Genetic population structure of *Gyrodactylus thymalli* (Monogenea) in a large Norwegian river system

RUBEN ALEXANDER PETTERSEN[1]*, TOR ATLE MO[2], HAAKON HANSEN[2] *and* LEIF ASBJØRN VØLLESTAD[1]

[1] *Department of Biosciences, Center for Ecological and Evolutionary Synthesis, University of Oslo, P. O. Box 1066 Blindern, 0316 Oslo, Norway*
[2] *Norwegian Veterinary Institute, P.O. Box 750 Sentrum, 0106 Oslo, Norway*

(*Received 15 April 2015; revised 17 September 2015; accepted 21 September 2015; first published online 14 October 2015*)

SUMMARY

The extent of geographic genetic variation is the result of several processes such as mutation, gene flow, selection and drift. Processes that structure the populations of parasite species are often directly linked to the processes that influence the host. Here, we investigate the genetic population structure of the ectoparasite *Gyrodactylus thymalli* Žitňan, 1960 (Monogenea) collected from grayling (*Thymallus thymallus* L.) throughout the river Glomma, the largest watercourse in Norway. Parts of the mitochondrial dehydrogenase subunit 5 (NADH 5) and cytochrome oxidase I (COI) genes from 309 *G. thymalli* were analysed to study the genetic variation and investigated the geographical distribution of parasite haplotypes. Three main clusters of haplotypes dominated the three distinct geographic parts of the river system; one cluster dominated in the western main stem of the river, one in the eastern and one in the lower part. There was a positive correlation between pairwise genetic distance and hydrographic distance. The results indicate restricted gene flow between sub-populations of *G. thymalli*, most likely due to barriers that limit upstream migration of infected grayling. More than 80% of the populations had private haplotypes, also indicating long-time isolation of sub-populations. According to a molecular clock calibration, much of the haplotype diversity of *G. thymalli* in the river Glomma has developed after the last glaciation.

Key words: Population structure, gene flow, Salmonid fish, ectoparasite, networks, isolation by distance.

INTRODUCTION

The study of genetic diversity and gene flow can provide fundamental insights into the demographic and evolutionary history of populations (Wright, 1943; Slatkin, 1987). In order for a population to adapt to environmental change, genetic variation is required (Slatkin, 1987). Gene flow due to individuals dispersing among sub-populations may increase genetic variation. However, the opportunity for individuals to disperse depends on the geographical topography and presence of potential barriers to movement. A lack of gene flow may lead to loss of genetic diversity and increased genetic drift (Grenfell *et al*. 2004). In general, there are a large number of studies on genetic diversity in a wide range of study systems. However, studies on the genetic variation and geographical distribution of parasites on larger scales are still rare. Host specific monoxenous fish parasites are interesting for studies of genetic diversity at local and regional scales as their occurrence and genetic diversity depend on the distribution and migration of their hosts as well as transmission probabilities. In rivers and streams, these parasites can disperse downstream by passive drift of detached parasites or fish host migration, whereas upstream dispersal depends on host migration (Criscione and Blouin, 2004). Such upstream migration may be restricted by barriers (waterfalls, dams) and may over time lead to a reduced genetic variation upstream due to genetic drift and bottleneck events in small populations. Further, the dominating downstream gene flow can over time result in increased genetic diversity in downstream areas.

Fish ectoparasites of the genus *Gyrodactylus* have a short generation time, give birth and have no specialized transmission stages. In the recent years, several molecular studies has been done on the taxonomy, systematics, phylogeography and genetic variation of species of *Gyrodactylus* based on mitochondrial DNA sequences (see e.g. Hansen *et al*. 2003, 2007*a*, *b*), but the genetic varation within river systems has not been studied in detail. One of the species that has been studied in some detail is *Gyrodactylus thymalli* Žitňan, 1960 (Monogenea) and several phylogenetic lineages and haplotypes of this species have been found on its main host grayling *Thymallus thymallus* (L.) in European rivers (Hansen *et al*. 2007*a*, *b*; Lindqvist *et al*. 2007; Anttila *et al*. 2008; Kuusela *et al*. 2009). *Gyrodactylus thymalli* occurrs frequently on grayling in the Glomma river system, the largest watercourse in Norway (Mo *et al*. 1998; Hansen *et al*. 2003, 2006, 2007*a*, *b*). Hansen *et al*. (2003, 2006, 2007*a*, *b*) found

* Corresponding author: R. A. Pettersen, Department of Biosciences, Center for Ecological and Evolutionary Synthesis, University of Oslo, P. O. Box 1066 Blindern, 0316 Oslo, Norway. E-mail: rubenap@ibv.uio.no

Parasitology (2015), **142**, 1693–1702.
doi:10.1017/S003118201500133X

8 haplotypes on a small number of examined grayling from a restricted part of the river.

The topography of the river Glomma is strongly influenced by the land rise after the last glacial period in Northern Europe that ended about 10 000 years ago. Following de-glaciation, the landscape has been influenced by many geological processes such as land rise, river capture and sea-level rise. As a result of these post-glacial geological processes several waterfalls act as barriers to fish upstream migration. This has influenced the immigration history of the various fish species found in the river Glomma (Huitfeldt-Kaas, 1918; Andersen and Borns, 1994; Nesbø *et al*. 1999; Koskinen *et al*. 2000; Bernatchez, 2001; Østbye *et al*. 2005). Koskinen *et al*. (2000) suggested that two immigration routes at different times during the post-glacial process have resulted in two different evolutionary lineages of grayling in the Glomma river system. This could also have affected the genetic variation and genetic population structure of *G. thymalli* in the river system.

In a recent paper Fromm *et al*. (2014) suggested that *G. thymalli* is a junior synonym to *Gyrodactylus salaris* Malmberg, 1957 described from Atlantic salmon *Salmo salar* L. However, Fromm *et al*. (2014) did not include specimens from the type localities for the two species which should be used in species synonymiation (zoological nomenclature). Furthermore, *G. thymalli* is not able to survive on Atlantic salmon while *G. salaris* can survive on grayling (Sterud *et al*. 2002). Thus, we have chosen to use the name *G. thymalli* for the parasites in this study.

Here, we present the first in-depth analysis of the genetic variation of a naturally occurring *Gyrodactylus* species, *G. thymalli*, from the largest river system in Norway. Given the large geographic scale (see below) and complex immigration history of the host species, we here investigate the level of genetic diversity and how it is distributed across this large geographic scale.

MATERIALS AND METHODS

The study system

The river Glomma covers has a large catchment area in Norway (Fig. 1), draining 13% (41 971 km^2) of the country. The river system has two major branches – the east branch (upper Glomma) and the west branch (the river Gudbrandsdalslågen). The two branches merge and form the lower main stem of the river (lower Glomma) about 136 km before draining into the sea. The upper Gudbrandsdalslågen branch includes Mjøsa (MJO), the largest lake in Norway (L'Abée-Lund *et al*. 2009).

Grayling is found throughout the Glomma river system (Hesthagen and Sandlund, 2004). Individual migrations over distances up to 154 km has been observed for grayling in this river (Heggenes *et al*. 2006), but upstream migrations are often prevented by natural waterfalls and dams (Østdahl *et al*. 2002; Heggenes *et al*. 2006). These barriers have led to population differentiation (Heggenes *et al*. 2006; Junge *et al*. 2014). Grayling has a life history with ontogenetic habitat shifts, with spawning and early juvenile life occurring in rivers and streams during spring and summer. The juveniles are territorial, but larger and older individuals may aggregate in schools (Northcote, 1995).

Sampling

Grayling were sampled at 20 localities using various fishing methods during 2007–2011 (Table 1, Fig. 1). The sampling sites were selected to cover the distribution of the grayling in the Glomma river system. The average length between collection sites is 42 km, and the difference between the northern population and the southern populations of the watercourse is 506 km.

After capture, fish were rapidly euthanized by a blow to the head. Individual fish <15 cm were preserved in 96% ethanol and examined whole. From larger fish, only the fins were excised, preserved and examined. In the laboratory, whole fish or fins were examined for the presence of *G. thymalli* under a stereomicroscope. Although many fish carried several *Gyrodactylus* specimens, only one specimen was sampled from each fish. The rationale behind this was that species of *Gyrodactylus* are viviparous (Malmberg, 1993) and most or all individuals on a fish are likely the offspring of one mother. Thus, sampling one parasite from each fish, rather than several parasites from the same fish, was an attempt to avoid pseudoreplication. In this study, all *G. thymalli* specimens from several fishes sampled at one locality will be referred to as one population.

Molecular methods

DNA was extracted from *Gyrodactylus* specimens using the Gene Mole® Robot with MoleStrips™ DNA Tissue kit, following the manufacturer's instructions. The mitochondrial dehydrogenase subunit 5 (NADH5) gene and cytochrome oxidase I (COI), the bar-coding gene (Hebert *et al*. 2003), were amplified. The protocol and primer pair described by Huyse *et al*. (2008) was used to amplify an 894-bp fragment of the NADH5 gene. For the COI gene the protocol and primer pair described by Meinilä *et al*. (2002) was used to amplify 820-bp. Polymerase chain reaction (PCR) was performed using PuReTaq ready-to-go PCR beads (GE Healthcare), 1 μmol L^{-1} of each primer, and 5 μL of the extracted DNA in a 25 μL reaction volume. PCR products were purified using 10× diluted exoSAP-IT (USB). Cycle sequencing, using the same primers as in the PCR reaction, was

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

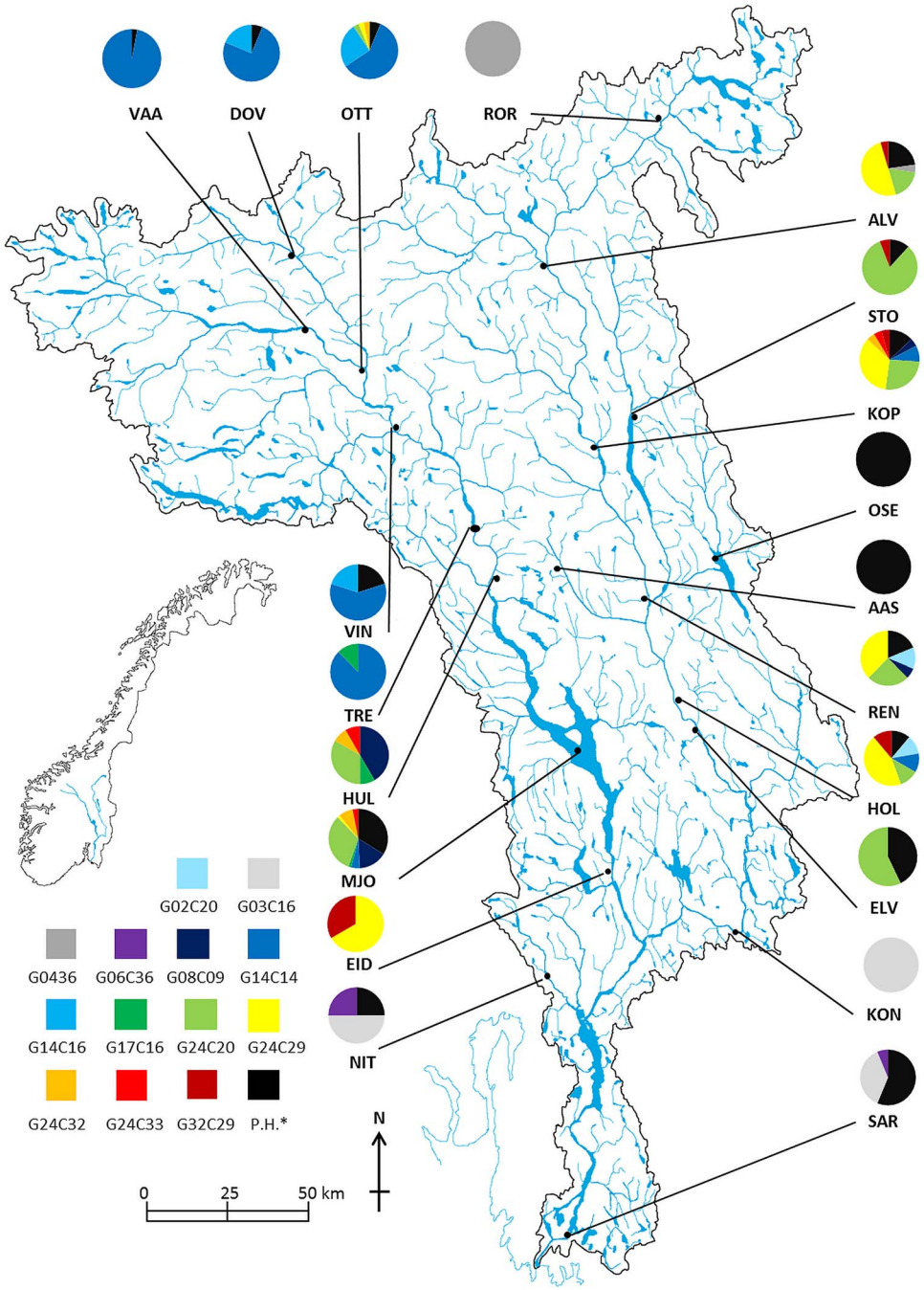

Fig. 1. The Glomma river system, with the locations where *Gyrodactylus thymalli* was sampled (the population codes refer to the names in Table 1). The colour codes refer to the concatenated mtDNA genes NADH5(G) and COI(C) haplotypes (see online Supplementary Table 2 and 3). P.H* indicates private haplotypes.

performed in 10 μL reactions using 2 μL BigDye terminator cycle sequencing ready reaction kit (Applied Biosystems), 2 μL 5× sequencing buffer, 10 pmol primer, and 3 μL cleaned PCR product. The sequencing was done on an ABI 3730 high-throughput capillary electrophoresis machine.

Data analysis

All sequences were proof-read and edited using Sequencher 5.0 and aligned using the Crustal W algorithm in MEGA 5.0 (Tamura *et al*. 2007), with default parameter settings. Genetic distances were calculated separately for COI and NADH5 according to the Kimura three-parameter estimates plus Gamma (T92 + G). This model gave the best fit in a model selection test of 24 different nucleotide substitution models, selected based on the corrected Akaike information criterion (Hurvich and Tsai, 1989) in MEGA 5.0 (Tamura *et al*. 2004). To test for positive selection, we estimated the numbers of synonymous (s) and nonsynonymous (n) substitutions (dN/ds) using maximum likelihood (ML) analysis of natural selection codon-by-codon by HyPhy

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

in MEGA 5.0 (Felsenstein, 1981; Muse and Gaut, 1994; Suzuki and Gojobori, 1999; Pond and Frost, 2005; Pond *et al*. 2005). A neutrality test (Tajima`s D) was used on each population (Tajima, 1989) to test for sequence deviation from equilibrium between mutation and genetic drift. To test for hierarchical population structure, as well as for a geographical pattern of population subdivision, analysis of molecular variance (MANOVA) in the program GenAlex 6.2 was used (Excoffier *et al*. 1992; Peakall and Smouse, 2012). Genetic variance (based on φ_{ST}) of *G. thymalli* was partitioned among individuals within populations and among populations (φ_{SC}) and between regions (φ_{CT}, region 1: Elverum (ELV), Holset (HOL), Rena (REN), Åsta (AAS), Osensjøen (OSE), Koppang (KOP), Storsjøen (STO), Alvdal (ALV), Røros (ROR), region 2: Eidsvoll (EID), Mjøsa (MJO), Hunderfossen (HUN), Trettenstrykene (TRE), Vinstra (VIN), Otta (OTT), Vågå (VAA), Dovre (DOV), region 3: Sarpsborg (SAR), Nittedalselva (NIT), Kongsvinger (KON)) using 9999 permutations. Further, the same program was used to calculate pairwise genetic distances (based on φ_{SC}) among populations and further used in a principal coordinate analysis (PCoA). The two axes explaining most of the variation in genetic distance were then extracted. PCoA plots give an opportunity to visualize potential population grouping in relation to geography (Peakall and Smouse, 2012).

Under migration-drift equilibrium, pairs of populations are expected to exhibit a significant correlation between their genetic and geographic distances, termed 'isolation by distance' (IBD, Wright, 1943). This means that populations in close proximity to each other should be genetically less differentiated, due to ongoing gene flow between them, compared with populations that are separated geographically. The occurrence of IBD was tested by correlating pairwise genetic distance with hydrographic distance (km) using a Mantel test implemented in GenAlEx 6.41 (Peakall and Smouse, 2012). The level of significance was estimated by performing 9999 permutations. To calculate the hydrographic distance between the sampling locations, the shortest distance over water was used (see online Supplementary Table S1).

A median-joining network analysis (Bandelt *et al*. 1999) with maximum parsimony calculation (Polzin and Daneshmand, 2003) of mtDNA haplotypes was performed using NETWORK 4.5.1.6 (http:www.fluxus-engineerin.com) Flexus Technology software. The phylogenetic relationship between haplotypes based on the mtDNA sequences was inferred using the Neighbour-Joining method (Saitou and Nei, 1987), ML (Felsenstein, 1981) and Maximum Parsimony (Felsenstein, 1978) analyses implemented in MEGA 5.0 with bootstrap estimates inferred from 1000 replicates (Felsenstein, 1985). Two relevant sequences extracted from Gen-Bank were used to compare with our *G. thymalli* data (*G. thymalli* (Gen-Bank acc.no EF 527269), *G. salaris* (DQ988931)), and a sequence from *G. derjavinoides* was used as an outgroup (EU293891).

To estimate the time of colonization of *G. thymalli* in the Glomma river system we used data from Hansen *et al*. (2003, 2007*a*, *b*) indicating that *G. thymalli* from the river Glomma are highly divergent from *G. thymalli* in a neighbouring river system (Trysilelva). Hansen *et al*. (2003, 2007*a*, *b*) calculated this divergence to be 2·32% based on the Kimura 2-parameter model (Kimura, 1980). These two lines have most likely not had contact after the post-glaciation immigration of grayling. The time of separation between these two lineages was estimated using the function Inferring Ancestral Sequences ML in MEGA 5.0. NADH5-sequences were not available from the river Trysilelva, therefore only COI sequences were used in this analysis. To be able to compare with previous result (Meinilä *et al*. 2004; Hansen *et al*. 2007*a*), the Kimura 2-parameter model was applied in this analysis and the following sequences from GenBank were used: *G. thymalli* (AY146612, AY146613 and AY486544) from Trysilelva, and *Gyrodactylus lavareti* (AY225306) as outgroup. This analysis is based on precise estimates of a molecular clock (i.e. mutation rates) (Kumar, 2005). We used three different mutation rates, one conservative estimate of 2·0% divergence per million years generally suggested for mtDNA (Irwin *et al*. 1991; Bermingham *et al*. 1997; Bernatchez, 2001), and one intermediate mutation rate of 13·7% and one high mutation rate of 20·3% divergence per million years suggested by Meinilä *et al*. (2004) based on studies of *G. salaris* and *G. thymalli*.

RESULTS

In total 687 grayling were examined for the presence of *G. thymalli*; 309 *G. thymalli* individuals were collected and analysed (Table 1). *Gyrodactylus thymalli* was found throughout the Glomma river system, with prevalence varying from 16 to 100% (Table 1).

The molecular analysis was based on a final alignment of 777 and 770 bp for the COI and NADH5 genes, respectively. The NADH5-sequences exhibited in total 35 distinct haplotypes (online Supplementary Table S2), defined by 38 polymorphic sites. These mutations resulted in 15 amino acid substitutions (dN/dS = 0·39). The COI-sequences exhibited 37 distinct haplotypes, defined by 28 polymorphic sites, leading to 3 amino acid substitutions (dN/dS = 0·07, online Supplementary Table S3). The nucleotide frequencies in NADH5 were 30·5% (A), 31·3% (T/U), 22·4% (C), and 15·9% (G) and on COI are 28·4% (A), 28·4% (T/U), 21·6% (C) and 21·6% (G). The nucleotide

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

Table 1. The number of sampled grayling (N_f) and number of infected fish with *Gyrodactylus thymalli* ($N_{i+G.t.}$). Locality name and code for each population, its location within the river system, and distance from the outlet to the sea (km) is given (see Fig. 1)

Locality	Code	River section	N_f	$N_{i+G.t.}$	km
Sarpsborg	SAR	Lower Glomma	15	15	19
Rotnes	NIT	Lower Glomma	25	4	126
Kongsvinger	KON	Upper Glomma	21	11	180
Elverum	ELV	Upper Glomma	22	7	285
Holset	HOL	Upper Glomma	30	18	289
Rena	REN	Upper Glomma	26	16	312
Åsta	AAS	Upper Glomma	24	14	347
Lake Osensjøen	OSE	Upper Glomma	15	3	352
Koppang	KOP	Upper Glomma	63	23	370
Lake Storsjøen	STO	Upper Glomma	31	17	383
Alvdal	ALV	Upper Glomma	45	17	450
Røros	ROR	Upper Glomma	7	6	525
Eidsvoll	EID	Gubrandsdalslågen	4	3	161
Lake Mjøsa	MJO	Gubrandsdalslågen	110	55	198
Hunderfossen	HUN	Gubrandsdalslågen	26	12	278
Trettenstrykene	TRE	Gubrandsdalslågen	17	8	300
Vinstra	VIN	Gubrandsdalslågen	6	6	345
Otta	OTT	Gubrandsdalslågen	66	32	378
Lake Vågå	VAA	Gubrandsdalslågen	41	27	411
Dovre	DOV	Gubrandsdalslågen	30	16	416

diversity (π) for the two genes ranged between 0·001 and 0·004 in the populations (online Supplementary material; Table S2 for NADH 5 and Table 3S for COI). The ML analysis of natural selection codon-by-codon was not significant for any of the polymorphic sites. Further, there was no divergence from neutrality tested by Tajima's D. For the concatenated NADH5 and COI sequences a total of 58 haplotypes were found, and 80% of the populations had private alleles (Fig. 1, Supplementary Table S2 and S3). The MANOVA using the concatenated NADH5 and COI sequences indicated that 39% of the variation was found among regions, 28% among populations and 33% within populations (MANOVA $P = 0{\cdot}01$, Table 2). A principal coordinate analysis (PCoA) of population genetic differentiation (φ_{ST},) and the Nei genetic distance showed a clear geographic structure with three main clusters (Fig. 2). One cluster consisted of all populations in the west branch (OTT, VIN, DOV, TRE) plus the ROR sample in the northern part of River Glomma. The second cluster contained populations found in KON and in the lower Glomma (river NIT, SAR near the outlet). The largest cluster contained all the populations from the east branch and from Lake MJO (Fig. 2).

There was a highly significant correlation between hydrographic distance, and population genetic differentiation ($\varphi_{ST,}$ Mantel test, $r = 0{\cdot}28$, $P = 0{\cdot}01$, Fig. 3), with the exception of a few outliers exhibiting high waterway distances and low genetic differentiation (φ_{ST}). The ROR population had haplotypes for the NADH5 gene that cluster to the west branch rather than the downstream Glomma river (see Fig. 2). A re-analysis, excluding the ROR population, gave a stronger correlation between waterway distance and φ_{ST} ($r = 0{\cdot}48$, $P = 0{\cdot}01$).

The geographic distribution showed more or less the same structure for both NADH5 (G), and COI (C). In the west branch the haplotype G14C16 was the most common, compared with the east branch, where the haplotypes G24C20 and G24C29 dominated. Haplotype G04C36 dominated in the lower Glomma (NIT, SAR) and at KON (Fig. 1).

The network analysis detected 11 possible unsampled ancestral nodes. All the three haplotype clusters were linked to one unsampled ancestral node (mv10) in three different branches (Fig. 4). This network indicates a clear geographic structure. In the lower Glomma; haplotype G04C36 grouped with G08C09. From G04C36 a star-shaped network structure grouped the less common haplotypes. All the haplotypes from west branch was found with a star-shaped network structure around haplotype G14C16. The rest of the haplotypes had more or less a star-shaped structure linked with G24C20 and G24C29. There was weak support (bootstrap estimates <80%) for phylogenetic relationships between the haplotypes within Glomma river system when Neighbour-Joining method, ML and Maximum Parsimony analyses were used.

Based on analysis of the COI-gene, the time of separation of the *G. thymalli* populations in the river Trysil and those in the river Glomma was estimated to be 5000 ± 1000 (95% CI) years ago using the conservative mutation rate (2·0% divergence per million years), 2200 ± 600 years ago when using the intermediate mutation rate and 680 ± 80 years ago using

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

Table 2. Analysis of molecular variance (AMOVA) is based on concatenated mtDNA genes NADH5 and COI of *Gyrodactylus thymalli*. Among regions; are tested under random permutation of individuals across regions (region 1: ELV, HOL, REN, AAS, OSE, KOP, STO, ALV, ROR, region 2: EID, MJO, HUN, TRE, VIN, OTT, VAA, DOV, region 3: SAR, NIT, KON, the population codes refer to the names in Table 1). Among populations; are tested under random permutation of individuals across populations. Within population; are tested under random permutation of individuals across populations without regard to either of their original populations

Variance component	Variance	% total	Sum of squares	Degrees of freedom	Φ-statistics
Among regions	1·60	39	331·80	2	0·385
Among population	1·18	28	314·83	17	0·461
Within population	1·38	33	398·82	283	0·688
Total	4·16		1045·55	308	

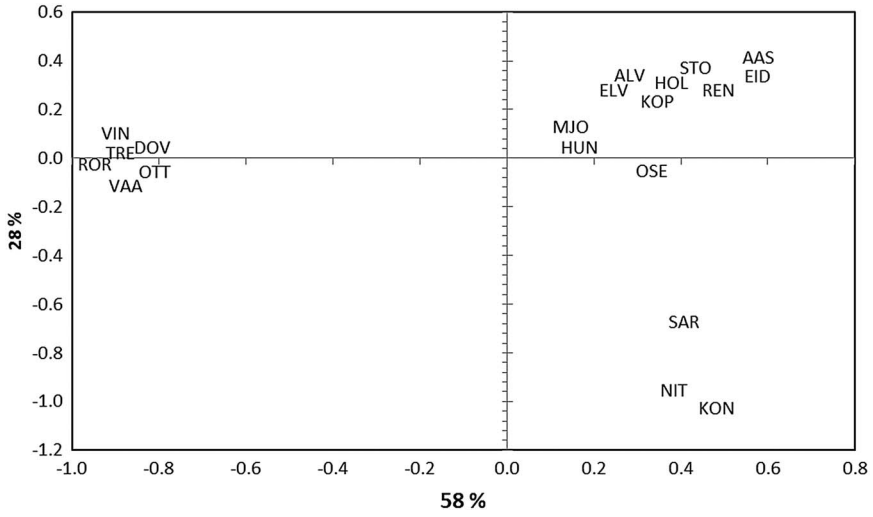

Fig. 2. Plot of the principal coordinate axis 1 (58%) and 2 (28%) of pairwise population genetic distances of *Gyrodactylus thymalli* populations within the Glomma river system. The analysis is based on the concatenated mtDNA genes NADH5 and COI. The population codes as in Table 1.

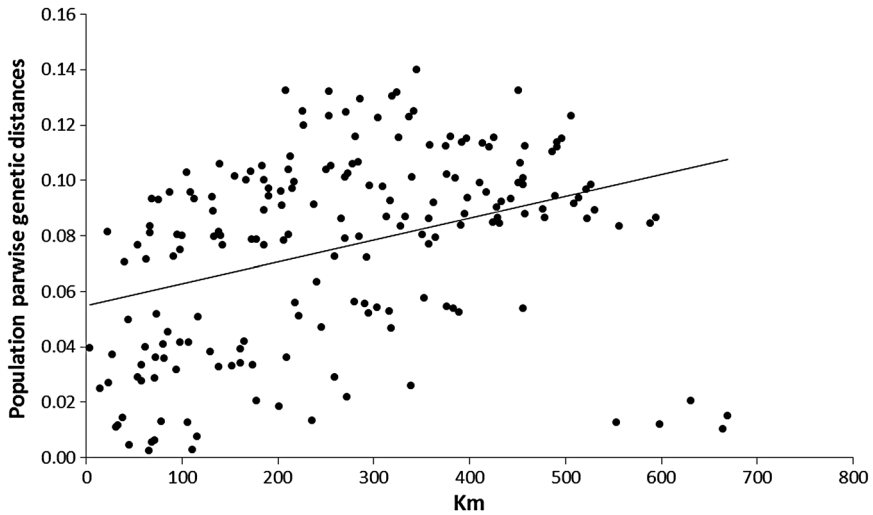

Fig. 3. Correlation between hydrographic distance (waterway distance, km) and *Gyrodactylus thymalli* population pairs genetic distances ($\varphi_{ST}/1-\varphi_{ST}$) within the Glomma river system (Mantel test, $r = 0{\cdot}28$, $P = 0{\cdot}01$). The analysis is based on the concatenated mtDNA genes NADH5 and COI. The outliers from the Røros population down to the right. A re-analysis, excluding the Røros population, gave a stronger correlation ($r = 0{\cdot}48$, $P = 0{\cdot}01$).

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

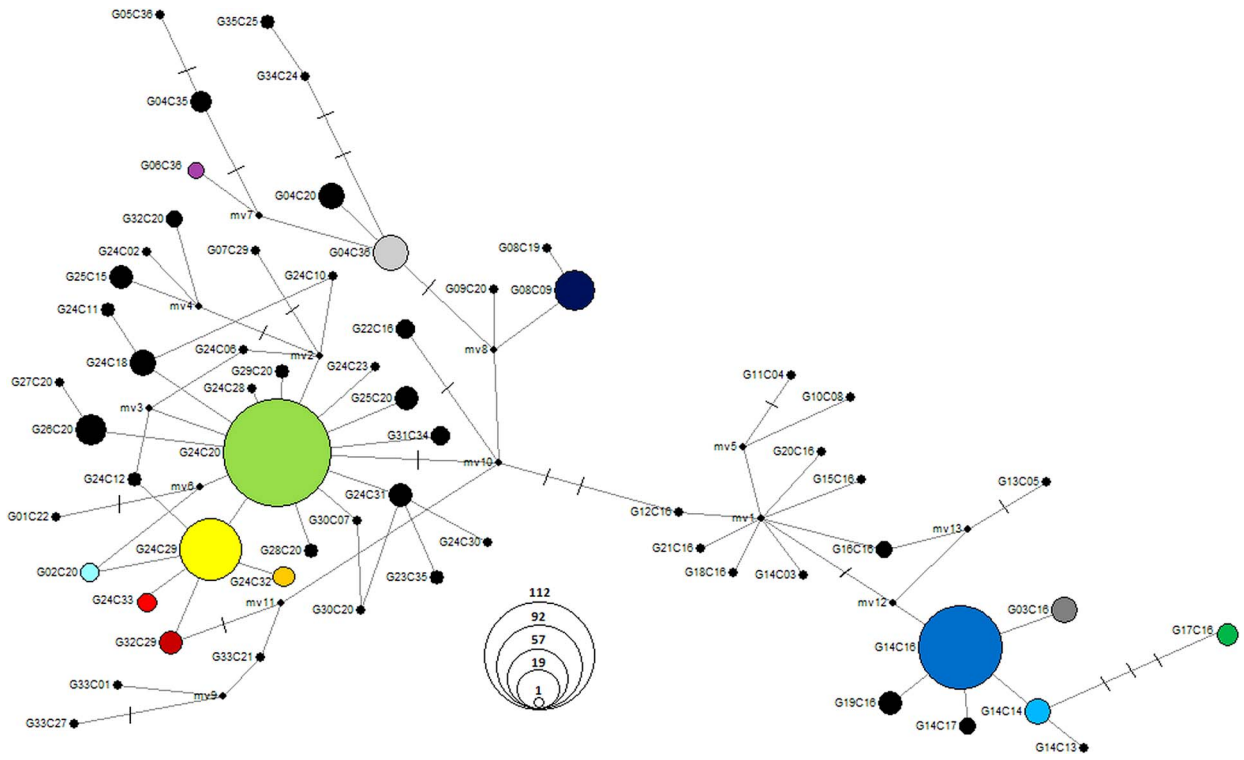

Fig. 4. Median-joining network is based on the concatenated mtDNA genes NADH5(G) and COI(C) haplotypes of *Gyrodactylus thymalli*, with haplotypes colour coded as in Fig. 1 (For the private haplotypes codes see online Supplementary Tables 2 and 3). The cross-lines indicate mutation steps more than one and the size of the circles represent haplotype frequency. The nodes mv1-11 is hypothetical unsampled haplotypes.

the high mutation rate suggested by Meinilä *et al*. (2004) for the genus *Gyrodactylus* (13·7–20·3% divergence per million years).

DISCUSSION

The monogenean ectoparasite *G. thymalli* and its grayling host have a wide distribution in the Glomma river system. Infected grayling were sampled at 20 localities and based on the analysis of two mitochondrial genes (COI and NADH5) we found large genetic variation with a high number of *G. thymalli* haplotypes throughout the river system. There were large genetic differences between populations and these differences were correlated with hydrographic distance. The 37 COI haplotypes revealed in our study included all the 8 haplotypes described earlier (Hansen *et al*. 2003, 2006, 2007*a*, *b*). By using the concatenated NADH5 and COI sequences, 58 haplotypes were found. The three most common haplotypes were differentially geographically distributed in the river system and represented nodes in a star-shaped network linked with several less-common haplotypes. These three main clusters of haplotypes were linked to three regions (the east branch, the west branch and lower Glomma) and described 39% of the molecular variation. This genetic structure may have evolved from one ancestral population after the last glaciation, or it may be the result of colonization by genetically different *G. thymalli* populations during separate invasion events. It may be impossible to resolve these two explanations, as earlier phylogenetic analyses were unable to resolve the relationship between the various clusters in Europe (Hansen *et al*. 2006, 2007*a*, *b*; Meinilä, *et al*. 2004). The estimated time of colonization of *G. thymalli* into the Glomma river system was after the last glacial period in Northern Europe. The different estimates for the divergence of the Glomma *G. thymalli* from the Trysil *G. thymalli* varied from 680 to 5000 years ago, varying from very recent divergence to several thousand years after the last glacial maximum.

Current available genetic methods have uncovered that the last ice age has strongly affected the genetic population structure of many species in the Northern hemisphere (Hewitt, 2000). A high haplotype diversity and low nucleotide diversity, together with a star-shaped haplotype network (single step mutations from central haplotypes), are indicative of a classic post-glacial expansion (Hewitt, 1996). This fits with our observations for *G. thymalli* in the Glomma river system. However, the distribution can also be due to different waves of post-glacial immigrant coming from different glacial refugia. Several studies propose that Mid-Norway and Sweden is such a mixture zones for many species

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

(Hewitt, 2000). Several evolutionary lineages of grayling have been identified, and these lineages seem to have invaded Norway during different time period and following different migration routes (Koskinen *et al*. 2000; Gum *et al*. 2005, 2009). It has been suggested that at least two grayling lineages meet in Lake MJO, the largest water body in the Glomma river system. This is supported by the observation of high numbers of *G. thymalli* haplotypes, and that all the three haplotype clusters are represented. The grayling and associated *G. thymalli* can have dispersed upstream or downstream, depending on how water flow has varied through time.

Dispersal and gene flow between populations of host specific fish parasites with a direct life cycle depends on dispersal of the fish host (Criscione and Blouin, 2004). Barriers to fish upstream migration will reduce the two-directional genetic exchange between parasite populations and result in separated populations (Blasco-Costa *et al*. 2012). The Glomma river system has numerous waterfalls and dams acting as migration barriers for grayling. When grayling colonized the river early after the last glaciation the geographic configuration of the river was different than today. However, in general upstream migration was prevented in a 'stepwise' manner as waterfalls arose when the ice melted and the land rose due to isostatic rebound (Huitfeldt-Kaas, 1918; Andersen and Borns, 1994).

For the entire river system, there was a strong correlation between geographic and genetic distance suggestive of a classic isolation-by-distance genetic structure, indicating restricted gene flow between the *G. thymalli* populations. In this large river system gene flow is limited by the presence of numerous migration barriers such as natural waterfalls and man-made dams. Such restricted gene flow might cause mutations to be sustained in isolated populations that have been separated for a long time (Mills, 2007). Our results show that 80% of the populations carry private haplotypes and in two of the populations examined we only observed such private haplotypes (AAS, OSE). These two populations are probably small, and are located where only downstream gene flow seems possible. Small populations of *G. thymalli* might lead to a strong genetic drift and an increased potential for allele fixation through genetic drift. Low genetic diversity might also be due to inbreeding (Mills, 2007). Isolated populations with private haplotypes as found in Glomma have also been found in a study of the genetic structure of the digenean *Crassicutis cichlasomae* (Razo-Mendivil *et al*. 2013). Their findings were that multiple colonization events and subsequent isolation are likely factors that shaped the genetic structure of the parasite (Razo-Mendivil *et al*. 2013). The grayling and the associated *G. thymalli* most likely colonized the mid parts of the Glomma river system early after the ice melted following the last glacial period. Following deglaciation when also the upper parts became available, the land also rose due to the isostatic rebound. At the same time the sea level rose, but at a different rate. Only grayling invading the river early were then able to colonize the upper reaches of the river system. These upper reaches later became isolated, allowing different mtDNA lineages of *G. thymalli* to evolve in isolation. Later, when the lower parts of the river, including the lake MJO, emerge from the sea and became accessible for fresh water fish, both upstream and downstream migrating grayling could colonize these areas (Koskinen *et al*. 2000). Koskinen *et al*. (2000) found three distinct lineages of grayling mtDNA in Scandinavia, presumably originating from three glacial refugia. Two of these lineages are found in MJO, indicating that grayling from two different evolutionary lineages live in sympatry in MJO. This can probably explain the occurrence of the most frequent *G. thymalli* haplotypes from all the three genetic clusters in the lake MJO. In addition, several *G. thymalli* haplotypes were found in the lake only. This observation can be the result of the relatively high sample size that we have from this lake. Or these private haplotypes may have evolved in geographically separated grayling populations spawning in the numerous tributaries draining into MJO (Kristiansen and Døving, 1996).

The genetic diversity of *G. thymalli* that we observed can be considered as high, whereas the genetic diversity of the grayling host is rather low (Koskinen *et al*. 2000). Earlier studies have shown similarity in genetic diversity of a fish host and its parasites (see e.g. Wu *et al*. 2009; Atkinson and Bartholomew, 2010; Razo-Mendivil *et al*. 2013) and parasites can be used as a proxy for understanding the evolutionary history of the host (Nieberding and Olivieri, 2007). In addition to the natural dispersal of fish and parasites following deglaciation, anthropochore relocation of grayling, within and/or between river systems, may have resulted in dispersal of *G. thymalli* haplotypes. The historically confirmed anthropochore relocation of grayling from OTT river to Lake VAA in 1906 (Huitfeldt-Kaas, 1918) is reflected in the genetic similarity of the *G. thymalli* haplotypes in Lake VAA and those in the OTT river as well as the genetic similarity between grayling above and below the migration barrier in the system (Junge *et al*. 2014). Human relocation of fish might also explain the genetic similarity between the haplotypes G03C16 and G14C16 in the east branch (ROR, ALV) and west branch (VAA, DOV, OTT, VIN, TRE), respectively. This was also supported by the isolation-by-distance analysis where ROR was removed, which showed a stronger overall relationship than when ROR was included. In a population genetic study

https://doi.org/10.1017/S003118201500133X Published online by Cambridge University Press

of grayling from the east branch of Glomma, the ROR population was more closely related to grayling from the neighbouring river Trysil than to populations downstream in the Glomma (ELV, ALV) (Heggenes *et al*. 2006). This result can be explained by a transient bypass between Glomma and Lake Femunden in the river Trysil in 1762. Fish movement to the Glomma from Lake Femunden is documented (Berg, 1986).

There is an ongoing discussion about if mtDNA is a good neutral marker, as selection on mtDNA genes has been demonstrated (Ballard and Pichaud, 2014). Our results show that mutations in the NADH5 gene cause amino acid substitutions to a greater extent than shown in previous studies of *G. thymalli* (Huyse *et al*. 2008). The differences on the NADH5 gene have often been associated with selection. However, a test for selection (using the ML model) was not significant. The COI gene contained less information at the population level than the NADH5 gene, even if both genes have similar estimated mutation rates. The NADH5 gene has also earlier been reported to have more variation than the COI gene in *Gyrodactylus* (Lindqvist *et al*. 2007; Huyse *et al*. 2008), and has been identified as being among the most variable regions of the mitochondrial DNA (Plaisance *et al*. 2007). By using both genes, we achieved stronger signals of population structure and a distinct isolation-by-distance genetic structure. In future studies, it is important to increase sample sizes, at the same time as using more sensitive molecular markers for more detailed studies of population structure. Since the whole genome of *G. salaris* is mapped (Hahn *et al*. 2014), these genetic resources clearly provide more opportunities for detailed studies into population structure by using markers such as single nucleotide polymorphism and microsatellites.

In conclusion, the observed mtDNA gene diversity of the *G. thymalli* populations in the Glomma river system is the result of postglacial processes that have created barriers to upstream migration of fish leading to isolation by distance structure. Reduced upstream gene flow, as well as potentially high levels of genetic drift, has resulted in reduced level of genetic diversity in upstream populations. *Gyrodactylus thymalli* has a higher genetic diversity compared with *G. salaris* which has been introduced into new rivers (Hansen *et al*. 2007*b*; Anttila *et al*. 2008). Low genetic variation may be an indication of newly introduced *Gyrodactylus* species. The results from this study should be taken into account when managing fish hosts and their parasites. Our study also indicates that sample size needs to be increased in such studies.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S003118201500133X

ACKNOWLEDGEMENTS

We are grateful to Ole Haakon Heier, Jan Teigen, Erik Lien, Finn Gregersen, Henning Pavels and Camilla Fossum Pettersen for their assistance during sampling of grayling. Nanna Winger Steen and Emelita Rivera Nerli are acknowledged for their technical assistance in the laboratory and Kjartan Østbye for helpful discussions.

FINANCIAL SUPPORT

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

REFERENCES

Andersen, B. G. and Borns, H. W. (1994). *The Ice Age World*. Scandinavian University Press, Oslo.

Anttila, P., Romakkaniemi, A., Kuusela, J. and Koski, P. (2008). Epidemiology of *Gyrodactylus salaris* (Monogenea) in the River Tornionjoki, a Baltic wild salmon river. *Journal of Fish Diseases* **31**, 373–382.

Atkinson, S. D. and Bartholomew, J. L. (2010). Spatial, temporal and host factors structure the *Ceratomyxa shasta* (Myxozoa) population in the Klamath River basin. *Infection Genetics and Evolution* **10**, 1019–1026.

Ballard, J. W. O. and Pichaud, N. (2014). Mitochondrial DNA: more than an evolutionary bystander. *Functional Ecology* **28**, 218–231.

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

Berg, M. (1986). *Det norske lakse- og innlandsfiskets historie: fiskeetaten 1855–1986*. Universitetsforlaget, Oslo.

Bermingham, E., McCafferty, S. S. and Martin, A. P. (1997). *Fish Biogeography and Molecular Clocks: Perspectives from the Panamanian Isthmus*. Academic Press, Inc., 1250 Sixth Ave., San Diego, California 92101, USA 14 Belgrave Square, 24–28 Oval Road, London NW1 70X, England, UK.

Bernatchez, L. (2001). The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution* **55**, 351–379.

Blasco-Costa, I., Waters, J. M. and Poulin, R. (2012). Swimming against the current: genetic structure, host mobility and the drift paradox in trematode parasites. *Molecular Ecology* **21**, 207–217.

Criscione, C. D. and Blouin, M. S. (2004). Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. *Evolution* **58**, 198–202.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analyse of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**, 401–410.

Felsenstein, J. (1981). Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution* **35**, 1229–1242.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Fromm, B., Burow, S., Hahn, C. and Bachmann, L. (2014). MicroRNA loci support conspecificity of *Gyrodactylus salaris* and *Gyrodactylus thymalli* (Platyhelminthes: Monogenea). *International Journal for Parasitology* **44**, 787–793.

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A. and Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332.

Gum, B., Gross, R. and Kuehn, R. (2005). Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. *Molecular Ecology* **14**, 1707–1725.

Gum, B., Gross, R. and Geist, J. (2009). Conservation genetics and management implications for European grayling, *Thymallus thymallus*: synthesis of phylogeography and population genetics. *Fisheries Management and Ecology* **16**, 37–51.

Hahn, C., Fromm, B. and Bachmann, L. (2014). Comparative genomics of flatworms (Platyhelminthes) reveals shared genomic features of Ecto- and Endoparasitic Neodermata. *Genome Biology and Evolution* **6**, 1105–1117.

Hansen, H., Bachmann, L. and Bakke, T. A. (2003). Mitochondrial DNA variation of *Gyrodactylus* spp. (Monogenea, Gyrodactylidae)

populations infecting Atlantic salmon, grayling, and rainbow trout in Norway and Sweden. *International Journal for Parasitology* **33**, 1471–1478.

Hansen, H., Martinsen, L., Bakke, T. A. and Bachmann, L. (2006). The incongruence of nuclear and mitochondrial DNA variation supports conspecificity of the monogenean parasites *Gyrodactylus salaris* and *G. thymalli*. *Parasitology* **133**, 639–650.

Hansen, H., Bakke, T. A. and Bachmann, L. (2007*a*). DNA taxonomy and barcoding of monogenean parasites: lessons from *Gyrodactylus*. *Trends in Parasitology* **23**, 363–367.

Hansen, H., Bakke, T. A. and Bachmann, L. (2007*b*). Mitochondrial haplotype diversity of *Gyrodactylus thymalli* (Platyhelminthes; Monogenea): extended geographic sampling in United Kingdom, Poland, and Norway reveals further lineages. *Parasitology Research* **100**, 1389–1394.

Hebert, P. D. N., Ratnasingham, S. and deWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, S96–S99.

Heggenes, J., Qvenild, T., Stamford, M. D. and Taylor, E. B. (2006). Genetic structure in relation to movements in wild European grayling (*Thymallus thymallus*) in three Norwegian rivers. *Canadian Journal of Fisheries and Aquatic Sciences* **63**, 1309–1319.

Hesthagen, T. and Sandlund, O. T. (2004). Fish distribution in a mountain area in south-eastern Norway: human introductions overrule natural immigration. *Hydrobiologia* **521**, 49–59.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247–276.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913.

Huitfeldt-Kaas (1918). Ferskvandsfiskenes utbredelse og indvandring i Norge, med et tillæg om krebsen. in Norwegian (Distribution and postglacial colonisation of freshwater fishes in Norway, including the cray fish). *Centraltrykkeriet, Kristiania (Oslo)*.

Hurvich, C. M. and Tsai, C. L. (1989). Regression and time-series model selection in small samples. *Biometrika* **76**, 297–307.

Huyse, T., Buchmann, K. and Littlewood, D. T. J. (2008). The mitochondrial genome of *Gyrodactylus derjavinoides* (Platyhelminthes : Monogenea) – A mitogenomic approach for *Gyrodactylus* species and strain identification. *Gene* **417**, 27–34.

Irwin, D. M., Kocher, T. D. and Wilson, A. C. (1991). Evolution of the cytochrom-b gene of mammals. *Journal of Molecular Evolution* **32**, 128–144.

Junge, C., Museth, J., Hindar, K., Kraabøl, M. and Vøllestad, L. A. (2014). Assessing the consequences of habitat fragmentation for two migratory salmonid fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* **24**, 297–311.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120.

Koskinen, M. T., Ranta, E., Piironen, J., Veselov, A., Titov, S., Haugen, T. O., Nilsson, J., Carlstein, M. and 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.

Kristiansen, H. and Døving, K. B. (1996). The migration of spawning stocks of grayling *Thymallus thymallus*, in Lake Mjosa, Norway. *Environmental Biology of Fishes* **47**, 43–50.

Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nature Reviews Genetics* **6**, 654–662.

Kuusela, J., Holopainen, R., Meinila, M., Anttila, P., Koski, P., Zietara, M. S., Veselov, A., Primmer, C. R. and Lumme, J. (2009). Clonal structure of salmon parasite *Gyrodactylus salaris* on a coevolutionary gradient on Fennoscandian salmon (*Salmo salar*). *Annales Zoologici Fennici* **46**, 21–33.

L'Abée-Lund, J. H., Eie, J. A., Faugli, P. E., Haugland, S., Hvidsten, N. A., Jensen, A., Melvold, K., Pettersen, V., Petterson, L. E. and Saltveit, S. J. (2009). *Rivers of Boreal Uplands*. Elsevier, Amsterdam.

Lindqvist, C., Plaisance, L., Bakke, T. A. and Bachmann, L. (2007). Mitochondrial DNA variation of a natural population of *Gyrodactylus thymalli* (Monogenea) from the type locality River Hnilec, Slovakia. *Parasitology Research* **101**, 1439–1442.

Malmberg, G. (1993). Gyrodactylidae and *Gyrodactylosis* of *Salmonidae*. *Bulletin Francais de la Peche et de la Pisciculture* **328**, 5–46.

Mäinila, M., Kuusela, J., Zietara, M. and Lumme, J. (2002). Brief report – Primers for amplifying not similar to 820 bp of highly polymorphic mitochondrial COI gene of *Gyrodactylus salaris*. *Hereditas* **137**, 72–74.

Mäinila, M., Kuusela, J., Zietara, M. S. and Lumme, J. (2004). Initial steps of speciation by geographic isolation and host switch in salmonid pathogen *Gyrodactylus salaris* (Monogenea : Gyrodactylidae). *International Journal for Parasitology* **34**, 515–526.

Mills, L. S. (2007). *Conservation of Wildlife Populations: Demography, Genetics, and Management*. Blackwell, Oxford.

Mo, T. A., Appleby, C. and Sterud, E. (1998). Parasites of grayling (*Thymallus thymallus*) from the Glomma river system, south-eastern Norway. *Bulletin of the Scandinavian Society for Parasitology* **8**, 6–11.

Muse, S. V. and Gaut, B. S. (1994). A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Molecular Biology and Evolution* **11**, 715–724.

Nesbø, C. L., Fossheim, T., Vøllestad, L. A. and Jakobsen, K. S. (1999). Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology* **8**, 1387–1404.

Nieberding, C. M. and Olivieri, I. (2007). Parasites: proxies for host genealogy and ecology? *Trends in Ecology and Evolution* **22**, 156–165.

Northcote, T. G. (1995). Comparative biology and management of Arctic and European grayling (Salmonidae, *Thymallus*). *Reviews in Fish Biology and Fisheries* **5**, 141–194.

Peakall, R. and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539.

Plaisance, L., Huyse, T., Littlewood, D. T. J., Bakke, T. A. and Bachmann, L. (2007). The complete mitochondrial DNA sequence of the monogenean *Gyrodactylus thymalli* (Platyhelminthes : Monogenea), a parasite of grayling (*Thymallus thymallus*). *Molecular and Biochemical Parasitology* **154**, 190–194.

Polzin, T. and Daneshmand, S. V. (2003). On Steiner trees and minimum spanning trees in hypergraphs. *Operations Research Letters* **31**, 12–20.

Pond, S. L. K. and Frost, S. D. W. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* **22**, 1208–1222.

Pond, S. L. K., Frost, S. D. W. and Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679.

Razo-Mendivil, U., Vazquez-Dominguez, E. and de Leon, G. P. P. (2013). Discordant genetic diversity and geographic patterns between *Crassicutis cichlasomae* (Digenea: Apocreadiidae) and its ciclid host, '*Cichlasoma*' *urophthalmus* (Osteichthyes: Cichlidae), in Middle-America. *Journal of Parasitology* **99**, 978–988.

Saitou, N. and Nei, M. (1987). The neighbor-joining method – a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792.

Sterud, E., Mo, T. A., Collins, C. M. and Cunningham, C. O. (2002). The use of host specificity, pathogenicity, and molecular markers to differentiate between *Gyrodactylus salaris* Malmberg, 1957 and *G. thymalli* Zitnan, 1960 (Monogenea: Gyrodactylidae). *Parasitology* **124**, 203–213.

Suzuki, Y. and Gojobori, T. (1999). A method for detecting positive selection at single amino acid sites. *Molecular Biology and Evolution* **16**, 1315–1328.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Tamura, K., Nei, M. and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11030–11035.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Wright, S. (1943). Isolation by distance. *Genetics* **28**, 114–138.

Wu, S. G., Wang, G. T., Xi, B. W., Xiong, F., Liu, T. and Nie, P. (2009). Population genetic structure of the parasitic nematode *Camallanus cotti* inferred from DNA sequences of ITS1 rDNA and the mitochondrial COI gene. *Veterinary Parasitology* **164**, 248–256.

Østbye, K., Bernatchez, L., Næsje, T. F., Himberg, K. J. M. and Hindar, K. (2005). Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Molecular Ecology* **14**, 4371–4387.

Østdahl, T., Skurdal, J., Kaltenborn, B. P. and Sandlund, O. T. (2002). Possibilities and constraints in the management of the Glomma and Lagen river basin in Norway. *Archiv fuer Hydrobiologie Supplement* **141**, 471–490.