespie 2001). Fourth, microsatellites can be informative among very closely related populations, even when mtDNA resolution is lacking (Brunner et al. 1998). Perhaps

due to these properties, microsatellites have become the marker of choice in many applications in evolutionary biology (reviewed in Estoup & Angers 1998). It is unquestionable that microsatellites are optimal loci e.g. for 'genetic tagging' of individuals, classification of the origins of specimens according to their multilocus genotypes, detection of admixture among populations, and studies of mating structure and relatedness (reviewed in Estoup & Angers 1998).

An overwhelming number of analytical methods for estimating evolutionary histories of populations based on relationships of microsatellite alleles has been proposed (reviewed in Goldstein & Schlötterer 1999). However, the relationships among alleles are determined by their exact model of mutation, which normally is unknown. Indeed, perhaps the most severe concern in the use of microsatellite loci for evolutionary studies is that their mutational properties appear very complex and remain largely unclear. It is widely accepted that microsatellites evolve due to 'slipped strand mispairing' according to which, during replication, the DNA polymerase detaches from the template, resulting in displacement of the replicated strand from the complementary strand, and subsequent generation of length variation via 'looping' during reannealing of the two strands (Kunkel 1993). This mutational mechanism primarily causes small changes in repeat number (Primmer et al. 1998). It follows that alleles more similar in length are expected to have diverged more recently from their common ancestral type than alleles which are more different in length (Primmer et al. 1998). The so-called stepwise mutation model (SMM) based methods make use of this information, therefore, arguably, being most appropriate for microsatellite data (e.g. Goldstein et al. 1995). However, also the SMM methods have many potential sources of bias. For example, the SMM methods assume that all mutations result in a change of one repeat unit, mutation rate is independent of a microsatellite's repeat length, and that a mutation is equally likely to cause an increase or decrease in repeat length; empirical evidence has revealed that all of these assumptions are often violated (e.g. Primmer et al. 1998). An alternative mutation model, the infinite allele model (IAM), assumes that each new mutation produces new and equidistantly related alleles. Also these assumptions are unrealistic in nature (Primmer et al. 1998). Consequently, it is important that a researcher is well aware of the exact assumptions of any particular method applied to analyse microsatellite data, and of the sources that can introduce bias.

In addition to the above-mentioned problems related to the application of microsatellites, the phylogenetic informativeness of these DNA loci has been suggested to diminish after only 3,000 to 30,000 generations since population divergence (Paetkau et al. 1997). This serious drawback has been suggested to result from frequent occurrence of 'allele homoplasy', i.e. alleles which are identical in state (i.e. have identical sizes), but are not necessarily identical by descent owing to convergent mutations (Paetkau et al. 1997). Accordingly, the use of microsatellites as phylogeographic markers has been directly criticized (Beaumont & Bruford 1999). Interestingly, however, some recent evidence implies that microsatellites can be informative for resolving phylogenies at substantial, and even at inter-specific divergence scales (Angers & Bernatchez 1998; Harr et al. 1998; Richard & Thorpe 2001; **II**; **III**), indicating that the issue requires further study.

Grayling (*Thymallus* spp.): natural distribution, life-history characteristics and cultural importance

This thesis focuses on genetic population studies of the salmonid genus *Thymallus* (Linnaeus 1758). The genus consists of four widely accepted species: European grayling (*T. thymallus*), Arctic grayling (*T. arcticus*), Mongolian grayling (*T. brevirostris*) and Amur grayling (*T. grubei*). *T. thymallus* occurs throughout most of Europe, whereas the Arctic, Mongolian and Amur grayling inhabit Lake Baikal and its surrounding great rivers. The distribution of *T. arcticus* also spans across North America.

Thymallus exhibit diverse life-history strategies. The majority of populations complete their entire life cycle in rivers or lakes (Northcote 1995), some grayling are adapted to a brackish environment and inhabit parts of the Baltic Sea (Nykänen & Huusko 1999), and also some limited anadromy has been reported (Northcote 1995). Unlike most other salmonids, grayling normally spawn in late March-June, when the water temperature reaches around 4-7°C (Northcote 1995). The spawning behaviour of grayling differs from that of other salmonids in that males acquire small spawning territories and defend access to the territory, rather than directly guarding females. During the spawning act, the eggs are deposited a few centimeters below gravel surface (Fabricius & Gustafson 1955).

Although less important for commercial fisheries than many other salmonids, *Thymallus* are considered culturally very important fishing species, and are highly valued among sports



Figure 1. Among other species of the genus *Thymallus*, European grayling (*T. thymallus*) is highly valued among sports fishermen. Overfishing of European grayling has led to its locally declining population sizes, and the species has been listed among the 'Threatened Fishes of Europe' (Lelek 1984). Photo: Rodney Calbrade. Reprinted with permission from The Grayling Society.

fishermen (Fig. 1; Broughton 2000). In fact, overfishing of *T. thymallus* has led to locally declining population sizes (Maitland & Campbell 1992), and the species was listed among the 'Threatened Fishes of Europe' in 1984 (Lelek 1984). In order to ensure the long-term survival of grayling, large-scale stocking practices are now commencing.

Outline of the thesis

This thesis applies mtDNA, microsatellite DNA and quantitative genetic methods to investigate how historical and contemporary evolutionary processes have influenced the spatial and temporal distribution of genetic diversity in grayling. The thesis begins with broad-scale phylogeographic studies of the genus. The phylogeographic studies focus on Lake Baikal and its surrounding rivers (**I**) and Europe (**II**, **III**), regions that cover the primary distribution of grayling. The thesis then focuses on microgeographic investigations in order to assess the level of natural inter-population migration across small spatial scales in *T. thymallus*, and to study the temporal effects of human-mediated gene flow in Lake Saimaa, eastern Finland (**IV**, **V**). The last chapter of the thesis further zooms in, both in evolutionary space and time, and applies a multifaceted analysis of quantitative traits and microsatellites, to unravel the relative importance of natural selection and random drift in contemporary *T. thymallus* microevolution (**VI**).

Summary of the original publications

I - Phylogeography of grayling within Lake Baikal and its surrounding watersheds

Lake Baikal, the world's oldest and deepest body of freshwater, lies near the headwaters of three great rivers: Enisey, Lena and Amur. Each of these rivers is thought to have been connected to Baikal in the past, providing colonization corridors for the ancestors of numerous zoogeographical enigmas, such as freshwater seals (*Phoca sibirica*), endemic families of freshwater sponges (Lubomirskiidae) and caddisflies (Limnephilidae), several endemic genera of leeches, and a monotypic family of turbellarians (Baicalarctidae) (Bănărescu 1991). Lake Baikal's paleo-hydrological development is of much interest for understanding the evolution of the Lake's curious organisms. However, due to their high level of

endemism, most taxa inhabiting Baikal provide little phylogeographic information. Thus, current theories on this ancient ecosystem's history are primarily based on geological evidence. One major issue of controversy is biologically very relevant: while some state that Baikal has never been glaciated (Kuz'min 2001), others speculate on cyclic patterns of glaciation paralleling Quaternary fluctuations elsewhere in Eurasia (Grosswald 1998). Such cataclysmic events should leave an imprint on the genetic architecture of fishes (Bernatchez & Wilson 1998).

Species of the genus *Thymallus* occur throughout Lake Baikal, its surrounding rivers, and isolated lakes of western Mongolia. Therefore, they provide an ideal model for investigating the phylogeography of this treasured ecosystem. In order to provide a biological perspective of the region's paleo-hydrological development and to assess how glaciations have influenced the genetic structure of *Thymallus*, we studied the demographic and genealogical history of grayling within Baikal and its surrounding watersheds (**I**). Our analyses were based on (i) mtDNA sequencing of the complete control region (circa 937 base pairs; bp) and 228 bp of flanking tRNA regions in 142 individuals, originating from 17 Asian populations and 6 locations in Europe and North America (Fig. 2); and (ii) ten nuclear microsatellite DNA loci genotyped in 346 individuals. These specimens were sampled from 11 populations in Asia, one from Europe, and one from North America (Fig. 2; **I**).

Phylogenetic reconstruction revealed six highly supported matrilineal clades, A-F (Fig. 2b). Lineage A contained all Baikal basin haplotypes as well as those from two populations in Baikal's outflow, Enisey (NTR and IRT; Figs. 2a, b). The Lena River haplotypes (LLL, LLO and KTR; Fig. 2a) formed a sister clade (B) to haplotypes from North America (C), whereas the *T. brevirostris* specimens (KN; D) were most closely related to a sample of *T. arcticus* captured in an extreme headwater region of the Enisey drainage, northwestern Mongolia (SHG; E). The most basal clade was formed by *T. grubei* haplotypes (ONG and SYP; F) from the Amur River (Figs. 2a, b; **I**).

Because dichotomously branching haplotype trees do not necessarily reflect population history (Goldstein et al. 2000), genealogy of *Thymallus* haplotypes was also evaluated using an unrooted network and a statistical parsimony criterion (Templeton et al. 1992). The network was used to

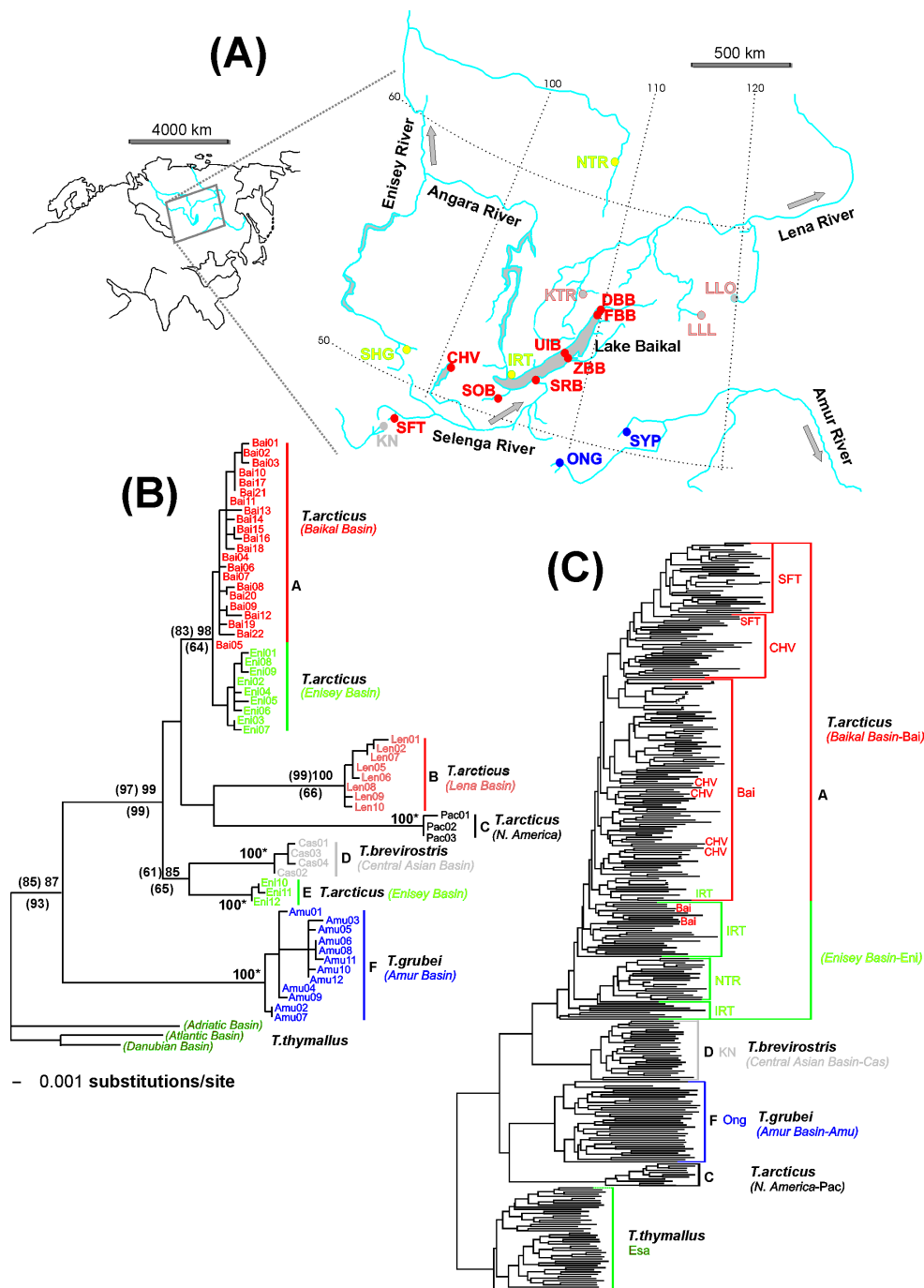


Figure 2. (A) A map indicating the origins of the populations utilized for constructing the demographic and genealogical history of grayling within Baikal and its surrounding watersheds (I). Complete population names and sample sizes are indicated in publication I. (B) The inferred mtDNA phylogeny. Node support is shown as the percentage of bootstrap replicates for maximum likelihood (100 replicates; below node) and Maximum Parsimony (1000 replicates; above node on the left) criteria and Neighbor Joining (1000 replicates; above node on the right) algorithm. 100* designates 100% support for all three methodologies. Six well-supported clades are labelled A-F. Haplotypes are color-coded to match their geographical origins on the map. The mtDNA haplotypes have been named according to the drainage basins. (C) A Neighbor Joining tree of individuals based on D_{AS} distances and ten microsatellite loci. Vertical bars indicate population clusters. Individuals that do not group within the correct clusters are marked with their population code beside the branches. Population codes and colors correspond to those on the mtDNA tree and the map. Population codes and colors correspond to those on the mtDNA tree and the map. Drainage basins have been indicated below the species designations.

apply Nested Clade Analysis (NCA; Templeton et al. 1995) for haplotypes occurring within Baikal and watersheds connected to Baikal (**I**). The nested design provided an inference of past fragmentation between the Baikal and Enisey haplotypes and supported a range expansion from Baikal into the Selenga River (**I**).

We further investigated the demographic history of Baikal mtDNA haplotypes using a pairwise mismatch distribution of haplotypes (Rogers & Harpending 1992). The distribution was uni-modal and provided a nearly perfect fit to a simulated model of sudden population expansion (Fig. 3). The coalescence estimates of the Baikal haplotypes implied that the expansion occurred approximately 110,000-450,000 ybp (**I**).

The mtDNA analyses revealed that old lineages of grayling (pre-Pleistocene, i.e. older than approximately circa 1,8 million ybp) currently inhabit Enisey, Lena and Amur (Fig. 2). For Lake Baikal, however, it appears that a mid-Pleistocene colonization or population size explosion has occurred (Fig. 3). This is interesting, because Baikal's previous outlet was to the Lena River via the ancient Manzurka River valley, existing through the Pliocene until a tectonic uplift resulted in the lake's isolation (Florensov 1978). If Lena basin grayling were abundant in Baikal [note that according to the geological record there should have been no barriers for dispersal until the isolation event (Florensov 1978)], they have either gone through an extreme bottleneck, or been completely extirpated, before the population expansion of Baikal's present inhabitants. The most

recent significant cooling periods of Baikal took place in the early Pleistocene, however, substantial temperature fluctuations have occurred throughout the last 800,000 years (Williams et al. 1997). It is therefore possible, that grayling experienced a bottleneck or were extirpated during any one of these paleoclimatic benchmarks. It should be noted, however, that salmonid species with complex life cycles, such as grayling, are sensitive to environmental perturbations (Allendorf & Waples 1996). Consequently, it is easy to envision less dramatic ecological changes in Baikal that could have substantially influenced the survival of *Thymallus*. For example, water-level fluctuations may have been detrimental for grayling by isolating or flooding tributaries suitable for successful spawning (**I**).

Interestingly, our estimated age of Baikal population expansion was fully concordant with a hypothesis regarding the formation of Baikal's contemporary outlet, the Angara river. Following Baikal's isolation from Lena, its water level has been presumed to have risen, until breaking through the Primorskiy mountain range approximately 200,000-300,000 years ago (Kozhov 1972). The outbreak of the Angara River provided a new hydrological corridor between the Enisey and Lake Baikal, which we hypothesize was associated with significant ecological changes in Baikal that provided an ideal setting for rapid population expansion. For instance, water level decrease resulting from formation of the Angara outlet may have substantially increased the area suitable for spawning of grayling (**I**).

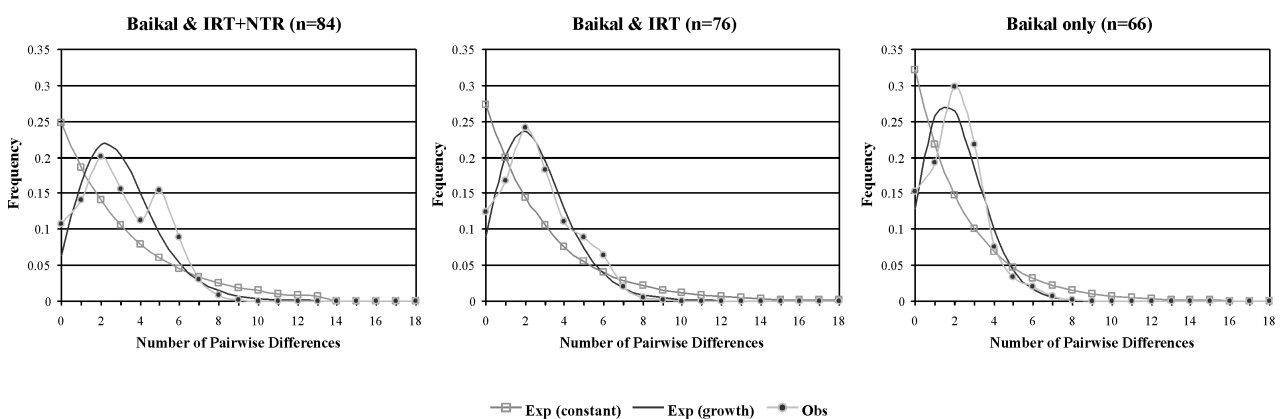


Figure 3. The pairwise mismatch distribution for the Baikal basin samples (**I**). The distribution is strongly bimodal when the IRT and NTR populations are included. When the distant NTR population is removed, the second mode nearly disappears, suggesting that it represented population subdivision. When both of the Enisey populations are removed, the leading wave completely disappears, and the remaining uni-modal distribution reflects a simulated model of sudden population expansion.

The overall population history of Asian *Thymallus* was well supported by the microsatellite D_{AS} tree of individuals (Fig. 2c). First, all individuals from Baikal and rivers presently connected to it clustered together. Second, the *T. brevivirostris* and *T. grubei* individuals were very diverged from all other grayling, *T. grubei* forming the most basal cluster with respect to the other Asian *Thymallus* (Fig. 2c). Interestingly, the Selenga river specimens (SFT and CHV) were most closely related to the Baikal grayling and formed the most internal branches of the tree (Fig. 2c). This observation was fully concordant with the results from the mtDNA based Nested Clade Analysis, implying that Selenga was colonized by a rapid range expansion from a single source, Lake Baikal. This expansion event apparently reached the extreme headwaters of Selenga, including the ancient Lake Chovsgul in Mongolia (I).

In order to investigate more recent demographic trends within Baikal, the microsatellite data were applied to test for variation in past population sizes, using a Markov Chain Monte Carlo (MCMC) simulation based method (Beaumont 1999; I). The simulations suggested that the contemporary Baikal *Thymallus* population size is roughly only 2-14% of its historical size. The most probable estimates for the commencement of this possible population size decline ranged from 20 to 50 years ago, depending on assumptions regarding microsatellite mutation rates (I). It is noteworthy, that such recent population declines are expected to have little or no effect on pairwise mtDNA mismatch distributions (Rogers 1996), i.e. the two marker types provided complementary, not contrasting, resolution. The Baikal grayling population is currently under a thriving sport and commercial fishery, and it has been recorded that overfishing has dramatically decreased grayling population sizes since the 1960's (Kozhova & Izmet'seva 1998). Microsatellites reflected the severity and timing of this anthropogenic influence strikingly well (I).

This study, involving one of the world's most enigmatic hydrological systems, revealed the dramatic influence that geological events have had on the contemporary genetic variation of the genus *Thymallus* inhabiting Lake Baikal and its surrounding rivers. MtDNA and microsatellites demonstrated that environmental perturbations have had a prevailing influence on the ability of *Thymallus* to survive in Baikal and supported the debated hypothesis that post-Pleistocene events have significantly altered the ecological conditions of the lake (Grosswald 1998). The evolutionary history of *Thymallus* implied an Enisey-Angara river route of colonization into Baikal,

corresponding to the paleo-hydrological event that led to the formation of Baikal's only contemporary outlet, and subsequent range expansion several thousand kilometers South into the Selenga River.

The comparative mtDNA and microsatellite data sets substantially add to the evidence that microsatellites can be used for resolving phylogenies even at inter-specific divergence scales (Harr et al. 1998; Richard & Thorpe 2001). Thereby, the general criticism against the utilization of microsatellites for assessing deep evolutionary histories (Paetkau et al. 1997; Beaumont & Bruford 1999) may be tenuous.

II, III - Phylogeography of European grayling

In contrast to many other salmonid species, phylogeographic information on European grayling remains very scarce. Phylogeographic studies of grayling would seem very important in order to: (i) help us further understand how European freshwater fish species were influenced by the Quaternary glaciations; and (ii) provide information that can be applied for developing appropriate conservation guidelines for this threatened salmonid. Indeed, *T. thymallus* is culturally an important fishing species and its population sizes have been severely declining for the past decade (Lelek 1984). Consequently, large-scale stocking programs are commencing, but the conservation actions are proceeding with little or no information regarding the distribution of abundant genetic variation in European grayling.

This thesis reports the first comprehensive phylogeographic data for *T. thymallus* by analysing a total of 623 individuals originating from across Europe (Fig. 4; II, III). The results are based on: (i) mtDNA polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) haplotyping of approximately half of the mitochondrial genome, and mtDNA sequencing of 529 bp of the ND5 gene (II); and (ii) 17 microsatellite DNA loci (III).

The mtDNA analyses revealed a total of 27 *T. thymallus* haplotypes which clustered into three phylogenetic lineages. The lineages exhibited nucleotide divergence estimates of 1.11-3.59%. Geographical distribution of the haplotypes indicated clear genetic distinction between: (i) northeastern (NE); (ii) northwestern/central (NW/C); and (iii) southern European (S) grayling (Fig. 4; II). Microsatellite based population phylograms supported these conclusions very well (Fig. 4; III). Collectively, our observations implied that postglacial colonization of northern Europe occurred from at least two diverged refugia isolated

well before the late Pleistocene period (defined as 130,000-10,000 ybp). We suggest that grayling colonizing Norway and mainland Sweden originated from a central European refugium, which possibly acted as a source also for postglacial dispersal into central Germany, Denmark and Poland. It can be speculated that the Elbe river system was utilized by the dispersing grayling, similarly as has been suggested for whitefish, *Coregonus clupeaformis* (Hansen et al. 1999).

Interestingly, microsatellites also indicated grouping of the NW/C and S populations with relatively high bootstrap support (Fig. 4; **III**). It therefore appears possible, that southern European grayling were the ancient ancestors of the NW/C populations. A feasible northward colonization route would have been via the Danube into the Elbe during ancient times, when the current headwaters of Elbe drained into the Danube (circa 1,8 million ybp; Hantke 1993). A Danubian origin has been previously suggested also for the central and northern European populations of perch (*Perca fluviatilis*; Nesbø et al. 1999) and bullhead (*Cottus gobio*; Englbrecht et al. 2000). Confident assessment of the inter-relationships and coalescence times of the NE, NW/C and S lineages requires further study. However, a hierarchical analysis of molecular variance (AMOVA) suggested that the split of the northern European assemblages (NE and NW/C) post-dates the divergence of the NW/C grayling from the southern populations (**III**).

European grayling populations that inhabit currently hydrologically connected locations can apparently be genetically very differentiated (**II**, **III**). For instance, the populations originating from Ulkokrunnit (Kru) and Tornionjoki (Tor; Fig. 4) exhibit similar microsatellite differentiation as populations of Arctic charr (*Salvelinus alpinus*) that are separated by distances 10 times as great (Primmer et al. 1999). There are two alternative explanations for the high level of genetic divergence between the sympatric populations: (i) they have originated from more than one already reproductively isolated glacial refugium, and persisted in isolation in postglacial times, despite potential for physical contact (i.e. the populations have become into 'secondary contact'). This has been recently argued to have occurred between some lake whitefish populations in St. John River basin, northeastern North America (Lu et al. 2001); or (ii) the populations have originated from the same panmictic refugium, but postglacial migration has been sufficiently limited to enable genetic drift

or mutations to cause the present divergence. Distinction between these alternatives may benefit from denser sampling across the natural range of grayling. At present, however, there is no reason to suggest that e.g. the abovementioned populations Kru and Tor would originate from more than one glacial refugium (**III**).

A recent meta-analysis of microsatellite polymorphism displayed by freshwater, anadromous and marine fishes reported freshwater populations to display, on average, 9.1 (± 6.1) alleles and expected gene diversities (H_E) of 0.54 (± 0.25 ; DeWoody & Avise 2000). Our microsatellite data indicated that European grayling populations exhibit, on average, only 3.5 (± 2.2) alleles and an H_E of 0.41 (± 0.27 ; **III**). Because the overall diversity of the 17 microsatellites was relatively high (5-58 alleles per locus across all populations), the low level of *T. thymallus* diversity was not an attribute of the loci, but of the populations.

The depauperate microsatellite diversity might relate to the life-history characteristics of European grayling. One possibility is, that *T. thymallus* display limited dispersal from their natal sites. This could result in low long-term effective population sizes and, consequently, low levels of within-population variation. An alternative explanation may relate to the tendency of grayling juveniles to favor shallow habitats, presumably due to their poor swimming performance (Nykänen & Huusko 1999). It has been argued that this tendency may result in massive juvenile mortality when water levels fluctuate, because the young grayling can not effectively switch habitat into deeper waters, unlike e.g. *Salmo trutta* or *S. salar* (Nykänen & Huusko 1999). Such year class 'drop-out' events could decrease effective population sizes and genetic diversity.

Populations with independent evolutionary histories, thereby contributing most to the overall intra-specific biodiversity, should arguably be prioritized for conservation (e.g. Moritz 1994). One alternative for direct application of genetic results in the development of appropriate conservation guidelines involves recognition of so-called management units (MU). Distinct MUs have been defined as populations connected by little or no contemporary gene flow, but not separated historically for very long periods of time (Waples 1991; Moritz 1994).

Our mtDNA and microsatellite results demonstrate that grayling populations separated by waterway distances in the range of few hundred kilometers appear to fulfill such criteria for being

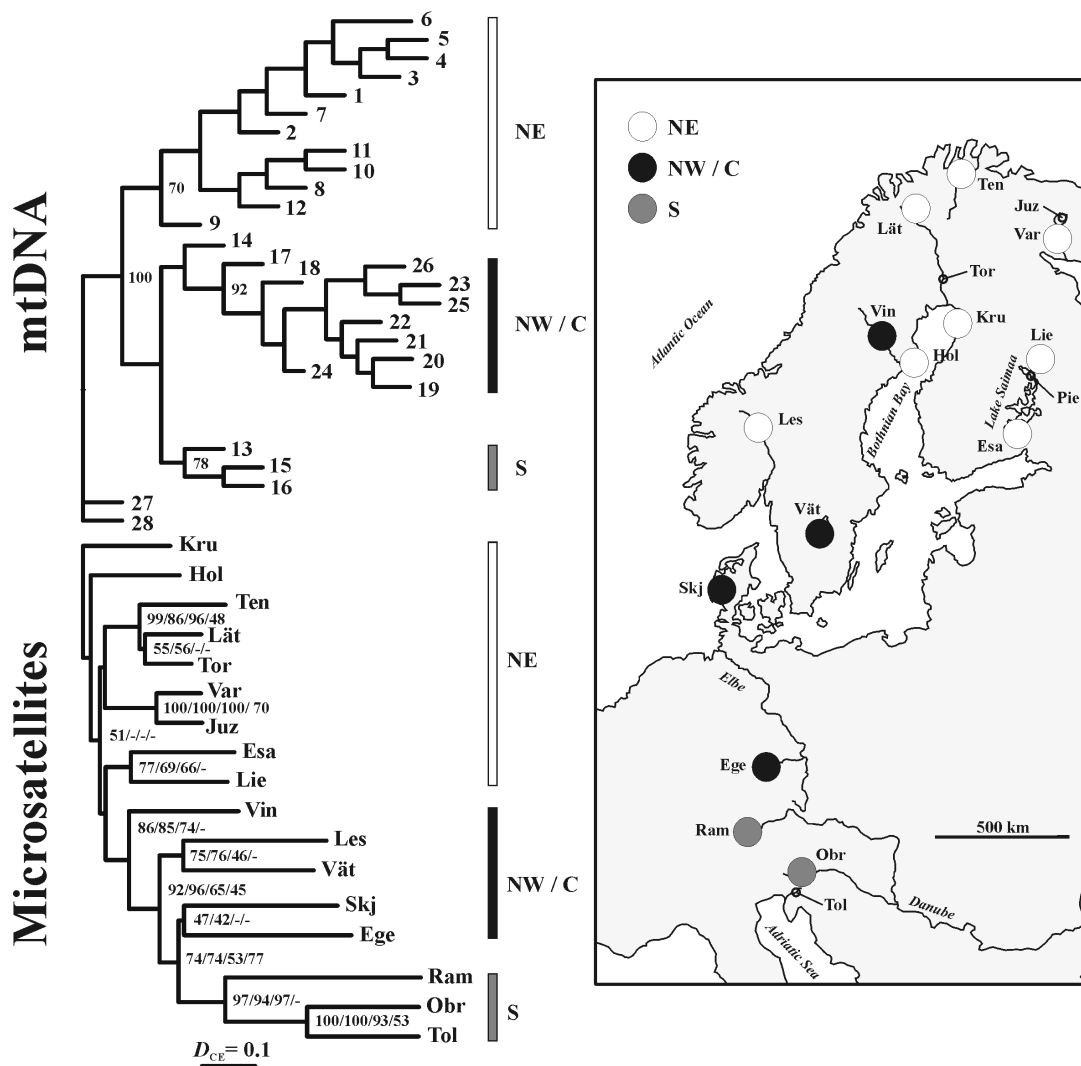


Figure 4. Sampling locations and genetic relationships of *T. thymallus* populations (II, III). The mtDNA tree is based on pooled PCR-RFLP site data, covering the ND5, ND6, *cyt b* and D-loop regions, and 529 bp of ND5 sequence data (II). The microsatellite population phylogram is based on analysis of 17 loci and D_{CE} genetic distance (III). The bootstrap support estimates are indicated beside the branches of the trees and the vertical bars denote the northeastern (NE; white), northwestern/central (NW/C; black) and southern (S) genetic lineages. The pie charts of the map indicate the distribution of the resolved mtDNA haplotype lineages. The shades of grey of the pie charts correspond to those of the vertical bars denoting the main mtDNA and microsatellite phylogram lineages.

recognized as separate MUs (II, III). It also seems possible that substantial genetic differences may occur at much smaller geographical scales (II, III).

To conclude, northern European grayling populations have originated from at least two refugia that have been genetically isolated for hundreds of thousands of years. Southern Europe harbours additional differentiated grayling lineages, which may have served as the ancient ancestors of the NW/C (and possibly also the NE) populations. The population groups have remained largely allopatric also in postglacial times, and contemporary inter-population gene flow has been

very limited, even in cases where possibilities for physical contact exist. A large proportion of the total *T. thymallus* genetic diversity is distributed among populations, intra-population diversity being atypically low. We suggest that European grayling populations be considered as separate MUs even across relatively short geographical distances.

IV, V - Microgeographic evolutionary relationships

Phylogeographic studies of *T. thymallus* provided preliminary implications that the species

may exhibit substantial genetic differentiation across short waterway distances (II, III). This was suggested to result from restricted migration between spawning sites, leading to the generation of divergence via random drift (II, III). Detailed microgeographic investigations would seem useful to further address this issue. Of particular interest would be to apply individual multilocus genotype-based approaches to study reproductive dispersal rates between populations (Bowcock et al. 1994; Rannala & Mountain 1997; Cornuet et al. 1999; Pritchard et al. 2000).

Grayling populations inhabiting Lake Saimaa (eastern Finland) have been severely declining in recent years (Sundell 2000). Consequently, the Finnish fisheries authorities and water owners have shown much interest towards creating hatchery broodstocks, which are utilized to supplement the declining populations with an aim of ensuring the long-term sustainability of the species within Lake Saimaa. However, the creation of the hatchery stocks, as well as the introductions, are proceeding with insufficient genetic information (Koskiniemi & Kilpinen 1987).

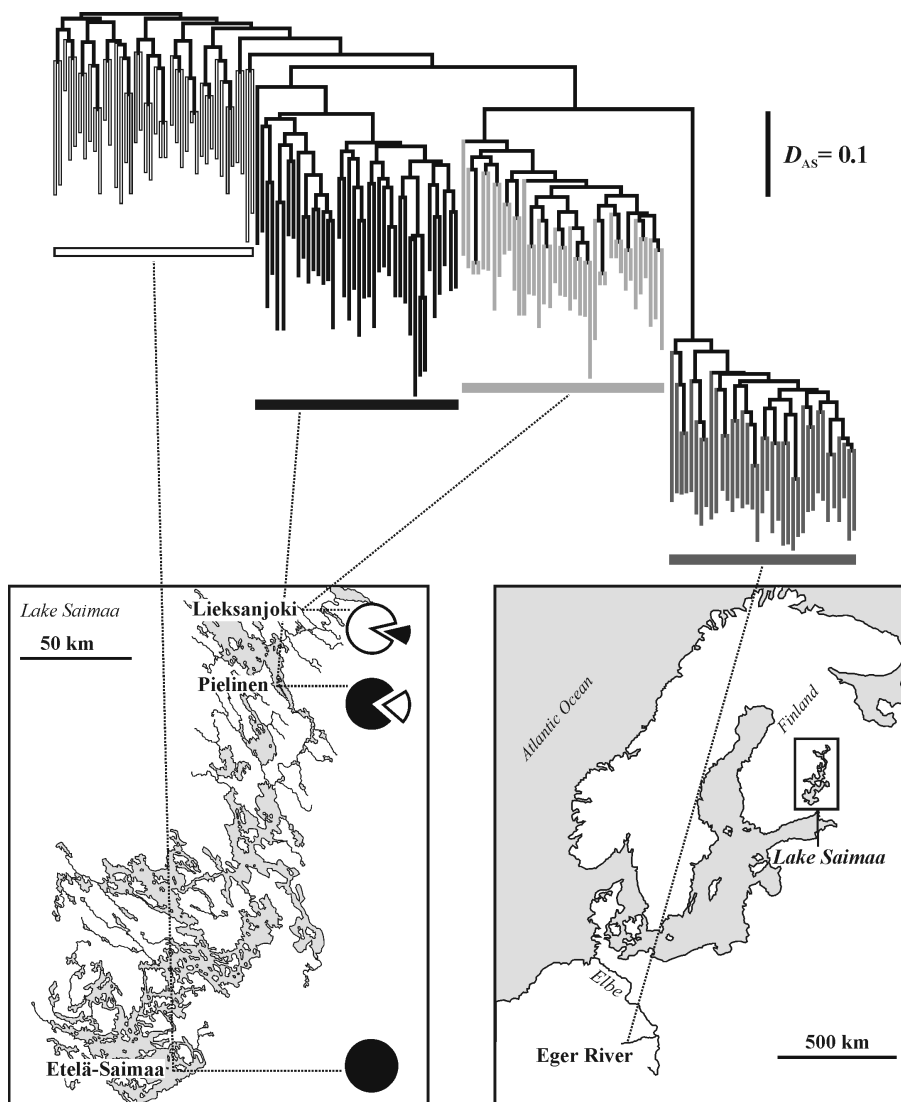


Figure 5. A Neighbor-Joining D_{AS} phylogram constructed for *T. thymallus* individuals based on data from 17 microsatellite loci, maps presenting the study regions and frequencies of two mtDNA composite haplotypes within the populations (IV). Shadings on the branches of the tree correspond to the origins of the specimens as follows: white, Etelä-Saimaa; black, Pielinen; light grey, Lieksanjoki; dark grey, Eger River (Germany). The dashed lines and horizontal bars indicate the sampling locations of the respective individuals. The pie charts denote frequencies of the populations for the mtDNA composite haplotypes aaaa (black) and bbbb (white). PCR-RFLP divergence of the two haplotypes was 1.57% (see publication II for further information on the haplotypes).

In order to gain further insights to the level of gene flow among *T. thymallus* populations across small spatial scales, and to provide information useful for the development of appropriate conservation guidelines, we conducted a microgeographic study of Lake Saimaa grayling populations. We applied: (i) mtDNA PCR-RFLP analyses with ‘diagnostic’ restriction endonucleases (the informativeness of the mtDNA region-restriction enzyme combinations was revealed in a previous study; **II**); and (ii) population and individual based microsatellite analyses (**IV**, **V**).

Two rather diverged (pairwise divergence=1.54%) mtDNA composite haplotypes were detected within Lake Saimaa. The divergence estimate of the haplotypes suggested that their separation pre-dates the postglacial colonization of Lake Saimaa, which occurred circa 10,000 ybp (**IV**). However, it remains to be determined whether: (i) Lake Saimaa was colonized from two diverged sources; or (ii) Lake Saimaa was invaded from a single refugium that was already polymorphic for the two haplotypes. The different spawning sites exhibited clear frequency differences for the two haplotypes (Fig. 5), which gave preliminary implications for limited gene flow between the populations across small spatial scales.

Microsatellites revealed highly significant ($P < 0.001$) genic differentiation between the grayling populations originating from different regions of Lake Saimaa (**IV**, **V**). The level of divergence was often substantial even across short waterway distances. For instance, the populations Lieksanjoki and Pielinen, separated by 55 kilometers (Fig. 5), exhibited a pairwise F_{ST} estimate of 0.30 (**IV**). Further, the Pudaskoski (LieP-96) and Naarakoski (LieN-97-98) populations, separated by only 10 kilometers (**V**), were highly significantly differentiated and had a pairwise F_{ST} of 0.11 (**V**). An NJ dendrogram based on the D_A genetic distance revealed five relatively well supported clusters, which coincided with the sampling origins of the populations (Fig. 6; **V**). The high levels of inter-population divergence and the resolution of the NJ dendrogram implied restricted (or no) contemporary gene flow among the Lake Saimaa grayling populations.

An individual based D_{AS} dendrogram clustered all of the specimens according to their sampling origins (Fig. 5; **IV**). Various individual assignment tests revealed 100% assignment success of individuals into their correct population of origin and 100% exclusion success ($P \leq 0.05$) of individuals from all alternative reference populations (**IV**). Hence, the individual level

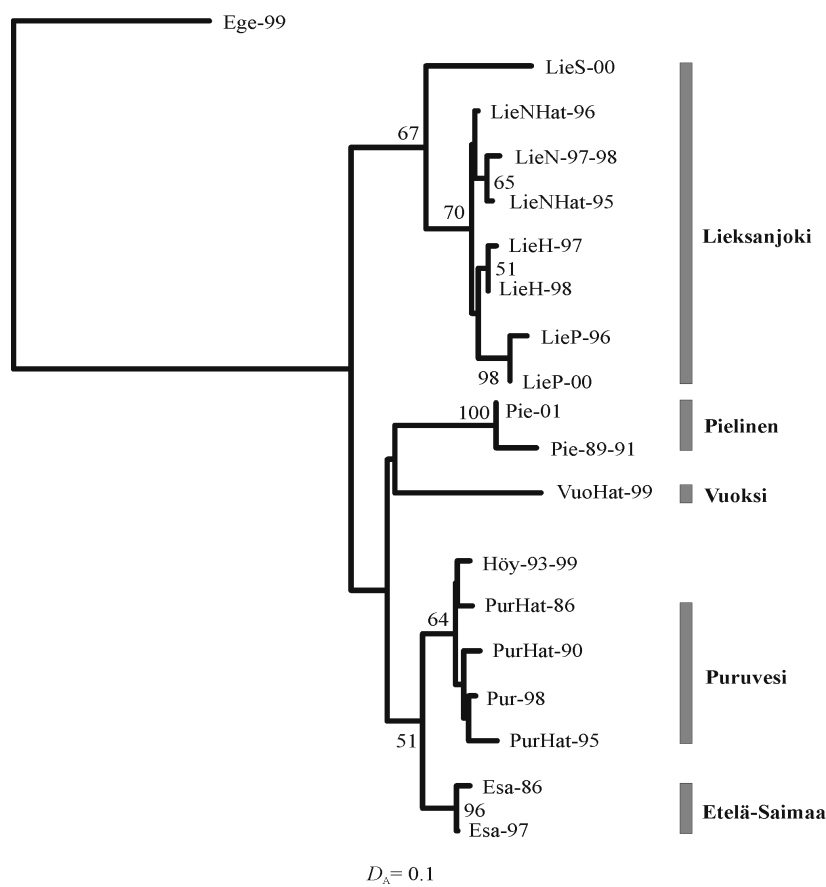
analyses further indicated that migration between different grayling populations is very limited.

We have demonstrated that Lake Saimaa is inhabited by a number of substantially genetically differentiated grayling populations that form five main groups coinciding relatively well with the geographic origins of the samples (**IV**, **V**). Our population and individual level mtDNA and microsatellite data suggest that grayling may exhibit very limited inter-population dispersal even across short waterway distances, and assist in the development of appropriate conservation guidelines for Lake Saimaa *T. thymallus*. It appears that the locations corresponding to the five main groups (Fig. 6) should be exclusively managed with hatchery populations created using locally caught founders. Moreover, consideration of population structuring within the main regions may be warranted.

V - Spatiotemporal assessment of stocking effects

Gene flow from domesticated populations into the wild, whether accidental (escapes from domestically reared populations) or deliberate (introductions), has become an important factor shaping contemporary evolutionary relationships in nature. It has been suggested that non-native specimens pose serious threats to natural populations via decreasing the overall inherited diversity, reduction of fitness due to outbreeding depression and loss of local adaptations, especially in salmonids owing to their complex life-history strategies (Allendorf & Waples 1996). In light of these concerns, it would seem important to monitor how human-mediated gene flow has affected the indigenous evolutionary relationships. Such results can be applied, e.g. for prioritizing the conservation of populations that have retained most of their natural genetic structure. A common approach has been to assess genetic differences between contemporary wild and introduced populations and to infer their admixture levels based on ‘diagnostic’ marker alleles (e.g. Gottelli et al. 1994). An alternative strategy has been to utilize individual assignment tests (e.g. Cornuet et al. 1999; Pritchard et al. 2000) to identify indigenous and introduced specimens and, subsequently, estimate their proportions in the wild (reviewed in Hansen et al. 2001). An optimal way to overcome this limitation would be to make use of archival genetic material collected prior to introductions as baseline data. However, such studies remain scarce (Nielsen et al. 1997; reviewed in Hansen et al. 2001; Nielsen et al. 2001; Hansen 2002).

Figure 6. A Neighbor-Joining phylogram of Lake Saimaa grayling populations based on 10 microsatellite loci and D_A genetic distance (V). Numbers beside the branches represent bootstrap support across 2,000 replicates of resampled loci (only values above 50% are indicated). Grey vertical bars indicate geographical origins of the populations. The complete names of the populations and their precise sampling locations are indicated in publication V and Fig. 7.



In this study, we used 10 microsatellite loci to assess the effects of introductions on the genetic structure of Lake Saimaa grayling populations (V). Stocking has been carried out in Lake Saimaa since 1986 using hatchery populations founded with spawning individuals originating from the Puruvesi region (Fig. 7). The Puruvesi individuals have been used for supplementing the Pielinen (from 1989 to 1998) and Etelä-Saimaa (from 1986 to 1999) spawning grounds (Fig. 7). In order to enable a rigorous assessment of the potential effects of stocking, we analysed DNA samples from: (i) Pielinen and Etelä-Saimaa collected prior to the introductions; (ii) the Puruvesi populations used for the introductions; and (iii) wild populations sampled after the introductions had occurred (V).

Despite that the Pielinen and Etelä-Saimaa populations had been extensively supplemented using the Puruvesi hatchery broodstocks, genic differentiation between the historical and contemporary samples was statistically insignificant ($P > 0.05$) within Pielinen and Etelä-Saimaa (V).

Accordingly, the D_A genetic distance based phylogram revealed highly supported grouping of the historical and contemporary populations within Pielinen (Pie-89-91 and Pie-01) and Etelä-Saimaa

(Esa-86 and Esa-97; Fig. 6; V). Hence, the grayling populations had retained much of their indigenous genetic structure throughout the introductions. Although requiring further study, it seems possible that many of the hatchery individuals stocked into the wild did not survive. This suggestion parallels the results of other salmonid studies, which have hypothesized that natural selection may be acting against introduced individuals (e.g. Skaala et al. 1996; Poteaux et al. 1998; Hansen et al. 2000). Potential reasons for such fitness differences remain largely unknown. It has been suggested, however, that non-native individuals have poor predator avoidance skills (Einum & Fleming 1997).

Approaches based on individual multilocus genotypes provided further resolution of stocking effects by enabling confident assessment of the origins of grayling within the Pielinen and Etelä-Saimaa regions. We were able to show that substantial proportions of the contemporary wild grayling most likely represented the indigenous populations unaffected by the introductions: 73% of the contemporary Pielinen individuals could be assigned to the local reference populations and excluded from all of the Puruvesi hatchery broodstocks. Furthermore, 48% of *T. thymallus* presently inhabiting Etelä-Saimaa could be

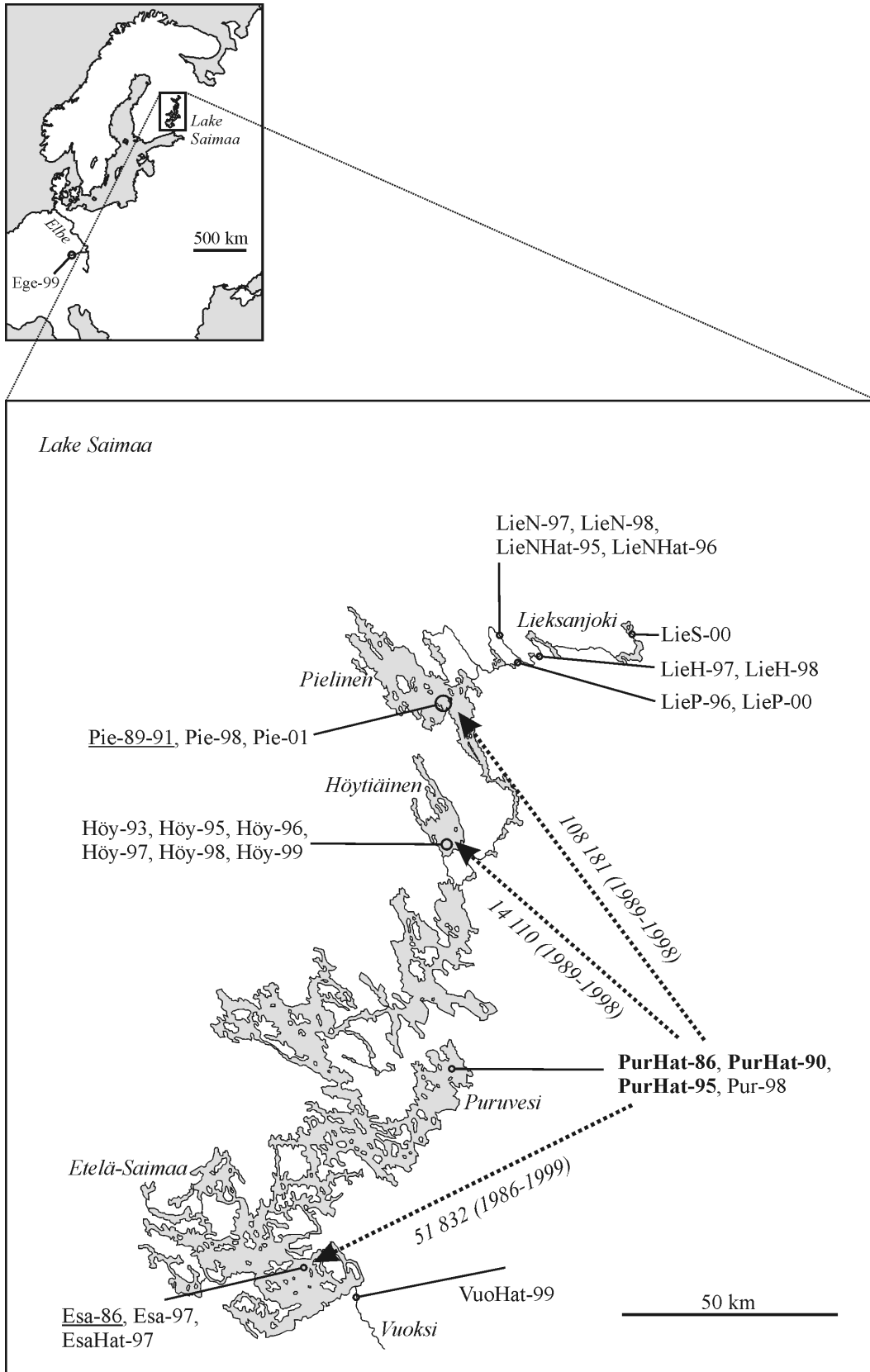


Figure 7. A map presenting the origins of the Lake Saimaa populations studied in publication V. The Puruvesi hatchery broodstocks (marked in bold) were used for introducing grayling into the Pielinen, Höytiäinen and Etelä-Saimaa regions. The numbers of stocked specimens and the years when the introductions occurred are indicated. The historical Pielinen and Etelä-Saimaa populations (underlined) were sampled prior to the commencement of the stocking activities.

excluded from all of the Puruvesi populations and assigned to Etelä-Saimaa (Table 1). These findings reinforced the population-level inferences that Lake Saimaa grayling had retained much of their indigenous genetic structure (V).

Interestingly, however, the level of differentiation between the wild populations and populations used for their supplementation had decreased across time (Fig. 8). The temporal decrease in F_{ST} averaged 0.06-0.03 units across all loci, corresponding to 15.0-18.2% in the Pielinen vs Puruvesi comparisons and 38.5-44.4% in the Etelä-Saimaa vs Puruvesi comparisons (Fig. 8; V).

Individual assignment tests revealed that 4% and 15% of the contemporary Pielinen and Etelä-Saimaa individuals, respectively, were excluded from all of the local populations but not from the Puruvesi broodstocks, indicating that these individuals were most likely of hatchery origin or pure descendants of the introduced specimens.

In contrast, none of the historically sampled Pielinen or Etelä-Saimaa grayling were excluded from the local wild populations while being assigned to the hatchery populations (Table 1), an observation very well in line with the idea that

Puruvesi hatchery individuals now inhabit Lake Saimaa due to stocking (V).

Although the population-level analyses and individual assignment tests indicated that stocking had resulted in accumulation of hatchery genomes into the indigenous populations, they did not indicate whether *inherited* population structure changes had occurred. Can introgressive hybridization between the hatchery and indigenous grayling be revealed in the wild? A clustering method introduced by Pritchard et al. (2000) was specifically applied to address this question (V).

Interestingly, many of the Pielinen grayling, captured in 2001, exhibited an intermediate admixture coefficient (q), but this was not the case with the historical *T. thymallus* (Fig. 9). This is exactly the pattern that one would expect if hybridization had occurred following the introductions (Pritchard et al. 2000).

It was also interesting that, although a proportion of the Pielinen individuals caught in 1998 were clearly of hatchery origin ($q \sim 0$), none of the specimens in this population had intermediate estimates of q , suggesting that hybridization between Puruvesi and Pielinen grayling had occurred only after 1998 (Fig. 9; V).

Table 1. Percentage of individuals assigned to, and excluded from ($P \leq 0.05$), Lake Saimaa grayling populations (V). (A) Pielinen (Pie) and Puruvesi (Pur) individuals assigned into at least one of the Pie populations while being excluded from all of the Pur populations, and assigned into at least one of the Pur populations while being excluded from all of the Pie populations; (B) Etelä-Saimaa (Esa) and Puruvesi individuals assigned into at least one of the Esa populations while being excluded from all of the Pur populations, and assigned into at least one of the Pur populations while being excluded from all of the Esa populations (V).

(A)

Population being classified	n	Assigned to Pie and excluded from Pur	Assigned to Pur and excluded from Pie
Pie-89-91	30	83	0
Pie-01	79	73	4
PurHat-86	19	0	84
PurHat-90	35	0	80
PurHat-95	39	0	82

(B)

Population being classified	n	Assigned to Esa and excluded from Pur	Assigned to Pur and excluded from Esa
Esa-86	37	41	0
Esa-97	27	48	15
PurHat-86	19	0	32
PurHat-90	35	0	54
PurHat-95	39	0	62

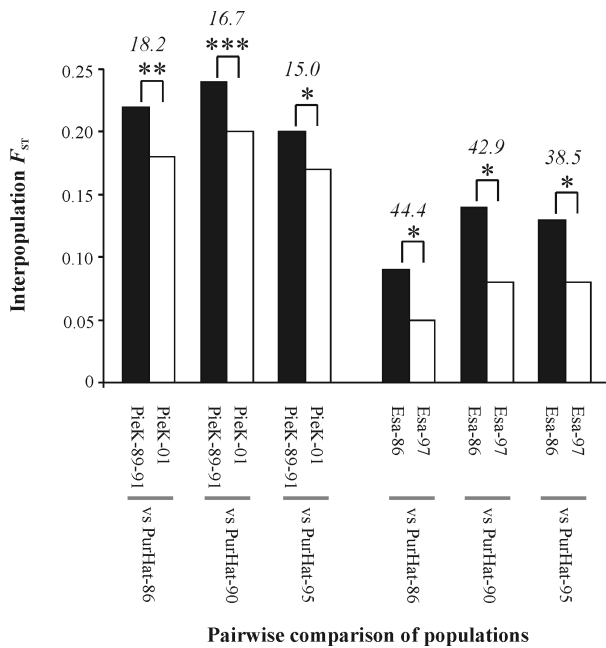


Figure 8. Temporal changes in F_{ST} estimates between Lake Saimaa Pielinen and Etelä-Saimaa grayling populations versus the Puruvesi hatchery broodstocks used for stocking these two regions (V). Pie-89-91 and Esa-86 are historical baseline samples collected prior to the commencement of the stocking activities. Pairwise F_{ST} estimates for each historical (black bars) and contemporary (white bars) population versus each of the three Puruvesi broodstocks are indicated. The numbers in italics show how much the pairwise wild versus hatchery F_{ST} estimates had decreased over time. The probability values for statistical significance of change in F_{ST} were obtained using Wilcoxon's signed rank test across loci. *, $0.01 < P \leq 0.05$; **, $0.005 < P \leq 0.01$; ***, $P \leq 0.005$ (V).

Averaging the estimates of q across individuals suggested that the level of hatchery x indigenous admixture, at the population level of resolution, had increased 7-11% within Pielinen (between 1989 and 2001) and 8-16% within Etelä-Saimaa (between 1986 and 1997; V).

This study demonstrated that analyses of historical baseline samples can enable rigorous quantification of stocking effects. Although our results provided clear detection of stocking-related genetic changes, the wild populations had retained much of their indigenous genetic composition. These data demonstrated that natural populations can maintain substantial proportions of their evolutionary relationships under intensive introductions, and aid in the development of management strategies for Lake Saimaa grayling. The Finnish fisheries managers now have a tool to

identify the origins of grayling inhabiting Lake Saimaa, enabling the selection of pure indigenous fish for creating hatchery broodstocks.

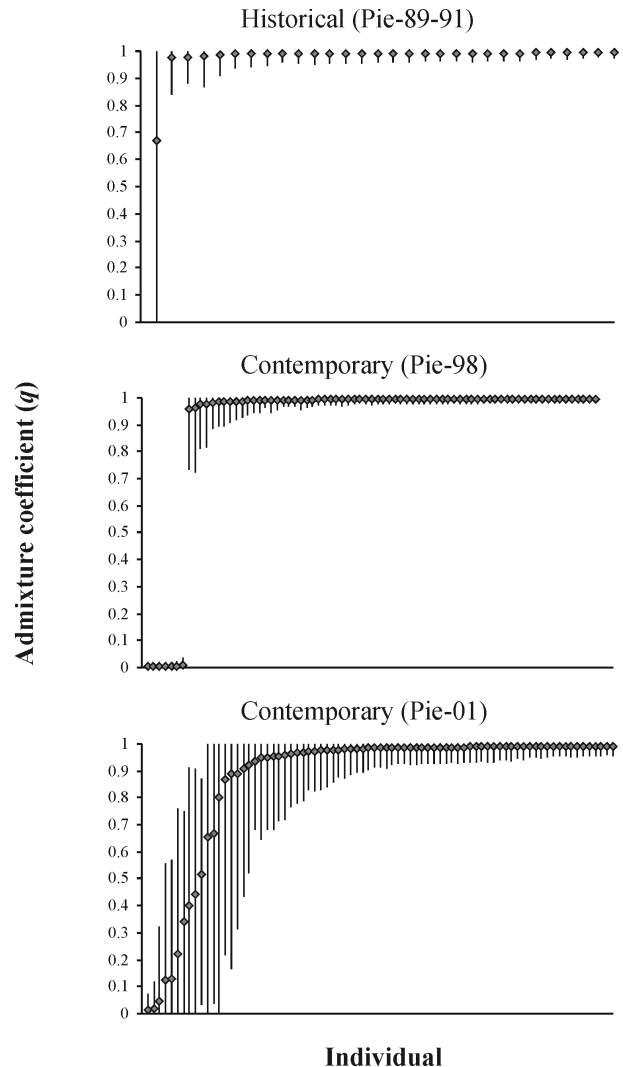


Figure 9. Posterior distributions of individual admixture coefficients (q) and their 90% probability limits in the historical Pielinen (Pie-89-91) and the contemporary Pielinen (Pie-98 and Pie-01) populations (V). The PurHat-86 broodstock was used in the analyses, however, the results were similar regardless of which Puruvesi hatchery stock was included as a reference population. An estimate of $q=1$ denotes pure indigenous genotype, $q=0$ pure hatchery genotype and $q=0.5$ may indicate first generation hybrid between the indigenous and hatchery populations. The individuals have been ordered from left to right with increasing estimates of q (V).

Microsatellites and the application of historical baseline material should also prove efficient for monitoring gene flow between domesticated and wild populations in other species. Such studies may be most relevant in the field of fisheries (Nielsen et al. 1997; Hansen 2001; Nielsen et al. 2001; Hansen 2002). However, similar applications may prove important, e.g. for studies of introgression between domesticated and wild canidae (reviewed in Wayne 1996), felidae (Beaumont et al. 2001), game animals (Thulin et al. 1997) and plants (Ellstrand 1992).

VI - Contemporary life-history evolution in grayling

The relative importance of fisherian natural selection (Fisher 1930) and random genetic drift (Wright 1931) in phenotypic evolution has been an enduring subject of discussion since the introduction of the first population genetic models (Kimura 1983; Gillespie 1991; Schluter 2000). The evolutionary theories warrant further study because: (i) formal tests for neutral divergence (Lande 1976; Lande 1977; Carson & Templeton 1984; Lynch & Hill 1986) have been rarely applied to wild organisms; (ii) sensitive attempts to separate the effects of selection and drift are scarce

(Baker 1992; Merilä & Crnokrak 2001); and (iii) studies linking the relative importance of selection and drift with well-known population demography are non-existent (VI; Schluter 2000). Here we present an investigation in an exceptional natural model system that provides a sensitive framework for determining the mechanisms driving the very early stages of phenotypic evolution (VI).

Grayling inhabiting Lake Lesjaskogsvatn (Les), Norway, originate from a single human-mediated introduction in 1880 (Fig. 10). Thirty years later, a local fisherman carried a small number of individuals from Les into two nearby mountain lakes, Hårrtjønn (Ht) and Øvre Mærrabottvatn (ØM). The exact number of individuals in the fisherman's bucket remains unknown, but was very small because Ht and ØM are separated from Les by five hours of demanding uphill hiking, making it impossible for the fisherman to transfer large volumes of water (J. Nordsletten, N-2667 Lesjaverk, Norway, pers. comm.). The founders of Ht and ØM (or their offspring) subsequently dispersed downstream into Aursjøen (Aur), where grayling were first observed in 1920 (Fig. 10). All *T. thymallus* in the system (Fig. 10) are known to originate from these stocking and dispersal events (J. Nordsletten, N-2667 Lesjaverk, Norway, pers. comm.).

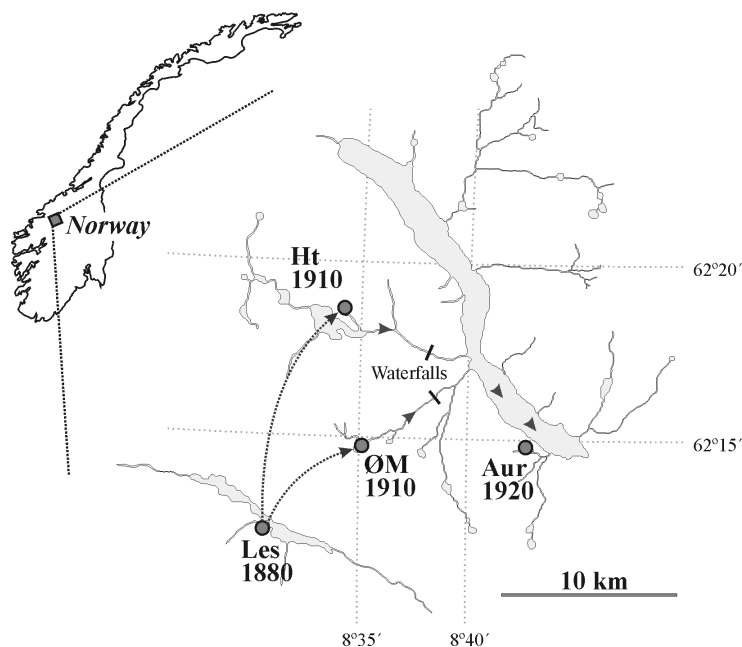


Figure 10. A map presenting the study region and location of the *T. thymallus* populations included in publication VI. The population Lesjaskogsvatn (Les) was founded in 1880 and used for introducing a small number of grayling into the lakes Hårrtjønn (Ht) and Øvre Mærrabottvatn (ØM) in 1910. These two populations were the sources for colonization of Lake Aursjøen (Aur), where grayling were first observed in 1920. All grayling in the system originate from these introductions. Land or impassable waterfalls prevent migration between Les-Ht, Les- ØM and Ht-ØM (VI).

When the divergence times of the populations are combined with their generation intervals in Les, Ht, ØM and Aur, it follows that the populations coalesce to a common ancestor only 11.8-22.0 generations ago (Haugen & Vøllestad 2001). With so few generations, diversifying mutations are negligible causes of phenotypic differentiation [see discussion in Lynch & Walsh (1998), page 328-352]. Further, migration to three of the populations has been impossible since their founding, either because of barriers formed by land (Les; Fig. 10) or impassable waterfalls (Ht and ØM; Fig. 10).

We tested the null-hypothesis of neutral evolution for additive genetic variances of the following seven life-history traits obtained from a 'common-garden' experiment: (i) incubation time (number of days from fertilization to hatching); (ii) fork length at hatching; (iii) yolk-sac volume; (iv) fork length at swim-up; (v) fork length at

termination of the experiment; (vi) specific growth rate; and (vii) survival rate during 180 degree-days of exogenous feeding. The null-hypothesis was tested using the F -statistic developed by Lande (1977):

$$F=(N_e\sigma_{GB}^2)/(h^2\sigma_{GWT}^2), \text{ where:}$$

$-N_e$ is harmonic mean effective population size, estimated using 17 microsatellite loci and a coalescent theory maximum likelihood based approach (O'Ryan et al. 1998).

$-\sigma_{GB}^2$ and σ_{GWT}^2 are additive genetic variances between and within populations, respectively.

$-h^2$ is narrow-sense heritability of a trait in a given population and environment.

$-t$ is divergence time of the populations in generations. Further details regarding these variables are provided in publication VI.

Table 2. Evidence for non-neutral evolution of seven life-history traits in three pairwise population comparisons (VI).

Trait	Pairwise comparison	h^2 *	σ_{GWT}^2 †	σ_{GB}^2 ‡	$F_{1,\infty}$ §	P	$N_e(\text{sign})$ ¶
length at termination	Les-Ht	0.10	0.08	1.47	384	***	0.25
	Les-Aur	0.06	0.18	0.03	10.7	*	22.6
	Ht-Aur	0.06	0.27	0.72	134	***	0.93
yolk-sac volume	Les-Ht	0.05	0.17	2.40	647	***	0.14
	Les-Aur	0.03	0.44	5.55	1.78×10^3	***	0.12
	Ht-Aur	0.03	0.46	0.45	99.3	***	1.54
growth rate	Les-Ht	0.18	6.57×10^{-5}	6.55×10^{-4}	111	***	0.84
	Les-Aur	0.22	7.23×10^{-5}	9.72×10^{-5}	28.2	***	7.54
	Ht-Aur	0.22	1.73×10^{-5}	2.22×10^{-4}	194	***	0.79
survival	Les-Ht	0.09	0.08	0.23	65.7	***	1.42
	Les-Aur	0.13	0.12	0.08	22.0	***	8.84
	Ht-Aur	0.13	0.27	0.10	9.43	*	16.6
incubation time	Les-Ht	0.22	0.26	0.50	17.4	**	5.33
	Les-Aur	0.44	0.30	1.00×10^8	3.45×10^{-7}	>0.99	6.16×10^8
	Ht-Aur	0.44	0.24	0.58	18.0	**	8.49
swim-up length	Les-Ht	0.11	0.02	2.50×10^{-3}	2.50	0.12	37.1
	Les-Aur	0.32	0.04	3.13×10^{-4}	0.12	0.73	1.73×10^3
	Ht-Aur	0.32	0.11	9.28×10^{-10}	8.32×10^{-8}	>0.99	1.83×10^9
hatching length	Les-Ht	0.15	0.51	3.62×10^{-9}	9.53×10^{-8}	>0.99	9.76×10^8
	Les-Aur	0.08	0.27	2.37×10^{-9}	4.84×10^{-7}	>0.99	4.40×10^8
	Ht-Aur	0.08	0.03	6.09×10^{-4}	0.70	0.40	218

*Narrow-sense heritability for the youngest population in each comparison. †Additive genetic variance within populations. ‡Additive genetic variance between populations. § $F=(N_e\sigma_{GB}^2)/(h^2\sigma_{GWT}^2)$, where N_e is the maximum likelihood estimate of effective population size, and t equals 11.8, 11.8 and 12.3 generations between Les-Ht, Les-Aur and Ht-Aur, respectively. ||The P -values have been adjusted for multiple significance tests using a Bonferroni correction. ***, $P<0.0001$; **, $0.0001 \leq P < 0.001$; *, $0.001 \leq P < 0.01$. ¶Indicates the level of N_e required to produce an F -value leading to rejection of the null-hypothesis at the $P \leq 0.05$ significance level ($F_{1,\infty}=3.84$).

Neutral evolution was rejected for the majority of the life-history traits with a very high level of statistical confidence (Table 2; **VI**). Evidence against divergence at a neutral rate was compelling in all of the pairwise population comparisons for length at termination, yolk-sac volume, growth rate and survival rate during 180 degree-days of exogenous feeding (Table 2).

The neutrality test was also applied to estimate the extent of N_e required to produce an F -value leading to rejection of the null-hypothesis at the $P \leq 0.05$ significance level ($F_{1,\infty} = 3.84$). This analysis yielded unrealistically low N_e estimates for evolution to proceed due to drift alone (Table 2). For instance, strictly neutral divergence of yolk-sac volume would have required an average N_e of only 0.12-1.54 individuals (mean across populations = 0.60) throughout the evolutionary time periods of the grayling populations (Table 2; **VI**).

We compared among-population differences based on additive genetic variances from the common-garden experiment (Q_{ST} ; Spitze 1993)

and analogous measures based on microsatellites from analysis of 17 loci (F_{ST} ; Wright 1951). This is a well-established method for inferring the relative importance of selection and drift (Merilä & Crnokrak 2001; MacKay & Latta 2002; Hendry 2002). Divergent selection is expected to result in $Q_{ST} > F_{ST}$, whereas strictly neutral divergence should lead to equal estimates of Q_{ST} and F_{ST} (Spitze 1993; Merilä & Crnokrak 2001; MacKay & Latta 2002). This expectation will hold regardless of the underlying model producing the population structure (Whitlock 1999).

Accordingly with the neutrality tests, the Q_{ST}/F_{ST} comparisons revealed that the quantitative genetic differentiation often exceeded the drift-based differentiation (Fig. 11; **VI**). Hence, evolution of many of the traits was consistently driven by divergent natural selection. For instance, the pairwise Q_{ST} estimates for growth rate (0.36-0.88) were 4-10 times higher than the corresponding F_{ST} estimates (Fig. 11; **VI**).

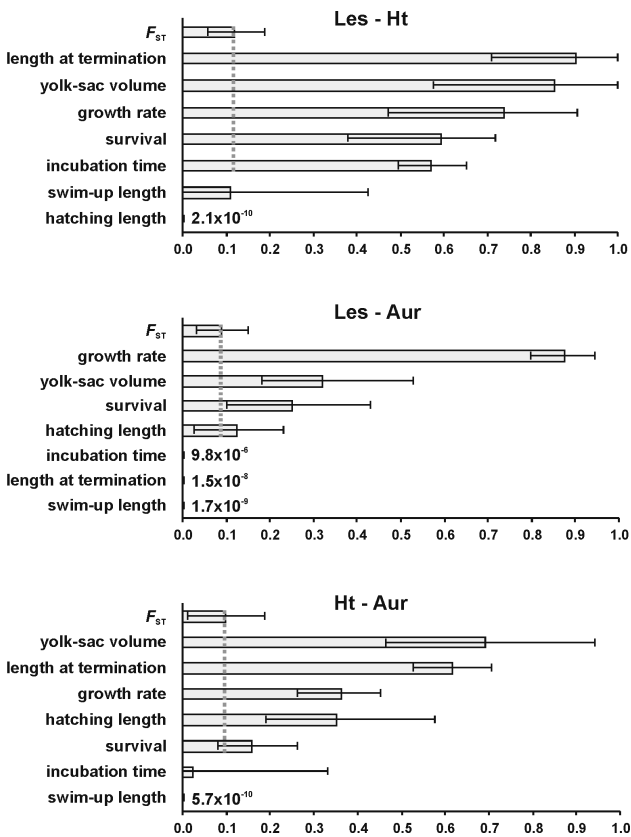


Figure 11. Among-population genetic differences based on additive genetic variances of seven life-history traits obtained from a common-garden experiment (Q_{ST}), and analogous measures based on 17 microsatellite DNA loci (F_{ST}). The horizontal bars indicate 95% confidence intervals for the estimates and the dashed vertical bars indicate the effect of random genetic drift (left-hand side) and natural selection (right-hand side) on quantitative traits that have been influenced by divergent selection.

We note that all of the traits evolving at a non-neutral rate (Table 2; Fig. 11; **VI**) are very important for fitness of *T. thymallus* (Haugen & Vøllestad 2000). When offspring of the populations were reared in three distinct July-August water temperatures naturally occurring within Lesjaskogsvatn, Hårtjønn and Aursjøen, they generally performed best in the experimental temperatures that mimicked the natural conditions (Haugen & Vøllestad 2000). For example, grayling originating from the naturally warm Lake Lesjaskogsvatn exhibited highest survival rate during 180 degree-days of exogenous feeding in warm experimental conditions, and lowest survival in cold conditions (Haugen & Vøllestad 2000). It therefore seems probable that selection acted in favor of local adaptation to the specific temperatures that the populations experience in nature.

An additional interesting aspect is that the pairwise Q_{ST} estimates between Ht-Les (mean across traits=0.53) and Ht-Aur (mean=0.39) were higher than among Les-Aur (mean=0.21; Fig. 11). This implies stronger selection among Ht-Les and Ht-Aur than between Les-Aur, which corresponds to the magnitude of differences in their ecological conditions: the spawning and nursery habitats of Les and Aur are cold and fast-running, while spawning and early-life of the Ht grayling takes place in relatively warm and still-water conditions.

In order to relate the neutrality test and Q_{ST}/F_{ST} results to the demography of the *T. thymallus* populations, the microsatellite data were used to estimate genetic diversity of the populations, effective population sizes and potential variation in the population sizes ('bottlenecks'; **VI**). We found that the populations exhibit low levels of microsatellite diversity, small N_e estimates, and show signatures of severe bottlenecks (Table 3). These observations were exemplified by the Hårtjønn grayling, which had an average of only

1.5 alleles per locus, a harmonic mean N_e of 24.2 individuals since the founding of the population and current population size of only 0.6% of its historical size (Table 3; **VI**). As a comparison, *T. thymallus* populations are generally much more diverse, exhibiting an average of 4.0 alleles per locus as revealed using the same set of 17 microsatellites (**III**). Such low levels of diversity, low effective sizes and severe bottlenecks appear striking for natural populations. However, they are very well in line with the known demographic history (small number of founders) of the populations. Furthermore, the N_e estimates seem realistic due to the small ecological niches characterizing the Norwegian mountain lakes investigated. For instance, the total area of Hårtjønn is only 0.35 km² and the lake has a single 10 m² *T. thymallus* spawning ground, thereby providing a suitable habitat for only a very small grayling population.

That evolution of *T. thymallus* was shown to be dominantly due to natural selection is very interesting in the light of the demographic history of the populations. This study provides an example that natural selection can be strong under circumstances commonly discussed to constrain selection and greatly enhance the influence of an alternative evolutionary mechanism: genetic drift (Wright 1931; Lande 1976; Kimura 1983; Carson & Templeton 1984; Lynch 1990; Kimura 1995).

Evolutionary models commonly predict notable divergence due to random drift in effectively small populations (Wright 1931; Lande 1976; Kimura 1983; Carson & Templeton 1984; Lynch 1990; Kimura 1995). Accordingly with this prediction, microsatellites revealed significant ($P < 0.001$) genic differentiation between all of the populations and relatively high pairwise F_{ST} estimates (0.05-0.21). Because microsatellite divergence was evident at multiple loci (**VI**), it is highly unlikely that the F_{ST} estimates were solely due to physical

Table 3. Genetic diversity indices, effective population sizes (N_e) and evidence for population bottlenecks.

	n	A (SD)*	H_o (SD)†	N_e (95% CI)	$N0 / N1$ (95% CI)§
Les	52	1.9 (±1.1)	0.17 (±0.23)	88.9‡	0.003 (0.0003-0.03)
Ht	48	1.5 (±0.9)	0.14 (±0.25)	24.2 (12.1-42.2)	0.006 (0.0003-0.05)
ØM	49	1.7 (±0.9)	0.17 (±0.24)	85.0 (36.0-170.5)	0.0006 (0.00008-0.01)
Aur	28	1.6 (±0.9)	0.14 (±0.20)	55.4 (24.8-110.0)	0.001 (0.0003-0.005)

*Mean number of microsatellite alleles within populations. †Mean observed heterozygosity within populations. ‡Indicates the lowest sampled point out of 100,000 Markov Chain Monte Carlo replicates. The maximum likelihood estimate of N_e could not be obtained most likely due near-zero level of drift in the Les population (M. A. Beaumont, pers. comm.). §Indicates the ratio of current population size ($N0$) and historical population size ($N1$; Beaumont 1999).

linkages to genes under selection, i.e. due to so-called 'genetic draft' processes (Gillespie 2000; Gillespie 2001). It can therefore be suggested that, while the populations evolved dominantly due to selection, also drift had an effect on the quantitative trait divergence (VI).

Because the F_{ST} estimates (up-to 0.21) seem reasonably high (given the short coalescence times), it is relevant to ask: are such high levels of drift theoretically feasible with a simple model of constant N_e , or must fluctuating population sizes or even selection via linkages to coding genes be invoked to explain the microsatellite divergence? Computer simulated genotypes were used to address this question (VI). When applying the known coalescence times and the microsatellite based N_e estimates, it was clear that the empirically obtained F_{ST} estimates were theoretically feasible (VI). While stable effective population sizes (along with many other assumptions of the simulations; VI) are unlikely to hold in nature, the simulations nevertheless served to demonstrate that the microsatellite based N_e estimates were sufficiently low to generate the observed levels of neutral divergence even without variation in the number of breeders (or selection) throughout the evolutionary histories of the grayling populations (VI).

To conclude, we have unambiguously shown that the early stages of grayling life-history evolution occurred dominantly due to natural selection. The power of selection is striking because the populations originate from a small number of founders, exhibit small effective sizes and show genetic imprints of bottlenecks, conditions expected to constrain selection and enhance the importance of drift. This study is among the first to evidence non-neutral evolution via formal neutrality tests, and to investigate the relative importance of selection and drift across a contemporary evolutionary time scale. The results reveal one of the clearest cases of fisherian evolution in small natural populations to date.

Conclusions

This thesis provided the first comprehensive biological evidence to evaluate how historical processes have influenced genetic variation within Lake Baikal and its surrounding great rivers (I). It was clear that ancient (pre-Pleistocene) grayling lineages currently inhabit Enisey, Lena and Amur. Lake Baikal, however, was likely re-colonized, or a grayling population size explosion occurred, only

110,000–450,000 ypb. Population genetic inferences supported an Enisey-Angara river route of re-colonization into Baikal, corresponding to the paleo-hydrological event that led to the formation of the Lake's only contemporary outlet, and a subsequent range expansion several thousand kilometers South into the Selenga River. The evolutionary history of Lake Baikal grayling was congruent with the controversial hypothesis of repeated glaciation. Alternatively, less profound but equally consequential environmental perturbations may have influenced the ability of *Thymallus* to survive in Baikal. Microsatellite data implied contemporary grayling population declines within Baikal, possibly reflecting the influence of an uncontrolled fishery on this treasured ecosystem (I).

This thesis reported the first European-wide phylogeographic data for a threatened and culturally important species *T. thymallus* (II, III). It was revealed that European grayling are divided into at least three diverged lineages, which currently inhabit northeastern, northwestern/central and southern Europe. The separation of these lineages most likely pre-dates the late Pleistocene period. It appeared likely that northern European grayling populations have originated from central and northeastern Pleistocene refugia and remained largely allopatric since colonizing their present habitats. Southern Europe was possibly the source for the colonization of the northwestern/central populations, but confident resolution of the inter-relationships of the three genetically distinct assemblages requires further study (II, III). A recent comprehensive mtDNA sequencing study of 316 grayling from 44 populations across western Europe has shed much new light on the glacial history of *T. thymallus* (Weiss et al. 2002). Several conclusions of this thesis (II, III) were supported by the new analyses, including ancient contribution of Danubian gene pools to the colonization of central and northern Europe via Elbe, and a long pre-Pleistocene history of grayling in central Europe. However, the results of Weiss et al. (2002) also demonstrate that central and southern Europe harbour much greater diversity of genetically distinct populations than revealed by the publications II and III, supporting the utmost importance of the southern and central populations in conservation of grayling diversity (Weiss et al. 2002).

The papers I-III clearly exemplified that microsatellite DNA markers can be very informative for resolving relatively old evolutionary relationships. This observation is in

stark contrast with generalizations that microsatellites are uninformative at inter-specific divergence scales (Paetkau et al. 1997). Future research should focus on resolving which factors of microsatellites or focal species influence the informativeness of these important genetic markers at deep divergence scales. In the case of *T. thymallus*, it appears that the larger than average number of loci utilized was an important factor (Koskinen et al. 2002; Landry et al. 2002).

It was revealed that European grayling populations can exhibit substantial levels of neutral genetic divergence across small geographical scales, and also within areas where populations can continuously exchange migrants (II-V). It is possible that *T. thymallus* is a species that exhibits extreme site fidelity even outside its spawning season, and the genetic differences across small geographical scales have originated due to post-glacial genetic drift.

The level of genetic diversity within *T. thymallus* populations was atypically low (II-V). This appeared not to be an attribute of the genetic markers utilized, but rather of the individual populations (III). It was concluded that the low levels of within-population diversity relate to the life-history characteristics of grayling, rather than reflecting historical population size fluctuations (III). It is possible that grayling display limited dispersal behaviour/capability from their natal sites, as was also suggested by the high levels of microgeographic among-population divergence (II-V). This could result in low long-term effective population sizes and, consequently, depauperate levels of genetic diversity.

In light of the emerging awareness that human-mediated mixing of genetically distinct populations may pose serious conservation problems, it has become important to monitor accurately whether non-native specimens have contributed to the gene pool of wild populations. Publication V presented perhaps the most rigorous quantification of stocking effects to date, through analysis of historical baseline samples in addition to contemporary genetic material. Significant decreases in genetic differentiation were detected between wild and introduced populations since commencement of the stocking activities. Accordingly, up-to 15% of the contemporary wild Lake Saimaa grayling were confidently identified to be of hatchery origin, and recent hybridization between the hatchery and indigenous individuals appeared likely. However, the contemporary populations exhibited evolutionary relationships congruent with the sampling locations, and up-to 73% of the individuals were identified to be of

pure indigenous origin. Thereby, despite the intensive introductions, the Lake Saimaa grayling populations have retained much of their natural genetic integrity. This observation indirectly suggested that the introduced specimens survived poorly in the wild. It would now seem interesting to further investigate whether fitness differences among wild, introduced and wild x introduced hybrid Lake Saimaa grayling indeed exist in nature. Because reliable genetic identification of the origins of individuals is now possible (V), Lake Saimaa grayling should provide an excellent model for undertaking such a challenging study.

Several findings of the papers II-V should prove useful for the development of appropriate conservation strategies for European grayling. First, it is noteworthy that substantial genetic differences exist among populations across short waterway distances (II-V). Hence, these studies imply that efficient conservation of the total genetic diversity in grayling may require utilization of distinct broodstocks created using locally caught founders even across relatively short geographical distances. In many cases, *T. thymallus* exemplifies a case with a clear need for multinational co-operation for managing and conserving biodiversity. For example, the Holmön grayling population inhabiting the northeastern coast of Sweden is very closely related to some Finnish *T. thymallus* populations, but highly diverged from populations occurring elsewhere in Sweden.

Second, a remarkably large proportion of the total grayling diversity resides between populations, the level of within-population variation being atypically low (II, III). This finding further reinforces the conclusion that mixing of populations can result in substantially decreasing overall *T. thymallus* diversity.

Third, papers IV and V revealed that reliable genetic identification of the origins of individual grayling is feasible even within continuous drainages, such as parts of Lake Saimaa. This possibility may offer the fisheries managers a convenient tool to separate representatives of introduced and indigenous specimens prior to founding new broodstocks. By enabling exclusion of non-native genomes from the created broodstocks, such a possibility may prove useful for conserving the natural genetic integrity of grayling populations in areas where human-mediated mixing of diverged stocks has already taken place.

The focal theme of this thesis was to investigate how historical and contemporary evolutionary forces have influenced the distribution of genetic

diversity in grayling. The enduring discussion on the relative importance of natural selection and random drift in phenotypic evolution is very relevant to this theme. Publication VI, conducted in a unique model system, provided an unequivocal evaluation of the determinants of phenotypic evolution in nature. It was shown that divergence of many *T. thymallus* life-history traits proceeded at a significantly greater rate than expected from pure drift processes. It was estimated that unrealistically small effective population sizes would have been necessary for evolution to proceed due to drift alone. In some cases, selection was estimated to be up-to an order of magnitude more important as a diversifying agent than drift. The power of natural selection was striking because the populations originated from a small number of founders, exhibited small effective sizes, and displayed genetic imprints of severe bottlenecks. All of these factors have been suggested to constrain selection and favor drift. Thus, while the generality of these results in other species and even in other grayling populations remains uncertain, this study nevertheless provided perhaps the clearest evidence to date that contemporary selection can greatly dominate over drift in nature, even under conditions where populations seem predisposed to the influence of drift.

Acknowledgements

I would like to express my warmest thanks and appreciation to Craig Primmer for his excellent guidance throughout my PhD studies. I am very grateful to Craig for his daily advice during the early stages of this thesis, for directing me to do more independent research at exactly the right moment, and for always being interested in my studies and offering help whenever it was needed.

I thank all of my co-authors for their contribution to this thesis. I am especially indebted to Steven Weiss for giving me the possibility to participate in his interesting studies on Asian grayling, and for commenting on the introductory section of this publication, and to Thronn Haugen for the excellent collaboration we have had over that past four years.

I am grateful to Esa Ranta and the entire Integrative Ecology Unit (IKP) for providing a very stimulating research environment, for always finding the time for helpful discussions, and for all of the memorable (?) times we have shared in the countless off-duty events.

Most of the studies included in this thesis may never have commenced without the generosity of numerous researchers providing grayling DNA samples. I particularly acknowledge Alexey Veselov, Sergei Titov, Alexandre Potutkin, Jorma Piironen and Pekka Sundell for willingly supplying hundreds of grayling samples, and also for letting me participate in exciting fishing expeditions throughout the Russian Karelia and Kola Peninsula. I also want to thank Jan Nilsson, Nina Peuhkuri, Timo Hartikainen, Teija Aho, Jaakko Lumme, Riho Gross, Eric Taylor, Mike Stamford, Michael Hansen, Thronn Haugen, Kjetil Hindar, Ross Gardiner, Simona Sušnik, Aleš Snoj, Bernhard Gum, Ralph Kühn and Pavel Brzuzan for providing samples, and Zöe Redenbach for the detailed information on PCR-RFLP patterns in *T. arcticus*.

I thank Michael Hansen and Eric Taylor for their pre-examination of this thesis, Louis Bernatchez for his constructive comments on articles I-III, Andrew Hendry for his thorough review of paper VI, and numerous anonymous referees who have much improved the articles included in this thesis.

Many people at the Division of Population Biology have been of much help during my PhD studies. I appreciate Sirkka-Liisa Nyeki for her assistance in finding all of the literature important for my research, and Hannu Pietiäinen and Ilkka Teräs for making sure that my studies remain in schedule. I also want to thank all of the people of the MES-lab for providing a very pleasant working environment. I am especially grateful to Leena Laaksonen and Anne Aronta for their excellent assistance in the lab.

I am grateful to my parents Jukka and Raija Koskinen for always supporting my decisions in life. Finally, I warmly thank my wife Laura and my son Santeri, the two most special people in my life. Laura has always been interested in my research and has enthusiastically listened to the good and bad turns of my work. Without her love and support this thesis may never have been completed. Santeri has been very successful in keeping my thoughts away from work in the evenings and has enormously influenced my perspective on life in many positive ways.

This thesis was funded by the Biological Interactions Graduate School, the Finnish Academy, the Finnish Game and Fisheries Research Institute, the Ministry of Agriculture and Forestry, the TE-centres of Kaakkois-Suomi and Pohjois-Karjala and the University of Helsinki.

References

- Allendorf FW & Waples RS (1996) Conservation and genetics of salmonid fishes. In: *Conservation Genetics: Case Histories From Nature*. Chapman & Hall, New York.
- Angers B & Bernatchez L (1998) Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. *Molecular Biology and Evolution* **15**, 143-159.
- Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge.
- Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489-522.
- Avise JC, Walker D & Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London B* **265**, 1707-1712.
- Baker AJ (1992) Genetic and morphometric divergence in ancestral European and descendent New Zealand populations of chaffinches (*Fringilla coelebs*). *Evolution* **46**, 1784-1800.
- Bănărescu P (1991) *Zoogeography of Fresh Waters Vol 2: Distribution and Dispersal of Freshwater Animals in North America and Eurasia*. AULA-Verlag Wiesbaden.
- Beaumont M (1999) Detecting population expansion and decline using microsatellites. *Genetics* **153**, 2013-2029.
- Beaumont M, Barrat EM, Gottelli D, Kitchener AC, Daniels MJ, Pritchard JK & Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology* **10**, 319-336.
- Beaumont M & Bruford MW (1999) Microsatellites in conservation genetics. In: *Microsatellites. Evolution and Applications*. Oxford University Press, New York.
- Bermingham E & Lessios HA (1993) Rate variation of protein and mitochondrial DNA evolution as revealed by sea urchins separated by the isthmus of Panama. *Proceedings of the National Academy of Sciences USA* **90**, 2734-2738.
- 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**, 351-79.
- Bernatchez L & Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* **7**, 431-452.
- Björklund M (1991) Patterns of morphological variation among cardueline finches (Fringillidae: Carduelinae). *Biological Journal of the Linnean Society* **43**, 239-248.
- Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR & Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455-457.
- Broughton R (2000) *The Complete Book of the Grayling*. Robert Hale, London.
- Brunner PC, Douglas MR & Bernatchez L (1998) Microsatellite and mitochondrial DNA assessment of population structure and stocking effects in Arctic charr *Salvelinus alpinus* (Teleostei: Salmonidae) from central alpine lakes. *Molecular Ecology* **7**, 209-223.
- Carson HL & Templeton AR (1984) Genetic revolutions in relation to speciation phenomena and the founding of new populations. *Annual Review of Ecology and Systematics* **15**, 97-131.
- Cornuet J-M, Piry S, Luikart G, Estoup A & Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.
- Darwin CR (1859) *On the Origin of Species by Means of Natural Selection*. John Murray, London.
- DeWoody JA & Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* **56**, 461-473.
- Dobzhansky T (1951) *Genetics and the Origin of Species, 3rd ed.* Columbia University Press, New York.
- Edwards AWF (1994) The fundamental theorem of natural selection. *Biological Reviews* **69**, 443-474.
- Einum S & Fleming IA (1997) Genetic divergence and interactions in the wild among native, farmed and hybrid Atlantic salmon. *Journal of Fish Biology* **50**, 634-651.
- Ellstrand NC (1992) Gene flow by pollen: Implications for plant conservation genetics. *OIKOS* **63**, 77-86.
- Englbrecht CC, Freyhof J, Nolte A, 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.

- Estoup A & Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: *Advances in Molecular Ecology*. IOS Press, Amsterdam.
- Fabricius E & Gustafson KJ (1955) Observations on the spawning behaviour of the grayling *Thymallus thymallus*. *Report of the Institute of Freshwater Research, Drottningholm* **36**, 75-103.
- Falconer DS & Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th ed. Longman, New York.
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
- FitzSimmons NN, Moritz C, Limpus CJ, Pope L & Prince R (1997) Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**, 1843-1854.
- Florensov NA (1978) *Istoriya ozera/Problemi Baikala*. Nauka Press, Novosibirsk.
- Frank SA & Slatkin M (1992) Fisher's fundamental theorem of natural selection. *Trends in Ecology and Evolution* **7**, 92-95.
- Gibbs HL & Grant PR (1987) Oscillating selection on Darwin's finches. *Nature* **327**, 511-513.
- Gillespie JH (1991) *The Causes of Molecular Evolution*. Oxford University Press, Oxford.
- Gillespie JH (2000) The neutral theory in an infinite population. *Gene* **261**, 11-18.
- Gillespie JH (2001) Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* **155**, 909-919.
- Goldstein DB, Roemer GW, Smith DA *et al.* (1999) The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics* **151**, 797-801.
- Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL & Feldman MW (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences USA* **92**, 6723-6727.
- Goldstein DB & Schlötterer C (1999) *Microsatellites: Evolution and Applications*. Oxford University Press, Oxford.
- Goldstein PZ, DeSalle R, Amato G & Vogler AP (2000) Conservation genetics at the species boundary. *Conservation Biology* **14**, 120-131.
- Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA & Wayne RK (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf, *Canis simensis*. *Molecular Ecology* **3**, 301-312.
- Grant PR & Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707-711.
- Grosswald MG (1998) New Approach to the Ice Age Paleohydrology of Northern Eurasia. In: *Paleohydrology and Environmental Change*. Wiley and Sons, Chichester.
- Hansen MM (2002) Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. *Molecular Ecology* **11**, 1003-1017.
- Hansen MM, Kenchington E & Nielsen EE (2001) Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries* **2**, 93-112.
- 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.
- Hansen MM, Ruzzante DE, Nielsen EE & Mensberg K-LD (2000) Microsatellite and mitochondrial DNA polymorphism reveals life-history dependent interbreeding between hatchery and wild brown trout (*Salmo trutta* L.). *Molecular Ecology* **9**, 583-594.
- Hantke R (1993) *Flußgeschichte Mitteleuropas*. Ferdinand Enke-Verlag, Stuttgart.
- Harr B, Weiss S, David JR *et al.* (1998) A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Current Biology* **8**, 1183-1186.
- Haugen TO & Vøllestad LA (2000) Population differences in early life-history traits in grayling. *Journal of Evolutionary Biology* **13**, 897-905.
- Haugen TO & Vøllestad LA (2001) A century of life-history evolution in grayling. *Genetica* **112**, 475-491.
- Hendry AP (2001) Adaptive divergence and the evolution of reproductive isolation in the wild: an empirical demonstration using introduced sockeye salmon. *Genetica* **112**, 515-534.
- Hendry AP (2002) $Q_{ST} > = \neq < F_{ST}$? *Trends in Ecology and Evolution* **17**, 502.
- 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) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**, 87-112.

- Hewitt GM (2000) The genetic legacy of quaternary ice ages. *Nature* **405**, 907-913.
- Hey J (1997) Mitochondrial and nuclear genes present conflicting portraits of human origins. *Molecular Biology and Evolution* **14**, 166-172.
- Hoelzel AR, Halley J, O'Brien SJ *et al.* (1993) Elephant seal genetic variation and the use of simulation models to investigate historical population bottlenecks. *Journal of Heredity* **84**, 443-449.
- Hutchison DW & Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**, 1898-1914.
- Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- Kimura M (1995) Limitations of Darwinian selection in a finite population. *Proceedings of the National Academy of Sciences USA* **92**, 2343-2344.
- Klicka J & Zink RM (1997) The importance of recent ice ages in speciation: a failed paradigm. *Science* **277**, 1666-1669.
- Koskinen MT, Hirvonen H, Landry P-A & Primmer CR (2002) The benefits of increasing the number of microsatellites utilized in genetic population studies: an empirical perspective. *Heredity*. Submitted.
- Koskiniemi J & Kilpinen K (1987) Harjuskantojen perinnöllisten erojen selvitys. *Suomen Kalastuslehti* **94**, 424-427.
- Kozhov MM (1972) *Ocherki po Baikalovedeniyu*. East-Siberian Press, Irkutsk.
- Kozhova OM & Izmet'eva LR (1998) *Lake Baikal. Evolution and Biodiversity*. Backhyus Publishers, Leiden.
- Kunkel TA (1993) Slippery DNA and diseases. *Nature* **365**, 207-208.
- Kuz'min MI (2001) *Vo l'dakh Baikala*. SB RAS Publishers, Novosibirsk.
- Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314-334.
- Lande R (1977) Statistical tests for natural selection on quantitative characters. *Evolution* **31**, 442-444.
- Landry P-A, Koskinen MT & Primmer CR (2002) Deriving evolutionary relationships with microsatellites and $(\delta\mu)^2$: all loci are equal, but some are more equal than others... *Genetics* **161**, 1339-1347.
- Lelek A (1984) Threatened fishes of Europe. In: *The Freshwater Fishes of Europe* **9**. AULA-Verlag, Wiesbaden.
- Linnaeus C (1758) *Systema Naturae*. Laurentius Galvius, Stockholm.
- Litt M & Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**, 397-401.
- Losos JB, Warheit KI & Schoener TW (1997) Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature* **387**, 70-73.
- Lu G, Basley DJ & 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.
- Lynch M (1990) The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *American Naturalist* **136**, 727-741.
- Lynch M & Hill WG (1986) Phenotypic evolution by neutral mutation. *Evolution* **40**, 915-935.
- Lynch M & Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland.
- Majerus MEN (1998) *Melanism: Evolution in Action*. Oxford University Press, Oxford.
- Maitland PS & Campbell RN (1992) *Freshwater Fishes of the British Isles*. Butler & Tanner, Somerset.
- McKay JK & Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution* **17**, 285-291.
- Merilä J (1997) Quantitative trait and allozyme divergence in the Greenfinch (*Carduelis chloris*, Aves: Fringillidae). *Biological Journal of the Linnean Society* **61**, 243-266.
- Merilä J & Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* **14**, 892-903.
- Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution*, **9**, 373-375.
- Moritz C, Dowling TE & Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics* **18**, 269-292.
- Nesbø CL, Fossheim T, Vøllestad 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.
- Nielsen EE, Hansen MM & Bach LA (2001) Looking for a needle in a haystack: Discovery of indigenous Atlantic salmon (*Salmo salar* L.) in stocked populations. *Conservation Genetics* **2**, 219-232.
- Nielsen EE, Hansen MM & Loeschcke V (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology* **6**, 487-492.
- Northcote TG (1995) Comparative biology and management of Arctic and European grayling (Salmonidae, *Thymallus*). *Reviews in Fish Biology and Fisheries* **5**, 141-194.
- Nykänen M & Huusko A (1999) Harjuksen elinympäristövaatimukset virtavesissä. *Riista- ja Kalatalouden Tutkimuslaitos, Kalatutkimuksia*, **156**. Helsinki.
- O'Ryan C, Harley EH, Bruford MW *et al.* (1998) Microsatellite analysis of genetic diversity in fragmented South African buffalo populations. *Animal Conservation* **1**, 85-94.
- Paetkau D, Waits LP, Clarkson PL *et al.* (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147**, 1943-1957.
- Pamilo P & Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution* **5**, 568-583.
- Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE & Blot M (1999) Genomic evolution during a 10,000-generation experiment with bacteria. *Proceedings of the National Academy of Sciences USA* **96**, 3807-3812.
- Poteaux C, Bonhomme F & Berrebi P (1998) Differences between nuclear and mitochondrial introgressions of brown trout populations from a restocked main river and its unstocked tributary. *Biological Journal of the Linnean Society* **63**, 379-392.
- Price GR (1972) Fisher's 'fundamental theorem' made clear. *Annals of Human Genetics* **36**, 129-140.
- Primmer CR, Aho T, Piironen J, Estoup A, Cornuet J-M & Ranta E (1999) Microsatellite analysis of hatchery stocks and natural populations of Arctic charr, *Salvelinus fontinalis*, from the Nordic region: implications for conservation. *Hereditas* **130**, 277-289.
- Primmer CR & Ellegren H (1998) Patterns of molecular evolution in avian microsatellites. *Molecular Biology and Evolution* **15**, 997-1008.
- Pritchard JK, Stephens M & Donnelly PJ (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Rannala B & Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences USA* **98**, 9197-201.
- Remold SK & Lenski RE (2001) Contribution of individual random mutations to genotype-by-environment interactions in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **98**, 11388-93.
- Reznick DN, Shaw FH, Rodd FH & Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275**, 1934-1936.
- Richard M & Thorpe RS (2001) Can microsatellites be used to infer phylogenies? Evidence from population affinities of the Western Canary Island lizard (*Gallotia galloti*). *Molecular Phylogenetics and Evolution* **20**, 351-360.
- Rogers AR (1996) Population structure and modern human origins. In: *Progress in Population Genetics and Human Evolution*. Springer Verlag, New York.
- Rogers AR & Harpending HC (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552-569.
- Rundle HD, Nagel L, Boughman JW & Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306-308.
- Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- Simpson GG (1944) *Tempo and Mode in Evolution*. Columbia University Press, New York.
- Simpson GG (1953) *The Major Features of Evolution*. Columbia University Press, New York.
- Skaala Ø, Jørstad KE & Borgstrøm R (1996). Genetic impact on two brown trout (*Salmo trutta*) populations after release of non-indigenous hatchery spawners. *Canadian Journal of Fisheries and Aquatic Sciences* **53**, 2027-2035.
- Slatkin M (1993) Isolation by distance equilibrium and non-equilibrium populations. *Evolution* **47**, 264-279.

- Spitze K (1993) Population structure in *Daphnia obtusa*: Quantitative genetic and allozyme variation. *Genetics* **135**, 367-374.
- Sundell P (2000) *Kalastuskuntien Käsitykset Harjuksesta ja Harjuskannan Tilasta Pielisellä*. Jyväskylän yliopiston raportti **18**, Jyväskylä.
- Svendsen JJ, Astakhov VI, Bolshiyarov DYU *et al.* (1999) Maximum extent of the Eurasian ice sheets in the Barents and Kara Sea region during the Weichselian. *Boreas* **28**, 234-242.
- Taberlet P, Fumagalli L, Wust-Saucy A-G & Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.
- Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **12**, 4127-4138.
- Templeton AR, Crandall KA & Sing CR (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619-623.
- Templeton AR, Routman E & Phillips CA (1995) Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140**, 767-782.
- Thulin C-G, Jaarola M & Tegelström H (1997) The occurrence of mountain hare mitochondrial DNA in wild brown hares. *Molecular Ecology* **6**, 463-467.
- Turgeon J & Bernatchez L (2001) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedii* (Teleostei: Coregoninae). *Evolution* **55**, 2274-2286.
- Waples RS (1991) Pacific salmon, *Oncorhynchus* spp. and the definition of a 'species' under the Endangered Species Act. *Marine Fisheries Review* **53**, 11-22.
- Wayne RK (1996) Conservation genetics in the canidae. In: *Conservation Genetics: Case Histories From Nature*. Chapman & Hall, New York.
- Weber JL & May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388-396.
- Weiss S, Persat H, Eppe R, Schlotterer C & Uiblein F. Complex patterns of colonization and refugia revealed for European grayling *Thymallus thymallus*, based on complete sequencing of the mtDNA control region. *Molecular Ecology* **11**, 1393-1407.
- Whitlock MC (1999) Neutral additive genetic variance in a metapopulation. *Genetical Research* **74**, 215-221.
- Williams DF, Peck J, Karabanov EB *et al.* (1997) Lake Baikal record of continental climate response to orbital insolation during the past 5 million years. *Science* **278**, 1114-1117.
- Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.