THE UNIVERSITY OF HULL

Genetic factors affecting establishment during invasions: the introduction of the topmouth gudgeon *(Pseudorasbora parva)* and the rainbow trout *(Oncorhynchus mykiss)* in Europe

being a Thesis submitted for the Degree of

Doctor of Philosophy

in the University of Hull

by

Andrea Simon MSc

July 2012

Abstract

The study of biological invasions is a major research topic, both because of the ecological and economical damage caused by invasive species and also as a great natural experiment to study evolutionary responses of non-native populations to their new environment, and the factors influencing invasions. Introduced species often evolve rapidly, despite the assumed loss of genetic variation associated with bottlenecks during the invasion process. In order examine the processes and mechanisms affecting the outcome invasions I studied two non-native fish species, the topmouth gudgeon (*Pseudorasbora* parva) is an Asian cyprinid that is found in most European countries as a result of accidental introductions. Rainbow trout (*Oncorhynchus* mykiss) has been introduced from the United States for aquaculture and angling, however, despite numerous introductions, it has only been able to establish in few European waters.

I used mitochondrial DNA and microsatellite markers to understand the invasion history of these species and the factors that influence their establishment success/failure.

Part of the *cytochrome b* gene was analysed in European and native Asian *P. parva* populations and microsatellite markers were used to investigate the source populations of the species. The analyses elucidated the colonisation pattern of *P. parva* in Europe and supported the hypothesis that the species spread through long-distance and stepping-stone methods and originate from admixed source populations.

In *O. mykiss*, part of the *d-loop* region of the mitochondrial genome was analysed to compare the phylogeographic structure of native US and introduced European populations to examine the spread of the species outside its native range, as well as to find out whether the resistant Hofer strain is the source population of the European rainbow trout populations. I found that European populations are likely to originate from various sources, mainly from California. The Hofer strain is likely to have contributed to some of the wild European populations.

Assessing the role of these processes is fundamental in understanding invasive species and finding suitable management practices to control them. From an evolutionary point of view, I was able to detect some of the processes that are important during invasions, in these studies particularly the role of multiple introductions and introduction from genetically admixed source populations.

Acknowledgements

I am extremely grateful to my supervisors, Dr. Bernd Hänfling, Dr. Cock van Oosterhout and Dr. Lori Lawson-Handley for their advice and support throughout my studies and I would like to thank them for their comments and suggestions for my thesis and to Dr. Bill Hutchinson for his help with my work.

My parents and family have always encouraged me (köszi drága Anyucikám, kérek mákostésztát!) and a big thank you to my girls, Kat, Eda, Manci, Krinya, Isabel, Violaine, Caitriona, Sandra - could not have done it without the chats and countless phonecalls, I am very lucky to have such great family and friends.

Thanks to my extended family, Nancy and Kenneth Stuart for letting their house be my home away from home...and for the G&Ts.

In my daily work, I have been very fortunate to be part of the Evolutionary Biology Group in Hull - what a great group to work in, thank you everyone!

I would also like to thank NERC for funding my work and the many people who helped me with samples or advice: Professor Robert Behnke and Mrs Sally Behnke, Susan Rivers, Péter Takács, Gary Kelley, Kevin Gadsby, Eva Zahorska and Stephen Moores.

I am also grateful to my examiners, Dr. Stefano Mariani and Dr. Domino Joyce for reading and commenting on my thesis, thank you!

Declaration

I declare that this thesis has been compiled by myself and is the result of my own investigations.

It has not been submitted for any other degree, and all other sources of information have been duly acknowledged.

Andrea Simon

Contents

CHAPTER V: General discussion

CHAPTER I

General introduction

1.1 Introduction: Biological invasions

Many populations change their range distribution naturally over time. However, substantial natural range shifts usually take place over geological time scales and sudden natural expansions are rare events. In contrast, human mediated dispersal during the last two centuries has led to large-scale species translocations resulting in increased levels of biotic homogenisation across continents.

Species that are introduced by humans either intentionally or unintentionally to geographic areas where they do not naturally occur are defined as Non-Indigenous Species (NIS) (Delach 2006; Shigesada and Kawasaki 1997). Only those NIS that establish populations in the non-native area and spread outside a human-dominated environment, are considered to be invasive (Levine, 2008).

Biological invasions are a constant ecological threat and are now considered to be the second most important cause for the loss of biodiversity after habitat loss (Sala *et al* 2000). Apart from the ecological costs, non-native species have been associated with extremely high economic costs: in the United States the cost of invasive species in the past few years on average were estimated to be around $137 billion per year (Pimentel *et al* 2000) and vary greatly between taxa (Figure 1).

Human activities, such as travelling, worldwide commerce between countries continue to increase, and these activities are greatly encouraging and promoting the spread of non-native species. As the number of invasive species continues to increase, predicting invasions has become more

relevant in research. Consequently, understanding the effects of biological invasions and comprehending the mechanisms supporting the invasion process is an increasingly important area of research. The study of invasive species is emerging as a new biological discipline and research area (Davis 2009) and has developed into a subfield within ecology (Lodge 1993) with its own theoretical and conceptual framework.

In invasion biology theory it is widely accepted that only a small proportion of invasive species become established in the new environment. This is known as the 'tens rule', according to which only one out of ten imported species escapes to the wild, one out of ten species spreads and one out of these ten becomes a pest (Vander Zanden 2005).

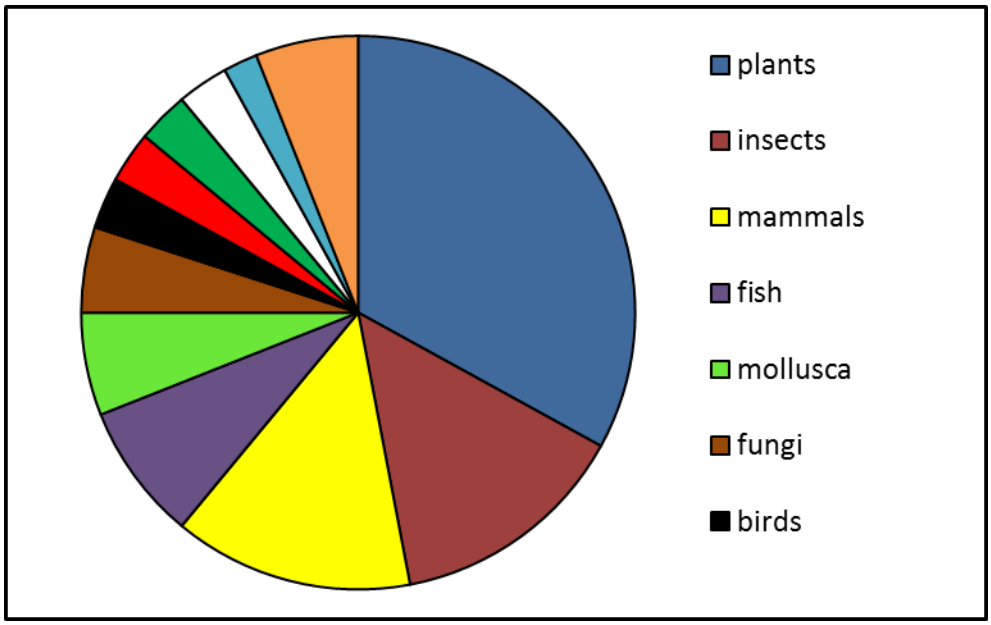


Figure 1. Taxonomic distribution of 100 most invasive species. Source: Global Invasive Species Database

However, although this rule has been widely accepted both in policy making and research thinking, research done on vertebrates by Jeschke and Strayer

(2005) found that around quarter of all invasions are successful as the establishment success of invasive species depends on both ecological and evolutionary factors, as well as on the characteristics of the target community.

1.1.2 The history of invasion biology

Accounts of invasive species have been around for centuries (Davis 2009), although formal records did not start until the 19th century. First botanists, then zoologists began to note the effects of human activities on the redistribution of various species (Drude 1896; Howard 1897) and we also find Darwin stating that 'if invading species were more similar to native species, it would be more difficult for them to establish, because of increased competition from ecologically similar species' (Darwin 1859). By the beginning of the 20th century, there was a shift in attitudes towards non-native species, as people realised that new species very often meant new pests (Froese and Pauly 2008). By the 1940s biologists wrote about human involvement in changing environments and Elton (1958) created the field of invasion biology. In his 1958 book, 'The Ecology of Invasions by Animals and Plants' he emphasised the damage caused by non-native species and in 1964, H.G. Baker and G.L. Stebbins held a conference on 'The Genetics of Colonising Species'. In the 1970s the general public started to show more interest in environmental issues and this interest has been increasing steadily ever since (Callaway and Maron 2006), so research on invasive species in the past 50 years has been well documented. In recent years the topic has

been getting further attention, helping invasion biology to become a significant part of modern biological research.

1.1.3 Stages of the invasion process

Biological invasions are multi-stage processes (Suarez and Tsutsui 2008), although there is some controversy whether there are three stages: arrival establishment, spread (Freckleton *et al* 2006) or four stages within the invasion process, which are arrival, establishment, spread and adjustment. Adjustment is a static or decline phase and can happen as a result of transmission and spread of local pathogens into the population (Reise *et al* 2006). Allendorf (2003) groups introduction together with establishment and colonisation (spread), where the invasive species initiating new populations is the first stage of invasion. During introduction, individuals must be moved outside their natural geographic range by various modes of transport, which include ships or accidental transfer of species with other species. This stage differentiates invasions and natural range expansions (Richardson *et al* 2000), as it occurs with human assistance and allows species to cross biogeographic barriers that they otherwise would not be able to. The second stage is the spread and possible replacement of native species by introduced species (Kolar and Lodge 2002; Rosecchi *et al* 2001). In this stage, invasive species might continue to spread from long-distance dispersal from foreign sources or from short-distance dispersal (Sakai 2001).

1.1.4 Factors facilitating the invasion process

One of the key questions in invasion biology has been why most introduced species fail to establish and what distinguishes the few successful invasions. Several explanations have been proposed, such as specific traits of the invader, the release from natural biotic constraints but also the potential to adapt quickly to the new environment (Kolar and Lodge 2001).

Previous studies on traits which distinguish successful invaders have identified a number of life history and/or ecological traits that enable invasive species to establish in new areas, such as short generation time, fast initial growth, being a good competitor (Newsome and Noble 1986; Sakai 2001), or tolerance to a wide range of environmental conditions or aggressiveness (Moyle 1986).

Another potential factor facilitating invasion is the loss of natural enemies during the translocation stage (the enemy release hypothesis; ERH; Keane and Crawley 2002, Colautti *et al* 2004). The ERH puts forward that the abundance and/or impact of an invasive species is related to the scarcity of natural enemies in the introduced range compared to the native range. This has been reported in the spotted knotweed (*Centaurea maculosa*) (Callaway *et al* 2004) or in the European starling (*Sturnus vulgaris*) (Torchin *et al* 2003).

Various models have been put forward to describe the how the factors described above interact during invasions. They differ in their focus is on the characteristics of the invading species, of the species in the invaded area (non-relational) or on the relationship between these two factors (key-lock

model) Heger and Trepl (2003). While certain characteristics are thought to be beneficial for invasive species, on their own not one characteristic is responsible for the success of an invader. The key-lock model assumes that all invading species have to suit certain conditions in their new environment before they are able to become established (Heger and Trepl 2003) and if these conditions are met, the species is able to establish. According to the Theory of fluctuating resource availability (Figure 2), Davis *et al* (2000) hypothesized that a plant community is more likely to be invaded if the amount of unused resources increases, as the intensity of competition is inversely correlated with the amount of unused resources (Davis *et al* 1998).

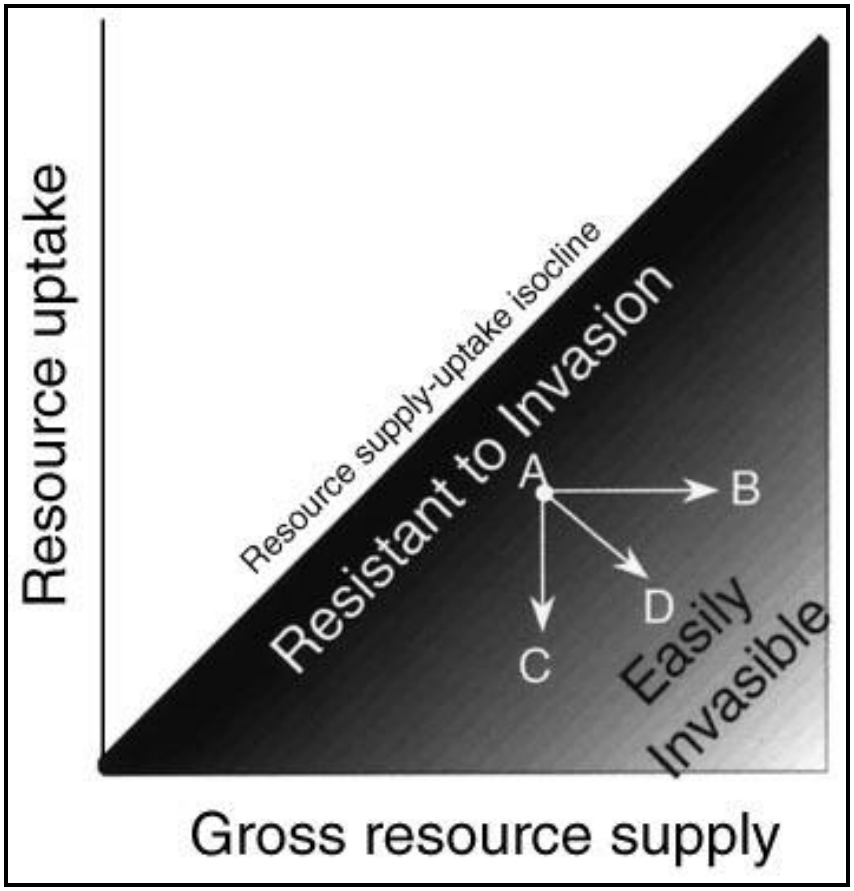


Figure 2. Theory of fluctuating resource availability as adapted from Davis *et al* (2000)

The model of steps and stages takes it further by stating that it is not the characteristics of the invaders and the invaded environments, or their interaction as separate steps that make an invasion a success or a failure, but the whole process as one and the invader has to come over each step individually, facing different threats at each step, such as unfavourable abiotic conditions, environmental stochasticity, or new areas in which to spread after colonisation (Heger and Trepl 2003).

1.1.5 Evolutionary aspects of invasions

The evolutionary aspects of biological invasions were somewhat overlooked for a long time, because the effects of invasions were thought to have been too fast for evolution to play any important part in the dynamics of invasions (Sax 2005). However, recent studies have shown that evolutionary factors are involved in both establishment success and impact on native species. An important characteristic of invasions is the time lag between colonisation and range expansion, during the establishment phase, which can be explained as a phase during which adaptive evolution to the new environment takes place (Suarez and Tsuitsui 2008). It is during this time that populations can be exposed to strong selection for local adaptation and the rate of this evolutionary response can have an effect on the speed of the invasion (Colautti *et al* 2010). Therefore, populations which invade areas that are similar to their native environments should make adapting easier. Ecological conditions might, however, be very different to those in the original area, which poses an additional challenge for the invading

population. In the new environment they might have to face a new range of selective pressures and they can also become selective agents on native taxa in the new environment and rapid evolution for both the invaders and the invaded species becomes possible, as both species have to quickly adapt to the new conditions (Suarez and Tsutsui 2008). This has been shown in various studies, such as the impact of the cane toad (*Bufo marinus)* on the morphology of native snakes (Phillips and Shine 2004) or the change in beak length in the soapberry bug (*Jadera haematoloma)* after colonising newly introduced host plants (Carroll and Dingle 1996). Invasive species continue to evolve during the invasion process, in particular if they are faced with new selection pressures produced by the new environment (Lau 2008; Suarez and Tsutsui 2008). However, according to population genetics theory, high dispersal ability and the ability to quickly adapt to the new environment are somewhat contradictory, as high dispersal ability and high migration between different populations hinders local adaptation; and the ability to rapidly adapt to new environments assumes spatial variation in selection pressures (Duckworth 2008). The first stages of invasion, the introduction and establishment are often associated with reduced genetic variability when compared to the source population, as a result of a small number of founders and small population size (Dlugosch and Parker 2008). The genetic bottleneck results in genetic drift that is thought to reduce genetic variability and control the ability of a population to adapt to a new environment and can be particularly strong when the invasive population originates from the same source. Decreased genetic diversity can increase

inbreeding depression that reduces the fitness of the offspring and the population's ability to evolve (Allentoft and O'Brien 2010; Ellstrand and Elam 1993). However, it has also been shown that intermediate size bottlenecks during invasions can purge deleterious alleles and the invasive population is able to persist despite being inbred (Facon *et al* 2011). When invasive species are introduced into new environments, they face novel selection pressures and the adaptive evolution following the initial colonisation is very important. This also offers a great opportunity to study evolution in the wild as a great natural experiment to understand the evolutionary responses of invasive species (Holway 1998; Yoshida 2007).

Population genetic studies are an increasingly significant tool in understanding biological invasions (Lee 2002) and they are carried out to understand more about the native as well as the introduced populations or about the number of introductions that took place, as the genetic structure of the invasive population can also influence its ability to become established.

Propagule pressure, the rate at which a species is introduced into an ecosystem (Roman and Darling 2007), was mentioned in an ecological context but it is also used in a genetic context to explain why some populations persist and others do not. It has a significant effect on genetic diversity in introduced populations as a complex measure of inoculum size and the number of introduction events (Roman and Darling 2007). Large inoculum size means the release of numerous individuals into a new area, which increases the probability that an introduced population will retain representative samplings of source genetic diversity and experience less of

the negative consequences of inbreeding as genetic drift takes over as the main force in driving evolution. Multiple introduction events can increase the size of the inocula and the genetic diversity of the invasive populations compared to the native range. In terms of the evolution of invasiveness it is important to look at how the introduction of new lineages affects the success of invasion (Roman and Darling 2007).

The genetic diversity of the population therefore might be crucial to a successful invasion. Some reviews in recent years have shown that the loss of genetic variation in invading populations is not as wide-spread as previously thought and many species manage to maintain genetic variation during the invasion process (Roman and Darling 2007).

This presents the question of the 'genetic paradox': how do invasive populations become established despite their expected low genetic diversity? Recent research has shown that loss of genetic diversity is not always found in these populations, such as in the common ragweed *Ambrosia artemisiifolia* (Genton *et al* 2005) or despite genetic bottlenecks, such as in the St John's wort (*Hypericum perforatum*) (Prentis *et al* 2008).

1.2. Population genetic approaches to study biological invasions

1.2.1 History of population genetics and theoretical background

Population genetics studies the genetic composition of populations, and the evolutionary mechanisms that affect it, such as natural selection, genetic drift, mutation and gene flow. Darwin highlighted the importance of natural selection and Mendel provided a mechanism for heredity, thereby both establishing the foundation of population genetics. Followers of Mendel believed that discrete traits are encoded by alleles and were opposed to the gradual changes proposed by Darwin's natural selection. While in the beginning of the 20th century - due to the difficulties of obtaining empirical data - population genetics was a theoretical science, the pioneers of this new branch of biological sciences described how the idea of natural selection can exist together with Mendelian inheritance. R.A. Fisher introduced the 'fundamental theorem of natural selection' (1930) and showed that natural selection can change gene frequencies and the continuous variation observed is a result of various discrete genes operating together, while J.B.S. Haldane applied statistical analysis to natural selection. Sewall Wright established his 'shifting balance theory' in 1932, suggesting that an interaction between genetic drift, migration and natural selection is more important than any of these forces alone. Wright also researched the ways in which different genes interact to produce a specific phenotype and concluded that the effects of gene combinations is more important than the effects of individual genes, known as the non-additive model or epistasis

(Wright 1978). The term epistasis itself was used by Bateson in 1909 to describe where one the alleles of one gene mask the effects of the alleles of another gene. The study of epistasis is central to research in evolutionary biology, and after Wright and Bateson, Fisher worked on it as well after discovering deviations form expected additive effects on quantitative traits of alleles at the same loci, demonstrating dominance (de Visser *et al* 2011).

The modern synthesis of evolution united genetics with Darwin's ideas and its most important advocate, Dobzhansky argued (1937) that mutation is the most important source of genetic variation but natural selection is the main process driving adaptive evolution and that natural selection not only could act on small genetic differences but that it actually did.

The molecular era of population genetics began when Lewontin and Hubby introduced protein electrophoresis as a population genetic method in 1966 (Lewontin and Hubby 1966). While electrophoresis had been used since the 1930s, they were the first to apply this method to show the amount of genetic variation that exists in a natural population, using *Drosophila pseudoobscura* as study species. Their initial thought was that the observed variation is maintained by some form of natural selection. As an alternative hypothesis Kimura (1968) proposed the neutral theory of molecular evolution. According to this theory, most amino acid substitutions are selectively neutral, and substitutions occur at a nearly constant rate and the maintenance of polymorphism is possible as there is no cost of selection and that random genetic drift has an important role in forming genetic structure.

Scientists in the 1960s expanded on the idea of natural selection incorporating the new discoveries of genetics, as well as the ideas of kin selection based on the idea of inclusive fitness by Hamilton (1964). By the 1980s it has become possible to analyse DNA using restriction enzymes and Kreitman and Akashi (1995) showed that variation is even more pervasive at the DNA level than at the protein level.

In the last few years, molecular techniques, such as sequencing, cloning (Brown 2006), Real-time PCR (Kubista *et al* 2006) have become much more powerful and accessible, a number of genome projects have taken place, which made the study of comparative genomics possible by comparing genes in different organisms, bridging the gap between genetics and ecology to answer questions about the processes that cause populations to change and diverge from each other over time, the kind of genetic variation is present and the causes responsible for it.

1.2.2. Genetic markers

1.2.2.1 Mitochondrial DNA

Mitochondrial DNA was introduced as a molecular tool in population biology in the 1970s (Avise 1979) and phylogenetic studies on mtDNA accelerated novel perspectives on the study of evolution (Avise 1987). The mitochondrial genome is maternally inherited and it demonstrates no recombination events but exhibits high mutation and mutation fixation rates. It is easy to isolate, has a simple genetic structure without repetitive elements, pseudogenes and introns and gene arrangement seems to be very

stable The evolutionary rate of mtDNA is 5 to 10 times faster than nuclear DNA (Hickey *et al* 2007), possibly because mitochondria do not have repair enzymes for errors in replication and this way a high level of transversions and transitions are produced, which leads to a high degree of variability between individuals (Castro 1998). Also, as the evolutionary rate of mtDNA is clock-like, the levels of mtDNA divergence should reflect divergence times (Galtier *et al* 2009). These properties make mtDNA a useful genetic marker to monitor divergence in animal populations and presently most molecular studies are likely to include the study of mtDNA genes.

Clonal inheritance (lack of paternal leakage), as well as germ-line bottleneck prevent effective recombination (Galtier *et al* 2009), as despite genetic variation found within species in mtDNA, individuals – with a few exceptions - are mostly homoplasmic (Cao *et al* 2009). In recent years however, some of these assumptions associated with mtDNA have been questioned and the suitability of mtDNA as a marker for population genetic analyses has come under scrutiny (Rand and Kann 1996; Fry 1999). For example, a study on human mitochondrial lineages by Eyre-Walker (1999) found levels of homoplasy that seem to suggest recombination in the mitochondrial DNA. This and other similar studies encouraged further research on this topic and it was empirically shown by Lunt and Hyman (1997) in nematodes, by Hoarau *et al* (2002) in fish, in butterflies by Andolfatto *et al* (2003) and by Filipowitz *et al* in mussels (2008). While the human studies later proved to be a likely result of mutation hotspots, when studying mtDNA, we still have to be aware of the possibility of within-

species homoplasy (Galtier *et al* 2009). The route, by which mitochondria becomes non-clonal has not been identified (Eyre-Walker and Adawalla 2001); the question remains open for future debate. Another questionable property was the constancy of mutation rates within the mitochondrial genome. It is difficult to measure mutation rates directly; hence the number of substitutions accumulating during a certain time period is often used as an indicator of mitochondrial mutation rate (Ellegren 2009). The mutation rate of mtDNA is thought to be 2% per million years (Moritz 1987); although studies have shown large amounts of variation in mtDNA substitution rates: 30-fold between birds by Galtier (2009), 100-fold between mammals by Nabholz *et al* (2008), although the causes of these variations are still unknown. Nabholz *et al* (2009) suggested that substitution rate is positively correlated with metabolic rate and negatively correlated with longevity.

Genetic variation within the mitochondrial genome was thought to be selectively neutral. This and other advantages associated with this marker have made it a very popular marker for population genetic studies. However, in recent years research has shown that mitochondrial genome is being shaped by selection, similarly to the nuclear genome (Dowling *et al* 2008) and that mitochondrial DNA also has a role in determining fitness. As a result of these issues, the use of mitochondrial DNA as a genetic marker in phylogenetic studies have been a constant source of discussion in recent years and due to the dilemmas presented, these debates about the suitability of mtDNA in genetic studies will continue (Baker *et al* 2009; Ballard 2000).

1.2.2.2. Microsatellite markers

Microsatellites are one of the most widely used genetic markers, consisting of tandem repeats of 1-6 nucleotides, also known as short tandem repeats (STR). The length of a microsatellite locus is between 5 and 40 repeats and di-, tri- or tetranucleotide repeats are the most common in mammalian genomes (Wright 1994; Lee *et al* 1999). In coding regions tri-, or hexanucleotide repeats are more common as they do not cause a frameshift. In microsatellites found in coding DNA, selection against frameshift mutations operates, which effectively prevents anything other than trinucleotide (or hexanucleotide) repeats. However, most of the satellite DNA is found in non-coding regions of the genome, so they are more likely to be neutral with respect to natural selection than coding regions and are thought to have evolved neutrally. Microsatellites show codominance, have high mutation rates (around the order of 10^{-2} to 10^{-6}, per locus/per generation (Hancock 1998) that generate high levels of allelic diversity and polymorphism, which makes microsatellites very suitable for the study of evolutionary processes (Selkoe and Toonen 2006). As most alleles are relatively rare, individuals tend to be mainly heterozygous at the majority of loci and this makes it a well-suited marker for looking at genetic structure in fish as they are characterised by high levels of heterozygosity and polymorphism, as well as high mutation rates and selective neutrality (Selkoe and Toonen 2006; Xiao-Gu *et al* 2006). Due to these characteristics, microsatellite loci have been used to investigate population structure when protein-coding loci and/or mtDNA lack sufficient resolution to reveal fine-

scale genetic differentiation (Hughes and Queller 1993; Forbes *et al* 1995; Tessier *et al* 1995). They are usually found to be variable even within small populations and microsatellites have been an important tool in both population genetic studies and in research into human diseases. While most microsatellite loci do not code for proteins, in some cases, a large number of repeats within a microsatellite sequence have been associated with various human diseases, such as Fragile-X syndrome (Jin and Warren 2000; Zhong *et al* 1999).

Classification of microsatellites is possible based on their association with coding sequences as this is related to the selective forces operating and the mutational rate. Most microsatellites are in non-coding DNA and are thought to evolve neutrally, whereas in coding DNA frameshift mutations are selected against – this stops the expansion of everything other than trinucleotide repeats (Ellegren 2004). The variability of these markers derives from differences in length rather than variation in the actual sequence. Microsatellite repeat sequences mutate by slippage, during DNA replication that can change the number of repeats (Dieringer and Schlötterer 2003). Two models have been proposed to explain the high rate of mutation found in microsatellite DNA: DNA slippage that involves slip-strand mispairing errors during the DNA replication and unequal recombination between DNA molecules (Schlötterer and Tauz 1992; Eisen 1998).

1.2.3 Developments in population genetics

As a discipline population genetics has undergone significant changes in the past few decades. The introduction of DNA sequencing technologies makes it possible to gather large amounts of data in a relatively easy and quick way (Wakeley 2004). Analysis of multiple loci make it possible to describe genome-wide patterns of SNP and haplotype variation, population structure and genomic diversity.

Until recently, mostly neutral markers (microsatellites, AFLP) were used to assess genetic diversity but recent advances in genomics studies make it possible to compare the genomes of different organisms that are closely related but live in different environments (Ouborg *et al* 2010). Functional genomics includes the analysis of large datasets, one such analysis is gene expression analysis, the study of the expression of various genes under certain conditions, made possible by microarray technology.

Advances in computational and analytical techniques, such as high throughput sequencing of DNA (Fahlgren *et al* 2007, Huang *et al* 2011) allows us to sequence entire genomes, and microarray analyses of gene expression make it possible to study gene expression.

The incorporation of genomic and transcriptional approaches and bioinformatics tools make it possible to understand what proteins are particularly important in understanding the role of stress, pollution or other environmental factors and what mutations in an organism enable it to survive in the stressed environment. Expressed sequence tags (EST) are used to identify candidate genes associated with various traits. Genomic

resources are important for evolutionary studies and these genetic and population genomic methods are used within an ecological framework.

The large datasets and the computational tools available have promoted the use of integrative approaches, such as the combination of population genomics and quantitative genetics (Pauwels *et al* 2008).

After the completion of the Human Genome Project in 2003 (IHGSC 2004) researchers started looking at variation among individuals and started the HapMap Project to map single nucleotide polymorphisms (SNP). SNPs are a suitable marker to map genes, and existing technologies make it possible to simultaneously analyse thousands of SNPs, gathering large amounts of data (Sanchez *et al* 2009). They are used to understand what variations are most common in certain populations and geographic areas as well as explain disease risks and by the application of genome-wide association studies (GWAS), it is possible to pinpoint which SNPs occur together with certain diseases.

Genomic data can support various areas of biological sciences, invasion biology, as well as medicine, so new areas, such as metagenomics makes it possible to study DNA directly from environmental samples and so gain a better and wider understanding of the structure of microbial communities on a much larger scale, while SNPs and GWAS can be used to track the inheritance of disease genes.

1.3. Study species

1.3.1 Rainbow trout (*Oncorhynchus mykiss*) as a study species

1.3.1.1 Life history of the species

The family Salmonidae contains about 175 species and includes the subfamilies of Coregoninae (whitefishes, ciscos and inconnu), Thymallinae (graylings) and Salmoninae (trouts, salmons and charrs) (Behnke 2002). The family Salmonidae contains about 75 extant species that are native throughout the Northern hemisphere (Crisp 2000).

Inland from the Pacific Basin, the appearance of salmonid-like fishes was thought to occur with the rise of the Rocky Mountains during the Pleistocene ice ages that created the cold mountain streams of the area (Cannings 2007). Eventually, the ancestors of modern trout travelled up the Columbia River and its largest tributary, the Snake River about a million years ago and about 10,000 years ago to the basins of the Platte, Arkansas rivers and the Rio Grande (Roberts 2001).

Oncorhynchus mykiss (Walbaum, 1972) occurs in two ecotypes

a. anadromous form, emerging in rivers and covering hundreds of kilometres migrating between the ocean and returning to the rivers only for spawning, known as steelhead and

b. non-anadromous form, the rainbow trout

but none of the species is solely marine (Augerot 2005).

Anadromy is an innate instinct that in various populations appears to different degrees, although in many anadromous populations some individuals or even whole breeding stocks refrain from migrating, spending

their whole life in freshwater (McDowall 2001). There are costs associated with being anadromous, such as greater mortality from predation during migration, or the greater osmoregulatory demands on the organism. On the other hand, it is accepted as a trait with selective advantages, through access to better food resources or beneficial environmental conditions, resulting in larger and more fecund individuals. Therefore, anadromy continues to survive as the benefits associated with it outweigh the costs (McDowall 2001).

The native region of rainbow trout is the pacific coast of North America (MacCrimmon 1971). As the species has been widely distributed all over the world, several local forms have developed, as well as several different strains through mass selection and cross-breeding (Pillay and Kutty 2005). The species is found in fast moving waters, cold-water streams with rocky bottoms. Rainbow trout has a fusiform body shape, the back, sides, head and fins of the fish are covered in small black spots, but the colouration can change with sexual condition, as lake residents have brighter colours and are more silvery and stream residents are darker (FAO 2012). Rainbow trout is a resilient and tolerant fish, able to occupy many different habitats and to withstand a range of temperatures, but spawning mainly occurs at a narrower range, normally around 9-14 degrees Celsius (Matthews and Berg 1997). *Oncorhynchus mykiss* is a seasonal spawner, the spawning season lasts from two to four months between autumn and spring (Roberts 2001). They tend to choose different gravel locations to spawn when compared to

wild brown trout that prefer the typical clam shell of fluffy gravel at the lift to the riffle in a pool/riffle sequence. (pers. comm. Stephen Moores).

The species has a relatively long reproductive cycle, lasting between 2-3 years. The female produces about 2000 mature ova per kilogram bodyweight, the eggs measure around 3-7mm, these hatch about 4-7 weeks post fertilisation, depending on water temperature, e.g. eggs will hatch within 30 days at 10°C (Springate 1984). The optimum temperatures for growth are between 12°C and 18°C as higher temperatures may cause considerable egg mortality (Lucas and Southgate 2003). The female does not guard her eggs, though it will cover them over with gravel (Roberts 2001). This always happens in flowing river water or in lake near a feeder river or a stream leading into the lake to provide oxygen rich water for the eggs as trout eggs have a relatively thick membrane, which requires oxygen-rich water to keep them alive, so slow moving streams make a poor breeding ground for trout (Stenberg 1988). Depending on temperature and food availability, maturation usually occurs at two or three years (Su *et al* 1999), though males usually mature earlier than females at 2 years old as the majority of females mature at 3 years old (Pillay and Kutty 2005).

In the wild, rainbow trout are opportunistic feeders, adult trout feed on aquatic and terrestrial insects, like mayflies, caddisflies, molluscs, crustaceans, fish eggs (particularly during spawning season), minnows, and other small fishes, but the most important food is freshwater shrimp, which contains the carotenoid pigments responsible for the orange-pink colour in the flesh (Bjerkeng 2000).

The rainbow trout is an intensively studied species in genetic, ecological and physiological research. As a result of the large-scale cultivation of the species, a substantial amount of information is available of the species' biology, which makes it a suitable species to address questions in toxicology, immunology and genomics (Thoorgard *et al* 2002).

Also, populations in different watersheds can have independent evolutionary origins, providing a suitable replication of adaptive patterns (Huey 2005) and after the initial introductions several new populations formed, presenting an opportunity for comparison between these new populations and native populations of salmonids.

1.3.1.2 The evolutionary history of the rainbow trout

The evolutionary history of *O. mykiss* (Figure 3) has presented biologists with an exciting research topic for a long time with regard to the on-going debate about the possible marine or freshwater origin of salmonid fishes.

Salmonids are believed to share an ancestry with Osmeridae and Retropinnidae (McDowall 2002) as all members of these families are diadromous and diadromy is an ancestral character state (McDowall 1997).

The evolutionary history of Salmonidae begins with the origin and diversification of bony fishes in the Devonian period (408 to 380 million years ago), during the 'Age of Fishes' (Behnke, 2002). The relationship of this family to other families is not well understood, and reconstruction of the family is further complicated by hybridization and introgression (Weigel 2003; Young *et al* 2001).

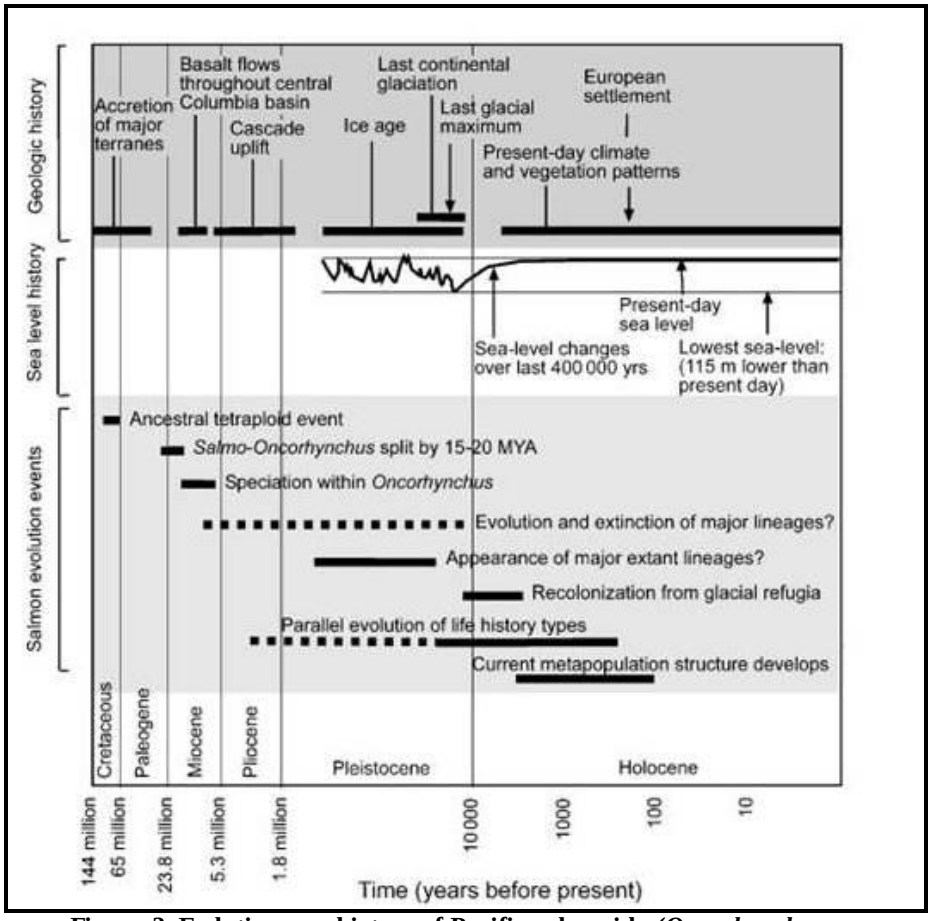


Figure 3. Evolutionary history of Pacific salmonids (*Oncorhynchus* species), adapted from Waples *et al* 2008

Previous research has shown that that salmonid fishes have undergone a whole genome duplication event 50-100 million years ago, and this event aided the rapid adaptive evolution that took place (Verspoor *et al* 2007) in the family. The earliest known salmonid fossil, *Eosalmo driftwoodensis* originates from the Eocene epoch about 37-55 million years ago and was found in the Pacific Northwest (Sutterby and Greenhalgh 2005). Initially, rainbows at this time were restricted south of the Columbia River, and about a million years ago the species spread into the basins of the Platte, Arkansas and the Rio Grande rivers (Behnke 2002) as the Pleistocene glaciations concealed large landmasses under ice. The most recent glaciation event, the

Wisconsian took place 23,000-18,000 years ago and was followed by a deglaciation 15,000-8,000 years ago. After the glaciation, large lakes were formed, which greatly aided the spread of freshwater species. The indirect impacts of glaciations were a significant impact on the genetic structure of both freshwater and anadromous species. When the glacial retreated, fish were able to move from glacial refugias and inhabit previously disused habitats. Geographical distribution of species today has been affected by glaciations during the Pleistocene period, which restricted the distribution of both terrestrial and aquatic fauna to southern refugia. The differences in these habitats have enabled trout to adapt to the specific conditions of their particular habitat and to develop differing morphological characteristics, such as colouring, spawning time, temperature tolerance (Bagley 1998).

McCusker *et al* (2000) reviewed the origin of rainbow trout in North America using mtDNA to understand the species' expansion throughout North America. His results showed that the source of rainbow and steelhead (the anadromous form) trout is California and found more variation in the Californian samples than in samples from British Columbia and Alaska. This is not unexpected, as when comparing fish species in glaciated and non-glaciated areas, we find that species that were displaced show much lower levels of genetic diversity than non-displaced species as a result of the species' restriction to glacial refugia.

1.3.1.3 Genetics of the rainbow trout

Polyploidy has been important in the evolution of both animal and plant species (Soltis and Soltis 1999). Salmonid fishes have long been thought to be derived from polyploidy evolution, and this view is supported by genome size estimates and studies of chromosome counts (Gharbi 2006). Allendorf and Thorgaard's study in 1984 showed that polyploidy in salmonids is a result of an intraspecific whole-genome duplication event about 25-100MYA, which produced duplicate genes allowing them to develop new functions and hence increasing phenotypic plasticity (Allendorf and Thorgaard 1984). After this event, a rediploidization took place, which reduced chromosome numbers for some trout populations and resulted in chromosome number variation of 58-64 chromosomes (Thorgaard 1983). Genome duplication is thought to be the most important evolutionary force since the emergence of the universal common ancestor (Ohno 1970). Rainbow trout is one of the most widely used study species in the world for research focusing on carcinogenesis, disease ecology, evolutionary genetics (Palti 2009). Estimates of genome size for salmonids are thought to be between 2.4 to 3.0 x 10^9 bp, which is about twice the size of the genome of related fishes.

1.3.1.4 Aquaculture

Aquaculture has been around for thousands of years, in Europe the farming of fish started in the Middle Ages and to this day aquaculture has a great significance in socio-economic development (Pillay and Kutty 2005).

Commercial trout farming was first introduced in Denmark about 50 years ago. Monoculture is the practice mostly used in trout culture, and many farms develop their own strain of fish with specific characteristics, such as quicker growth, early or late maturity, larger egg size (Sigler and Sigler, 1990). Trout production is similar to restocking, the fish are raised and stocked in waters, where anglers are allowed to fish them. Triploid fish are also frequently used in aquaculture, these fish are also infertile, which eradicates the risk of cross-breeding with wild stocks (Dunham 2004) and as there is no reproductive effort required, they grow larger in size as well. Only females are used in triploid stocking, as males may still develop functional gonad tissue. Developing eggs are heat treated during meiosis, which causes the development of three sets of chromosomes (Crozier and Moffatt 1989). Crossing tetraploid and diploid fish is another way of producing triploid fish, if the use of fertile diploid fish is unacceptable (Sheehan *et al* 1999). Reduced gonadal growth allows triploid fish to allocate resources to somatic growth but generally it is difficult to differentiate between triploid and diploid fish visually.

1.3.1.5 The Hofer strain

Chapter 4 explores the population origin of the rainbow trout and the role of the Hofer strain. The Hofer strain originates from the Hofer Farm in Germany, which started importing rainbow trout eggs in the 1880s and anecdotal evidence tells us that eggs from the Gunnison River in Colorado were shipped to Germany. This strain is thought to have contributed to the

existing European populations and the limited establishment success observed in *O. mykiss* is thought to be a result of the species' susceptibility to whirling disease. Whirling disease is a disease of the brown trout (*Salmo trutta*) and is caused by the myxozoan parasite *Myxobolus cerebralis*, resulting in severe problems for rainbow trout populations. The disease was first described in Europe at the end of the 19th century and was then imported to the United States in the 1950s. As the disease mainly attacks fish when they are young, if a population is susceptible to the disease, and if no fry survive, the population will not survive. However, it is also important that fish retain some of the wild characteristics to enable them to survive in rivers and streams and to be able to establish self-sustaining populations in the wild. As the Hofer strain has been domesticated, they lost certain wild characteristics, such as being able to escape from predators, as well as the ability to find food and lay the eggs in a safe way to help them to survive (Schisler 2009).

1.3.1.6 The Wye Population

The population of rainbow trout in the Wye River in Derbyshire is the only self-sustaining population of rainbow in the United Kingdom. The species was initially introduced into the UK to Howietiun Hatchery in Scotland in 1884 (Palmer 1996) but this consignment of fish failed to establish. A second delivery arrived a year later from the United States, and after this until 1905 a shipment of rainbow ova arrived from the U.S. Bureau of Fisheries every year. According to this Bureau, these fish were *Salmo*

Shasta, originating from the Shasta Mountain area of California and subsequent shipments comprised of rainbow trout from the McCloud River area of California, with these two subspecies hybridising. There is no exact date for the arrival of the Wye population but during a 1941 survey the population already existed (Palmer 1996).

1.3.2 Topmouth gudgeon (*Pseudorasbora parva*) as a study species

1.3.2.1 Life history of P. parva

Topmouth gudgeon (*Pseudorasbora parva)* (Temminck and Schlegel 1846) is a cyprinid fish, originating from Asia. Cyprinidae are among the most widely spread freshwater fish families in the world (Dubut *et al* 2010), which together with various sociological and economic factors fuelled their rapid spread. In appearance the males are darker than the females. *P. parva* reaches early sexual maturity (1 year) and can breed up to four times a year. Other life characteristics, such as its small size and quick dispersal ability, as well as being a batch spawner (between April and August) and being able to adapt to and tolerate different environments (Pinder and Gozlan 2003) all contribute to the colonisation success of the species. Topmouth gudgeon grow to about 8 cm in length and occupy the bentho-pelagic zone. Its diet consists of algae, zooplankton, and the eggs of other fish, as well as feeding on terrestrial insects. *P. parva* is able to spawn on any surface although they show a preference for cavities as they are easier to protect from other fish.

The species was inadvertently introduced into European waters in Romania in the 1960s, probably via ornamental fish trade escapes or releases (Gozlan

et al 2002), originating from Japan and China. Since then it had rapidly spread all over Europe (Figure 4) and in the past 40 years it has colonised most European waters, in the UK it was first reported in 1996, when the species became established through connected river networks and as a result of accidental and intentional human means (Pinder 2003).

Topmouth gudgeon are now found in at least 32 countries with very diverse climates (e.g. Algeria, Austria, Poland, and Spain) and has been able to invade habitats with a wide range of ecological conditions. *P. parva* has had severe impacts on native fish through inter-specific competition by out-competing native species, aided by its life history characteristics (Pinder *et al* 2005). Previous research has mainly concentrated on these aspects of invasion, but facultative parasitism and pathogen transfer is a very important area of research on the species, as topmouth gudgeon is also a healthy carrier of the RLA (rosette-like agent) (Gozlan 2009), a pathogen similar to *Sphaerotheca destruens*, which affects North American salmonids.

1.3.2.2 Control of the species

Control of the species is extremely difficult given its size and the security of the fisheries sites known to contain the species, as the mechanical screening devices can only prevent the escape of larger fish. Also, movement of fish between fisheries has also contributed to unintentional stocking together with other fish species (Pinder and Gozlan 2003). As a result, *P. parva* has been classified as an international pest species (Pinder 2005). In England, the fisheries where topmouth gudgeon was found were treated with a

rotenone-based piscicide after removing all other fish (Britton *et al* 2007). This, together with other measures, such as preventing dispersal was somewhat effective in removing the species from these waters. However, as the pesticide is not host-specific, this method carries high risk. These techniques have a low success rate and together with poor legislative background and public opposition (Britton *et al* 2007) the eradication of topmouth gudgeon has been extremely difficult. The broad range of habitats in which *P. parva* is found and the species' high dispersal capacity highlights the considerable threat this species poses to endemic fish in Europe. The control of invasive species is very often difficult, and depending on the species, elimination should be instigated as soon as possible, as after the species has become established and shown adaptation to its novel environment, eradication becomes strenuous and often impossible, and often after the initial establishment of the species, secondary introductions take place. In some waters, however, topmouth gudgeon are co-habiting with other fish causing no obvious decline - such a marked geographic difference in impact suggests that differences in competitive ability or parasite load exist (Gozlan *et al* 2005).

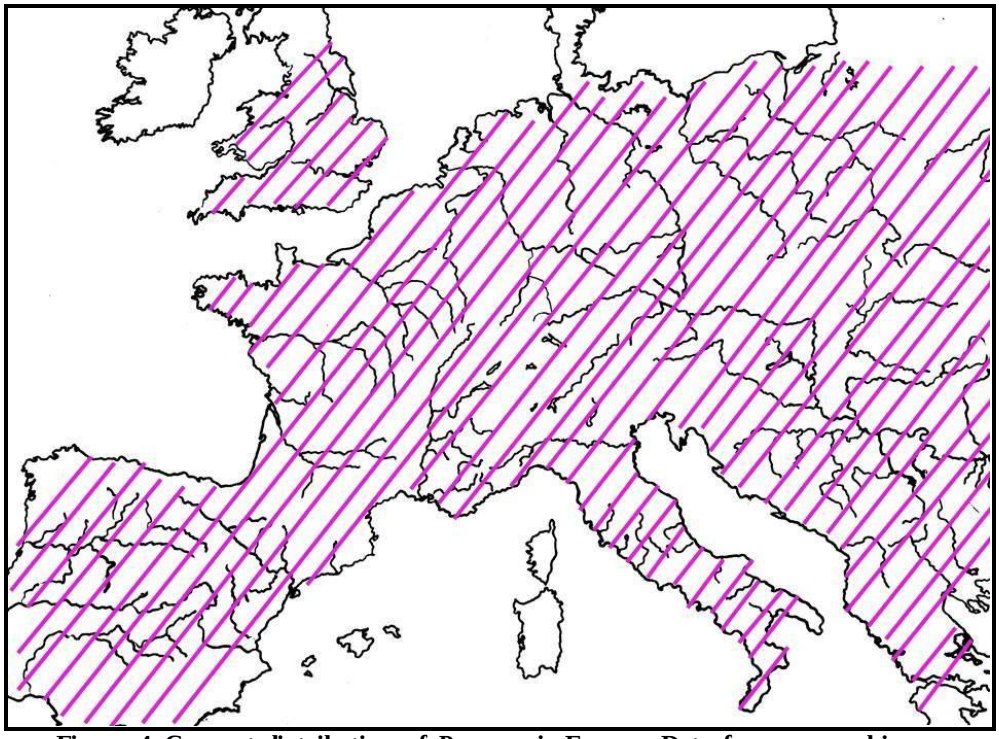


Figure 4. Current distribution of *P. parva* in Europe. Data from www.cabi.org and www.habitas.org.uk

1.4. Aims and Objectives

The overall objective of this PhD thesis was to understand the evolutionary aspects, population genetics and introduction history of two invasive fish species, the topmouth gudgeon (*P. parva*) and the rainbow trout (*O. mykiss*). *P. parva* was introduced into Europe in the 1950s and has since spread to over 30 countries on the continent. *O. mykiss* is one of the most popular fish species used in aquaculture, still, despite repeated introductions; the species has only been able to become established in a few European waters. I wanted to understand the genetic factors that influence the establishment success or failure of non-native fish species.

The findings of this thesis will have implications on further studies of invasive species, evolutionary biology and the understanding of genetic factors that play a role in enabling species to invade new areas.

The thesis is organised into five chapters. The first chapter is the general introduction of the project, the study species, the genetic markers used, as well as the objectives of the thesis.

Chapter two presents and discusses the population genetic analysis of topmouth gudgeon using mitochondrial DNA across the native and invasive range of the species. The analyses are carried out in order to test specific hypotheses regarding the colonisation models that shape the genetic structure of *P. parva* and to determine the levels of genetic variation found across the invasive range of the species. Mitochondrial DNA provides an objective framework for the analysis of the link between biogeographic patterns and evolutionary processes by understanding the invasion history of topmouth gudgeon (*P. parva*).

In Chapter three I further investigated the introduction and spread of *P. parva* and the evolutionary factors that helped the species to establish in Europe by applying microsatellite makers to assess fine-scale diversity and to interpret current phylogeographic patterns found in the species to test the hypothesis that a single introduction event took place in the invasive range of the species.

The fourth chapter explores the population origin of the rainbow trout and my main objective was to find the source populations that contributed to the European populations and to understand the introduction history and the

phylogeographic patterns found in the introduced range of the species. This information is essential towards the development of appropriate management strategies for the conservation of the species. The objective of this chapter was to elucidate contemporary patterns of movement between European populations.

Chapter five presents a synthesis of the findings and discusses the evolutionary aspects of fish invasion.

The thesis concludes with the Reference list and the Appendices. Chapter two of the thesis has been published in PLoS One, as is indicated on the last page of the chapter.

CHAPTER II

Invasive cyprinid fish in Europe originate from the single introduction of an admixed source population followed by a complex pattern of spread

2.1 Introduction

Population genetic studies of invasive species have become an instrumental component in the study of biological invasions (Geller *et al* 2010; Le Roux *et al* 2009; Hänfling 2007). The application of neutral molecular markers can elucidate demographic processes during the invasion process and identify colonization pathways and source populations (Guillemaud 2010; Muirhead *et al* 2008). Such information not only facilitates management and prevention of further invasions but also provides a framework for studies on adaptive evolution during the invasion process (Lee 2002). An issue which has recently received much attention but remains poorly understood is the role of genetic diversity in determining the outcome of introductions of non-native species. Introductions of non-native species are often based on the release of a low number of founding propagules containing only a fraction of the genetic variation of the source populations (Brown and Stepien 2008). Such reduced genetic diversity theoretically limits a species' ability to establish invasive populations invoking a 'genetic paradox' (Roman and Darling 2007, Nei *et al* 1975; Williamson 1996; Frankham 2005; Poulin *et al* 2005; Ficetola *et al* 2008), asking how invasive populations overcome this reduced genetic diversity and become adapted to their new environment?

Although many successful invasive species show reduced genetic diversity, recent research suggests that the effects of such bottlenecks are often counteracted by admixture among genetically divergent source populations (Hänfling 2007; Roman and Darling 2007). For example, multiple

introductions have resulted in high genetic diversity of invasive crustaceans (Kelly *et al* 2006), fish (Hänfling 2007; Haynes *et al* 2009; Zidana 2009), lizards (Kolbe 2004) and plants (Rosenthal 2008). Nevertheless, it is currently unknown whether such admixture is merely a side-effect of the invasion process or is actually facilitating the establishment process. Additional population genetic case studies, in combination with studies on ecologically significant traits are crucial in providing answers to this question.

One of the most compelling fish invasions in the world today is arguably the topmouth gudgeon *Pseudorasbora parva* (Temminck and Schlegel 1846). This small cyprinid species originating from East Asia was accidentally introduced into Europe in the 1960s in several countries around the Black Sea as part of contingents of Chinese carps for aquaculture (Gozlan *et al* 2002; Gozlan *et al* 2010a). Since then, they have proved highly invasive through a combination of sociological, economic and ecological factors that enabled their rapid human-assisted and natural dispersal throughout the continent. On introduction into a new water body, colonisation is facilitated by their tolerance of degraded aquatic ecosystems and their reproductive traits of early sexual maturity, batch spawning, high reproductive effort and paternal nest guarding that provide a high degree of invasive vigour (Gozlan *et al* 2010a; Rosecchi *et al* 2001; Gozlan *et al* 2010b). Their capacity for subsequently forming high density populations can then result in sharing of common food resources with native fishes resulting in overlaps in trophic niche (Britton *et al* 2010), with additional concerns over egg predation,

disease transmission and facultative parasitism (Gozlan *et al* 2010b). Whilst this *P. parva* invasion has been traced from the initial point of introduction towards the northern and western parts of Europe, as well as the south towards Turkey and Iran (Gozlan *et al* 2010b), its exact demographic scenario is currently unclear. They are now found in at least 32 countries with contrasting climates (e.g. Algeria, Austria, Poland, Spain), have invaded habitats with a wide range of ecological conditions and their life history traits differ considerably among invasive populations (Gozlan *et al* 2010b). Possible (non-mutually exclusive) explanations of such variability are: (1) the existence of considerable phenotypic plasticity in life history traits and tolerance to environmental conditions, (2) a rapid evolutionary response, or (3) multiple independent introductions from divergent source populations (Gozlan *et al* 2002; Gozlan *et al* 2010b; Britton *et al* 2010). Molecular markers have previously been employed to study such questions in other freshwater fish invasions in Europe (Benejam *et al* 2005; Vidal *et al* 2010) and North America (Brown and Stepien 2008). For example, using mitochondrial DNA, Vidal *et al* (2010) showed that the mosquitofish (*Gambusia holbrooki*) was introduced into Europe multiple times from USA. Some *P. parva* populations have also been identified as healthy carriers of pathogens, such as *Anguillicola crassus* (Cesco *et al* 2001) and the rosette agent *Sphaerothecum destruens* (Gozlan *et al* 2005, Gozlan *et al* 2009). It is currently unknown whether other invasive populations or native populations show a similarly low susceptibility to the rosette agent.

Consequently, *P. parva* appear to be a model fish well suited to studying the evolution of ecologically significant traits, disease resistance and the role of genetic diversity in establishment success. Thus, we perform a population genetic analysis of *P. parva* across their native and introduced ranges in order to test different models of colonisation and to determine levels of genetic variation across the invasive range of the species (see Material and Methods for specific hypothesis). This will provide a first population genetic framework for further evolutionary studies on the species.

2.2 Materials and methods

2.2.1 Sampling scheme and hypothesis testing

Samples were collected at a total of 22 sites, 14 in Europe and 8 in Asia (Table 1; Figure 1). Sample size was 15 for the majority of sites with the exception of three sites where 6-10 individuals where sampled. There was also a single sample (Japan) that comprised three individuals; it was excluded from all population-based analyses. The native range of the species is the East Asian sub-region, including the basins of the Huang He, Yangtze, Hai He and Amur Rivers, as well as some Japanese islands, Taiwan and the southern part of Korea (Berg 1949, Bănărescu 1999) and the sampling scheme covers most of the latitudinal space in this range, as well as spanning across the largest part of the European invasive range.

The density of the coverage in the native range was appropriate to test some general demographic processes but not the identification of the exact location of potential source populations.

Code	Population in China	N	Geographical coordinates	
CG	Guangdong, River, Zhuijang River basin, China	6	23° 07′ 53″ N	113° 15′ 59″ E
CH	Huairou Reservior, Hai He River basin, China	15	40° 18′ 46″ N	116° 36′ 36″ E
CK	Kinmen Island population, China	6	24° 26′ 11″ N	118° 21′ 27″ E
CRH	River Hai He, Hai He River basin, China	15	39° 07′ 15″ N	117° 12′ 54″ E
CY	Wuhan, Yangtze River Basin, China	10	29° 58′ 20″ N	113° 53′ 29″ E
JB	Lake Biwa, Yodo River basin, Japan	3	32° 20′ 44″ N	136° 10′ 15″ E
TI	I-lan county, I-lan River, Lanyang River Basin. Taiwan	15	24° 45′ 00″ N	121° 45′ 00″ E
TT	Dajia River, Taichung county, Dajia River basin, Taiwan	15	23° 09′ 00″ N	120° 38′ 34″ E
European Populations				
BS	Slangebeek nean Hasselt, Belgium	15	50° 55′ 48″ N	05° 15′ 00″ E
EB	Byland Abbey, Yorkshire, UK	15	54° 12′ 10″ N	01° 09′ 35″ W
FG	Grand Lieu, France	15	47° 05′ 45″ N	01° 43′ 46″ W
G	River Ammer, Wielenbach, Germany	15	47° 52′ 11″ N	11° 09′ 00″ E
HA	Hortobagy, Hungary	15	47° 36′ 00″ N	21° 06′ 00″ E
HE	Ederecsi-patak, Hungary	15	46° 48′ 04″ N	17° 23′ 16″ E
HG	Gic, Hungary	15	47° 25′ 32″ N	17° 44′ 44″ E
HS	Salyi-patak, Hungary	15	47° 56′ 06″ N	20° 39′ 58″ E
IN	Nestore, Italy	15	43° 21′ 14″ N	12° 14′ 10″ E
PU	Utrata River, Poland	15	50° 35′ 50″ N	18° 09′ 32″ E
SC	Vrakuna, Slovakia	15	47° 49′ 24″ N	18° 49′ 16″ E
SE	Ebro Basin, Spain	15	40° 43′ 12″ N	00° 51′ 47″ E
SWS	Sylen Lake, Llanelli, South Wales, UK	15	51° 40′ 42″ N	04° 09′ 47″ W
T	Blanice River, Vodnany, Czech Republic	15	49° 08′ 52″ N	14° 10′ 32″ E

Table 1. Sample locations and sample sizes and geographical coordinates for native and invasive populations

Thus, the aim was to test three non-mutually exclusive models that were proposed to explain the spread of *P. parva* in Europe: i) 'multiple source-sink' model where several independent introduction events from genetically differentiated native source populations to separate European locations would have occurred without involving admixture; ii) 'stepping-stone' model (Gozlan *et al* 2002) where introduction into a single geographical area would have been followed by gradual expansion from the original introduction; and iii) ‘long-distance’ model (Gozlan *et al* 2010b) where

introduction into a geographical area would have been followed by long-distance translocation within Europe. Furthermore, it was tested whether iv) the invasive populations show signs of a genetic bottleneck or v) might have resulted from an admixture between divergent source populations.

Population genetic theory predicts that these demographic processes will result in different patterns of genetic population structure and therefore molecular approaches can be used to test the likelihood of alternative models. A number of phylogenetic and population genetic analyses were carried out in order to test the results against the theoretical expectations for the scenarios outlined above. Note that some of these tests assume that a relatively clear phylogeographic subdivision exists in the native range. Therefore the first step was to carry out a network analysis in order test this assumption. Genetic distances and *F*-statistics were used to quantify the degree of differentiation between populations and nucleotide diversity, and haplotype diversity at a standard sample size was used to estimate within population variability. These analyses were complemented by coalescent simulations and a Bayesian estimation of effective population size. The results were then compared with theoretical expectations from the various models and scenarios:

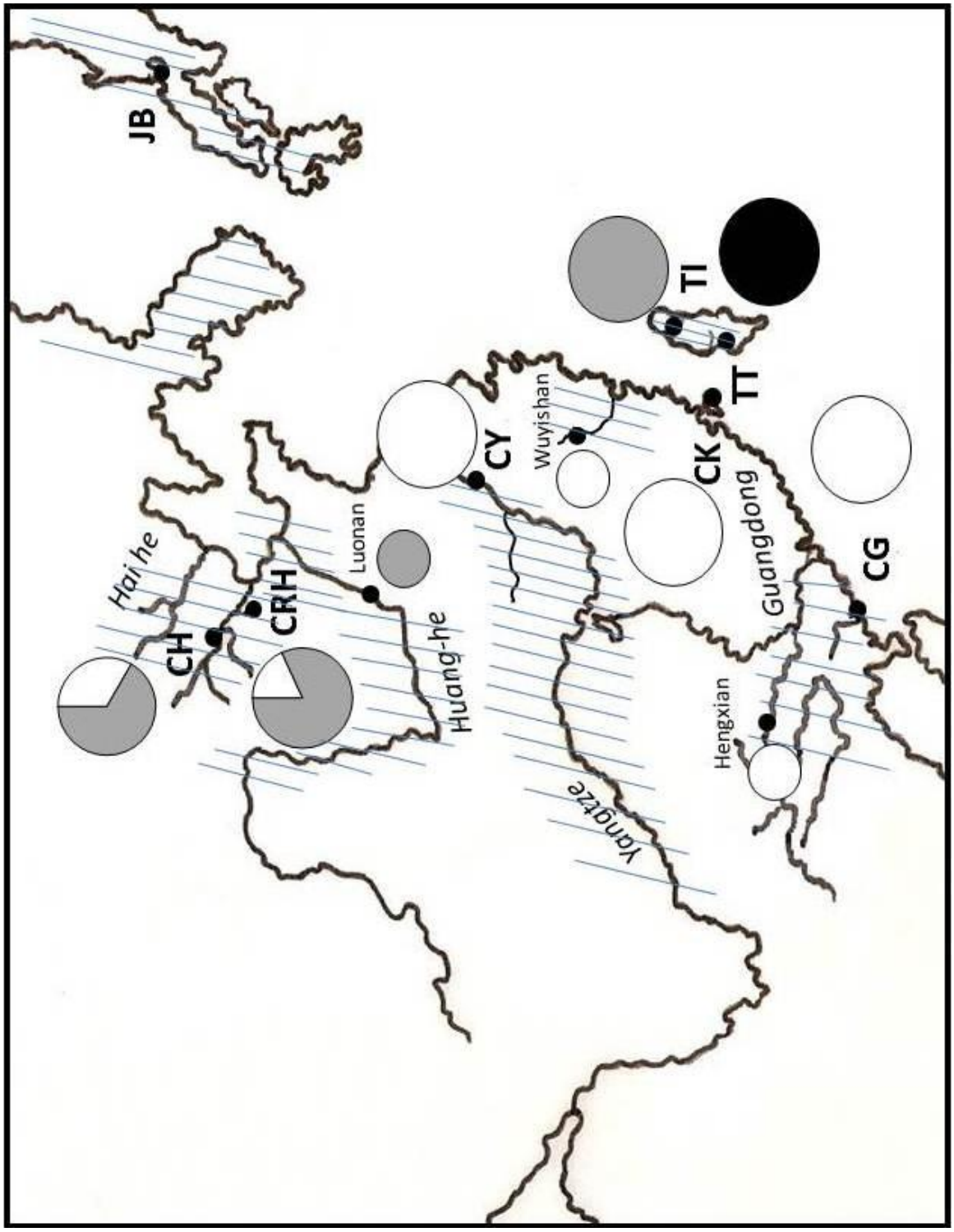


Figure 1a. Distribution of Pseudorasbora parva samples sites in Europe (Figure 1a) in the native range (Figure 1b). Pie charts represent the geographical distribution of major mtDNA lineages (see Figure 2).Lineage 1 = white, Lineage 2 = black, lineage 3= grey. See Table 1. for population codes. Large pie charts represent samples collected in this study, small pie charts samples from Liu *et al* 2010

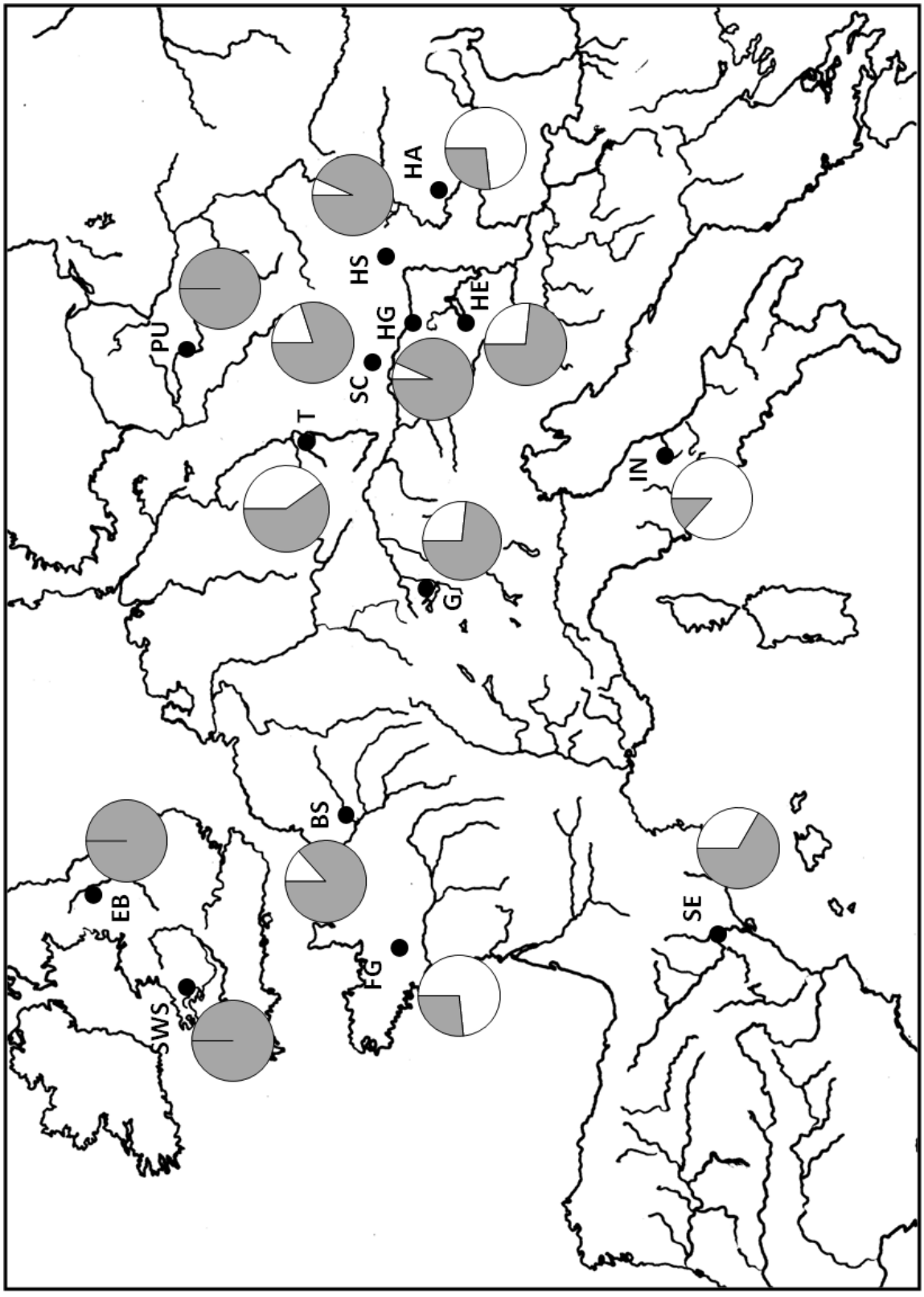


Figure 1b.

i) ‘multiple-source-sink’ model: genetic differentiation among invasive populations is high and similar to that found in the native range;

ii) ‘stepping stone’ model: genetic differentiation in the invasive range is lower than that in the native range, and there is a significant pattern of isolation-by-distance;

iii) ‘long-distance’ model: genetic differentiation in the invasive range is lower than that in the native range, and there is no pattern of isolation-by-distance;

iv) ‘genetic bottleneck’ scenario: genetic diversity of invasive populations, in particular haplotype diversity, is lower than that of the source populations; and

v) ‘genetic admixture’: genetic scenario: genetic variation expressed in nucleotide diversity is higher than that of the source population. Furthermore, recent admixture increases the nucleotide diversity above that expected under equilibrium conditions.

2.2.2 Molecular procedures

The fish were collected and stored in 98% ethanol. Genomic DNA was extracted from the caudal fin tissue using the HotShot method (Truett *et al* 2000). An approximately 700bp long section of the mtDNA genome, containing the partial *cytochrome b* gene was amplified applying standard PCR techniques using Verity Thermal Cycler. Primers L15267 and_H15891Ph, previously described by Briolay *et al* (1998), were used. Thermal cycle amplifications were performed in 15 μL reactions, containing 1.5 μL 160 mM NH_4, 1.5 μL 100 mM dNTPs, 0.4 μL 50 mM $MgCl_2$, 0.075 μL *Taq* polymerase, 0.3 μL each of primers L15267 and H15891Ph, 9.425 μL PCR water and 1.5 μL of template DNA. Cycle parameters were as follows: 2 min at 95°C; 45 s at 94°C, 45 s at 48°C, 1 min at 72°C; 10 min at 72°C. PCR products were directly sequenced in both directions using the PCR primers by Macrogen Inc. Forward and reverse sequences were aligned and edited using CodonCode Aligner (Ewing *et al* 1998), (GenBank accession numbers: JF489575-JF489887). Consensus sequences were imported into MEGA v. 4.1 (Tamura *et al* 2007) and aligned with ClustalW (Thompson *et al* 1997).

2.2.3 Phylogenetic analyses and haplotype network

Phylogenetic relationships of haplotypes were reconstructed using the maximum composite likelihood method (Tamura *et al* 2004) in combination with Neighbour-Joining as implemented in MEGA v. 4.1 (Stamatakis *et al* 2008). Furthermore we created a Maximum Likelihood tree, using the

RaxML programme (Stamatakis *et al* 2008) using the GTR model optimised for each codon position. Branch support of both was obtained using non-parametric bootstrapping as percent of 1000 repeats and ML support values over 70% were added to tree nodes. Our aim using the phylogenetic tree approach was to show how distant haplotypes relate to major clades, rather than to provide definite resolution within clades.

In order to increase the geographic coverage, GenBank sequences from five *P. parva* individuals (Liu *et al* 2010) sampled in the Minjiang River at Wuyishan (EU934500), the Pearl River at Hengxian (EU934501 and EU934502) and the Yellow River at Luonan (EU934503 and EU934504) were included in the phylogenetic analysis. Representatives of the main lineages of the cyprinid subfamily Gobioninae according to Tang *et al* (2010) were included as an outgroup using the same GenBank sequences as Tang *et al* (2010).

A haplotype network was constructed using a median-joining algorithm in Network v. 4.5.10 (Bandelt *et al* 1999). Possible homoplastic sites (153, 195, 300, 462, and 585) were weighted down to 1 and all other nucleotide positions were weighted at 50 and we used an ε value of 0. Furthermore, transversions were weighted three times higher than transitions to decrease the likelihood of homoplastic substitutions (Broughton *et al* 2000). A BLAST search of nucleotide sequences (Altschul *et al* 1990) was performed in order to confirm that all sequences belonged to *P. parva*.

2.2.4 Population genetic data analysis

DNaSP v. 4.5 (Rozas *et al* 2003) was used to estimate within population diversity (nucleotide diversity, π; haplotype diversity, Hs). Standardised measures of genetic diversity were calculated by resampling data sets 1000 times using a bootstrapping procedure (Nei and Jin 1989; Lynch and Crease 1990) based on the size of the smallest sample (6 individuals). Differences in genetic diversity between native and invasive populations were tested using a Mann-Whitney test. The invasive population PU was excluded from the comparison of π because it contained one highly divergent haplotype which is suspected to be derived from hybridisation with *Gobio gobio*. Both *P. parva* and *G. gobio* have similar life history traits in generation time, body size, spawning season (Rosecchi 2001) and hybridization has been observed in closely related species in cyprinid fishes (Yang *et al* 2006).

Coalescent based simulations as implemented in DNaSP were used to predict the expected relationship between haplotype diversity (*H*) and nucleotide diversity (π) under drift-mutation equilibrium and constant population size (Hudson, 1990). Effective population size of native populations assuming mutation-drift equilibrium and absence of migration among watersheds was estimated using MIGRATE-n v. 2.5 (Figure 5) (Beerli, 2006). The option Bayesian inference was used with the default search strategy settings. The rationale of this analysis was to estimate the populations size required to maintain the amount of genetic diversity found in the each population assuming mutation-drift-equilibrium.

Pairwise genetic differentiation among samples was computed as F_{ST} and D_{XY} using Kimura two-parameter method (Kimura, 1980) (Table 1, 2, Appendix I, Appendix II) using DNaSP v. 4.5 (Rozas *et al* 2003). A multi-dimensional scaling (MDS) analysis based on F_{ST} was carried out in order to visualise the genetic relationship between samples. The average pairwise differentiation between native populations was compared to the average pairwise differentiation of invasive populations using a Mann-Whitney test. Isolation by distance (IBD) analysis was then used to test whether the 'stepping-stone' model could explain the spread of *P. parva* within Europe. Pairwise geographic distances among European sites were calculated as Euclidean distances. The theoretical expectation is that a significant correlation should only occur under the 'stepping-stone' model (Ramachandran *et al* 2005; Herborg *et al* 2007). Three different approaches were used. First, a 'classical' IBD analysis (Wright 1943) was carried out to test the relationship between matrices of geographical distance and genetic differentiation (F_{ST}) using a Mantel test (1000 permutations) as implemented in the software IBDWS v. 3.16 (Jensen *et al* 2005). The genetic F_{ST} values were log-transformed to achieve a normal distribution. Second, a general linear model (GLM) was used to test the relationship between the geographic distance and genetic differentiation from the putative site of introduction. Third, a GLM was used to test the relationship the geographic distance from the putative site of introduction and genetic diversity of populations. Under a 'stepping stone' model, genetic diversity is expected to decrease with geographic distance to the original site of

introduction, and hence, the genetic distance is expected to increase. The putative site of introduction was Nucet-Dombovita, Romania in the early 1960s (Bănărescu and Nalbant, 1965), however around this time several other introductions took place into Hungary (Gozlan *et al* 2010a), so this population (HA) was used as reference population.

2.2.5 Approximate Bayesian Computation (DIY ABC)

Approximate Bayesian Computation (DIY ABC) (Cornuet *et al* 2008) was used to estimate the relative likelihood of alternative scenarios of the initial introduction of the species into Europe. In the programme, reference tables (containing parameters based on known values) were used to compare the scenarios and the simulated datasets were then compared to the true values (Cornuet *et al* 2008). DIY ABC is a computationally intensive approach and therefore only three simplified scenarios where chosen, which appeared most feasible after the initial population genetic analysis. An explicit rationale for choosing specific models will therefore be given in the Results section. The prior distribution of the coalescence time in the evolutionary scenario was partially informed by historical data, such as the date of the first introduction (Appendix III). The effective population size was set as uniform, 10 and $5x10^4$ individuals. No ecological data exist on the effective population size of topmouth gudgeon or in specific population samples, so priors were chosen covering the full range of biologically feasible values using the Kimura 2 parameters (Kimura 1980) mutation model. For each scenario 10^6 datasets were simulated with the parameter values drawn from

the prior distribution (Appendix III). The relative likelihoods of the three scenarios were compared by using logistic regression on 1% of the closet simulated data sets.

2.3. Results

2.3.1 Phylogenetic and network analysis and distribution of haplotypes

A total of 30 haplotypes were identified using 310 sequences (608bp) from 8 native and 14 introduced populations (Table 1). The phylogenetic relationship among haplotypes is shown in Figure 2. Both NJ and ML methods yielded the same topology, hence only the NJ tree is displayed but ML support values were added to tree nodes. The two Japanese haplotypes, H23 and H24 were closely related to each other and the phylogenetic analysis (Figure 3) showed that they formed a highly divergent sister group to the remaining *P. parva* haplotypes (sequences divergence ~5-6%).

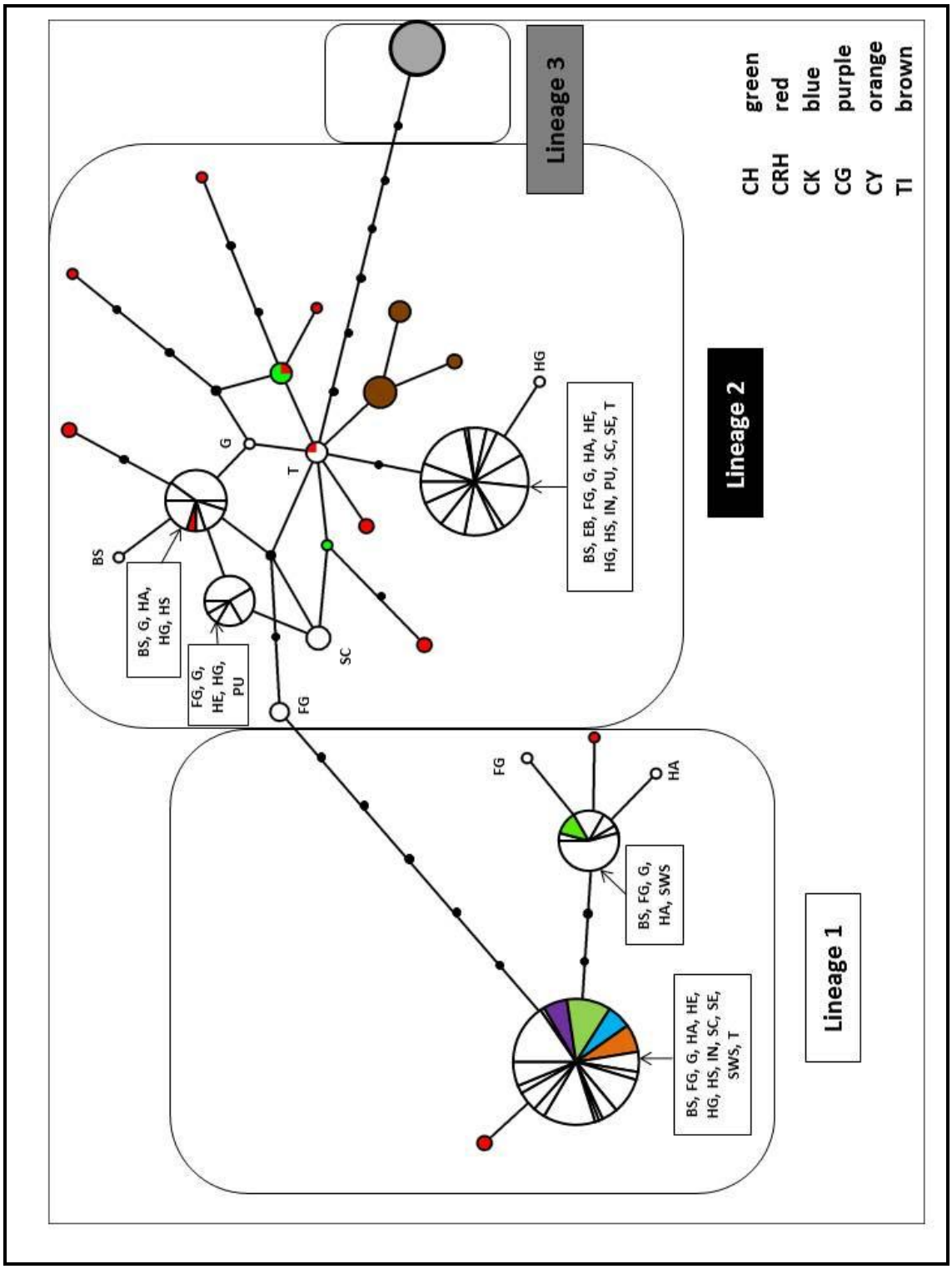


Figure 2. Medium joining network of cytb haplotypes from native and introduced populations of P. parva, excluding H22, H23 and H24. Adjacent haplotypes are connected through a single point mutation. Each circle represent a single haplotype and its diameter is is proportional to the number of individuals with that haplotype. The colour codes represent the locations in which the haplotypeis found, filled cricles (N) represents unsampled haplotypes.

One highly divergent haplotype found in the invasive Polish populations clustered closely to a sequence of *G. gobio*, this was confirmed by a BLAST search of these sequences. This haplotype and the Japanese haplotypes were therefore not included in the network analysis.

Thirteen haplotypes were found in the invasive populations, five of which were found in more than one invasive population and will be subsequently referred to as common haplotypes. Three of the common haplotypes and two of the rare haplotypes were also found in at least one native population.

Three main lineages of *P. parva* haplotypes can be recognised outside of Japan (Figures 2 and 3); a highly diverse central lineage (lineage 2) and two peripheral lineages (lineages 1 and 3) that are separated from the central lineage by 6 and 7 mutations, respectively. Lineage 3 consists of a single haplotype which is fixed in one of the native Taiwanese populations. One native population (TI) sampled in this study and the yellow river sample from Liu *et al* (2010) are restricted to lineage 2 but do not share haplotypes with invasive populations. Three native populations (CG, CK, CY) sampled in this study and the Minjiang sample from Liu *et al* (2010) are restricted to lineage 1 and these populations also share a common haplotype with most introduced populations. Furthermore the Pearl River samples from Liu *et al* (2010) fall into lineage 1 but do not share haplotypes with native populations. Two native populations (CRH, CH, Figure 1), however, contained haplotypes from both lineage 1 and 2 but share few haplotypes with the invasive populations. These two populations are from the Hai He

River basin at the northern margin of the species distribution. The invasive populations are widely scattered across the network and most populations

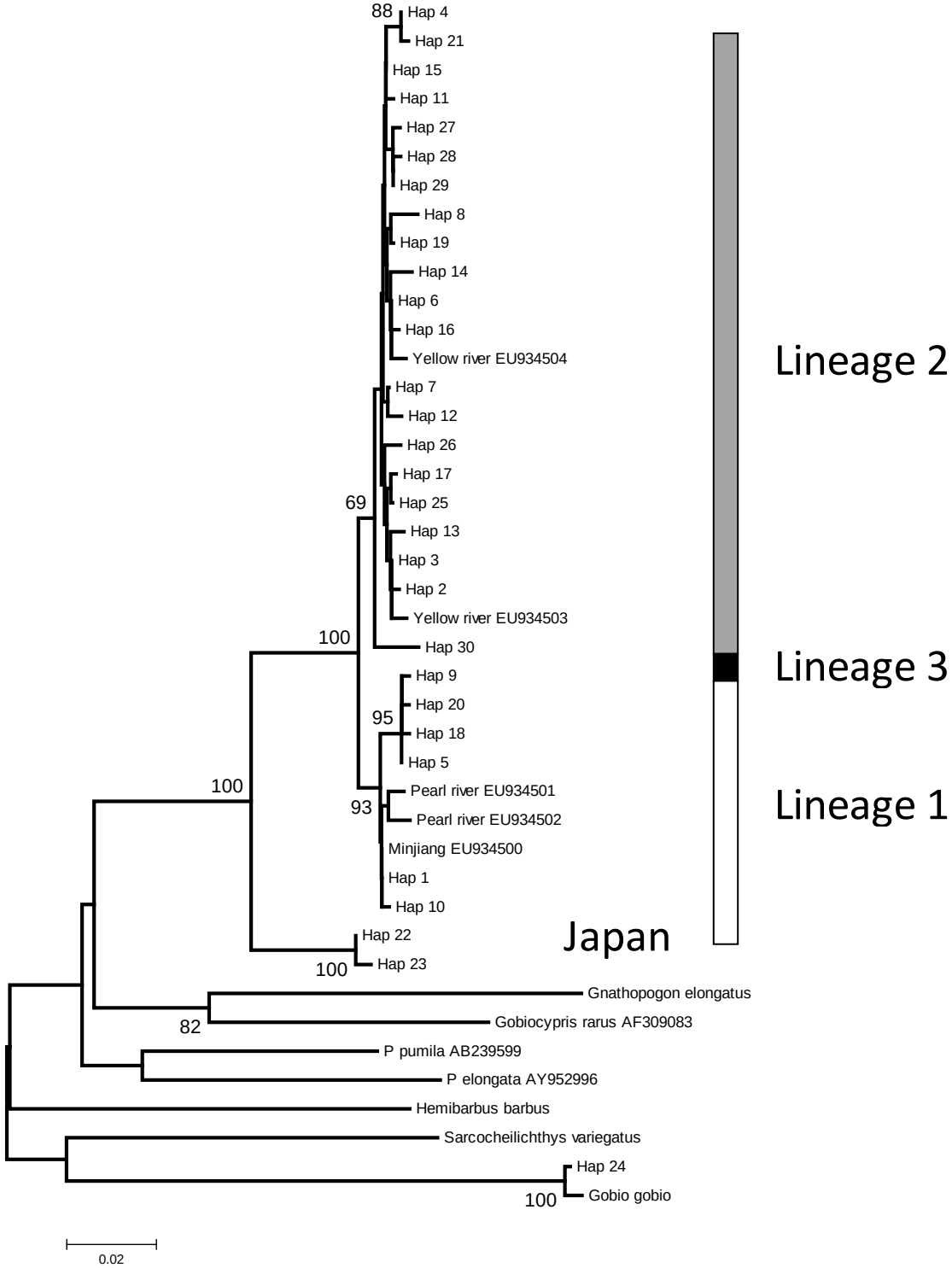


Figure 3. Phylogenetic relationship of haplotypes based on NJ analysis. Values on branches indicate non-parametric bootstrapping of the NJ-tree. ML values are only given for support values was > 70%.

contain highly divergent haplotypes from both lineage 1 and 2. Among the native populations, the Taiwanese and Japanese (TI, TT, JB) populations do not share haplotypes with any native or invasive populations.

2.3.2 Diversity within populations

After bootstrapping to account for differences in sample size, the genetic diversity of the native populations varied widely among geographical regions. Whereas the two populations of the Hai He drainage showed relatively high diversity (H = 0.34, 0.76; π = 0.006, 0.010), the populations from other drainages of mainland China and Taiwan where much less variable (H = 0.00-0.46; π = 0.000-0.002). (Table 2, Figure 4).

Genetic variation in introduced populations also varied considerably. The two recently established British populations showed low levels of variability (H = 0.00, 0.20; π = 0.000, 0.001) whereas the populations from continental Europe showed relatively high levels of variation (H = 0.20-0.66); π = 0.003-0.026). The highest nucleotide diversity was found in the Polish (PU) population (π = 0.026); this population contained one extremely divergent haplotype that clustered with a *G. gobio* haplotype, suggesting hybridisation and so was excluded from further comparisons. Overall genetic diversity in native populations (mean±SD; Hs = 0.27±0.29; π = 0.003±0.004) and invasive populations (mean±SD; Hs = 0.43±0.19; π = 0.008±0.006) was not significantly different (H, P = 0.108; π, P = 0.068). However, a more detailed analysis revealed that there were significant differences among certain groups of native and invasive

populations. When the recently introduced UK populations were excluded from the analysis, both haplotype diversity and nucleotide diversity and were significantly higher in the invasive populations than native ($P = 0.043$ and, $P = 0.014$, respectively).

Population	Group	*N*h	*H*	H_6	π	code
CH	Native	3	0.44	0.34	0.0056	green
CRH	Native	11	0.96	0.75	0.0103	red
CK	Native	1	0.00	0.00	0.0000	blue
CG	Native	1	0.00	0.00	0.0002	purple
CY	Native	2	0.46	0.35	0.0023	orange
TI	Native	3	0.59	0.46	0.0011	brown
TT	Native	1	0.00	0.00	0.0002	grey
BS	invasive	5	0.68	0.54	0.0067	white
EB	invasive	1	0.00	0.00	0.0000	white
FG	invasive	5	0.78	0.61	0.0109	white
G	invasive	6	0.84	0.66	0.0098	white
HA	invasive	5	0.62	0.48	0.0078	white
HE	invasive	3	0.59	0.46	0.0077	white
HG	invasive	5	0.62	0.49	0.0049	white
HS	invasive	3	0.25	0.20	0.0028	white
IN	invasive	3	0.34	0.27	0.0049	white
SC	invasive	3	0.67	0.53	0.0073	white
SE	invasive	3	0.67	0.52	0.0080	white
SWS	invasive	2	0.24	0.20	0.0012	white
T	invasive	3	0.68	0.55	0.0076	white
PU	invasive	4	0.60	0.47	0.0257	white

Table 2. Genetic diversity of Pseudorasbora parva populations. Columns represent populations, origin (native or invasive) number of haplotypes found in each population(Nh),observed haplotype diversity (H), mean haplotype diversity after bootstrapping based on sample size of 6 and nucleotide diversity (π) and colour code used in Figure 2.

Next nucleotide (π) and haplotype (H) diversity were simulated as expected in a population that is in mutation-drift equilibrium with constant effective population size (N_e), and we compared this to the empirical data (Figure 4). The simulations show that with increased N_e, both H and π increase, which is predicted from theory, given that larger populations can harbour more

nucleotide and haplotype diversity (Figure 4). However, the observed values of π for the introduced populations fall consistently above the theoretically predicted relationship between π and H.

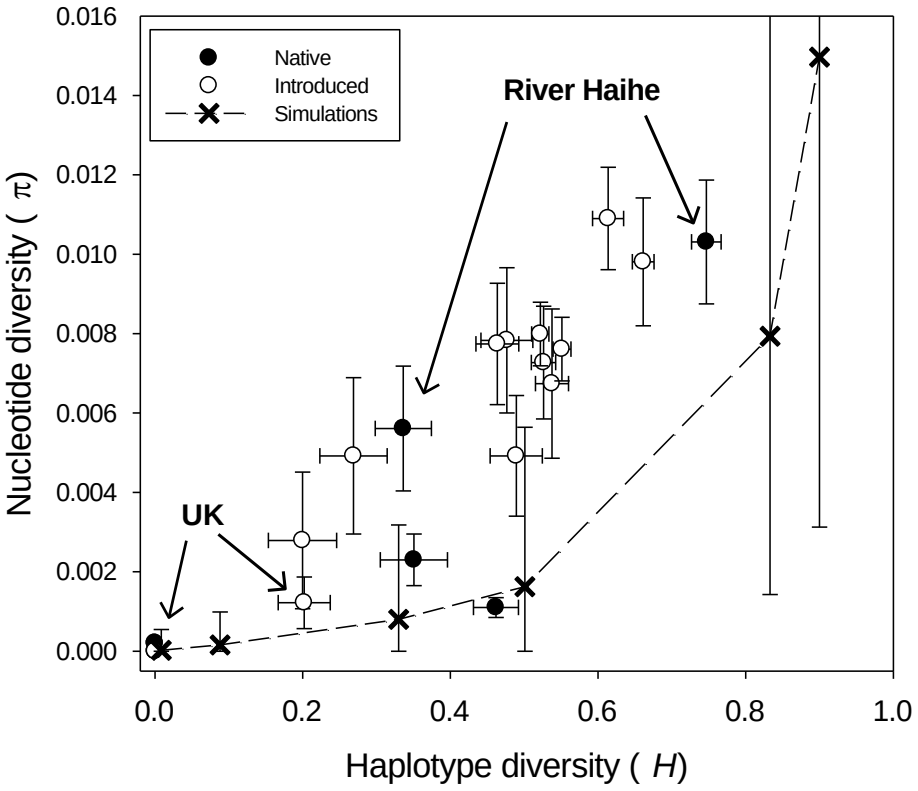


Figure 4. Plot of nucleotide (π) diversity versus haplotype (H) of the introduced (open circles) and native populations (solid circles). Also shown is the expected relationship between nucleotide diversity (5-95% CI) and haplotype diversity of simulated populations (crosses) under mutation-drift equilibrium for populations. Excluded is the Polish population (PU) because of its high nucleotide diversity.

Thus, the introduced populations showed a relative excess of nucleotide diversity, given the observed haplotype diversity and assuming mutation-drift equilibrium. Similarly, some native populations also showed a relative excess in π compared to H (Figure 4). This pattern was inconsistent with a mutation-drift equilibrium and can be explained by admixture of

populations with diverged nucleotide variation. Maximum likelihood estimates of theta ($N_e\mu$), using Migrate, differ by several orders of magnitude among native populations, ranging from 0.00006 for population TT to 0.01847 for population CRH (Figure 5).

Using an average mutation rate for mtDNA of 2% per MY (Bernatchez *et al* 1989), this translates into effective population size estimates between approximately 10^3 and 4.10^5 individuals.

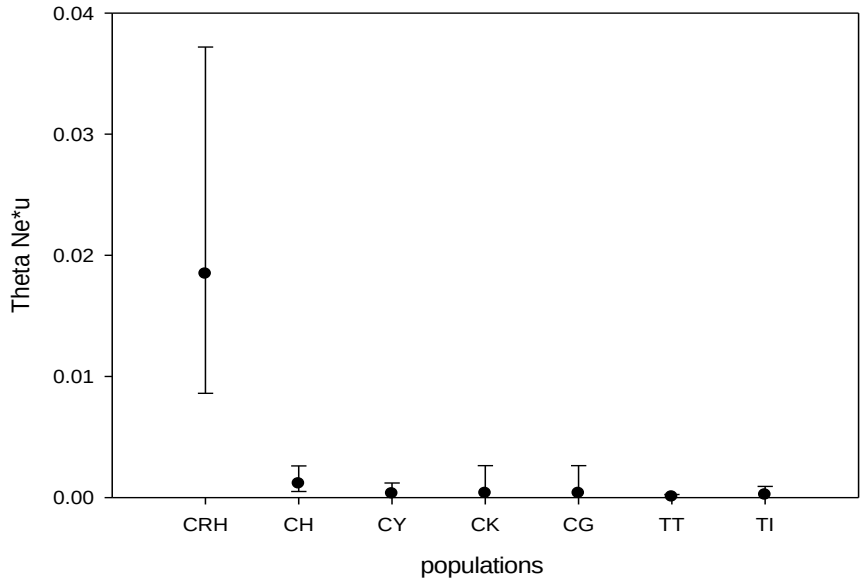


Figure 5. Estimates of effective population size (theta) of native populations based on equilibrium assumptions.

This analysis is consistent with the previous simulation study as it shows that the standing nucleotide variation in some populations can only be explained by an exceedingly large effective population size, or more plausibly, by population admixture.

2.3.3 Genetic differentiation and population structure

Pairwise genetic distance (D_{XY}) ranged from 0 to 0.02715 and pairwise genetic differentiation (F_{ST}) ranged from 0 to 1 (Appendix I, Appendix II), not including the Japanese (JB) population. The pairwise genetic distance among invasive populations (median $D_{XY} = 0.009$) was only marginally lower than that among native populations (median $D_{XY} = 0.012$), ($P > 0.3$). Similarly, the native-invasive pairwise comparison expressed in D_{XY} (median $D_{XY} = 0.012$) (Appendix II) was neither significantly different from the genetic distance among native populations ($P > 0.3$) nor from that among invasive populations ($P > 0.3$).

In contrast, genetic differentiation (F_{ST}) was considerably lower among the invasive populations (median F_{ST}=0.21) than among the native populations (median F_{ST}=0.58) ($P < 0.001$). Furthermore, the F_{ST} between the native-invasive pairwise comparison (median F_{ST}=0.53) was not significantly different from the genetic distance among native populations ($P = 0.27$), but it was significantly higher than that among invasive populations ($P < 0.001$). This result is inconsistent with the 'multiple-source-sink' model, and supports both the 'long-distance' and 'stepping stone' models.

The multidimensional scaling analysis of the F_{ST} matrix (Figure 6) showed that most of the invasive populations cluster together with two native populations (CRH and CH). This cluster is surrounded by the remaining native populations and two introduced populations (EB and SWS).

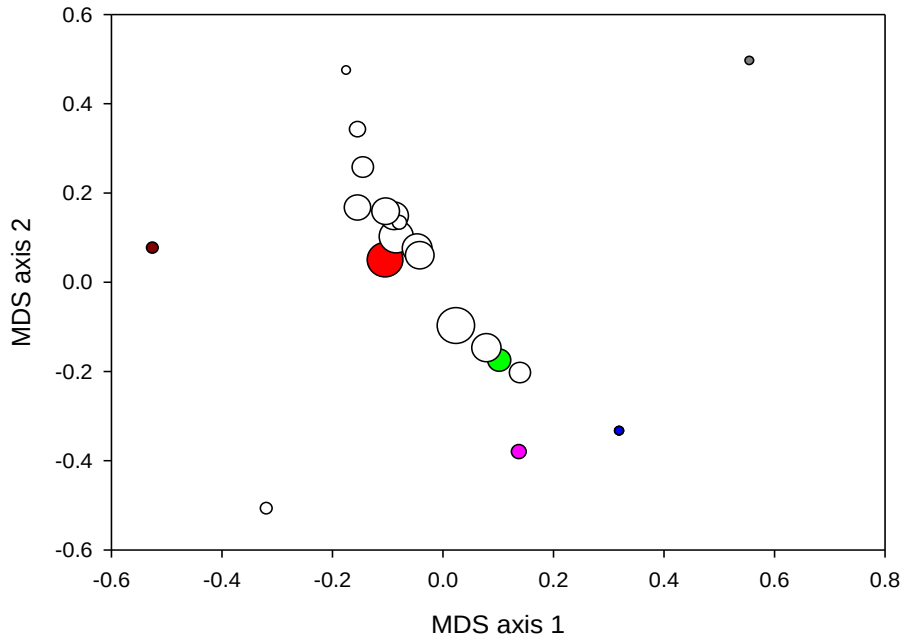


Figure 6. Plot of the first and second axis of a multidimensional scaling analysis based on pairwise F_{ST} values among populations. Size of symbols is proportional to the nucleotide diversity of populations. Native populations are colour coded according to Table 1.

Overall, there appears to be a pattern that nucleotide diversity increases towards the centre of the plot, i.e. intermediate populations have the highest nucleotide diversity, which again indicates that these populations (invasives and the samples from the river Hai He) are genetically admixed.

A Mantel-test showed a significant relationship between genetic and geographic distance among the European populations (Z = -106685; r = 0.28, one sided $P < 0.05$). However, when the recently introduced English (EB) and Welsh (SWS) populations were removed, there was no significant genetic isolation-by-distance relationship (Z = 9201, r = 0.06; one sided $P >$ 0.30). Regression analysis revealed no significant relationship between distance from source and genetic differentiation (R^2 = 0.05; P > 0.03) or

genetic diversity ($R^2 = 0.018$; $P > 0.03$) respectively. This reveals that the English and Welsh populations are bottlenecked, resulting in the spurious isolation-by-distance signal obtained when including these samples in the Mantel test. However, across continental Europe, topmouth gudgeon does not show evidence of isolation-by-distance and so we conclude that the 'long-distance' model is most consistent with these data.

2.3.4 Approximate Bayesian Computation (DIY ABC)

Based on the geographic distribution of the haplotype lineages, samples were pooled into three native and one invasive population for which we considered three feasible evolutionary scenarios (Figure 7): (i) pop 1 (native populations of haplotype lineage 1; CG, CK, CY, Minjiang), (ii) pop 2 (admixed native populations from the river Hai He; CH, CRH), (iii) pop 3 (all invasive Hungarian populations; HA, HE, HG, HS), pop 4 (native populations of lineage 2; TI, Yellow River). The Hungarian populations were chosen to represent invasive populations because they were located in close proximity to the original site of introduction. In order to account for the unsampled variation in the native range in lineage 2, one or two ghost population (GH1, GH2) were included in the scenarios (represented as branches without terminal ends in Figure 7). All three scenarios assumed that a founder event of size NF that lasted DB generations had taken place after introduction into Europe:

Scenario 1: The source of the invasive population (Pop 3) is the admixed Chinese population (Pop 2) which originates from an admixture of Pop 1 and a ghost population which split from Pop 4 at time t3.

Scenario 2: The invasive population (Pop 3) is a result of an admixture between pop 1 and an unsampled ghost population which split from pop 4 at time t4. Pop 2 evolved as in scenario 1.

Scenario 3: same as Scenario 2 but the admixture of the pop 1 and GH2 populations took place before the admixture of Pop 1 and GH1.

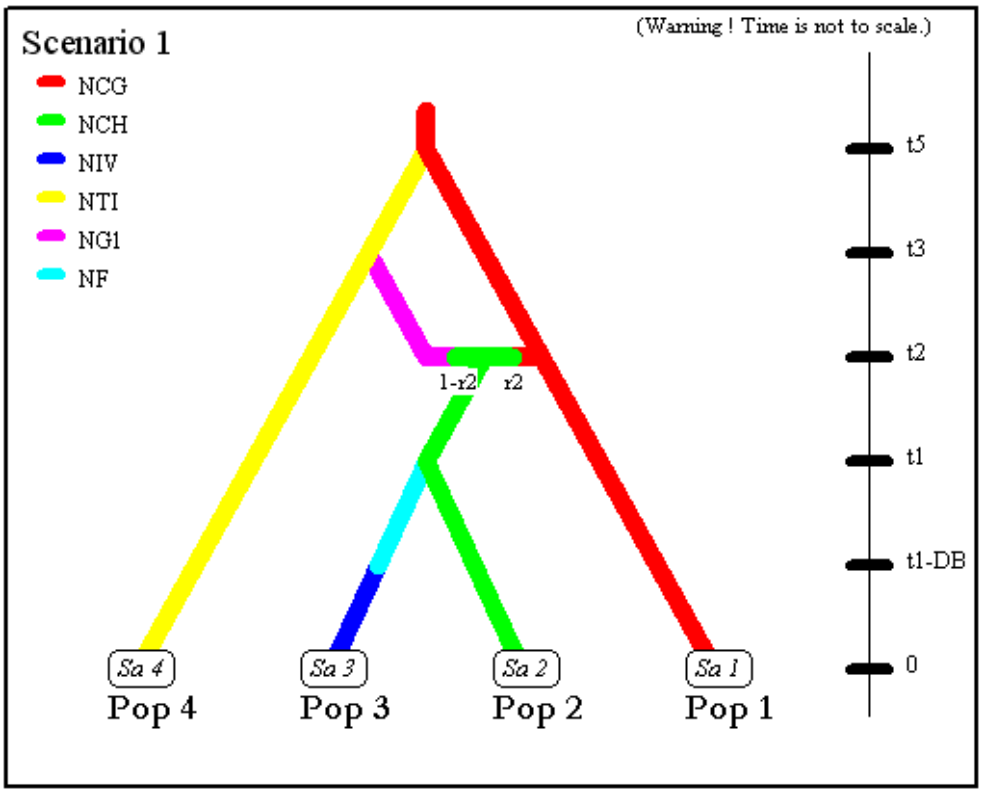


Figure 7a. Scenario 1

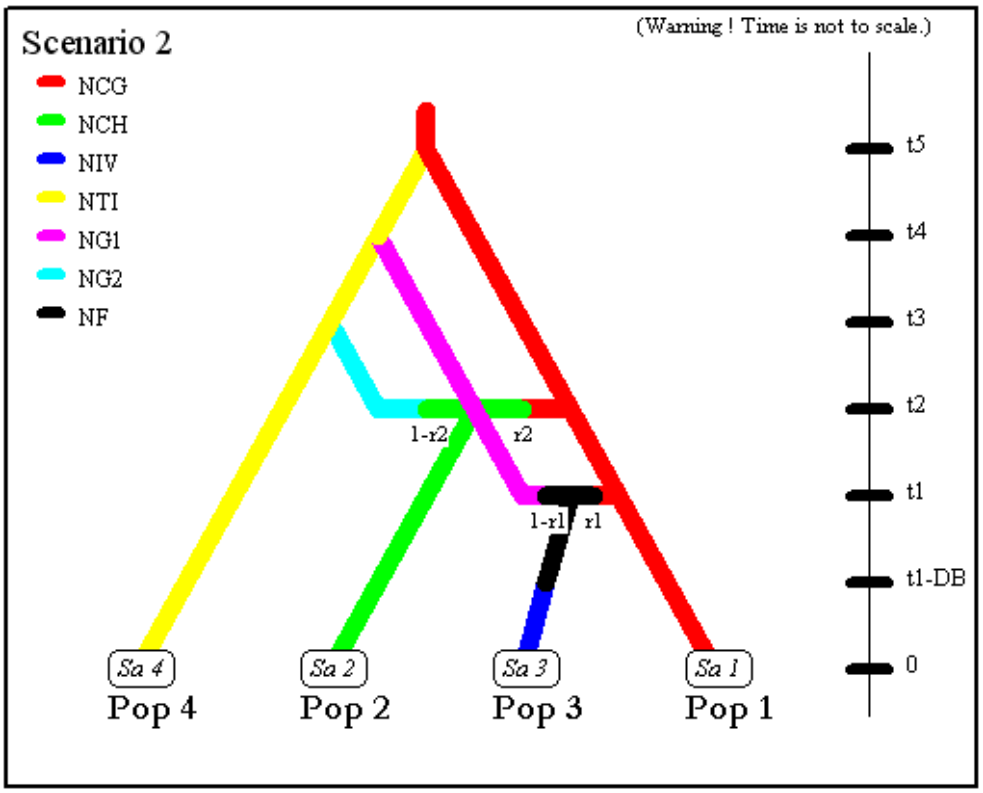


Figure 7b. Scenario 2

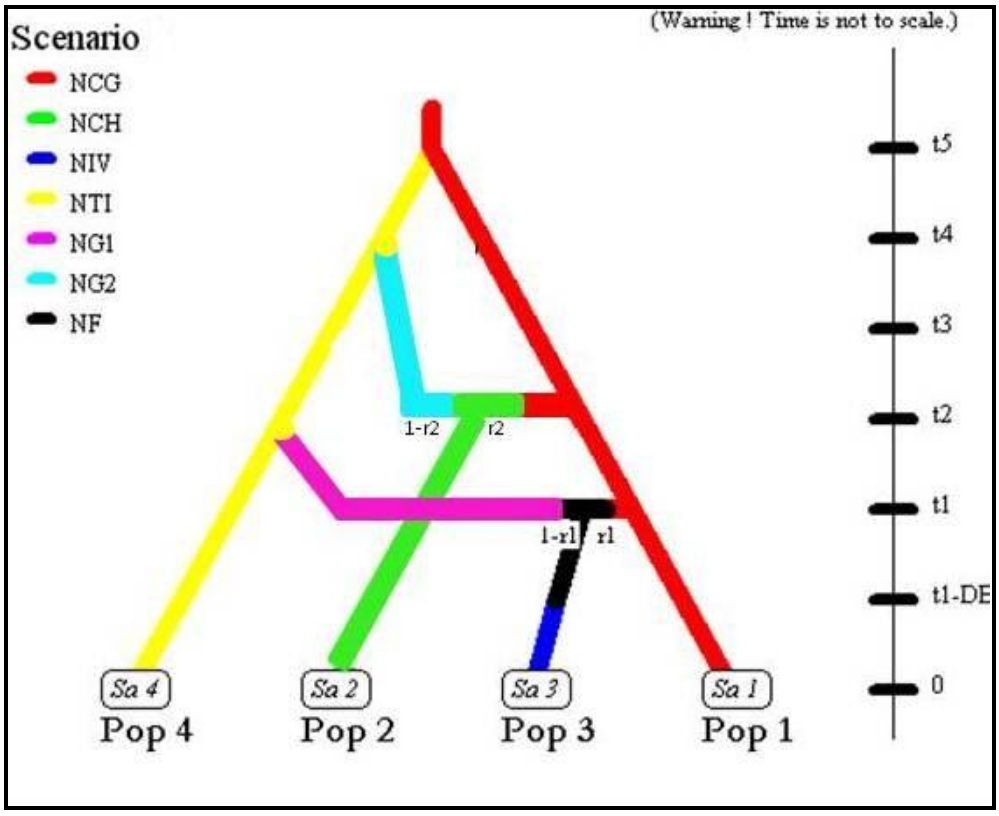


Figure 7c. Scenario 3

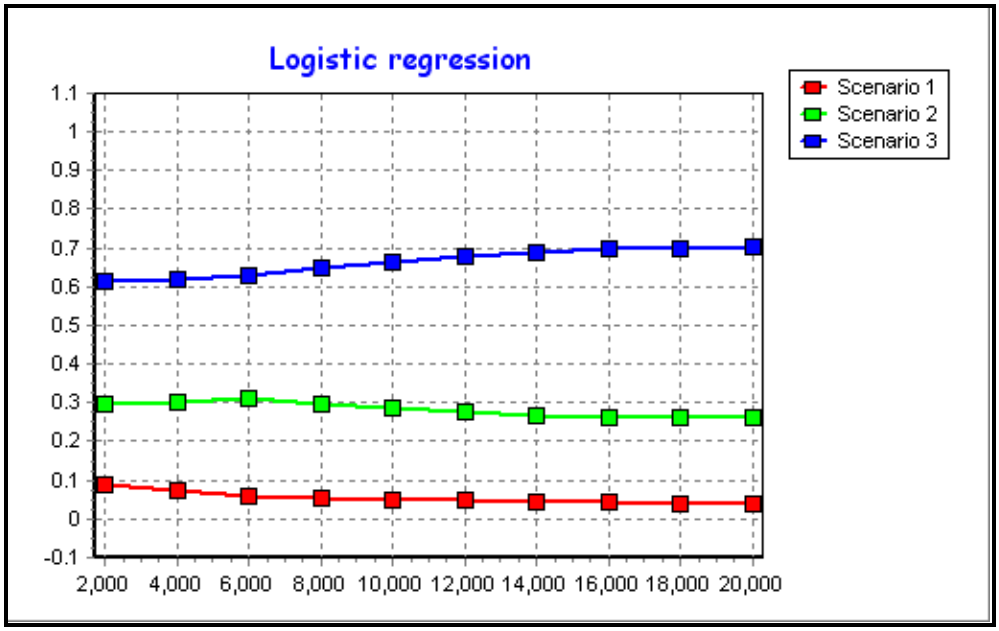


Figure 7d. Graphic representation of the three competing invasion scenarios considered in the DIY ABC analysis. Posterior probabilities of scenarios are obtained through a logistic regression, computed every 10% (between 10 and 100%) of the number of selected data sets (y axis). X axis represents the number of simulations.

A comparison of posterior probabilities of the three scenarios using logistic linear regression (Figure 7, Appendix III) showed that scenario 1 showed the lowest support with probabilities lower than 0.1. The highest probability was shown for scenario 3. The posterior distribution of model parameters under the most likely scenario was used to make inferences about the timing of events during the colonisation process assuming a generation time of one year. The posterior density of the time of first introduction ($t1$) agrees with historical records (median = 47 generations, 95% credibility interval (CI) = 30-60). Full table of posterior distributions are given in the Appendix III.

2.4. Discussion

The outputs of these analyses revealed that i) there are three evolutionary lineages of the topmouth gudgeon (*P. parva*) in the native range, two of which contributed to the colonisation of Europe; ii) most invasive populations have a higher genetic diversity than their native counterparts and a higher genetic diversity than expected under equilibrium conditions; iii) most native populations have a low genetic diversity typical for riverine fishes, an exception being samples from the Hai He river system which showed very high levels of genetic diversity, which under equilibrium conditions predict extremely high effective population sizes; and iv) the differentiation among invasive populations is much lower than among native populations.

2.4.1 Population genetics of native populations

The existence of four highly divergent haplotype lineages indicates a long isolation among geographic populations of *P. parva*. An approximate estimation of divergence times using a standard molecular clock rate of 1% MY (Durand *et al* 2002) suggests a separation of the Japanese from the Chinese and Taiwanese populations during the Miocene (5-6 MYA) which is consistent with Watanabe *et al* (2000), as well as Sakai *et al* (2002), as the areas around the Sea of Japan are thought to have formed a a 'speciation center' for freshwater and marine organisms throughout the Tertiary and Quarternary, hence the Japanese populations can be viewed as separate species. Accordingly, the remaining lineages will have formed during early

Pleistocene (1-1.5 MYA), which implies that multiple glacial refugia must have existed during the ice ages. Although the sampling scheme limits detailed phylogeographic inferences, it is apparent that there is a clear geographic association of each lineage across most of the range, but also an area in Northern China where two lineages are found in sympatry. This becomes apparent when analysing genetic diversity within populations. Most *P. parva* populations from their native range showed low haplotype and nucleotide diversity and high levels of differentiation among river systems, which is consistent with the pattern found in many other small freshwater fishes of similar size, such as the European bullhead (*Cottus gobio* L.; Hänfling and Brandl 1998; Hänfling and Kollman 2002) and guppies (*Poecilia reticulata*) (Barson *et al* 2009). The native populations from the northern range of the distribution were, however, characterised by extremely high genetic diversity, particularly the population CRH. Such high diversity is unusual among freshwater fish populations as they are usually highly structured and show low effective population sizes. Indeed, the effective population size was estimated as requiring approximately 400,000 individuals to maintain the levels of diversity observed in the CRH population and 24,000 individuals in population CH, based on a coalescence approach that assumes mutation-migration-drift equilibrium. Published estimates of effective population size in other freshwater fishes and our own estimates from the remaining native populations (N_e < 7000) are several orders of magnitude lower; for example, other cyprinid fishes range around 500 to 1000 individuals (Saillant *et al* 2005), guppies range from 100 to 900

(Barson *et al* 2009) and European bullheads between 80 and 500 (Hänfling and Weetman 2006). This suggests that the populations CRH and possibly CH are not at equilibrium but represent relatively recent secondary contact between divergent populations.

It is possible that the geographic area around the Hai He River basin represents a natural secondary contact zone between divergent phylogeographic lineages. Although the literature on this subject is relatively limited, it seems clear that high tectonic activity and sea level changes during the Pleistocene have created a complex phylogeographic pattern with little concordance among species (Yang and He, 2008). Nevertheless, studies on other freshwater fish, such as *Hemibarbus lameo* (Lin *et al* 2010) and *Salanx ariakensis* (Hua *et al* 2009), found evidence that secondary contact between diverged populations from different major river systems took place during low sea levels at the end of the Pleistocene. Furthermore, the geographic area around the Hai He River basin represents a natural secondary contact zone between divergent phylogeographic lineages of the estuarine, flathead mullet (*Mugil cephalus)* (Jamandre *et al* 2008, Liu *et al* 2009).

Alternatively, recent human translocations associated with aquaculture might have caused such an admixture; this may not be considered surprising given that freshwater aquaculture in this area of China is intense (Naylor *et al* 2000). According to Gozlan *et al* (2010b), a high volume of *P. parva* translocations have occurred in China prior to introduction in Europe. These

cyprinid translocations coincided with the end of the Chinese civil war and the need for additional sources of animal proteins (Gozlan *et al* 2010b).

2.4.2 Colonisation history

Our data showed that all invasive populations shared at least one of the four common haplotypes and that levels of genetic differentiation were low compared to native populations. Such a pattern would be expected if the invasive populations had spread from a single source. The alternative explanation of high levels of gene flow among initially differentiated invasive populations is extremely unlikely given that this would involve regular gene flow across watersheds. Therefore we reject the possibility that different European populations were independently colonised from divergent source populations ('multiple-source-sink model'). However, a number of results indicated that the introduced populations represented an admixture of divergent source populations. First, the levels of nucleotide diversity of populations in continental Europe were, on average, higher when compared with native populations. Second, the nucleotide diversity of invasive populations was higher than expected from coalescent theory. Finally, the population structure analysis showed that the majority of the invasive populations and two (admixed) native populations occupied central positions in the MDS plot between divergent native populations. The main exceptions to this pattern were the two British populations, which showed a low genetic diversity and high levels of differentiation from other invasive populations, but were fixed for one or two common invasive haplotypes.

These populations were founded relatively recently, most likely from sources in Germany (Gozlan *et al* 2002, Gozlan *et al* 2010b). We suggest that this pattern is a result of secondary bottlenecks during spread and translocation within Europe. Our data do not enable us to distinguish whether the admixture event has happened before the introduction into Europe or shortly after the introduction, before the large scale expansion across Europe, but based on the assumption that a single introduction to the same geographical location is more parsimonious than two independent introductions we suggest that it is more likely that the admixture event has happened in the native range.

Although the sampling coverage in the native range was not comprehensive enough to pinpoint the exact location(s) which acted as a source of invasive European populations, some more general inferences can be drawn. The data outputs suggest that the invasive populations originate from mainland China rather than Taiwan or Japan. The haplotype distribution of invasive populations and populations from northern China raise the possibility that this area is the source of introduction. However, the DIY ABC analysis suggests that this is much less likely than a scenario where the invasive populations in Europe originate from an admixture between populations from lineage 1 (such as the Yangtze) and an unsampled population from lineage 2. Anecdotal reports suggest that *P. parva* were initially translocated to Romania and Hungary from the Yangtze River at Wuhan which is geographically close to our CY sample (Gozlan *et al* 2010b) and most likely

originate from an aquaculture pond. Given our genetic results we suggest that these aquaculture populations consisted of a mixture of the local Yangtze population and fish wish were introduced from a different more northern river system possibly a tributary of the Yellow River.

The isolation-by-distance analysis indicated that both 'stepping-stone' and 'long-distance' processes might have contributed to the spread of *P. parva* in Europe. The weak but significant pattern of IBD across the whole data set was mainly caused by the highly bottlenecked British populations at the margin of the distribution. After excluding these two populations, none of the tests was significant. The 'stepping-stone' colonisation is therefore not likely to be the predominant process for the spread of the species in Europe. We suggest that long-distance dispersal must have played a major role, possibly as a consequence of fish transport associated with aquaculture, as the introduction of non-native fish species is an important part of aquaculture, including intentionally introduced fish, as well as species introduced and cultured illegally, further confirming the need for appropriate water management. The finding of long-distance dispersal as a colonisation model is in agreement with Gozlan *et al* (2010b), who suggested a *P. parva* dispersal model showing dispersal distances of approximately 250 km from the 1970s to the end of the 1990s, followed by shorter dispersal of 20 km on average since 2000. Additional genetic analyses at the country level with greater resolution of the geographical

pattern of haplotypes are likely to confirm this two-stepped invasion process.

2.4.3 Evidence of hybridisation

A single individual from the Polish population contained a highly divergent haplotype. The phylogenetic analysis revealed that the sequence is very closely related to a published GenBank sequence of *G. gobio*. The genus *Gobio* belongs to the same cyprinid subfamily as Pseudorasbora, the Gobioninae and is a close European relative of *P. parva* (Tang *et al* 2010). Despite the close phylogenetic relationship, the two species show very different phenotypic appearances and misidentification is extremely likely given that only adults were sampled. Although laboratory experiments have not confirmed this, based on these results we therefore conclude that this indicates mitochondrial introgression and this suggests that the invasive *P. parva* is able to hybridise with at least one native European species. This raises further concerns about the threat which *P. parva* poses to native European fish fauna and corroborates experimental evidence that hybrids between *P. parva* and another European cyprinid *Leucaspius delineatus* are possible (Gozlan and Beyer, 2006).

2.5 Conclusion

The European introduction of *P. parva* resulted from accidental releases from a human-induced faunal translocation (Gozlan *et al* 2010b). Their European colonisation was initiated by the introduction to a single location

or small geographic area it was preceded by, or associated with, the admixture of genetically diverse source populations. This adds to the existing evidence that many invasive populations show the genetic signature of admixture or of multiple introductions (Hänfling 2007; Roman 2006). Although the data available did not fully allow us to disentangle the source populations of the invasive populations, we now have a better perspective of the spread of the species within the native range and the introduction of the species into Europe. It remains to be tested how much of the observed phenotypic variation can be attributed to phenotypic plasticity, but the single origin model supported by our data makes it more likely that the disease resistance reported in some populations of *P. parva*, that potentially will lead to devastating consequences for native fishes (Gozlan *et al* 2005, 2009) is an ubiquitous feature of the invasive populations.

This chapter has been published as a research article in PLos ONE in March, 2011: Simon A, Britton R, Gozlan R, van Oosterhout C, Volckaert FAM, *et al* (2011) Invasive Cyprinid Fish in Europe Originate from the Single Introduction of an Admixed Source Population Followed by a Complex Pattern of Spread. PLoS ONE 6(6): e18560. (DOI:10.1371/journal.pone.0018560).
This work was supervised by Dr Hänfling and Dr van Oosterhout, however, all laboratory and analytical work, as well as writing of the chapter was carried out by me. Dr Gozlan and Dr Britton are collaborators working on the ecology of the topmouth gudgeon *P. parva* and supplied samples and co-funded parts of the molecular work.

CHAPTER III

Understanding the spread and invasion history of topmouth gudgeon *(Pseudorasbora parva)* populations in Europe based on microsatellite data analysis

3.1 Introduction

Invasive species pose a major threat to agriculture, economy and native species. However, invasions also provide an opportunity to study contemporary evolution, as the rapid range expansions during invasions can expose invaders to selection and local adaptation (Maron *et al* 2004), when the species is introduced into a new environment. Reconstructing the invasion history of non-native populations is important to understand the evolutionary processes that underlie these invasions (Prentis *et al* 2008; Okada *et al* 2009), such as bottlenecks and other founding effects (Nei 1975). Due to these founding events, invasive populations are likely to be genetically less variable than the source populations they originate from. Invasive species are often responsible for the loss of species diversity and the extinction of native species (Pimental *et al* 2000) but not enough is known about the patterns of genetic differentiation in these populations and the geographical history of invasions.

Pseudorasbora parva (Figure 1) is a cyprinid fish that was introduced into Europe from China in the 1950s and has since spread to many countries on the continent. My aim was to understand the distribution of the species within Europe and to uncover the sources of the invasive populations.

A previous study has looked at the origin of the species (Simon *et al* 2011), using mtDNA markers and this study showed that all invasive populations share at least one of the four common haplotypes found and the level of genetic differentiation was found to be low compared to native populations. This confirmed that the likely introduction method of the species into

Europe was through a single introduction from divergent source populations within the native range. An analysis of population structure also showed that most European populations are more closely related to each other when compared to native populations and long-distance method of spread was important for the species' introduction in Europe.

Figure 1. ***Pseudorasbora parva*****, adapted from www.silurus.acnatsci.org**

While mitochondrial DNA is a suitable marker for phylogenetic studies, a lot of variation found in mtDNA can be lost during bottlenecks making this a less suitable marker to study demographic processes after a species introduction. In contrast, hyper-variable microsatellite markers preserve variability longer (Neigel and Avise 1986, Villablanca *et al* 1998). Consequently, in this chapter I used microsatellite markers to quantify the genetic effects of a population invasion and to describe the genetic relationships of populations in a multilocus analysis.

Previous studies have successfully used microsatellite markers to address such questions. For example Herborg *et al* (2007) used Bayesian assignment methods to demonstrate that rates of gene flow among invasive populations of the Chinese mitten crab *(Eriocheir sinensis)* are significantly correlated with shipping routes indicating that ballast water transport might have

played an important role for dispersal and a study on the brown mussel (*Perna perna*) by Holland (2001) revealed similar results about the importance of ballast water in the spread of invasive species. Walters *et al* (2008) used microsatellite markers to measure the amount of hybridization between the invasive cyprinid, the red shiner (*Cyprinella lutrensis*) and the native blacktail shiner (*C. venusta stigmatura*) and Boyer *et al* (2008) investigated hybridization and invasion patterns between native cutthroat trout (*Oncorhynchus clarkii lewisi*) and rainbow trout (*Oncorhynchus mykiss*). Other studies using microsatellites include examining gene flow and by showing the lack of genetic bottlenecks the adaptive plasticity in two Lessepsian rabbitfish species (*Siganus rivulatus* and *Siganus luridus*) (Hassan 2003), the presence of a founder event in Trinidadian guppies (Barson *et al* 2009) and the role of bottlenecks in sea lamprey populations (Bryan *et al* 2005).

Due to high mutation rates (μ=10^{-6} to 10^{-2}) (Schlötterer, 2000) microsatellite loci show generally high gene diversity and large number of alleles in pre-bottlenecked populations. Therefore these markers are particularly suited to study recent evolutionary events and are a powerful tool to assess the geographical structuring of *P. parva* across its invasive range and test specific hypotheses of dispersal in the invasive range.

The objectives of this study were threefold: (1) to confirm and further understand my previous findings with regard to the pathways and modes of colonisation of the species into Europe and the number of introduction events in the invasive range (2) to test for the presence of founder events

and to test the hypothesis that genetic bottlenecks had an effect on the genetic diversity found in these populations (3) to test for admixture in the European populations.

3.2 Materials and methods

3.2.1 Sampling scheme and *laboratory procedures*

A total of 378 individuals were collected from 15 locations, 14 European and one Chinese population (Table 1, Figure 2.) to ensure geographical coverage of most of the invasive range; I also included one Chinese population (CRH), which we believe to have greatly contributed to the European populations.

Finclip tissue was used to extract DNA using the Puregene DNA extraction kit (Qiagen). Initially ten pairs of primers from Konishi & Takata (2004) were tested for polymorphism and consistency of amplification. After preliminary tests of running 10 primer pairs, the 5 loci that amplified the best were chosen to be included and were amplified separately using Verity Thermal Cycler the following PCR conditions (Table 2). Thermal cycle amplifications were performed in 10 µL reactions, containing 1.0 µL 160 µM NH4, 1.0 µL 100 µM dNTPs, 0.3 µL 50 µM MgCl2, 0.05 µL Taq polymerase, 0.2 µL each of primers, 6.25 µL PCR water and 1.0 µL of template DNA. Cycle parameters were as follows: 5 min at 95uC; 15 s at 95uC, 20s at 58 or 60C, depending on which primer was used for the reaction, 30s at 72C; 5 min at 72uC.

Code	Population	N	Geographical coordinates	
CG	Guangdong, River, Zhuijang River basin, China	6	23° 07′ 53″ N	113° 15′ 59″ E
CH	Huairou Reservior, Hai He River basin, China	15	40° 18′ 46″ N	116° 36′ 36″ E
CK	Kinmen Island population, China	6	24° 26′ 11″ N	118° 21′ 27″ E
CRH	River Hai He, Hai He River basin, China	15	39° 07′ 15″ N	117° 12′ 54″ E
CY	Wuhan, Yangtze River Basin, China	10	29° 58′ 20″ N	113° 53′ 29″ E
JB	Lake Biwa, Yodo River basin, Japan	3	32° 20′ 44″ N	136° 10′ 15″ E
TI	I-lan county, I-lan River, Lanyang River Basin. Taiwan	15	24° 45′ 00″ N	121° 45′ 00″ E
TT	Dajia River, Taichung county, Dajia River basin, Taiwan	15	23° 09′ 00″ N	120° 38′ 34″ E
BS	Slangebeek nean Hasselt, Belgium	15	50° 55′ 48″ N	05° 15′ 00″ E
EB	Byland Abbey, Yorkshire, UK	15	54° 12′ 10″ N	01° 09′ 35″ W
FG	Grand Lieu, France	15	47° 05′ 45″ N	01° 43′ 46″ W
G	River Ammer, Wielenbach, Germany	15	47° 52′ 11″ N	11° 09′ 00″ E
HA	Hortobagy, Hungary	15	47° 36′ 00″ N	21° 06′ 00″ E
HE	Ederecsi-patak, Hungary	15	46° 48′ 04″ N	17° 23′ 16″ E
HG	Gic, Hungary	15	47° 25′ 32″ N	17° 44′ 44″ E
HS	Salyi-patak, Hungary	15	47° 56′ 06″ N	20° 39′ 58″ E
IN	Nestore, Italy	15	43° 21′ 14″ N	12° 14′ 10″ E
PU	Utrata River, Poland	15	50° 35′ 50″ N	18° 09′ 32″ E
SC	Vrakuna, Slovakia	15	47° 49′ 24″ N	18° 49′ 16″ E
SE	Ebro Basin, Spain	15	40° 43′ 12″ N	00° 51′ 47″ E
SWS	Sylen Lake, Llanelli, South Wales, UK	15	51° 40′ 42″ N	04° 09′ 47″ W
T	Blanice River, Vodnany, Czech Republic	15	49° 08′ 52″ N	14° 10′ 32″ E

Table 1. Sample locations and sample sizes and geographical coordinates for native and invasive populations

As these conditions still did not yield any amplification for some of the populations/primers, touchdown PCR was performed, increasing the annealing temperature from 58°C/60°C to 63°C and 65°C respectively, which allowed better amplification of the target gene. The amplified DNA was then visualised on a 1.5% agarose gel (100ml 1X TE buffer, 1.5g agarose) and the amplified PCR products were analysed

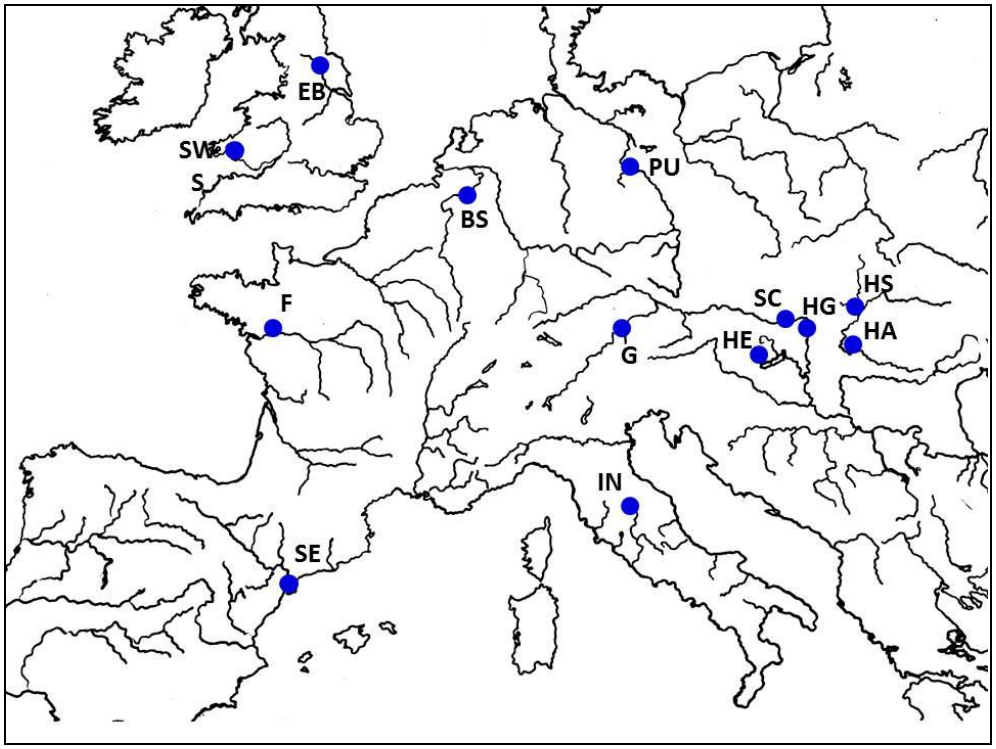


Figure 2. Sample sites in Europe

using Beckman CEQ800 sequencer (Beckman Coulter) using two multiplex mixes: Loci PA02, PA06, PA22 and either PA05 or PA22 as the fourth locus.

Locus	Repeat unit	Sequence	Tm	Allele size range (bp)
PA02	(CA)18+ (CA)4+(CA)4	F:ACATGCTGCCATCGTAACTC R:GCAAATCTTCTCCAATCCTA	60ºC	225-401
PA05	(CA)11	F:CACAGCATAAACCCTCCTCT R:GGTGTGGTTTATTAGACAGA	58ºC	95-118
PA06	(CA)10	F:CTTCACACAGTCCACATCTG R:AACCGATTACAGTGCTCCAT	60ºC	172-370
PA22	(AC)2 + (AC)20	F: GCGGGATGGGAGGGATGTA R: GGGGTCAGACGGTGCTAAC	60ºC	114-186
PA24	(CA)14 + (CA)5 + CA)10	F:GACAGGATCACTCGCTTCTC R:CTCCTGCGTGTCTGTTTATG	60ºC	118-334

Table 2. Microsatellite primers data used in the study, as adapted from Konishi & Takata (2004).

3.2.2 Data analysis

Sized alleles were manually inspected for correctness of size scoring and the data was checked for stuttering, large allele drop-outs and null alleles using the software Micro-Checker (van Oosterhout *et al* 2004). As some of the loci appeared to contain null alleles, the genotypes were adjusted using the Brookfield 1 method (van Oosterhout *et al* 2004) for null allele correction. Tests for linkage disequilibrium and deviations from Hardy-Weinberg equilibrium (HWE) were carried out using Genepop v 4 (Raymond and Rousset 1995) (number of batches: 20, Iterations per batch: 5000 by Markov chain method (Appendix IV, Appendix VI). Some causes of possible departure are migration, mutation, non-random mating or other evolutionary processes, but most populations are expected to deviate from HWE because of genotyping errors. This software was also used to calculate allele frequency. F_{IS} was calculated with FSTAT genetic software (Appendix V) (Goudet 2005) and observed (H_o) and expected heterozygosity ($H_{e;}$) (Nei 1978) for individual populations were calculated using Cervus 3.0.3 software (Kalinowski *et al* 2007).

In order to analyse geographic population structure, pairwise F_{ST} values (Weir and Cockerham 1984) across loci were calculated using Genepop v 4 (Raymond and Rousset 1995). A multi-dimensional scaling analysis (MDS) based on F_{ST} values was used to visualise the genetic differentiation between populations.

The individual based Bayesian cluster analysis implemented in STRUCTURE v2.3 (Pritchard *et al* 2000) was used to determine the number

of genetically different groups among the European samples and the proportion of membership of each individual to each group. STRUCTURE was run 5 times independently for a range of *K* values (as the assumed number of populations for sample assignment as described by Evanno *et al* 2005) to determine the number of clusters between 1 and 10 with 100,000 MCMC repetitions and a 'burn-in' of 10,000 iterations using the admixture model of population structure, the correlated allele frequency model with Brookfield corrections (Appendix VIII).

The minimum number of cluster was determined by investigating the second order rate change in the likelihood distribution for different values of *K* (Evanno *et al* 2005). This analysis was conducted using the online tool STRUCTURE Harvester (Earl 2011).

The programme Bottleneck version 1.2.02 (Cornuet and Luikart 1996) was used to test for the presence of a recent population bottlenecks. The principle of this analysis is that populations that have experienced a bottleneck will exhibit a reduction the number of rare alleles. A Wilcoxon test with 10,000 permutations was used to evaluate for a significant excess for heterozygosity as it is the most suitable test to use when only a few number of loci are used (typically less than 20) (Piry *et al* 1999). The TPM (Two Phase Model) mutation model was used, which incorporates the mutational process of SMM but also allows for larger jumps in repeat number (di Rienzo *et al* 1994) as well. I also looked at whether the allele frequency distribution produced a mode shift, indicating a recent bottleneck event. The programme LDNe (Waples and Do 2008) was used to assess

estimates of the effective population size (N_e) from each population based on the LD method. Calculations of N_e and the confidence intervals (CI) were estimated considering alleles with a frequency of c ≥ 0.05 and c ≥ 0.02 and ≥ 0.01, respectively. The number of alleles specific to a population is important, especially with the use of highly variable microsatellites. HP-Rare (Kalinowski 2005) was used to calculate the number of private alleles found in each population per locus, as a measure of genetic distinctiveness across loci. This programme uses rarefaction to correct for sample size bias, as the probability of finding rare alleles increases with increased sample size (Brown and Greene 1987).

A Mantel-test was carried out in XLStat in Excel to evaluate the correlation between the genetic and geographic distances within the data.

Using Minitab version 12.1 (Minitab Inc., 2010), a regression analysis was carried out to investigate the relationship between the mean H_e and the invasion level of the European populations. For this analysis, all populations were assigned a rank ranging from 0 (Chinese population) to 5 (the most recent introductions in the UK), depending on the order of the first sighting of the species in each location (based on Gozlan *et al* 2002, Gozlan *et al* 2010b). These ranks were then correlated with the mean H_e values. This pattern was further analysed by investigating whether the invasion history could explain variation in gene diversity. The rationale was that populations that were invaded first would have had more time to regenerate gene diversity through new mutations, as well as have fewer consecutive founder events.

In order to tie the results of the mitochondrial and microsatellite data analysis together and to quantify the level of correspondence between these markers, a regression analysis was performed using the variables mean expected heterozygosity (H_e) for the microsatellite data and haplotype diversity (H) for the mtDNA data.

3.3 Results

Gene diversity and genetic differentiation of the invasive *P. parva* populations were investigated using five microsatellite markers. Microchecker detected the presence of null alleles in loci PA02, PA05, PA06 and departure from HWE, the high number of null alleles is due to low DNA template quality. Within-population HWE tests (H_0=random union of gametes) found 19 out of the 75 locus-population combinations (5 loci x 15populations) to be statistically different from HWE. Exact tests for deviation from HWE showed deviation in all five loci. PA02, PA05, PA22 and PA24 each had 3 populations showing departure from HWE and the PA06 locus deviated from HWE in no less than 7 populations.

This deviation from HWE is considerably more than expected from a type I error (false positive rate), and can be explained by null alleles segregating at these loci, as was indicated also by Micro-Checker.

Linkage disequilibrium (LD) tests for pairs of loci across the populations were overall not significant, except for 4 out of 150 (≈2.6%) pairings, these were P22-P24 in population FG, P5-P24 in population HA, P5-P22 in

population HS and P2-P6 in population PU (Appendix VII). LD can be generated by demographic processes such as population expansion and migration. However, the proportion of pairs (2.6%) that show LD is below the 5% type I error rate, which suggests that these processes have not resulted in LD in the invasive *P. parva* populations. The observed heterozygosity (H_o) ranged from 0.151 - 0.954 and the expected heterozygosity (H_e) ranged from 0.54 - 0.81 (Table 3).

Pop	Location	Sample size	Mean H_e	SD H_e	Mean H_o	SD H_o	p-value	Allelic richness	Private alleles	Invasion level
BK	Belgium	30	0.6796	0.17758	0.6832	0.12659	0.549	2.89	0	4
BSa	Belgium	30	0.6762	0.18186	0.6878	0.16819	0.37436	2.87	0	4
CRH	China	30	0.8116	0.12574	0.7386	0.13244	0.09354	3.6	19	0
EB	England	23	0.562	0.27639	0.4904	0.26669	0.46186	2.5	0	5
FG	France	30	0.6916	0.17981	0.5948	0.18942	0.18906	2.98	5	3
G	Germany	25	0.6618	0.19587	0.5588	0.23449	0.28092	2.88	3	3
HA	Hungary	30	0.7048	0.17784	0.6198	0.13799	0.2035	3.05	5	1
HE	Hungary	30	0.6686	0.17108	0.641	0.14827	0.4459	2.9	4	2
HG	Hungary	24	0.6994	0.18345	0.5212	0.21024	0.16842	2.99	2	2
HS	Hungary	17	0.7018	0.15345	0.6588	0.16309	0.20262	3.01	2	2
IN	Italy	20	0.6492	0.26527	0.5166	0.24889	0.13828	2.91	5	3
PU	Poland	30	0.68	0.19884	0.6382	0.23266	0.22568	2.94	2	2
SC	Slovakia	18	0.7242	0.11097	0.5982	0.12922	0.4003	3.03	3	2
SE	Spain	13	0.5414	0.24446	0.4368	0.24223	0.24218	2.36	3	5
SWS	Wales	28	0.7428	0.13857	0.686	0.19233	0.0547	3.04	1	5

Table 3. Summary of microsatellite allelic variability for each locus, mean observed and expected heterozygosity for each population, mean p-values for each population and invasion levels used in regression analysis, this was assigned based on date of first sighting (Gozlan, 2010b). Also shown are allelic richness and number of private alleles across all loci per population.

The analysis of likelihood for different values of *K* from the STRUCTURE analysis showed that there is a signal of two main clusters but also a slightly weaker signal for seven clusters (Figure 4). An assignment into two clusters

basically separates the Chinese individuals which all belong to one cluster from European individuals which fall either into the second cluster or represent an admixture between the two clusters. Using an assignment into seven clusters also assigns all Chinese individuals to a single cluster. In contrast most European individuals represent an admixture between a number of different clusters. Individuals from three European populations (EB, SE and SWS) belong to just a single cluster each.

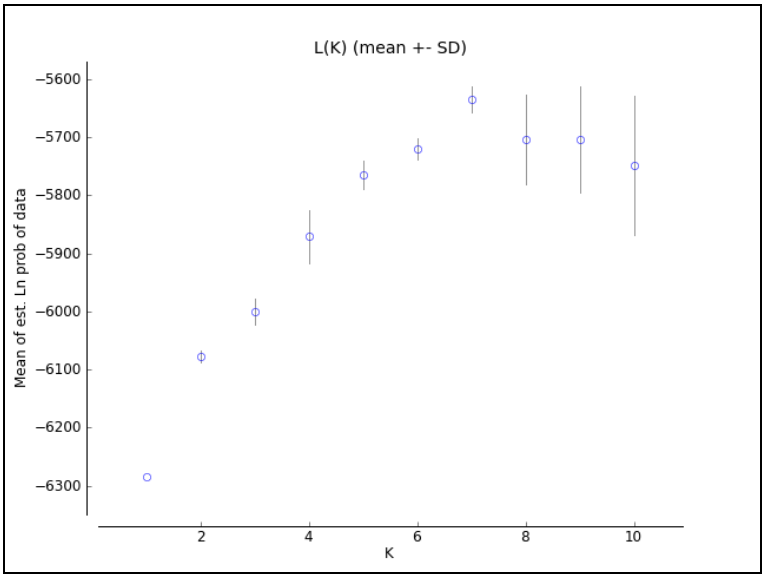


Figure 3a

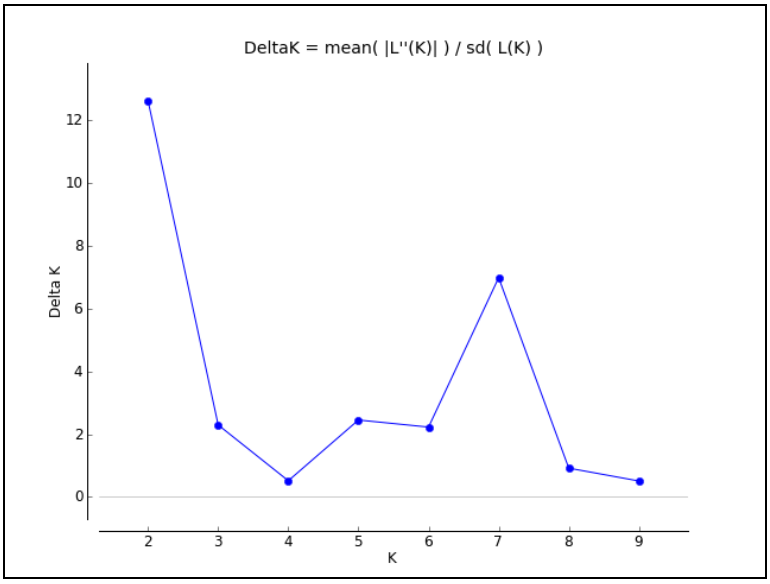


Figure 3b

Figure 3a+b. Analysis of likelihood for different number of genetic clusters (*K*) based on an individual based cluster analysis implemented in STRUCTURE v2.3. (Pritchard *et al*. 2000). Mean likelihood for different values of *K* across independent runs (n = 5) (a); and the second order rate change (delta *K*) of the likelihood values based on the method of Evanno *et al* (2005) (b)

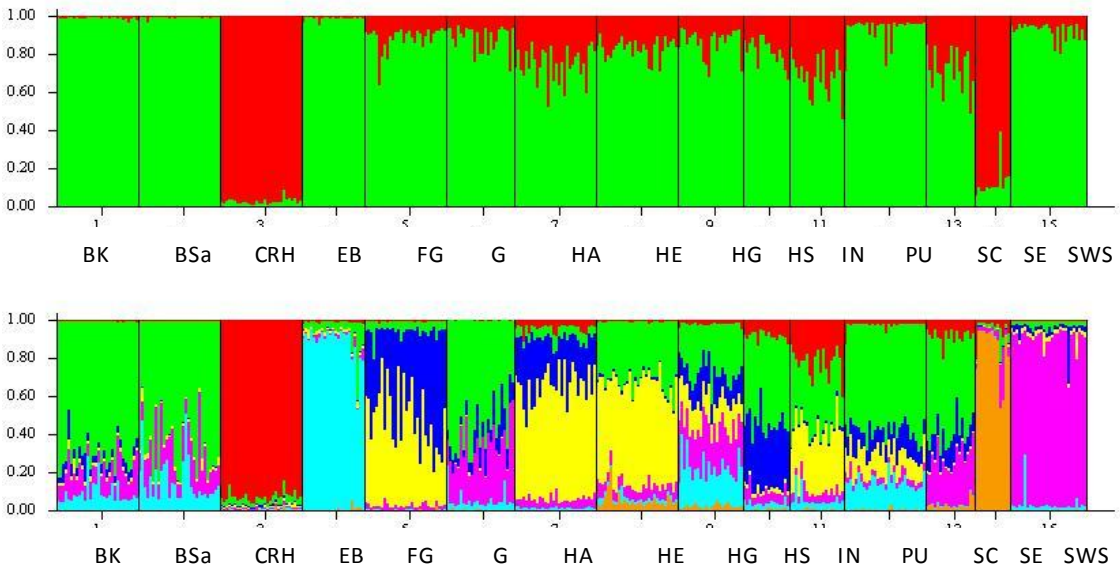


Figure 4. Assignment of individuals into 2 and 7 genetic clusters respectively based on an individual based cluster analysis implemented in STRUCTURE v2.3. (Pritchard *et al* 2000)

The bottleneck analysis under the TPM model (Wilcoxon rank test $p<0.05$) showed that 4 of the populations have undergone recent declines but all populations showed an L-shaped distributions (Table 4).

The LDNe method produced several estimates of negative (7 point estimates) and also the measures of confidence intervals indicated infinity for all populations. An infinite estimate is a result of not enough information, so the genetic signal found in the data is due to sampling error rather than genetic drift (Waples and Do 2010). The number of alleles per locus ranged from 5 to 30 and allelic richness ranged from 2.36 (SE) to 3.6 (CRH), the CRH population also had the most private alleles (N=19) (Table 3).

	TPM	**Mode-shift**
BK	0.10938	L-shaped distribution
BS	**0.04688**	L-shaped distribution
CRH	**0.03125**	L-shaped distribution
EB	0.3125	L-shaped distribution
FG	**0.01563**	L-shaped distribution
G	0.6875	L-shaped distribution
HA	0.6875	L-shaped distribution
HE	0.96875	L-shaped distribution
HG	0.89063	L-shaped distribution
HS	0.59375	L-shaped distribution
IN	0.3125	L-shaped distribution
PU	0.3125	L-shaped distribution
SC	0.3125	L-shaped distribution
SE	0.6875	L-shaped distribution
SWS	**0.03125**	L-shaped distribution

Table 4. Bottleneck test of European *P. parva* populations, each entry in the table represents the probability that bottleneck was detected using the Wilcoxon test under TPM model. Wilcoxon-test P-values represent one-tailed probabilities for heterozygosity excess. Significant values are indicated in bold.

A Mantel-test showed there is a strong and statistical significance association between genetic and geographic distances, with 46.2% of the variation in H_e being explained by population geography (Mantel test: P<0.0001, with 10000 randomizations) (Figure 5, Appendix IX). The correlation between the two matrices of pairwise F_{ST} values of the two markers using a Mantel-test showed a significant positive relationship between the divergence estimates. In order to test whether geography also explains the genetic divergence between populations, I conducted a MDS analysis based on F_{ST} values. Figure 6 shows the genetic relationship between the populations.

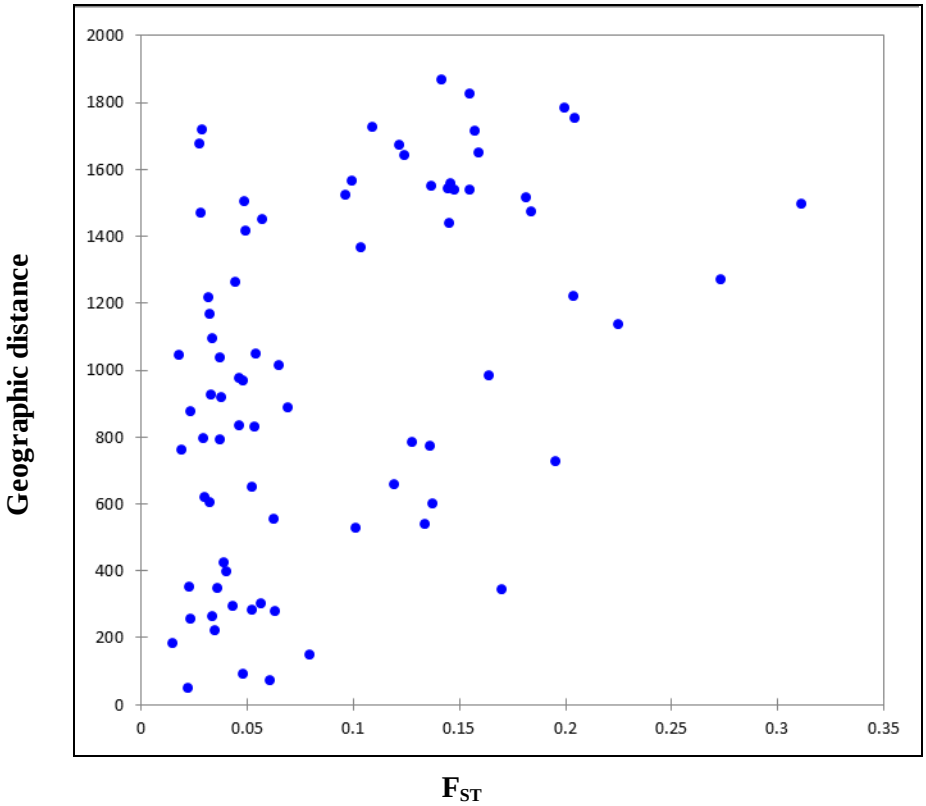


Figure 5. Mantel test showing correlation between F_{ST} and geographic distance

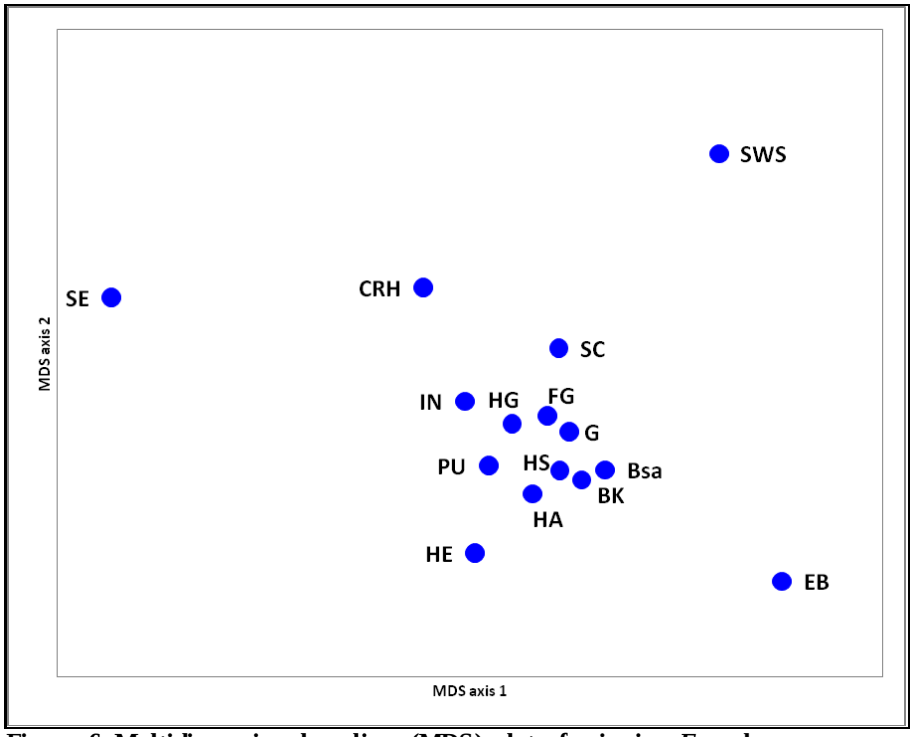


Figure 6. Multidimensional-scaling (MDS) plot of pairwise F_{ST} values among populations, labels are according to Table 1 and Figure 2.

Invasion history was therefore given a rank number (based on Gozlan 2002, 2010). Significant variation in *He* was explained by the invasion history expressed as rank numbers based on date of first sighting in a country (Regression: R^2=44.0%, $F_{1,13}$=10.23, p=0.007). This analysis showed that 44% of variation in gene diversity is due to the invasion history (Figure 7).

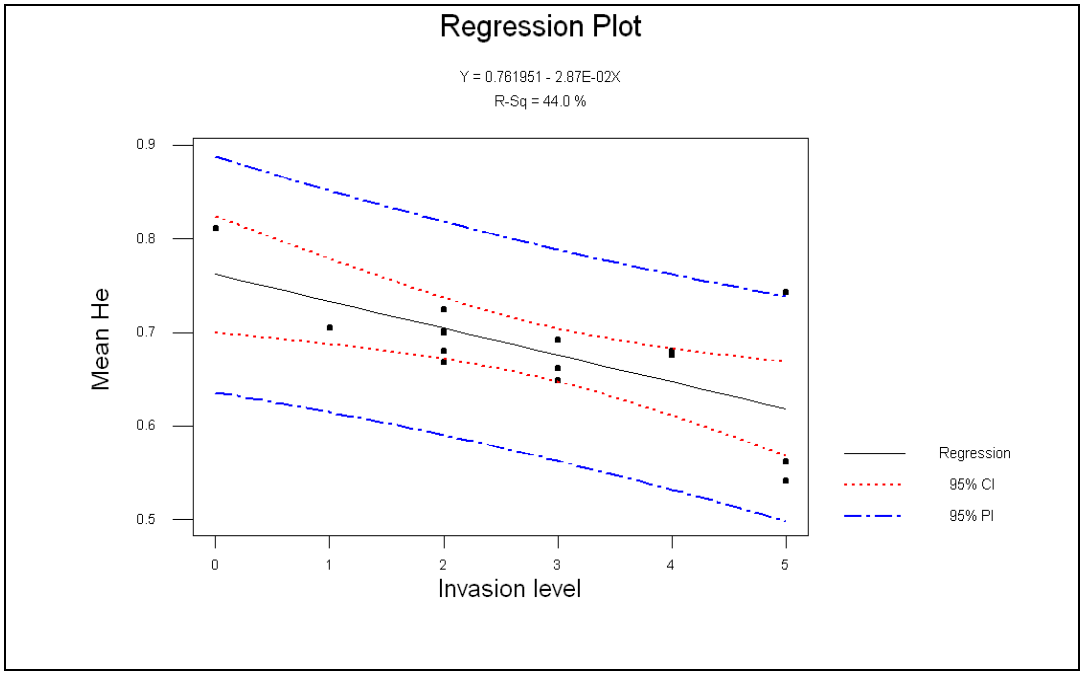


Figure 7. Regression analysis of mean *He* versus the invasion level, 95% Confidence interval, 95% Prediction interval (based on Gozlan 2002, Gozlan 2010).

In the MDS plot most European populations cluster together, except for the two UK populations (EB, SWS) and the Spanish population (SE). Two of these populations (SE and EB) showed the lowest level of gene diversity (H_e=0.54 and 0.56 for the SE and the EB populations respectively), which seemed to indicate that these populations are severely bottlenecked,

However, the SWS population had the highest gene diversity in Europe (H_e =0.74), and the H_e of this population was in fact close to that found in the Chinese population (H_e =0.81). While this is likely to be a result of multiple introductions, it has to be noted that there is a large difference between H_o and H_e in this population - H_e can be overestimated if this discrepancy is a result of stuttering, null alleles or large allele dropouts (DeWoody *et al*

2006; Bjöerklund 2005).The bottleneck test however, only showed the presence of bottleneck in the SWS population.

Although the genetic diversity of nuclear and mtDNA markers was positively correlated, this relationship was not significant when tested with a regression analysis ($F_{1,13}$=2.49, p=0.139) (see Figure 8). A linear regression analysis between the F_{ST} estimates based on mtDNA and microsatellite markers showed a significant positive relationship between the divergence estimates of both markers

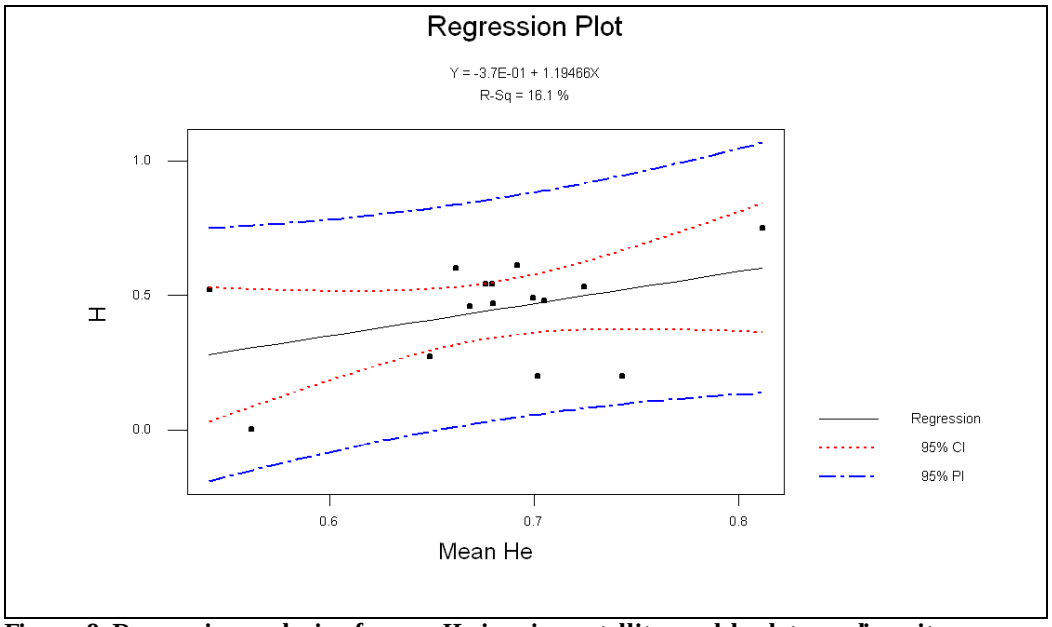


Figure 8. Regression analysis of mean H_e in microsatellites and haplotype diversity (H) in mtDNA

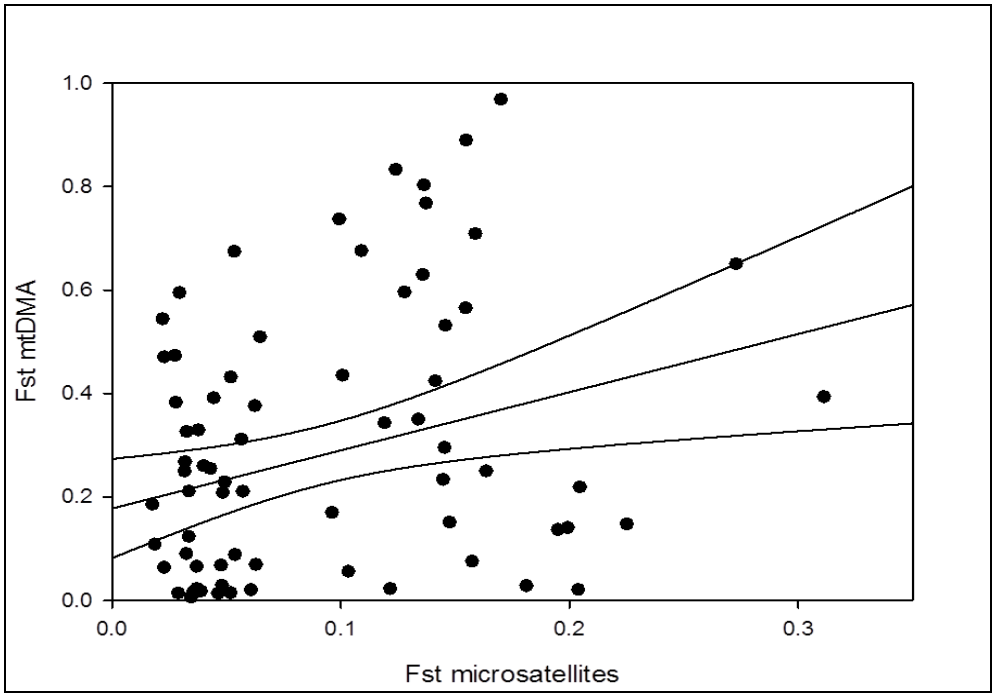


Figure 9. Linear regression analysis between the F_{ST} estimates based on mtDNA and microsatellite markers. There is a significant positive relationship between the divergence estimates of both markers (Regression: R2=8.3%, F1,76=6.85, p=0.011).

3.4 Discussion

I analysed five microsatellite loci of 14 European and one Chinese *P. parva* populations. The data available did not make it possible to confirm whether the Chinese population studied here is the actual source population of the invasion in Europe but it is one of the populations that contributed to several of the European populations.

3.4.1 Genetic signal of invasion history

The level of gene diversity and genetic differentiation across an invasive range is affected by two processes. Firstly, populations closest to the founder site will have the highest level of gene diversity (and lowest level of genetic differentiation) (Kang *et al* 2007) because they are expected to have experienced fewer consecutive bottlenecks or founder events (Genton

2005). These populations are therefore expected to be least affected by random genetic drift, and hence, they should be genetically most similar to the original source population. In addition, such populations are predicted to show the highest level of genetic similarity to other invasive populations that have been founded through these initial "bridgehead" populations (Facon *et al* 2010; Handley *et al* 2011).

The second process is the effect of novel mutations that can replenish the gene diversity (Amos *et al* 1998; Prospero *et al* 2011) that is inevitably lost during a founder event. Populations that were founded first are older and have had more time to regain new gene diversity through mutations. Both processes predict that the earliest founded populations, and populations closest to the founder site should have the highest level of gene diversity (McCauley *et al* 2003), and show the lowest level of genetic differentiation.

The empirical data from this study are largely consistent with these theoretical predictions and overall, gene diversity and the level of genetic differentiation appear to be well explained by geography and invasion history.For example, populations that are closest to the founder site in Hungary show high levels of gene diversity and little genetic differentiation. In the F_{ST} and isolation by distance analysis (Rousset 1997; Hardy 1999; Broquet 2006), a clear pattern of Isolation-by-Distance was found, which suggest an equilibrium and indicates that the invasion was relatively recent and that the gene pools in Europe have not yet had chance to homogenize by on-going migration and mutation, this pattern is consistent with low migration rate. Indeed, support for such "time-lag" was found when I

explained variation in expected heterozygosity (H_e) by invasion level. This showed that 44% of the variation in H_e was explained by the invasion history of the species but other factors, such as previous bottlenecks and the number of introduction events have also influenced variation in H_e.

The number of private alleles was between 0 and 5 in all the populations, except for the CRH population (19 private alleles). Although the number of distinct alleles is affected by sample size (Szpiech *et al* 2008) the low number of private alleles in these samples also indicates recent migration and gene flow between these populations.

The STRUCTURE (Figure 3) analysis suggests the admixed nature of the European populations. This indicates that the Chinese populations have been admixed a long time ago and that the linkage disequilibrium caused by this admixture has been eroded over time. In contrast the admixture history of the European populations is much more recent and a signal is still present in the nuclear genome of individuals. However, in some populations the admixture signal has been overridden by strong genetic drift: all bottlenecked populations show low levels of admixture and are in contrast assigned to a separate cluster. This is consistent with the late colonisation of these habitats, and reflects the effects of serial bottlenecks during the "stepping-stone" colonisation.

Also consistent with the population genetic theory on invasion history is the observation that two of the most recently founded populations (EB and SE) are characterised by low gene diversity, and that they are significantly differentiated from the other European populations, as also shown by

STRUCTURE. In addition, because the EB and SE populations are relatively recent, novel mutations have not had the opportunity to restore the gene diversity in these populations.

3.4.2 Evidence for multiple invasions of P. parva *in the UK*

In contrast to theory there is one notable exception, the SWS population in the UK. Unlike the other two recently invaded populations (EB and SE), the SWS population was a marked outlier in both the MDS analysis and the regression analysis of gene diversity versus invasion history; it showed signs of a bottleneck and had only one private allele as a result of recent introduction. Although this population is also recently established, its level of gene diversity is the highest of all European populations analysed in this study.

P. parva is thought to be introduced into the UK from Germany in the late 1980s. The introduction site was the River Test in Hampshire (Pinder and Gozlan 2003), however, according to the literature (Britton *et al* 2006), only one introduction took place and the species spread from this introduction. Given that this population is also an outlier in the earlier Regression analysis of mean H_e versus the invasion level (see Figure 4), this current analysis suggests there may have been another invasion from a genetically distinct (and diverse) source population. I propose that this population may have been founded by multiple introductions from distinct Chinese source populations that are distinct from the original *P. parva* invasion that

colonised the rest of Europe but with contributions from the already diverse German population.

3.4.3 What facilitates successful invasion?

Genetic diversity is an important factor that may affect invasion success and together with ecological traits it affects the outcome of invasions, and explains why some invaders become successful after a few generations (Kirk *et al* 2010; Therriault 2005). *P. parva* has a very short generation time of 2 years (Rosecchi *et al* 2001), this has been previously suggested as one of the possible reasons as to the success of the species as an invader (Musil and Adamek 2007). The high level of reproduction means *P. parva* is able to quickly gather higher levels of heterozygosity through novel mutations. It is possible that other species that fail to establish in new environments have a longer generation time which means they require a longer time to produce the same number of generations (Dukes and Mooney 1999) and species with brief generation time are able to invade new areas, like the mosquitofish (*Gambusia holbrooki*) (Alemadi and Jenkins 2008) . Also its small size and ubiquity may help to increase its invasion potential, because habitats may be invaded by multiple (genetically distinct) source populations, fuelling genetic variability and facilitating novel evolutionary adaptations (cf. Fisher's Fundamental Theorem, Fisher 1930).

3.4.4 Incongruence between markers

There appears to the some incongruence between the microsatellite and mtDNA diversity, a regression analysis between the diversity of both marker loci (Figure 8) showed a statistically non-significant positive association. This phenomenon is not unique, the difference between the two markers in recombination, effective population size, mutation rate means that often the two markers do not show corresponding results. As mtDNA is used for resolving phylogeographic relationships and microsatellites are utilised to reveal processes that affected the genetic structure of populations more recently. Similar inconsistencies between the two markers have also been observed in the Carpathian red deer (Feulner *et al* 2004), Northern fur seals (Dickerson *et al* 2010) and the whitefish (Lu *et al* 2001). These inconsistencies may be due to the low power, given there are only 14 data points or can be a result of differing mutation rates between the markers (Lu *et al* 2001), as well as the differing modes of inheritance. The results yielded from these markers may also depend on the effective population size (Moore 1995). Another possible interpretation is that single markers, such as mtDNA genes, might be more sensitive to stochastic variation, highlighting the need for using multiple independent markers.

3.5 Conclusion

The introduction of *P. parva* into Europe was thought to have happened through a single introduction from an admixed population in China. As predicted from population genetic theory, I found that the populations'

genetic diversity and differentiation is strongly correlated to the spread of the species across its invasive range in Europe, as shown by the initial high levels of diversity in the populations HA and HE populations. However, the SWS population appears to have been founded by a genetically distinct source population, which could explain the extremely high level of gene diversity and genetic differentiation of this population and it is further evidence for the significance of multiple introductions and the role of genetic diversity in biological invasions.

CHAPTER IV

Mitochondrial DNA analysis of population structure of rainbow trout (*Oncorhynchus mykiss*) in its native and introduced range

4.1 Introduction

The introduction of species beyond their historical range by humans, whether intended or accidental, is extremely widespread (Richardson *et al* 2000) and global trade expansion has greatly increased the number of invasive species (Pysek *et al* 2010). Apart from the economical and ecological consequences, the process of invasion can introduce novel selective pressures, even if the new environment is similar to the native environment of the invader. Successful invasions might be the result of the species' ability to evolve in response to this new environment (Mooney and Cleland 2001), thus understanding the role of the evolutionary factors and their influence on the outcome of the invasion is extremely important (Hänfling 2002, 2007). Processes often associated with biological invasions, such as hybridization (Huxel 1999), multiple introductions (Dutech *et al* 2010), bottlenecks, genetic drift (Dlugosch and Parker 2008) can all increase genetic variation and thus play a role in the establishment of new species. A prerequisite of studying such processes is a sound understanding of the phylogeographic structure of the native and introduced populations which can help us identify source populations and colonisation pathways of invasive species. While many introduced species fail to invade, others are able to establish in new areas which show marked environmental differences to their native habitat and lack their recent evolutionary context (Verhoeven 2011), sometimes even outcompeting local species (Gevrey *et al* 2006; MacDougall and Turkington 2005; Dukes and Mooney 2004).

By their very nature successful invasions are easy to observe, while unsuccessful introductions often go unnoticed. As a result, most empirical studies and meta-analyses have previously concentrated on analysing data of successful invasions (Levine *et al* 2004; Gaertner *et al* 2004, Cornell and Hawkins 1993). It has therefore been suggested that empirical studies, which include invasions that have occurred but have not been successful are best suited to investigate what functional characteristics make a good invader and which evolutionary processes might facilitate biological invasions (Hänfling 2007; Gurevitch *et al* 2011). Ideally suited for such studies would be a single species which has invaded many areas but with mixed success. Fish might be among the best suited taxonomic groups to identify such systems, since introductions are often recorded by fishery authorities and angling clubs irrespective of establishment success. Many non-native fish species are introduced intentionally for sport fishing, or unintentionally through escaping from aquacultural facilities (Naylor *et al* 2005). The establishment of these species is then aided by their ability to utilize empty niches and changes in genetic composition (Urban 2007; Shea and Chesson 2002). One of the most popular introduced fish species in Europe is the North American rainbow trout (*Oncorhynchus mykiss*). First introduced in the late 1800s (Behnke 2002), *O. mykiss* is now an important species for freshwater aquaculture and recreational fisheries in Europe.

Figure 1. Rainbow trout, source www.epa.gov

However, despite repeated introductions to rivers throughout Europe (Copp *et al* 2005a), the species has established few self-sustaining populations in European rivers, with only five or six of these in the UK (Walker 2003). In contrast, rainbow trout has successfully colonised large freshwater systems in South America, New Zealand, Japan and on the East Coast of North America (Fausch 2007). The reasons for the low establishment success in Europe remain poorly understood, but an interaction between various biotic (competition, parasites and predators) and abiotic (temperature, flood disturbance) factors (Fausch *et al* 2001) are thought to play a role.

Rainbow trout is native to the Pacific coast of the United States from Alaska to Mexico, as well as the Pacific Ocean and the east coast of Asia. The evolutionary history of *O. mykiss* has presented biologists with an exciting research topic for a long time with regard to the ongoing debate about the possible marine or freshwater origin of salmonid fishes (Behnke 2002). The deglaciation of the Wisconsin Glacial Episode (Robinson 2005) created

large lakes in the western US (Catto *et al* 1996), greatly aiding the spread of freshwater species. This also had a significant impact on the genetic structure of both the freshwater and the anadromous forms, as due to the retreating glaciers, fish were able to inhabit new areas (Cannings *et al* 2011). The differences in these environments have enabled salmonids to adapt to the specific conditions of their particular habitat and to develop differing morphological characteristics, such as differences in colouring, spawning time, and temperature tolerance (Bagley 1998).

Currently, six subspecies of rainbow trout are recognised: Columbia River red band trout (*Oncorhynchus mykiss gairdnerii*), California golden trout (*Oncorhynchus mykiss aquabonita*), Kern & Little Kern golden trout (*Oncorhynchus mykiss whitei*), Sacramento red band trout (*Oncorhynchus mykiss stonei*), Coastal rainbow trout (*Oncorhynchus mykiss irideus*) and the Kamchatka rainbow trout, as well as the anadromous form of the species, known as the steelhead (Docker and Heath 2003).

According to Behnke (2002), the first rainbow trout propagated in the United States around 1871 came from diverse sources around the San Francisco Bay area – using both resident rainbows and steelheads. In 1875 these fish were shipped to Northville, Michigan and to Caledonia, NY in order to establish brood stock, followed by McCloud River rainbows from California in 1878. During the 1880s the two stocks were maintained separately, as two different populations but later on mixing of the two stocks occurred. In Europe, rainbow trout was first introduced at the end of the 19th

century, the first shipment of the stock to the UK took place in 1885 and into the River Wye around 1910 (Palmer 1996).

Since then, the rainbow trout has been introduced in 97 countries (Fausch 2007), making it the most widely introduced fish species in the world (Fausch 2001). In the UK, the species has been repeatedly introduced but despite various introduction programmes and constant restocking it has rarely been able to establish in UK waters (Fausch 2007).

One of the possible reasons that prevents the species from becoming established in Europe is thought to be its high susceptibility to whirling disease (WD). WD is caused by the myxozoan parasite *Myxobolus cerebralis*, a parasite of the native brown trout (*Salmo trutta*). It interferes with the skeletal development of fingerlings, resulting in malformation and mortality (Staley 2000). Native rainbow trout populations in the United States declined rapidly after exposure to the parasite, which was thought to be introduced through the import of infected frozen fish (Hoffman 1990, Hedrick *et al* 1999). However, a strain of rainbow trout which was imported into Germany from Colorado in the 1880s was recently observed to have a natural resistance to the disease, which was confirmed by various exposure experiments. These experiments found myxospore concentrations similar to that in brown trout and stocks with up to 100 fold less spores than other, susceptible strains (Hedrick *et al* 1999, Hedrick *et al* 2003). Known as the 'Hofer strain', this strain was then reintroduced into Colorado, offering an opportunity for further studies on breeding disease resistant fish. A broodstock programme was started to cross these highly resistant hatchery

rainbows with natural populations to produce a resistant strain that also possesses the necessary physiological characteristics to survive in the wild (Schisler 2006, Fetherman *et al* 2011) Because of its unique resistance to WD the Hofer strain has also been proposed as a source for the few established European populations (Schisler 2006).

The aim of this study was to use a molecular based phylogeny to identify the sources of the European rainbow trout populations and to investigate the population structure and distribution of haplotypes. Specifically I predicted that established European populations originate from the same source and therefore share the same evolutionary history, and that this source is the WD resistant German Hofer strain. This will show whether evolutionary factors in general and disease resistance in particular is a factor in establishment success.

4.2 Materials and methods

Mitochondrial DNA was analysed to identify the possible origins of the European populations and to understand whether the Hofer strain is indeed the original introduced population from which other European populations have derived.

A total of 20 sites were sampled in the native and introduced range of the species and a number of Genbank sequences were also added for wider geographical coverage. All sequences, including the ones from Genbank were pooled into five groups: (1) California wild (2) Western US (Nevada, Idaho, British Columbia, Washington), (3) Eastern US (Maryland, West

Virginia), (4) European wild and (5) European hatchery populations (Figure 8), although in the sample map and the analysis graphs I also indicated the wild and hatchery populations in the United States.

DNA was extracted from the dorsal fin with Puregene DNA extraction method (Qiagen) and an approximately 700bp long section of the D-loop section of the mitochondrial genome was PCR amplified to examine the genetic structure of rainbow trout in its native range in the US and in its introduced range in Europe. D-loop is a rapidly evolving part of the animal mitochondria, which makes it a suitable marker for resolving relatively recent evolutionary history (Sumida *et al* 2000). Amplifications were performed in 15µl reactions, containing 1.5µl 160 mM NH4, 1.5µl 100 mM dNTPs, 0.4µl 50 mM MgCl2, 0.075µl Taq polymerase, 0.3µl each of primers Vio-F (5'-TTACCGGCCCTCTTAACCTT) and TroutR2 (5'-TTTTCTTTTCCTTTCAGCTTGC), 9.425µl PCR water and 1.5µl of template DNA. I designed the primers Vio-F and TroutR2 using the primer design programme Primer3 (Rozen and Skaletsky 2000). Cycle parameters were as follows: 2 min at 95°C; 35 cycles of 45 s at 94°C, 45 s at 48°C, 1 min at 72°C; 10 min at 72°C using a Verity Thermal Cycler (Applied Biosystems). PCR products were visualised on 1.5% agarose gel stained with ethidium bromide. PCR products were directly sequenced in the forward direction by Macrogen Inc. Sequences were aligned and edited using CodonCode Aligner v3.7.1. A BLAST search of nucleotide sequences was performed to confirm that all sequences belonged to *O. mykiss*.

In order to widen the geographical coverage of native populations, the following published Genbank sequences were included: AF044130, AF044132, AF044136-AF044145, AF044149, AF044151-AF044155, AF044158-AF044163, AF044165-AF044167 (Bagley 1998), AF312568-AF312570, AF312576, AF396658, AF396659 (Docker and Heath, 2003), AY032629-AY032633, (Brown 2002), HM229292-HM229295, HM229297-HM229299, HM229301, HM229302, HM229306-HM229308, HM229311-HM229313, HM229315-HM229319, HM229322-HM229325, HM229327-HM229330 (Brunelli, 2010).

Consensus sequences from the present study and GenBank sequences were imported into MEGA v. 4.1 (Tamura *et al* 2007) and aligned with ClustalW. After alignment, a 421bp part of the sequences were used, which made it possible to include shorter Genbank sequences as well.

A haplotype network was constructed using a median-joining algorithm in Network v. 4.5.10 (Bandelt and Röhl, 1999, Appendix XII), including a cutthroat trout (*Oncorhynchus clarkii*) sequence to be used as an outgroup (Figure 7). To create the network all sequences were cropped to equal lengths. Positions showing high mutation rates (positions 38, 123, 258, 293, 314, 324, 358, 378, 381, 415, 419, 421) were weighted at 5 (the default weighting is 10) and I used an e value of 0. The epsilon (e) parameter in the median-joining algorithm builds a sparse network when set to the default value of 0 or other small values - this is important when running large datasets, as a low value of epsilon can cut the run-time significantly.

Code	Population	N	Co-ordinates	H	Π
	California				
AM	American River	1	38°38'02.43"N 121°13'32.08"W	not included	
CC	Corralitos Creek	4	36°56'05.00"N 121°44'30.00"W	1	0.00974
CO	Coleman	4	38°38'20.67"N 121°13'36.74"W	0.5	0.00121
EL	Eagle Lake	4	60°36'30.45"N 151°00'03.49"W	0.5	0.00121
HC	Hot Creek	8	37°40'25.42"N 118°50'29.78"W	0.75	0.00355
KA	Kamloops Junction	5	50°44'43.60"N 120°42'26.47"W	0.9	0.00339
PC	Pine Creek	6	40°40'29.99"N 120°51'10.71"W	0.93333	0.00469
PP	Peppermint Creek	6	37°58'18.00"N 120°25'39.00"W	0.93333	0.00697
RW	Whale Rock Reservoir	1	36°26'59.99"N 120°52'27.31"W	not included	
SA	Sacramento	1	41°22'35.00"N 122°28'00.00"W	not included	
SC	Scott Creek	3	37°02'30.48"N 122°13'37.02"W	1	0.013
SH	Shasta	3	40°46'12.00"N 122°17'47.17"W	0	0
TH	Thomas Creek	3	41°21'55.00"N 121°56'40.00"W	1	0.00485
VC	Volcano Creek	3	36°40'25.42"N 118°51'30.78"W	1	0.0081
	Western US				
BC	British Columbia	5	52°38'60.00"N 122°50'00.00"W	1	0.0068
ID	Idaho	6	44°41'03.00"N 114°02'22.00"W	0.6	0.02802
NV	Nevada	4	41°30'22.00"N 116°23'15.00"W	0.5	0.00121
OR	Oregon	6	42°09'59.00"N 124°21'47.00"W	0.93333	0.00615
WA	Washington State	3	47°46'06.00"N 117°27'08.00"W	1	0.00483
	Eastern US				
APH	Maryland	5	39°35'00.00"N 77°35'00.00"W	0	0
LA	Maryland	3	39°41'00.00"N 77°35'00.00"W	0.66667	0.00161
ML	Maryland	3	39°32'00.00"N 79°24'00.00"W	0	0
VH	West Virginia	3	38°17'46.66"N 80°09'00.20"W	0	0
W	West Virginia	4	38°17'46.66"N 80°09'00.20"W	0	0
	Europe				
A	Liechtenstein	3	47°08'29.31"N 09°31'03.61"W	0	0
B	Brow Well Fisheries	5	54°03'53.68"N 01°59'27.25"W	0.7	0.00291
BB	Bibury Trout Farm	9	51°45'33.66"N 01°50'18.41"W	0.55556	0.00148
C	Liechtenstein	4	47°08'29.31"N 09°31'03.61"W	0	0
CF	Chirk Fisheries	15	52°55'56.47"N 03°05'51.62"W	0	0
DR	Drummond Fish Farm	2	56°23'34.04"N 04°14'27.76"W	1	0.00485
E	Liechtenstein	5	47°08'29.31"N 09°31'03.61"W	0.6	0.00145
H	Hofer, Germany	9	48°17'00.00"N 12°47'00.00"W	0.80556	0.00377
WF	Wild fish, Germany	12	48°17'00.00"N 12°47'00.00"W	0.5303	0.00128
YR	River Wye	9	53°13'33.68"N 01°42'57.23"W	0.55556	0.00134

Table 1. Code, name, number of sequences analysed, geographic coordinates, haplotype diversity (H) and nucleotide diversity (π) of all sample sites analysed in the present study. Hatchery populations are indicated in bold.

In the final network, to correct for events that are less likely to happen, transversions were weighted 3x as high as transitions. After creating the network, the MP option was applied to identify unnecessary median vectors and these links could be switched off to simplify the network.

In order to examine variation within and among populations, I looked at genetic diversity within populations, calculating haplotype (H) and nucleotide diversity (π) using DNaSP v.5.10.1. (Rozas *et al* 2003).

This programme was also used to calculate pairwise genetic differentiation as F_{ST} and D_{XY} (using Kimura two-parameter method; Kimura 1980) between native and introduced populations (Figure 10, Appendix X, Appendix XI). F_{ST} comparisons were performed in ARLEQUIN to test for significant population genetic structure within *O. mykiss* (Excoffier *et al* 2005), this showed significance in 86 out of 440 pairwise comparisons.

A multi-dimensional scaling analysis (MDS) based on F_{ST} values was carried out using XLStat in Excel (Addisonsoft 2008) to show the genetic relationship between the samples (Figure 11). This method simplifies data without losing much information, and is more suited than trees to describe data when there is extensive genetic exchange between close geographic neighbours (Cavalli-Sforza *et al* 1994).

4.3 Results

A total of 27 haplotypes were identified from the total data set of 167 individuals sequenced in this study (Table 1). 12 haplotypes are unique to California and one is unique to the Eastern US (H14). The California group

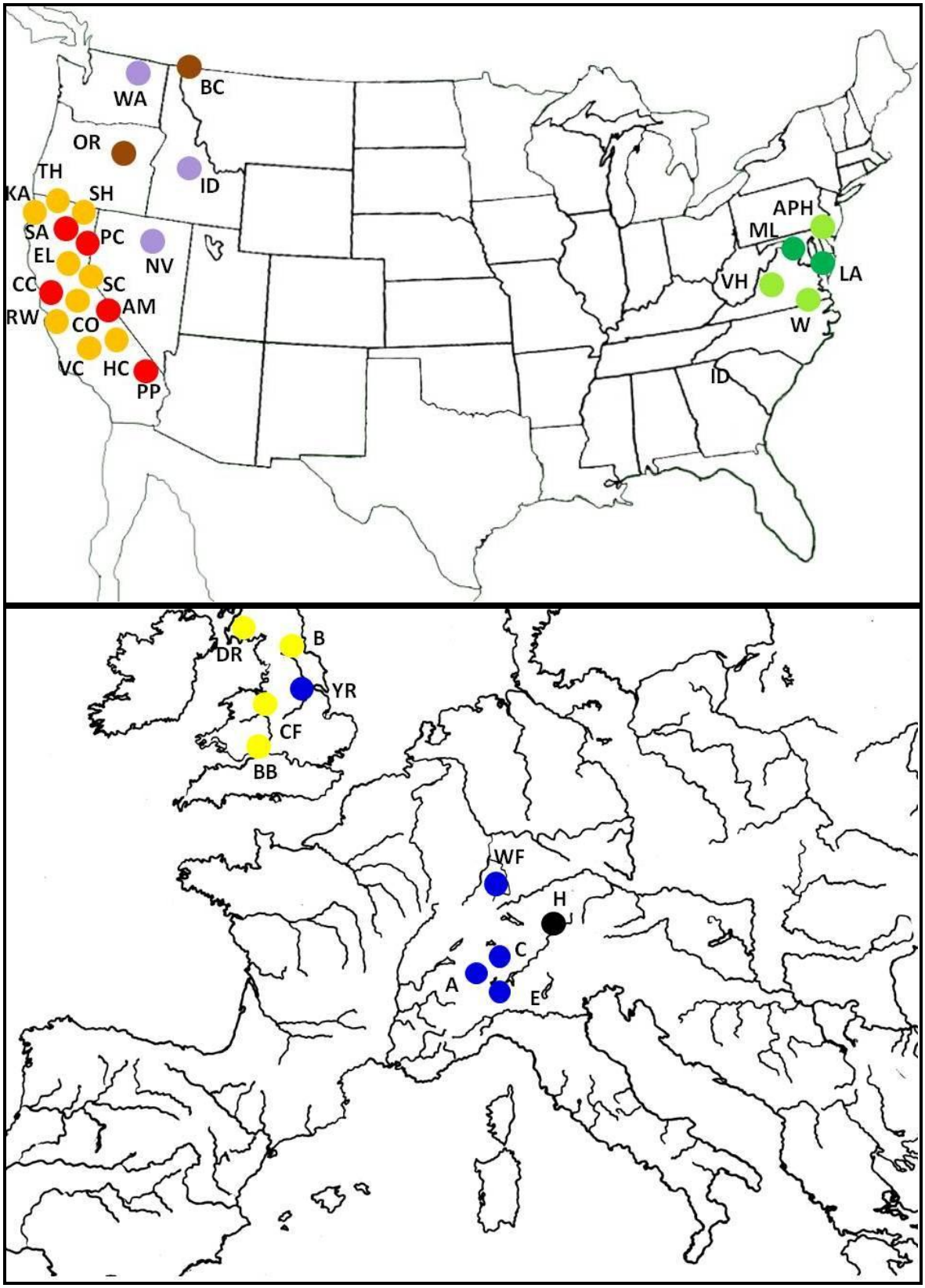


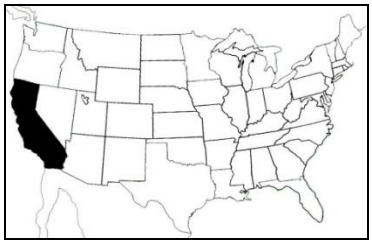
Figure 2. Map of US and Europe origin of samples CA wild CA hatch E US wild E US hatch W US wild W US hatch Eur wild Eur hatch Hofer

shares 4 haplotypes with the Western US samples, 3 with the Eastern US group and 4 with the European samples (H12, H13, H18, H23).

The two most common haplotypes are H12 and H18, these are found in 16 and 17 populations, respectively. H12 also contains all European populations, except for the Brow Well Fisheries (B) samples. Four haplotypes (H12, H18, H22, H23) were common and shared between native and invasive populations. Out of these haplotypes, the Hofer population contains three (H12, H22 and H23), sharing them with Californian, Eastern US and European hatchery and wild populations (Figure 7). In fact, the European populations only contain haplotypes also found in the Hofer population and two of the Liechtensteinian populations and a hatchery population from Wales were fixed for haplotype H12.

Genetic variation was estimated using haplotype (H) and nucleotide (π) diversities (Nei, 1989). Haplotype diversity values varied between 0-1 and nucleotide diversity between 0-0.02802.

In the grouped data, the highest level of haplotype diversity was observed in the Western US group, followed by the California group and the same was



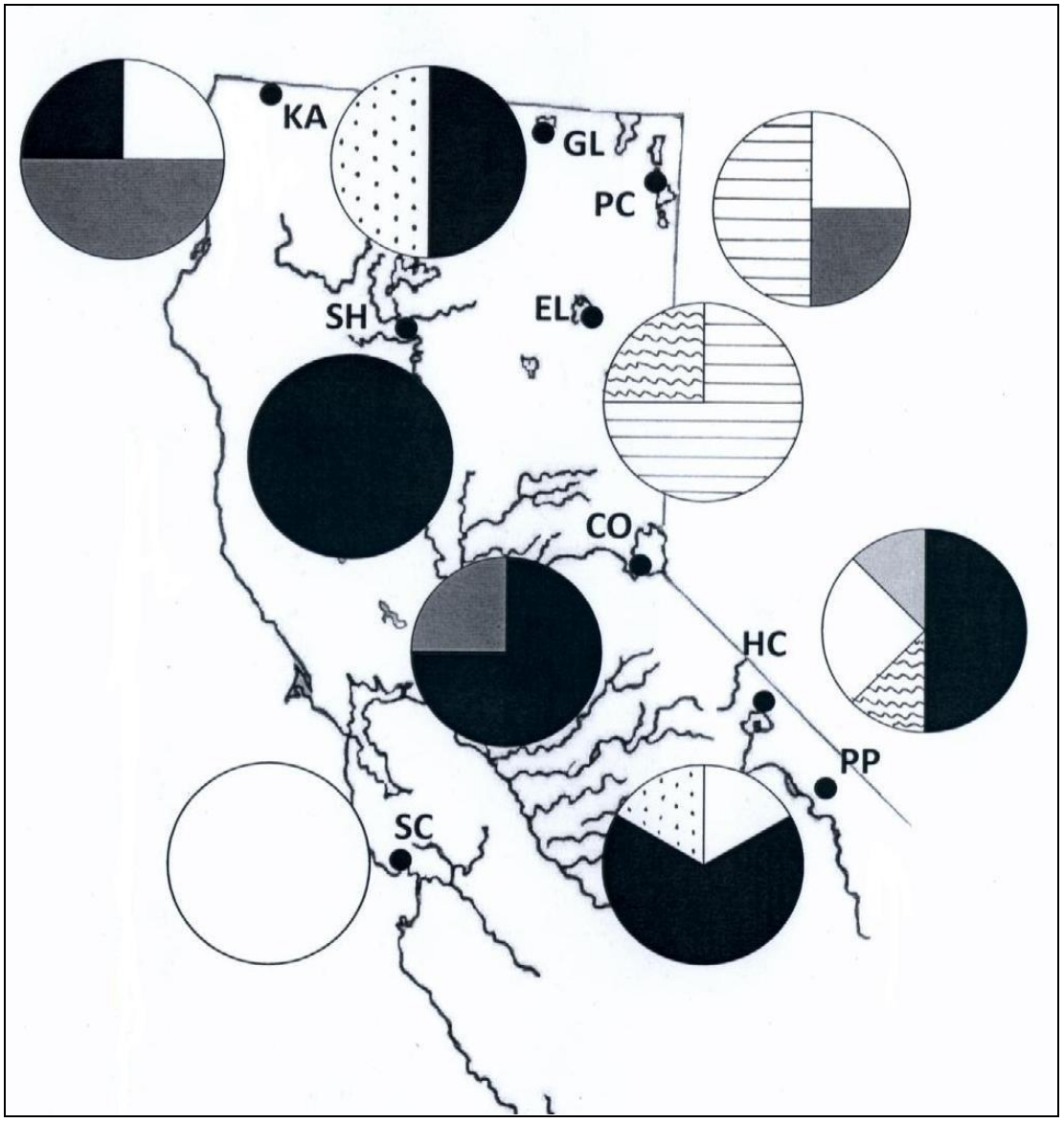


Figure 3. Sample sites and haplotype distribution of *O. mykiss* haplotypes in California

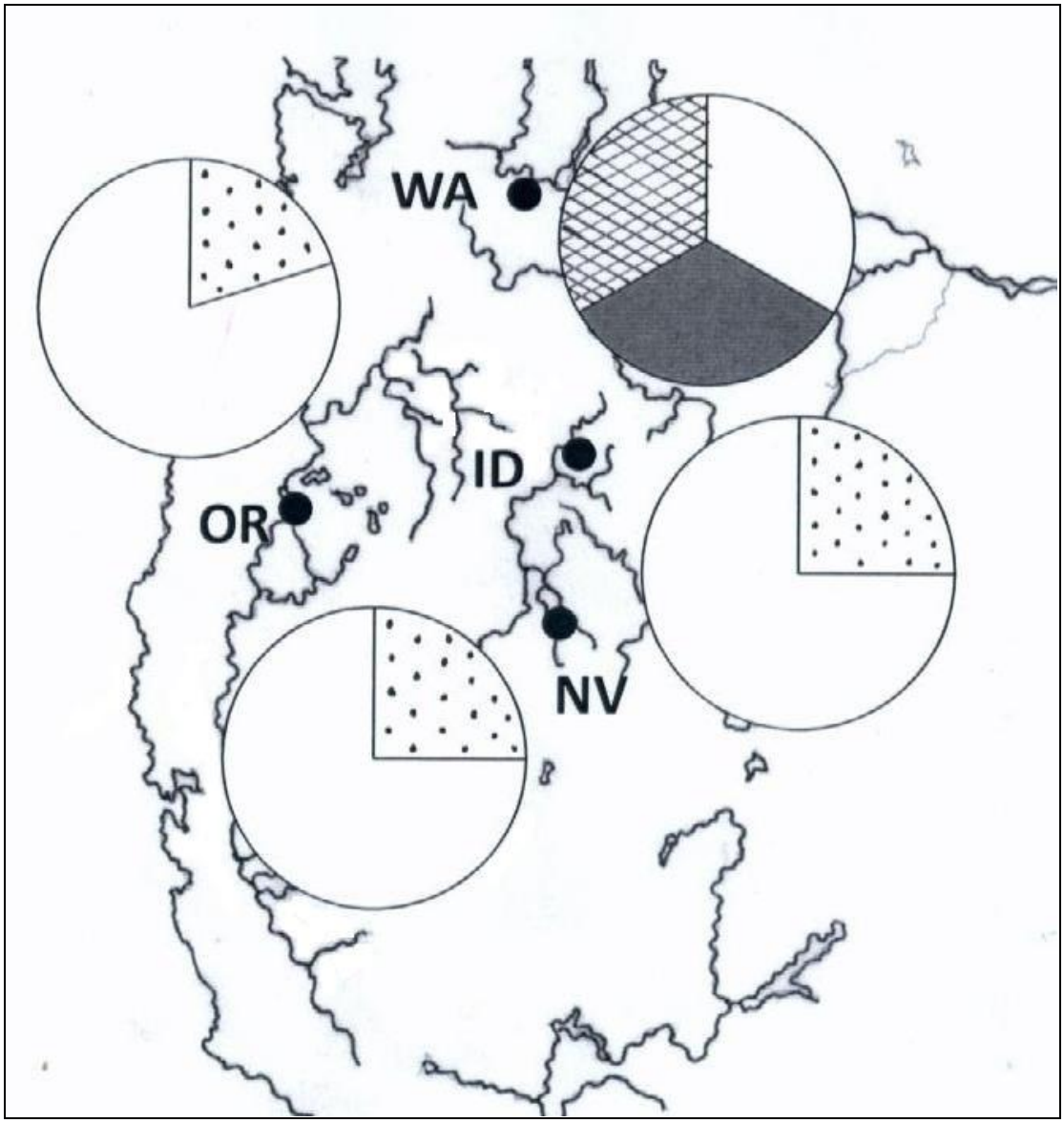
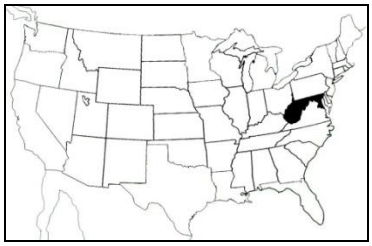


Figure 4. Sample sites and haplotype distribution of *O. mykiss* haplotypes in the North Western US



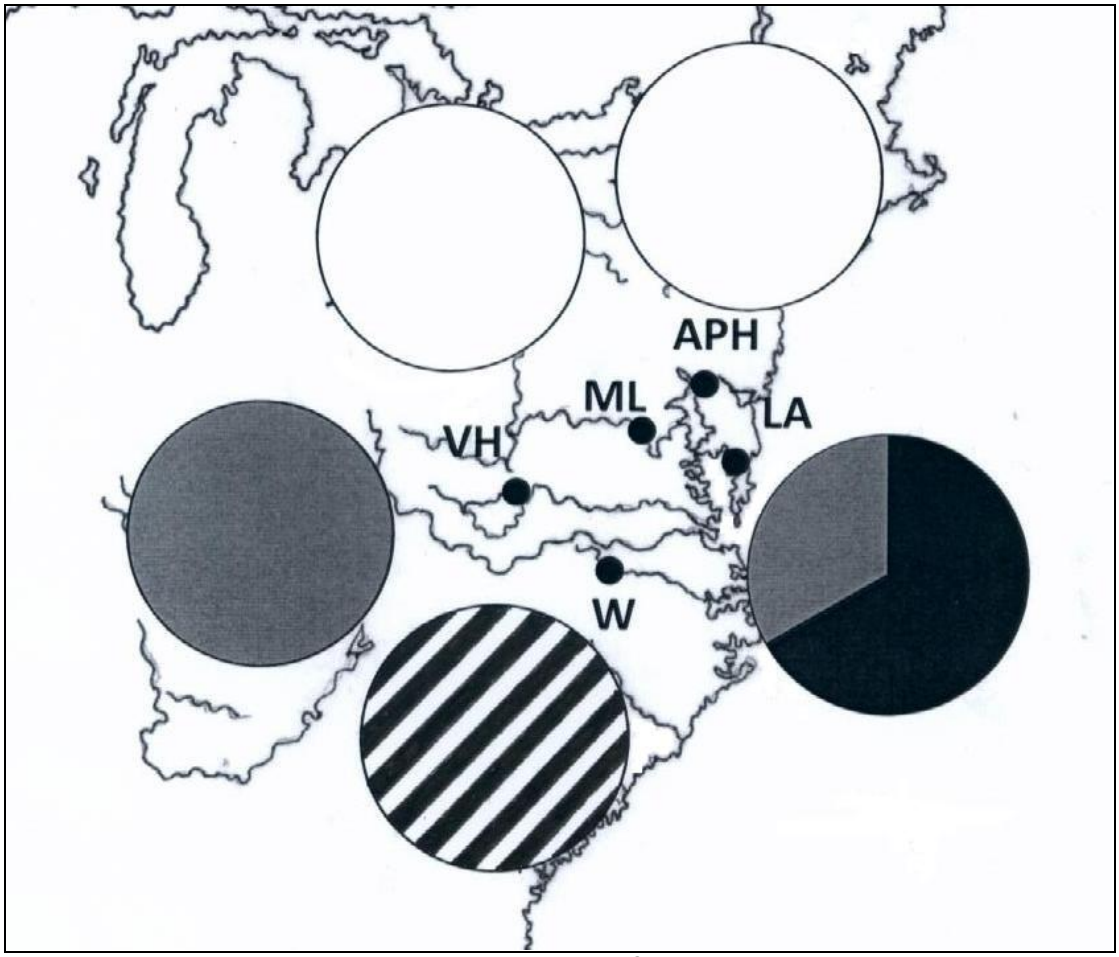


Figure 5. Sample sites and haplotype distribution of *O. mykiss* haplotypes in the Eastern US

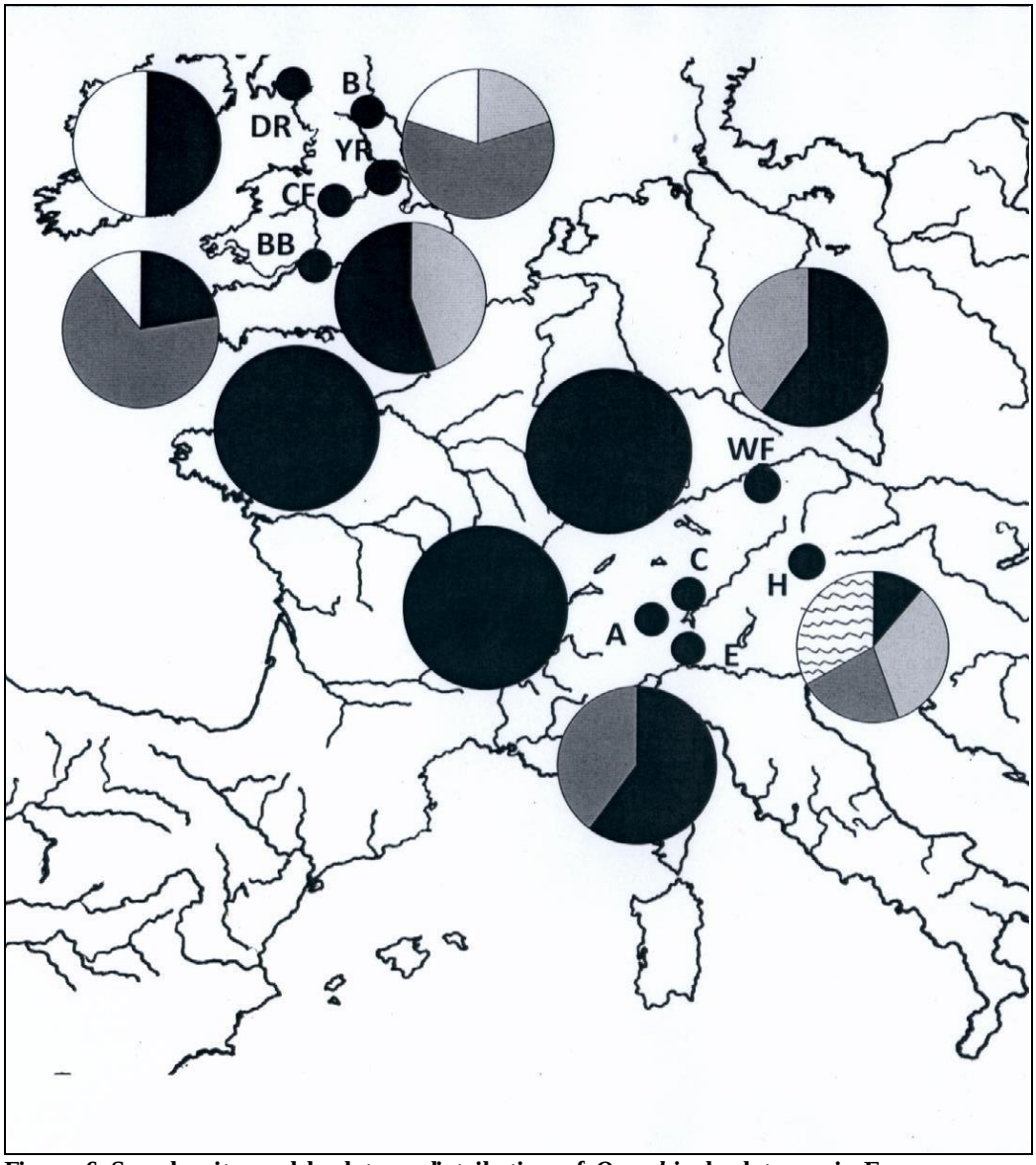


Figure 6. Sample sites and haplotype distribution of *O. mykiss* haplotypes in Europe

true for the nucleotide diversity as well. The European hatchery and wild groups, as well as the Eastern US group displayed much lower haplotype and nucleotide diversity.

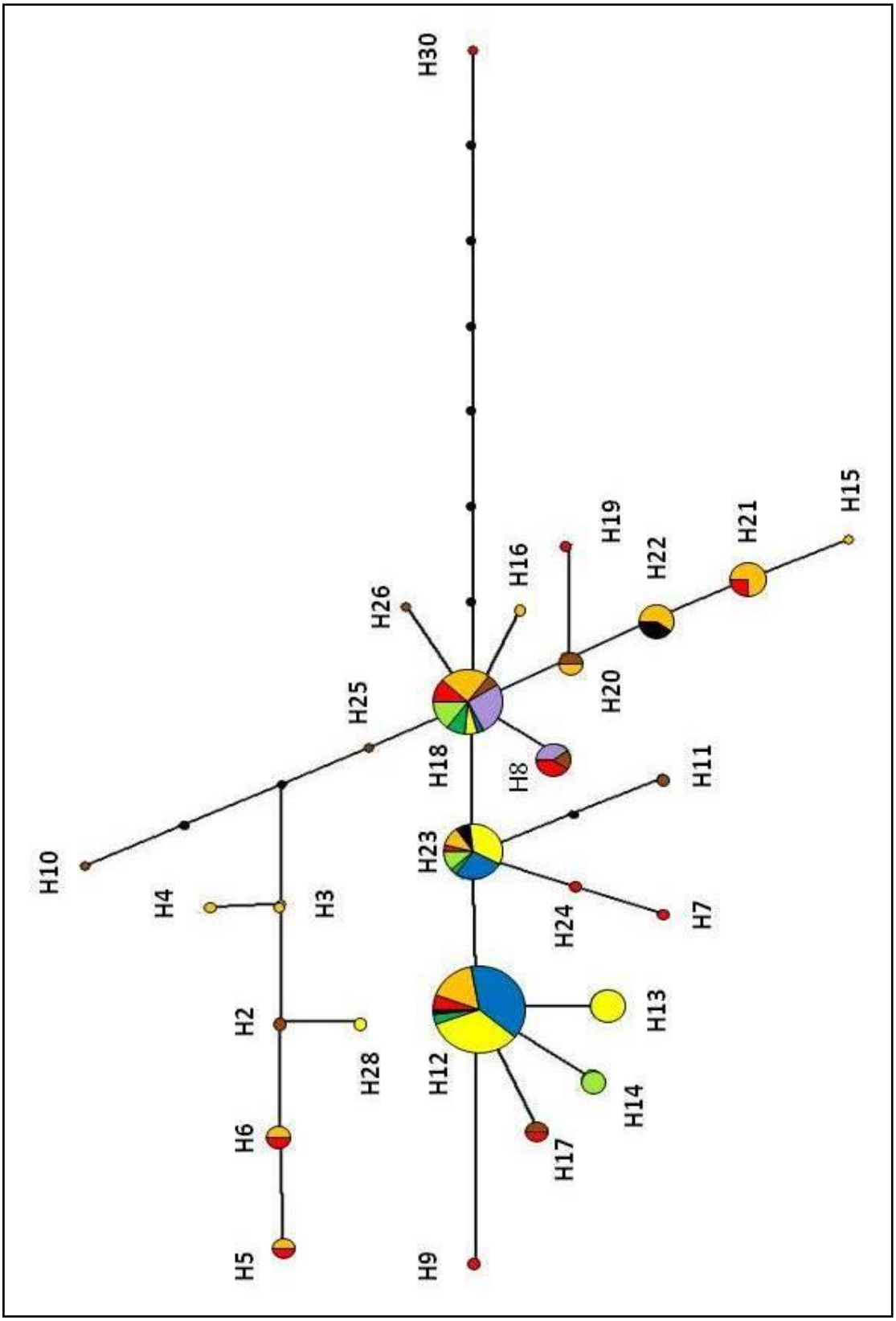


Figure 7. Median-joining network of D-loop haplotypes of native and introduced populations of *O. mykiss*. Sampled haplotypes are represented as pie charts, where the size of the pie chart is proportional to the total frequency of the haplotype and the coloured slices represent the contribution of different populations.

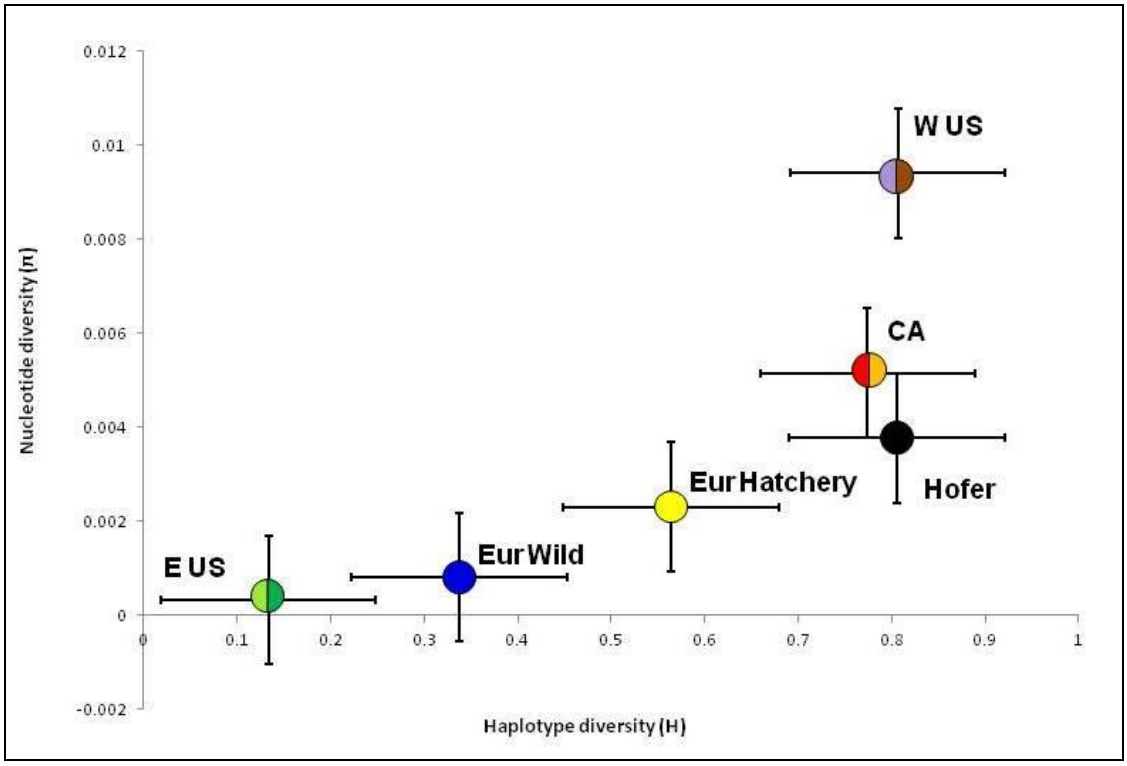


Figure 8. Plot of haplotype versus nucleotide diversity of grouped data, plus Hofer population.

On a population level (Figure 9), similarly high haplotype diversity was found in several of the Californian (SC, CC, VC) and Western US (BC) populations as well as the DR European hatchery population. By far the highest nucleotide diversity was found in the Idaho (ID) population (π=0.02802). Due to this high diversity, this population was excluded from the haplotype and nucleotide diversity graph, as this made it possible to achieve better resolution of other populations. In the Eastern US populations apart from the LA Maryland population, all other populations showed no diversity at haplotype and nucleotide level.

In continental Europe the Hofer strain is the most diverse (H=0.80556 and π=0.00376, Table 1) and two European wild populations (A, C) and one UK hatchery population (CF) showed no haplotype and nucleotide diversity.

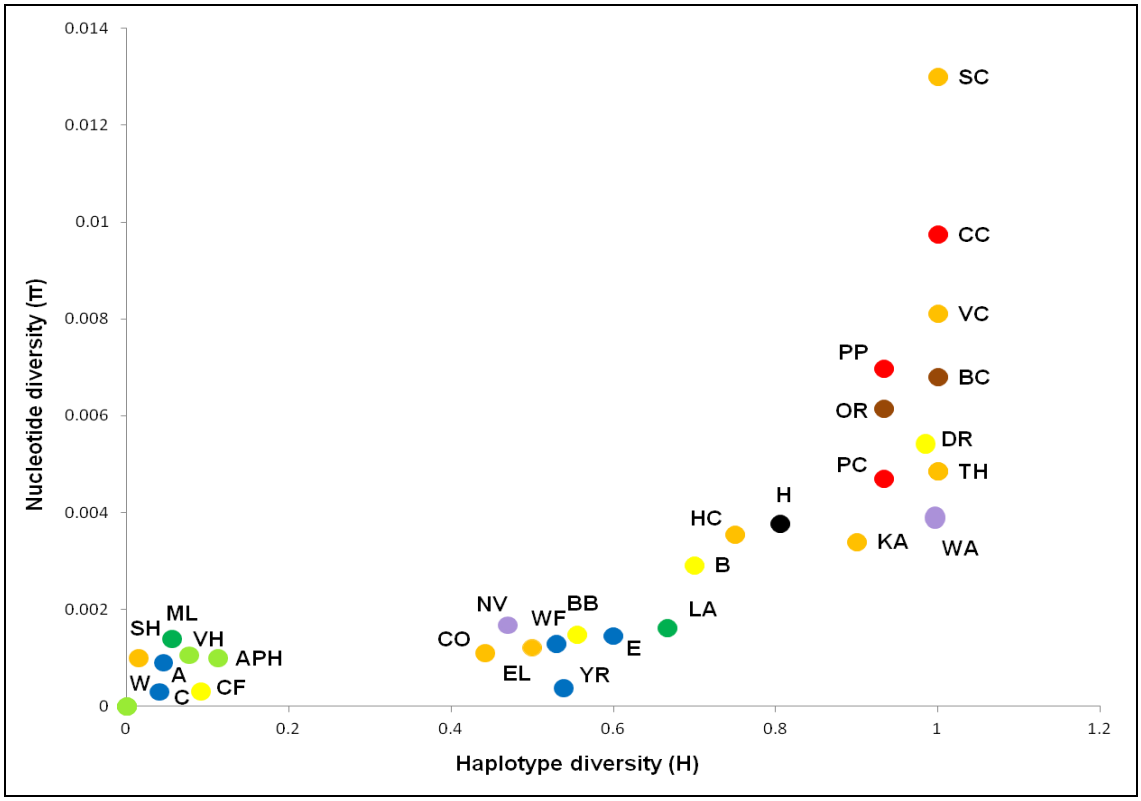


Figure 9. Haplotype and nucleotide diversity graph of all populations

Pairwise genetic distance (D_{XY}) ranged from ranged from 0 to 0.0219 and pairwise genetic differentiation (F_{ST}) ranged from 0 to 1(Appendix X, Appendix XI). Pairwise comparisons of native and invasive populations (Figure 10) showed that the pairwise genetic distance was the lowest among the European groups and the Eastern US group showed really high population structuring (F_{ST}=0.785).

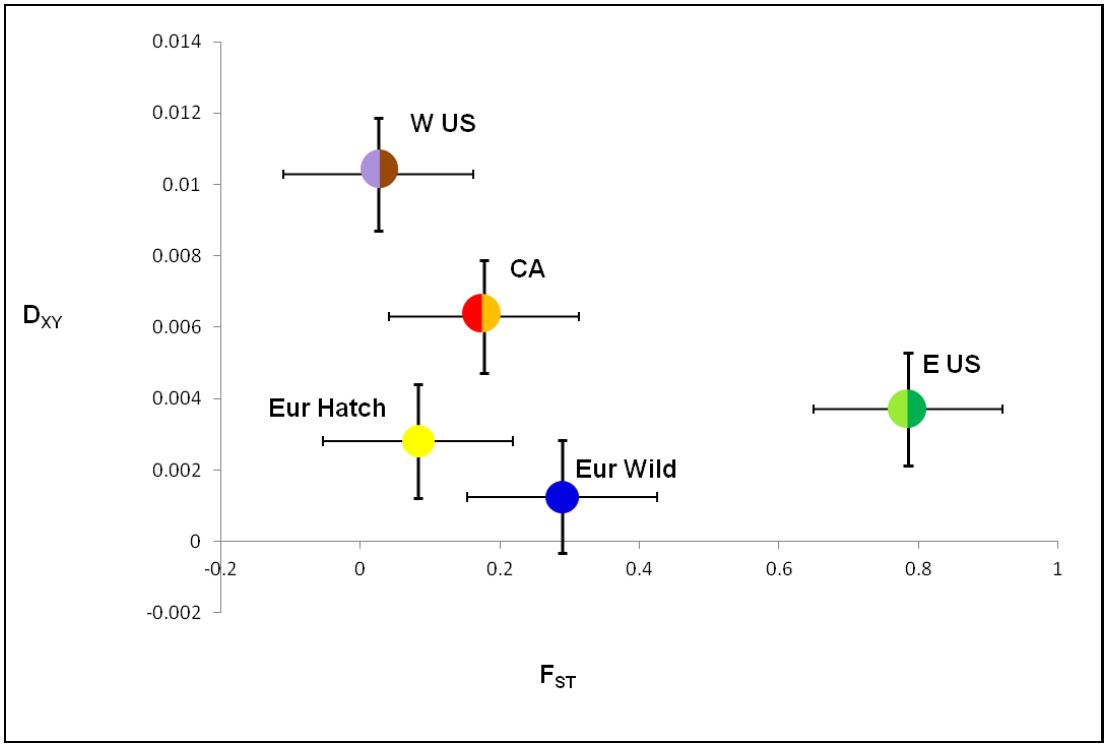


Figure 10. Pairwise comparisons of *Fst* and *Dxy* values of grouped data.

The MDS analysis based on F_{ST} values shows less scatter for the native populations than for the introduced ones and it also showed that most of the Californian populations group together and the Hofer population is relatively close to this group, whereas most European populations, apart from two hatchery populations (B, BB) and two of the European wild populations (E, WF) most European populations show more scattering. In the Eastern US populations, only the LA Maryland population groups near the centre of the plot (this was the only population to show haplotype and nucleotide diversity), all other populations are scattered near the edge of the plot.

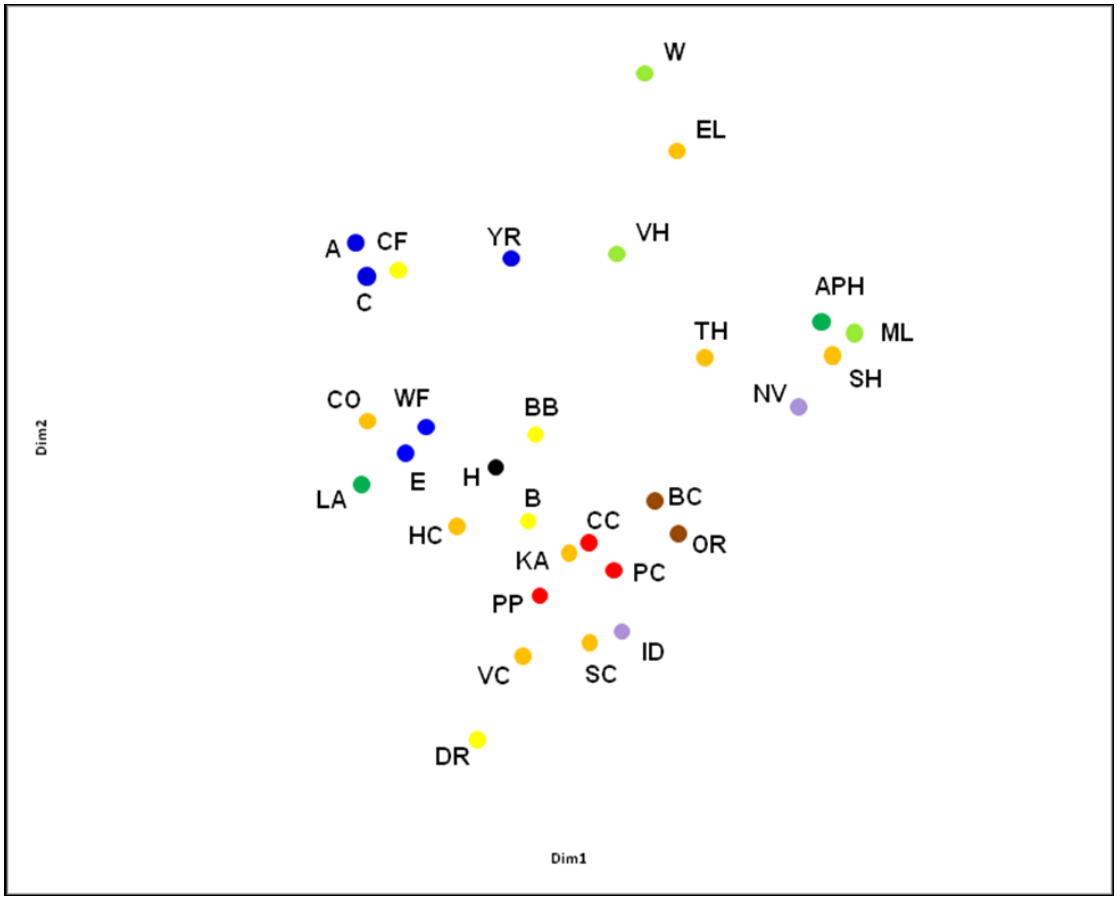


Figure 11. Plot of the first and second axis of a multidimensional scaling analysis (MDS) based on pairwise F_{ST} values among populations.

CA wild **CA hatch** **E US wild** **E US hatch** **W US wild**
W US hatch **Eur wild** **Eur hatch** **Hofer**

4.4 Discussion

Untangling the origin of *O. mykiss* is a difficult task as the species has been translocated many times to a range of environments over the past hundred years, both within its native as well as its introduced ranges. In this study I explored the phylogeographic patterns found in rainbow trout, comparing mtDNA diversity between the native and introduced ranges.

The retreating glaciers of the Pleistocene formed huge lakes, allowing the species to spread north into the Columbia River drainage, then as these lakes receded, subspecies of rainbow trout became isolated and the deeper

phylogenetic structure observed and the high levels of genetic diversity in California rainbow trout when compared to populations in British Columbia, Alaska or Kamchatka (McCusker 2000) suggested that the species has persisted longer in these areas. Research by Behnke (1992) also showed that *O. mykiss* found in the north western parts of North America form a very divergent lineage, in contrast to Californian rainbow trout, and the presence of such divergent lineages indicate that *O. mykiss* originates from California and the western Pacific coast.

This was further confirmed by the large genetic differences observed between coastal and inland groups in California (Currens *et al* 2009), as well as the differences observed in the upper Sacramento, Klamath and Columbia Rivers. Previous genetic analyses have found major subdivisions on the east and west side of the Cascade-Coast mountains (Tamkee *et al* 2010), demonstrating that the species is very diverse in its native range.

Apart from the natural diversity of the species, the effect of hatchery practices on free-living populations is not well-documented, as records of the origin of the fish planted into California during the first half of the century are scarce (Nielsen *et al* 1997) but the genetic diversity observed in these populations is likely to have contributed to introductions into local streams. Nielsen *et al* (1997) also showed based both on microsatellite and mtDNA data, that hatchery and natural populations of *O. mykiss* are closely related in California.

The high haplotype and nucleotide diversity of the Californian and Western US populations suggests the presence of admixture between different

populations in these regions, this is especially prominent in the Idaho (ID) population. All sequences from this area were downloaded from GenBank and grouped together, which means that these sequences might originate from a mixture of locations within the state, which would explain the high diversity observed in this population. Also, in this area introgressive hybridization with cutthroat trout has taken place (Weigel *et al* 2003), although a BLAST check of the sequences did not reveal any introgression in these samples.

In Europe the Drummond (DR) population was found to be the most diverse and it shares haplotypes with three Eastern US populations (APH, LA, ML). This population is stocked from several different sources (Kevin Gadsby, pers. comm.), such as Denmark, Tasmania, Norway and various hatcheries in England, as well as the Eastern US and this explains the high nucleotide diversity observed and indicates admixture in this populations. The low diversity observed in most of the Eastern Us populations might be due to that although in this area multiple introductions take place from the same sources, this in effect is only increasing the size of inocula (Roman and Darling 2007) but not the genetic diversity of these populations. Also, admixture between hatchery and stocked rainbow trout has been noted to have taken place (Cordes *et al* 2006), which could also explain the rare haplotype found in the Eastern US, alternatively the presence of haplotype H14 suggests that introductions into this area have also occurred from unsampled sources.

The divergent origin of the Drummond Farm samples is further confirmed by the position of this population in the MDS plot, as it is also split off from the main centre group, it is genetically more distant from the Californian, as well as the other European populations, However, most other populations in Europe showed low diversity, especially the wild populations, this was observed both in the grouped and population analysis, indicating founder effects and a loss of diversity in these populations.

Haplotype distribution showed that the most common haplotype is H18, found in 17 populations. The second most diverse haplotype (H12) is shared by most of the Californian populations and all European populations - with the exception of the Brow Well (B) population - suggesting that the introduction of these populations originate from California. Also, in the MDS plot, many of the Californian and Western US populations remain closely grouped spatially and most of the European hatchery populations, including the Hofer strain, together with two of the wild populations (E, WF) also group in the centre, implying that these populations are more closely related. This is further confirmed by the low F_{ST} and D_{XY} values that show low levels of population structuring in the European hatchery populations. On the other hand, there is a relatively high divergence among European wild populations, suggesting that they might originate from different sources and the wild European samples A, C and YR are scattered across the plot, further indicating diverse origins of these samples.

The Hofer (H) population showed high haplotype and nucleotide diversity, also sharing haplotypes with both two of the wild European population (YR,

WF) populations, although both these populations shows quite low levels of nucleotide diversity. The initial hypothesis of this study was that the Hofer strain is the likely source of the European populations, and in conclusion, my results demonstrate that the Hofer population originated from very diverse admixed sources before being introduced into Europe and is likely to have contributed to the European rainbow trout populations.

My findings are consistent in showing that as a particularly successful invasive population this strain had a bridgehead role in the introduction and invasion of the species in Europe. Historic evidence and these findings support the hypothesis that the Hofer strain has contributed to the European populations and is likely to be one of the sources.

The power of this current dataset is not high enough to identify the native population from which the European populations originate, however, the MDs analysis indicates that it is possible that the Hofer strain has been the source of some wild populations in particular WF and E.

It is known that the Hofer strain was originally introduced from the Gunnison River in Colorado, which consisted of fish of Californian origin, similarly to all US populations but a strain from Montana was also imported to Colorado, this strain exhibited strong resistance to WD, this thought to have occurred due to strong selection pressure in the Harrison Lake, as this population is a naturally reproducing and is exposed to the parasite. Although the original Gunnison rainbow trout population crashed but with restocking and crossing with the Hofer strain the population is recovering. The Hofer population was exposed to WD for many years after being

introduced into Germany and through artificial selection - allowing only the resistant fish to survive - the resistant strain evolved.

It has been widely assumed that introductions of *O. mykiss* took place from California. Previous records show that initially *Salmo shasta* that was being sent over form the US, which consisted purely of fish from the McCloud River in California but after 1890 it was a mixture of *Salmo shasta* and *Salmo irideus* (relating to non-migratory rainbow trout) and although the origin of this population is still open to debate (Palmer 1996), it probably provided the source for the Wye population.

In the wild populations, escapees from aquacultural facilities may have contributed to the diversity of wild populations, for example to the wild population resident in Germany (WF), as well as some introductions into the UK also taking place from the Eastern US, based on the shared haplotypes between these samples. Also, translocations of fish between hatcheries is very extensive both in the UK and in the United States, therefore it is likely that other strains from the United States have contributed to the genetic diversity of these populations.

Despite originating from such a diverse population, the European hatchery and especially the wild populations are not very diverse but loss of genetic diversity in introduced species does not mean that the species can't establish and flourish, especially if there is constant restocking on the species. While this helps the existing populations to survive, the level of continuity in restocking the species throughout makes it more difficult to resolve the invasion history of *O. mykiss*.

Detailed phylogeographic tests are beyond the scope of this study but the presented results are significant in understanding the invasion history of the species. Further work on both larger sample sizes, as well as more locations is needed to further identify the sources of the European rainbow trout populations but genetic processes, such as multiple introductions or introductions from genetically admixed sources together with ecological processes do aid the establishment success of introduced species.

CHAPTER V

General discussion

5. 1 Summary of findings

5.1.1 MtDNA analysis in P. parva

There are three evolutionary lineages of the topmouth gudgeon (*P. parva*) in the native range, two of which contributed to the colonisation of Europe.

Most invasive populations have higher genetic diversity than their native counterparts and higher genetic diversity than expected under equilibrium conditions.

Most native populations have low genetic diversity typical for riverine fishes, an exception being samples from the Hai He river system which showed very high levels of genetic diversity, which under equilibrium conditions predict extremely high effective population sizes.

The differentiation among invasive populations is much lower than among native populations.

5.1.2 Microsatellite analysis of European P. parva *populations*

Populations closest to the founder site in Hungary show high levels of gene diversity and little genetic differentiation, which is also confirmed by the isolation-by-distance analysis.

There was a low number of private alleles in all the European populations, indicating recent migration and gene flow between these populations.

Admixture was also observed in most European populations, however in some it has been overridden by genetic drift – all bottlenecked populations showed low levels of admixture, especially in the relatively recent EB and SE populations.

One outlier to this trend was the Welsh population (SWS), which showed very high levels of diversity and is likely to have been founded by multiple introductions from distinct Chinese source populations.

5.1.3 MtDNA analysis of O. mykiss in its native and introduced range

The species originates from California and is extremely diverse in its native habitat, and the haplotype distribution indicates that the European hatchery and wild populations are likely to originate from California.

The Californian and Western US native populations and the European hatchery samples showed much higher diversity both in the group and the individual population analysis, as well as a high number of shared haplotypes and low population differentiation due to continuous gene flow and admixture between different populations in these regions.

There was a loss of genetic variation in the European wild populations.

The Eastern US and the European wild groups on the other hand showed significantly lower levels of diversity, possibly due to founder effects.

Shared haplotypes between the Eastern US and some European hatchery populations suggest that some introductions into Europe took place from this area.

The Hofer population originated from very diverse admixed sources before being introduced into Europe and is likely to have contributed to some of the wild European rainbow trout populations, especially E and YR and this strain had a bridgehead role in the introduction and invasion of the species in Europe.

5.2 General discussion

5.2.1 Introduction

Invasive species are recognised as a pervasive threat to natural populations with serious impacts on native species. However, they are also of evolutionary interest due to the various evolutionary processes that take place as these populations become successful in the new environment. Theory predicts that a population's range is static if fringe populations cannot adapt to local environments because of constant gene flow from the central populations (Kirkpatrick and Barton 1997). Still many invasive species expand range margins and so they can provide an insight into the evolutionary mechanisms that drive adaptation and enable them to invade new areas. Theory also states that invasive species become successful despite experiencing genetic bottlenecks (Nei *et al* 1975), which are thought

to reduce adaptive potential and limit spread. Apart from bottlenecks, invasions are also affected by various factors, such as multiple introductions and introductions from genetically different source populations. These factors can introduce novel genotypes that are not present in the native range, and help the invasive species to overcome the negative effects of genetic bottleneck. Therefore understanding the origin and role of genetic diversity in the prediction of invasion success is essential in studying invading species, while identifying the sources responsible for these introductions is important for the development of effective prevention and management strategies (Fofonoff *et al* 2003). Also, studying the origins of non-native species can provide an insight into their biology and facilitate suitable management decisions.

The aim of this thesis was to look at the evolutionary processes that affect the establishment success/failure of two non-native species by studying their source populations, introduction history and genetic diversity.

5.2.2. P. parva *introductions*

The family Cyprinidae represents a species rich group of freshwater fishes, including over 2000 species (UEPA 2000), exhibiting substantial variation in morphology, as well as diet and use of habitat. *Pseudorasbora parva* is native to Asia and its ability to colonise European waters in a relatively short time makes it a very interesting subject to study ecological and evolutionary processes. My study investigated the population structure and colonisation history of the species with the help of various molecular

markers. The mitochondrial DNA study aimed to perform a population genetic analysis of *P. parva* across the native range and introduced range in order to test different models of introduction and uncover levels of genetic variation. A microsatellite study then was carried out to detect finer scale relationships between the European populations.

Genetic population structure differentiates in most species as a result of patterns of dispersal and gene flow (Avise 1994) and the structure reveals genetic drifts that took place over long periods of time and relate to major geological shifts, such as refugia and dispersal during Pleistocene. MtDNA studies are particularly useful in showing the influence of Pleistocene glaciation events and shaping the intraspecific genetic structure of freshwater fishes (Dodson *et al* 1995; Billington and Hebert 1991). This was shown in the native *P. parva* populations, where I found highly divergent haplotypes, indicating long geographic isolation. This was further confirmed by low haplotype and nucleotide diversity and high levels of differentiation between river systems (Chapter 2), as well as high F_{ST} values in the native range, indicating higher levels of population structuring (Chapter 2). This kind of structuring was shown in other fish species (*Cottus gobio* L.; Hänfling and Brandl 1998; *Poecilia reticulata*; Barson *et al* 2009) as well as other aquatic species, the stonefly (*Yoraperla brevis*) (Hughes *et al* 1999) and the waterstrider (*Aquarius remigis*) (Preziosi and Fairbairn 1992).

The native populations are spread over a large area and our sampling was somewhat restricted - the DIY ABC analysis confirmed that while we were able to locate some populations that contributed to the European

populations, there were also unsampled populations from other lineages and the European populations originate from an admixture of these (Chapter 2).

5.2.3. P. parva introductions in Europe

In the European populations specifically, populations that have been established longer showed higher nucleotide diversity than the more recent UK populations and this also suggested that they originate from admixed sources (Chapter 2). The microsatellite analysis carried out on these populations showed the highest level of gene diversity with the lowest level of genetic differentiation in populations that are close to the founder site, these resemble the native populations most closely (Chapter 3). The low F_{ST} values imply lower levels of population structuring, which is inconsistent with a multiple source-sink model of introduction. The IBD analysis of the populations showed that as the invasions are quite recent, the populations have not yet homogenised and that the stepping-stone method was likely to be the method of introduction of the species, however, this model assumes that the exchange of individuals is restricted to adjacent populations (Thibault *et al* 2009). While this a likely method of introduction for fish, as the mtDNA showed low levels of genetic differentiation compared to native species, it is probable that long-distance translocation of the species with other fish also played a part (Chapter 2).

The high diversity found in the Welsh (SWS) *P. parva* population is likely to be the result of secondary introduction taking place, from the native range rather than from mainland Europe. Secondary introductions have also been

observed during other invasion events, for example in the spiny waterflea (*Bythotrephes longimanus*) (Panov *et al* 2007), the European green crab (*Carcinus maenas)* (Darling *et al* 2008). However, this conclusion was not supported by the mtDNA analyses which showed that the SWS population does not have any unique mtDNA haplotypes not found elsewhere in Europe (Chapter 2) and only had one private allele.

The 'time-lag' analysis (Strayer *et al* 2006; Lockwood 2007; Williamson 1996) showed that *P. parva* was able to gather higher levels of genetic diversity as the species spread within Europe. This suggests that 'time-lag' (Sakai 2001) between the initial introduction and the spread of the species is an important factor in establishment success and in the management of endangered species. I was able to relate the gene diversity (H_e) of the populations to the invasion date and this showed that 44% of the variation in gene diversity could be explained by invasion history. Often invasive species do not appear have it any detrimental effects on native species, such as the Brazilian pepper (*Schinus terebinthifolius)* in Florida (Simberloff *et al* 1997) as the species that were able to go unnoticed for many years. This is a result of post-invasion evolution (Strayer *et al* 2006), which is common in invading species and helps to increase their population size or range.

5.2.4 O. mykiss *introduction*

The vicariance caused by the last ice age is thought to be a major influence in shaping present day biodiversity (Hewitt 1996; Pielou 1991) but after the Wisconsian Glaciation Episode, with the retreat of the glaciers, *O. mykiss*

was able to inhabit new areas. This post-glacial expansion is important in structuring genome distribution and mtDNA is a useful marker in tracing lineages and source populations (Hewitt 1999). However, post-glacial migration can be great enough to wipe out genetic signals of refugial origins if the population size is too small, for example in the caribou (*Rangifer tarandus*) in Canada (Flagsted and Røed 2003), so the signal relating to refugial separation may be stronger in species that are poor dispersals.

Rainbow trout is thought to have an extremely diverse life history pattern (Bartrand *et al* 1993), which can be attributed to the migratory tendencies of the species. The ecological and genetic reasons behind this are unclear but are thought to be influenced by the interbreeding between rainbow and steelhead trout populations. The native range of the species is California but the species migrates long distances, as well as being translocated with further dispersal to the north to establish the species' current distribution (McCusker *et al* 2000). The species was introduced into Europe from various different, genetically admixed sources, with the Hofer strain being one of the sources that the European populations, including the WF and E populations originate from. There is a high number of shared haplotypes and low population differentiation due to continuous gene flow and admixture between different populations in these regions.

5.2.5 Hybridization

Although in this study there was only one possible example of hybridisation in the form of the highly divergent Polish *P. parva* sample, it is interesting and important to note that hybridization happens frequently between native and invading species and it can also increase the invasion potential of certain species (Lucek *et al* 2010; Blair *et al* 2011) - often at the cost of the native species. Also, as a result of heterosis, hybrids might have increased fitness and recombination can create new genotypes and produce phenotypes that are better suited to colonising the new environment. Hybridization between invasive and native species can have major consequences, especially where endangered species are involved. Consequently, understanding the ecological and evolutionary processes that affect the risk of hybridization is crucially important for the conservation of native species. Apart from *P. parva*, which is able to hybridize with the native G. gobio, this has been observed in many other species, for example between the endemic pupfish (*Cyprinodon bovinus*) and sheepshead minnow (*C. variegatus*) (Echelle and Echelle 1997) and between native, threatened California Tiger Salamander (*Ambystoma californiense*) and the introduced Barred Tiger Salamander (*Ambystoma tigrinum mavortium*). Hybridization between rainbow and steelhead trout is well-observed in the wild (Pearsons *et al* 2007), mainly due to stocking wild waters with farmed fish –it is a common practice, and is widely used both for anglers, as well as for the rehabilitation of natural waters (Madeira *et al* 2005), even though it can increase anthropogenic hybridization, as local adaptations may not have

developed. While rainbow and steelhead trout are ecotypes, and rainbow trout is a suitable source of genes if natural steelhead populations are destroyed, long-term hybridization between them can also lead to a loss of fitness in the steelhead due to loss of adaptation.

5.2.6 Incongruence between the markers

Mitochondrial and nuclear markers differ in many characteristics, such as mode of inheritance, ploidy level, repair mechanisms and size (Nabholz *et al* 2009). Also, high mutation rate, which is also the reason behind the high levels of homoplasy can complicate the use of mtDNA in population genetic studies. However, discrepancy between the various markers is expected, it has been documented by several other studies (Ting *et al* 2007, Sota and Vogler 2001), it was observed when comparing microsatellite and mitochondrial DNA diversity in *P. parva* studies and has to be acknowledged.

5.2.7. Future genetic studies

Research on invasive species remains important, both due to continuing global trade but also as species continue to alter their distribution in response to climate change. Evolutionary factors were first recognised as important in biological invasions in the 1970s (Baker 1974) and investigating the genetic basis of traits that are important for adaptation to the new environment is a key aim in evolutionary and invasion biology.

The loss of variation expected during invasions resulted in most research concentrating on comparing the genetic diversity of source and invasive populations (Dlugosch and Parker 2008).

Until recently, mostly neutral markers (microsatellites, AFLP) were used to assess genetic diversity in invasive populations and to understand the relationship between population size and the level of genetic variation and only a few studies have attempted to understand the source of epigenetic variation underlying adaptation. While these markers provided a great deal of information, as a result of being selectively neutral, they are more prone to losing genetic variation than genes under selection. The evolutionary history of a species depends on its ability to adapt to its environment and it is a factor of the plasticity and the variability of the species' genome. Recent advances in functional genomics make it possible to compare the genomes of different organisms that are closely related but live in different environments, as gene function can only be defined in relation to the environment in which it is measured (Ouborg *et al* 2010). By comparing the levels of expression of genes between the different stages of development makes it possible to understand what proteins are particularly important in understanding the role of stress, pollution or other environmental factors and what mutations in an organism enable it to survive in the stressed environment.

The genomic resources needed to identify candidate genes in invasive species associated with invasiveness, such as genomic or expressed sequence tags (ESTs), have mostly not been developed and while they

might exist for some species – and these can be used for closely related species - in order to achieve real results, specific resources must be developed to identify genetic changes in invasive species.

Aquacultural genome projects include the US NRSP-8 project, which aimed to integrate the genetic and physical maps of aquaculturally important species, as well as to facilitate the incorporation of genomic and transcriptional approaches and bioinformatics tools to extract and evaluate information.

Genomic studies have great implications for aquacultural research as well, which will make it possible to improve existing stocks, and treat various diseases that currently present serious problems for aquacultural facilities. Apart from these issues, genetic markers are also used for the analysis of genetic resources, to identify strains and populations, and for conserving genetic diversity.

Genome research on farmed fish has so far mainly focused on increasing performance, growth rate or meat quality, however other important aspects, such as disease resistance need to be researched using these methods. Aquaculture will have to address questions relating to environmental impacts and combine research on benefiting aquaculture, while at the same time protecting the environment using metagenomics and ecogenomics methods.

Genomic resources about non-native species are important for evolutionary studies to resolve question about the phenotypes present in invasive species

and a better knowledge of quantitative trait loci (QTL) and adaptive traits are vital in the study of biological invasions (Booth *et al* 2007).

The discovery of the Hofer strain makes it possible to compare the genome of susceptible and resistant trout to further understand the genetic basis of the disease. Previous studies have identified certain genes that are thought to have a role in salmonid immune response to WD (Baerwald *et al* 2008) and the development of next-generation sequencing offers the opportunity for further research, using these candidate genes to further explore the relationship between host and parasite as well as to gain an understanding of disease resistance and its effect on establishment success.

In *P. parva* similar studies could be used to understand the genetic basis of being a carrier of the fish parasite, Sphaerothecum destruens, which causes mortality in salmonid fish (Gozlan *et al* 2009). By understanding the exact source locations of the European populations, as well as the colonisation history within Europe, and possible levels of infection or presence of parasite could be correlated with population origin.

In order to further understand the processes involved with these invasions, a more extensive sampling of both the native and invaded areas is necessary in both species. Sampling the native range of *P. parva* further would allow us to identify the evolutionary processes that took place before introduction. The inclusion of additional samples from other hatchery populations could provide a more detailed understanding of both the genetic structure of rainbow trout and the natural processes that helped shape their evolutionary history. Microsatellite analysis of the trout samples, as well as further

sampling could provide further information about more recent processes that took place in these populations as well as pinpointing the origin of these populations and population structure in the native range (Nielsen 2001).

In further research, finding the genes that are under selection during multiple introductions, increasing the invasive species' potential for evolution is an important area of future research, both in hatchery and wild populations of fish. For example Haidle *et al* (2008) looked at the genes responsible for sexual maturation in trout or Roberge *et al* (2008) investigated genes regulating behaviours relating to social dominance and behavioural plasticity.

The associations between the number of introduction events, genetic diversity, bottlenecks, which may leave their signature on the structure of the population and the success of the invasion are important and have to be understood and incorporated into the management of non-native species as well as the theoretical framework around it. It provides an insight into the evolutionary relationships of these two species between the native and the introduced range and can provide an exciting avenue for further research by using genetic and population genomic methods within an ecological framework. Developing databases and new methods and tools is the direction that invasion biology and conservation biology is taking and the comparative approach of functional genomics will improve our understanding of the processes affecting genetic variation and establishment success of invasive species.

5.3 Final conclusions and future work

Most invasion attempts are not successful (Williamson 1996) but many of the successful invaders are considered to be detrimental for ecosystems. The role of evolution was first recognised as a significant process in biological invasions in the 1970s (Baker 1974). Invading species evolve in order to adapt to the new environment in response to selection (Philips and Shine 2004), aided by multiple colonisation events that help them overcome bottlenecks and increase the evolutionary potential of the population (Lavergne and Molofsky 2007) by recombination and increasing phenotypic plasticity. However, the adaptive potential is a combination of may already be present in the invasive population before the introduction (Eales *et al* 2010), and can be a combination of structure components, i.e. existing as a result of new genotypes being introduced (Lavergne and Molofsky 2007). Prentis (2008) suggests that existing genetic variation (i.e. not new mutations) is also useful during the range expansion of the species, promoting rapid evolution.

Examining the establishment history of these species using various genetic markers provided us with a great deal of information about the processes and mechanisms that play a part in the invasion success of *P. parva* colonising Europe in a relatively short time as well as understanding more about the colonisation history of *O. mykiss*.

Working with these two fish species was interesting because of their different life history traits, and the differences in their methods of introduction and spread, as well as the differences in population structure in

the native range. Aquaculture also played a big part in the spread of both species, although while rainbow trout was purposely translocated with the intention of stocking, topmouth gudgeon was accidentally introduced with other fish intended for aquaculture.

Rapid increase in the population size after introduction can also improve the establishment success (Roman and Darling 2007), as well as the highly admixed source populations (Handley *et al* 2011), which are also likely to have aided the spread of the study species, although to be able to say this with absolute certainty, fitness experiments would need to be carried out.

Most European and UK trout populations shared similar haplotypes and there was little variation between these populations, which apart from indicating similar origins, can also lead to the homogenisation of these populations. If recurrent introductions take place from the same populations repeatedly, it could explain the establishment failure of the species as the genetic make-up of a population is important in stock management. Also, hatcheries can cause unintentional selection as a result of differing conditions between hatcheries and natural environments (Heath 2003) and it can 'relax' selection. While this may not matter in hatchery populations, it can have serious consequences for threatened species or species that are reared for stocking.

While my studies concentrated on the evolutionary processes involved in invasion, environmental factors are also important and can cause rapid evolution, for example anadromous trout will be under strong selection to

adopt a resident life-history strategy, when faced with an impassable barrier (Pearse *et al* 2009).

My thesis suggests that *P. parva* has high evolutionary potential while in the native range of *P. parva* it was not possible to sample all lineages; these lineages are probably significant in the evolutionary sense, as they can possess unique characteristics, which can aid invasion potential. Over the past few years next generation sequencing has greatly increased the speed of genomics studies and the amount of data that can be generated compared to previous sequencing methods. A more comprehensive sampling of the native and introduced ranges in both species would confirm the method of introduction, using next-generation sequencing technology to identify and compare particular sequences for genes that are involved in biological invasions (Wang *et al* 2011, Prentis *et al* 2010).

In conclusion I found that identifying genetic admixture the number and model of introductions that take place and the sources of these introductions are key components in understanding invasive species. The present study represents a first step at inferring the evolutionary history of *P. parva* and *O. mykiss* in their introduced range.

REFERENCES:

Addisonsoft (2008) XLStat 2008: Data analysis and statistics with Microsoft Excel. Paris, France.

Alemadi SD, Jenkins DG (2008) Behavioral constraints for the spread of the eastern mosquitofish, Gambusia holbrooki (Poeciliidae). *Biological Invasions*, **10**:59-66.

Allendorf FW, Thorgaard GH (1984) Tetraploidy and the evolution of salmonid fishes. In: Turner BJ, editor. Evolutionary genetics of fishes. New York: Plenum Press. p. 1–46.

Allendorf FW, Lindquist LL (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, **17**:24-30.

Allentoft ME, O'Brien J (2010) Global Amphibian Declines, Loss of Genetic Diversity and Fitness: A Review. *Diversity*, **2**:47-71.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, **215**:403–410.

American Fisheries Society http://www.fisheries.org/afs/

Amos W, Harwood J (1998) Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of the Royal Society*, **353**:177-186.

Andolfatto P, Scriber JM, Charlesworth B (2003) No association between mitochondrial DNa haplotypes and a female-limited mimicry phenotype in *Papilio glaucus*. *Evolution*, **57**(2):305–316.

Augerot, Xanthippe. Atlas of Pacific Salmon: The First Map-Based Status Assessment of Salmon in the North Pacific. Berkeley, CA: University of California Press, 2005.

Avise JC, Lansman RA, Shade RO (1979) The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus Peromyscus. *Genetics* 92: 279-295.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology, Evolution and Systematics*, **18**:489-522.

Avise JC (1994) Molecular Markers, Natural History, and Evolution. Chapman & Hall, New York (511 pp.).

Baerwald MR, Welsh AB, Hedrick RP, May B (2008) Discovery of genes implicated in whirling disease infection and resistance in rainbow trout using genome-wide expression profiling. *BMC Genomics*, **9**(37).

Bagley MJ, Gall GAE (1998) Mitochondrial and nuclear DNA sequence variability among populations of rainbow trout (Oncorhynchus mykiss). *Molecular Ecology*, **7**:945-961.

Baker HG (1974) The evolution of weeds. *Annual Review of Ecology and Systematics*, **5**:1-24.

Baker AJ, Tavares ES, Elbourne RF (2009) Countering criticisms of single mitochondrial DNA gene barcoding in birds. *Molecular Ecology Resources*, **9**:257–268.

Ballard JWO (2000) When one is not enough: introgression of mitochondrial DNA in *Drosophila*. *Molecular Biology and Evolution*, **17**:1126–1130.

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

Bănărescu P, Nalbant T (1965) Studies on the systematics of Gobioninae (Pisces Cyprinidae). *Rev Roum de Biol Ser de Zoologie*, **10**:219–229.

Bănărescu P (1999) Pseudorasbora parva (Temmnick et Schlegel 1846) In P.Bănărescu, ed. The freshwater fishes of Europe 5I Cyprinidae 2/I Aula Verl Wiesbaden.

Barson NJ, Cable J, van Oosterhout C (2009) Population genetic analysis of microsatellite variation of guppies (Poecilia reticulata) in Trinidad and Tobago evidence for a dynamic source-sink metapopulation structure founder events and population bottlenecks. *Journal of Evolutionary Biology*, **22**:485–497.

Bartrand, M. Fischer, and S.A. Leider. Yakima River species interactions studies. Annual Report for FY 1993. DOE/BP-99852-2. Bonneville Power Administration, Portland,Oregon

Bateson W (1909) Mendel's Principles of Heredity. Cambridge University Press, Cambridge.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*,**22**:341–345.

Behnke RJ, (1992) Native Trout of Western North America, *American Fisheries Society*, **6**:65-69.

Behnke RJ, Tomelleri JR (2002) Trout and salmon of North America. Chanticleer Press.

Benejam L, Carol J, Alcaraz C, Garcia-Berthou E (2005) First record of the common bream (Abramis brama) introduced to the Iberian Peninsula. *Limnetica*, **24**:273–274.

Berg LS (1949) Ryby presnych vod SSSR i sopredelnych stran 2 ss 477–925 Izd AN SSSR Moskva-Leningrad.

Bernatchez, L., J. J. Dodson, and S. Boivin. 1989. Population bottlenecks: influence on mitochondrial DNA diversity and its effect in coregonine stock discrimination. *Journal of Fish Biology*, **35**:233–244.

Billington N, Hebert PDN (1991) Mitochondrial DNA Diversity in Fishes and its Implications for Introductions. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**:(S1) 80-94.

Bjerkeng B (2000): Carotenoid pigmentation of salmonid fishes - recent progress. In: Cruz -Suárez, L.E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo, R., (Eds.). Avances en Nutrición Acuícola V. Memorias del V Simposium Internacional de Nutrición Acuícola. 19-22 Noviembre, 2000. Mérida, Yucatán.

Björklund M (2005) A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes*, **5**:676–679.

Blair AC, Blumenthal DM, Hufbauer RA (2011) An experimental test of the role of hybridization in the invasion of diffuse knapweed (Centaurea diffusa Lam.) 96th ESA Annual Meeting, 2011

Booth D, Provan J, Maggs CA (2007) Molecular approaches to the study of invasive seaweeds. Botanica Marina 50: 385–396.

Boyer MC, Muhlfeld CC, Allendorf FW (2008) Rainbow trout (*Oncorhynchus mykiss*) invasion and the spread of hybridization with native westslope cutthroat trout

(*Oncorhynchus clarkii lewisi*). *Canadian Journal of Fisheries and Aquatic Sciences*, **65**(4): 658-669.

Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular Phylogeny of Cyprinidae Inferred from cytochrome b DNA Sequences. *Molecular Phylogenetics and Evolution,* **9**: 100–108.

Britton JR, Brazier M (2006) Eradicating the invasive top- mouth gudgeon Pseudorasbora parva from a recreational fishery in Northern England. *Fisheries Management and Ecology,* **13**:329–335.

Britton JR, Davies GD, Brazier M, Pinder AC (2007) A case study on the population ecology of a topmouth gudgeon Pseudorasbora parva population in the UK and the implications for native fish communities. *Aquatic Conservation: Marine & Freshwater Ecosystems*, **17**:749–759.

Britton JR, Davies GD, Harrod C (2010) Trophic interactions and consequent impacts of the invasive fish Pseudorasbora parva in a native aquatic foodweb: a field investigation in the UK. *Biological Invasions,* **12**:1533–1542.

Broughton RE, Stanley SE, Durrett RT (2000) Quantification of homoplasy for nucleotide transitions and transversions and a re-examination of assumptions in weighted phylogenetic analysis. *Systematic Biology,* **49**:617–627.

Broquet T, Ray N, Petit E, Fryxell JM, Burel F (2006) Genetic isolation by distance and landscape connectivity in the American marten (Martes americana). *Landscape Ecology*, **21**:877–889.

Brown KL, Greene JM (1987) Genetic Variation in a Wintering Flock of the American Crow (*Corvus brachyrhynchos*). Transactions of the Kansas Academy of Science (1903-), **90**(3), 113-119.

Brown KH, Thorgaard GH (2002) Mitochondrial and nuclear inheritance in an androgenetic line of rainbow trout, Oncorhynchus mykiss. *Aquaculture,* **204**(3-4):323-335.

Brown T (2006) Gene Cloning and DNA Analysis: An Introduction, 5th ed. Blackwell Publishing.

Brown JE, Stepien CA (2008) Ancient divisions, recent expansions: Phylogeography and population genetics of the round goby Apollonia melanostoma across Eurasia. *Molecular Ecology,* **11**:2598–615.

BrunelliJP, Steele CA, Thorgaard GH (2010) Deep divergence and apparent sex-biased dispersal revealed by a Y-linked marker in rainbow trout. *Molecular Phylogenetetics and Evolution*, **56**(3):983-990.

Bryan MB, Zalinski D, Filcek KB, Libants S, Li W, Scribner KT (2005) Patterns of invasion and colonization of the sea lamprey (*Petromyzon marinus*) in North America as revealed by microsatellite genotypes. *Molecular Ecology*, **14**:3757–3773.

Callaway RM, Thelen GC, Rodriguez A, Holben WE (2004) Soil biota and exotic plant invasion. *Nature*, **427**:731-733.

Callaway RM, Maron JL (2006) What have exotic plant invasions taught us over the past 20 years? *Trends in Ecology and Evolution*, **21**(7):369-374.

Cannings R (2007): The Rockies: A Natural History. Greystone/David Suzuki Foundation. p. 5.

Cannings S, Cannings R, Nelson J (2011) Geology of British Columbia: A Journey Through Time. Greystone Books.

Cao L, Shitara H, Sugimoto m, Hayashi J, Abe K, Yonekawa H (2009) New Evidence Confirms That the Mitochondrial Bottleneck Is Generated without Reduction of Mitochondrial DNA Content in Early Primordial Germ Cells of Mice. *PLoS Genetics*, **5**(12).

Carroll SP, Dingle H (1996) The biology of post-invasive species. *Biological Conservation*, **78**:207–214.

Castro JA, Picornell A, Ramon M (1998) Mitochondrial DNA: a tool for populational genetics studies. *International Microbiology*, **1**:327–332.

Catto NR (1996) Richardson Mountains, Yukon-Northwest Territories: the northern portal of the postulated 'ice-free corridor'. *Quartenary International*, **32**:3-19.

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The History and Geography of Human Genes. Princeton, NJ: Princeton Univ. Press.

Cesco H, Lambert A, Crivelli AJ (2001) Pseudorasbora parva (te´le´oste´en Cyprinidae) espe`ce invasive nouvel agent du maintien et de la disse´mination de l'anguillicolose en France? *Parasite*, **8**:75–76.

Colautti RI, Ricciardi A, Grigorovich IA, MacIsaac HJ (2004) Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, **7**(8):721-733.

Colautti RI, Eckert CG, Barrett SCH (2010) Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. *Proceedings of the Royal Society of London*, **277**: 1799-1806.

Copp GH, Bianco PG, Bogutskaya NG, et al. (2005a) To be, or not to be, a non-native freshwater fish? *Journal of Applied Ichthyology*, **21**:242-262.

Cordes JF, Stephens MR, Blumberg MA, May BP (2006) Identifying Introgressive Hybridization in Native Populations of California Golden Trout Based on Molecular Markers. *Transactions of the American Fisheries Society*, **135**:110–128.

Cornell HV, Hawkins BA (1993) Accumulation of native parasitoid species on introduced herbivores: a comparison of hosts as natives and hosts as invaders. *American Naturalist*, **141**: 847–65.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**:2001-2014.

Cornuet J, Santos F, Beaumont MA, Robert CP, Marin J, et al. (2008) Inferring population history with DIY ABC a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**:2713–2719.

Crisp DT (2000) Trout and salmon: ecology, conservation and rehabilitation, Oxford : Fishing News Books

Crozier WW, Moffatt IJJ (1989) Experimental production of triploid brown trout, Salmo trutta L, using heat shock. *Aquaculture and Fisheries Management*, **20**:343-353.

Currens KP, Schreck CB, Li HW (2009) Evolutionary Ecology of Redband Trout. *Transactions of the American Fisheries Society*, **138**(4):797-817.

Darling JA, Bagley MJ, Roman J, Tepolt CK, Geller JB (2008) Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Molecular Ecology*,**17:**4992–5007.

Darwin C (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. London, John Murray.

Davis MA, Wrage KJ, Reich PB (1998) Competition between tree seedlings and herbaceous vegetation: support for a theory of resource supply and demand. *Journal of Ecology*, **86**:652-661.

Davis MA, Grime JP, Thompson K (2000) Fluctuating resources in plant communities: a general theory of invisibility. *Journal of Ecology*, **88**:528-534.

Davis MA (2009): Invasion biology, Oxford University Press, pp. 9.

Delach A (2006) Invasive species in the Northwestern United States: Threats to wildlife, and defenders of wildlife's recommendation for prevention policies. *Northwestern Naturalist*,**87**(1):43-55.

DeWoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*,**6**:951–957.

di Rienzo A, Peterson AC, Garzat JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences USA, **91**: 3166-3170.

Dickerson BR, Ream RR, Vignieri SN, Bentzen P (2010) Population Structure as Revealed by mtDNA and Microsatellites in Northern Fur Seals, *Callorhinus ursinus*,throughout Their Range, *Plos One*, **5**(5).

Dieringer D, Schlötterer C (2003) Two Distinct Modes of Microsatellite Mutation Processes: Evidence From the Complete Genomic Sequences of Nine Species, *Genome Research*, **13**: 2242-2251.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*,**17**:431–449.

Dobzhansky, Th. (1937) Genetics and the Origin of Species. Columbia University Press, New York. (2nd ed., 1941; 3rd ed., 1951)

Docker MF, Heath DD (2003) Genetic comparison between sympatric anadromous steelhead and freshwater resident rainbow trout

in British Columbia, Canada. *Conservation Genetics*, **4:**227-231.

Dodson JJ, Colombani F, Ng PKL (1995) Phylogeographic structure in mitochondrial DNA of a South-east Asian freshwater fish, *Hemibagrus nemurus* (Siluroidei; Bagridae) and Pleistocene sea-level changes on the Sunda shelf. *Molecular Ecology*, **4**: 331-346.

Dowling DK, Friberg U, Lindell J (2008) Evolutionary implications of nonneutral mitochondrial genetic variation. *Trends in Ecology and Evolution*, **23**(10):546-554.

Drude (1896) Manual de Geographie Botanique. Librairie Des Sciences Naturelles, Paris.

Dubut V, Sinama M, Martin J, Meglécz E, Fernandez J, Chappaz R, Gilles A, Costedoat A (2010) Cross-species amplification of 41 microsatellites in European cyprinids: A tool for evolutionary, population genetics and hybridization studies. *BMC Research Notes*, **3**:135.

Duckworth RA (2008) Adaptive dispersal strategies and the dynamics of a range expansion. *The American Naturalist*, **172**, S4–S17.

Dukes JS, Mooney Ha (1999) Does global change increase the success of biological invaders? *Trend in Ecology and Evolution*, **14**(4):135-139.

Dukes JS, Mooney HA (2004) Disruption of ecosystem processes in western North America by invasive species, *Revista Chilena de Historia Natural*,**77**:411-437.

Dunham RA (2004) Aquaculture and fisheries biotechnology: Genetic approaches. CABI Publishing.

Durand JD, Tsigenopoulos CS, Unlu E, Berrebi P (2002) Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA - Evolutionary significance of this region. *Molecular Phylogenetics and Evolution*, **26**(9):1963–1973.

Dutech C, Fabreguettes O, Capdevielle X, Robin C (2010) Multiple introductions of divergent genetic lineages in an invasive fungal pathogen, *Cryphonectria parasitica*, in France. *Heredity, **105**:220-228.*

Eales J, Thorpe RS, Malhotra A (2010) Colonization history and genetic diversity: adaptive potential in early stage invasions. *Molecular Ecology*, **19**, 2858–2869.

Earl DA (2011) Structure harvester v0.6.1. Available at http://taylor0.biology.ucla.edu/structureHarvester/ (accessed 19th September, 2011)

Echelle AA, Echelle AF (1997) Genetic Introgression of Endemic Taxa by Non-Natives:A Case Study with Leon Springs Pupfish and Sheepshead Minnow. *Conservation Biology*, **11**(1): 53–161.

Eisen JA (1998) A phylogenomic study of the MutS family of proteins. *Nucleic Acids Research,* **26**: 4291–4300.

Ellegren H (2004) Microsatellites: Simple sequences with complex evolution. *Nature Reviews,* **5**:435.

Ellegren H (2009) Is genetic diversity really higher in large populations? *Journal of Biology*, **8**:41.

Ellstrand NC, Elam DR. (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**:217–42.

Elton C (1958) *The Ecology of Invasions by Animals and Plants*, 1958, Methuen, London. Reprinted 2000 by The University of Chicago

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**:2611-2620.

Ewing B, Hillier L, Wendl MC, Green P (1998) 'Base calling of automated sequencer traces using PHRED I Accuracy assessment'. *Genome Research*, **8**:175–185.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**:47-50.

Eyre-Walker A, Smith NH, Maynard-Smith (1999) How clonal are human mitochondria? *Proceedings of the Royal Society*, **266**:477-483.

Eyre-Walker A, Awadalla P (2001) Does Human mtDNA Recombine? *Journal of Molecular Evolution*, **53**:430–43.

Facon B, Ruth A. Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, Vitalis R, Guillemaud T, Lundgren JG, Estoup A (2010) Report Inbreeding Depression Is Purged in the Invasive Insect Harmonia axyridis. *Current Biology,* **21**:424–427.

Facon B, Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, Vitalis R, Guillemaud T, Lundgren JG, Estoup A (2011) Inbreeding depression is purged in the invasive insect Harmonia axyridis. *Current Biology,* **21**(5):424-427.

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, et al. (2007) High-Throughput Sequencing of Arabidopsis microRNAs: Evidence for Frequent Birth and Death of MIRNA Genes. *Plos One* **2**(2): e219.

Fausch KD, Taniguchi Y, Nakano S, Grossman GD, Townsend CR (2001) Flood disturbance regimes influence rainbow trout invasion success among five holarctic regions. *Ecological Applications,* **11**:1438-1455.

Fausch KD (2007) Introduction, establishment and effects of non-native salmonids: considering the risk of rainbow trout invasion in the United Kingdom, *Journal of Fish Biology,* **71**:1-32.

Fetherman EH, Winkelman DL, Schisler GJ, Myirk CA (2011) The effects of Myxobolus cerebralis on the physiological performance of whirling disease resistant and susceptible strains of rainbow trout. *Journal of Aquatic animal Health,* **23**(4):169-177.

Feulner PGD, Bielfeldt W, Zachos FE, Bradvarovic J, Eckert I, Hartl GB (2004) Mitochondrial DNA and microsatellite analyses of the genetic status of the presumed subspecies Cervus elaphus montanus (*Carpathian red deer*). *Heredity,* **93**, 299–306.

Ficetola GF, Bonin A, Miaud C (2008) Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology*,**17**:773–782.

Filipowicz M, Burzynski A, Smietanka B, Wenne R (2008) Recombination in Mitochondrial DNA of European Mussels Mytilus. *Journal of Molecular Evolution*, **67**:377–388.

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon Press, Oxford

Flagsted Ø, Røed KH (2003) Refugial origins of reindeer (Rangifer tarandus L.) inferred from mitochondrial DNA sequences. *Evolution,* **57**: 658–670.

Fofonoff PW, Ruiz GM, Steves B, Carlton JT (2003) In ships or on ships? Mechanisms of transfer and invasion for non-native species to the coasts of North America. In: *Invasive Species:Vectors and Managements Strategies* (eds Ruiz GM, Carlton JT), pp. 152–182. Island Press, Washington, D.C.

Forbes SH, Hogg JT, Buchanan FC, Crawford AM, Allendorf FW (1995) Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Molecular Biology and Evolution*, **12**(6):1106-1113.

Frankham R (2005) Genetics and extinction. Biological Conservation, **126**:131–140.

Freckleton RP (2006) Census error and the detection of density dependence. *Journal of Animal Ecology*, **75**:837-851.

Froese R, Pauly D (2008) FishBase. Worldwide Web Electronic Publication.

Fry A (1999) Mildly Deleterious Mutations in Avian Mitochondrial DNA: Evidence from Neutrality Tests, *Evolution*, **53**(5):1617-1620.

Gaertner M, Den Breeyen A, Hiu C, Richardson DM (2009) Impacts of alien plant invasions on species richness in Mediterranean-type ecosystems: a meta-analysis. *Progress in Physical Geography*, **33**:319-338.

Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology*, **18**(22):4541–4550.

Geller JB, Darling JA, Carlton JT (2010) Genetic Perspectives on Marine Biological Invasions. *Annual Review of Marine Science*, **2**:367-393.

Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Mol Ecol* **14**:4275–4285.

Gevrey M, Worner S, Kasabov N, Pitt J, Giraudel J (2006): Estimating risk of events using SOM models: A case study on invasive species establishment. *Ecological Modelling*, **197**(3):361-372.

Goudet J. (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3), updated from Goudet 1995. Available from http://www.unil.ch/izea/soft-wares/fstat.html

Gozlan RE, Pinder AC, Shelley J (2002) Occurrence of the Asiatic cyprinid Pseudorasbora parva in England. *Journal of Fish Biology*, **61**:298–300.

Gozlan RE, St-Hilaire S, Feist SW, Martin P, Kent ML (2005) Biodiversity –disease threat to European fish. *Nature*, **435**:1046.

Gozlan RE, Peeler EJ, Longshaw M, St-Hilaire S, Feist SW (2006) Effect of microbial pathogens on the diversity of aquatic populations, notably in Europe. *Microbes and Infections*, **8**:1358-1364.

Gozlan RE, Whipps CM, Andreou D, Arkush KD (2009): Identification of a rosette-like agent as Sphaerothecum destruens, a multi-host fish pathogen. *International Journal for Parasitology*, **39**(10): 1055-1058.

Gozlan RE, Britton JR, Cowx I, Copp GH (2010a) Current knowledge on non-nativefreshwater fish introductions. (Review) *Journal of Fish Biology*, **76**:751–786.

Gozlan RE, Andreou D, Asaeda T, Beyer K, Bouhadad R, et al. (2010b) Pancontinental invasion of Pseudorasbora parva: towards a better understanding of freshwater fish invasions. *Fish and Fisheries*, **11**:315–340.

Gharbi K, Gautier A, Danzmann RG, Gharbi S, Sakamoto T, Høyheim B, Taggart JB, Cairney M, Powell R, Krieg F, Nobuaki O, Ferguson MM, Holm LE, Guyomard R (2006) A Linkage Map for Brown Trout (*Salmo trutta*): Chromosome Homeologies and Comparative Genome Organization With Other Salmonid Fish. *Genetics*, **172**(4):2405-2419

Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity*, **104**:88–99.

Gurevitch J, Fox GA, Wardle GM, Inderjit, Taub G (2011) Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters*,**14**:407–418.

Haidle L, Janssen JE, Gharbi K, Moghadam HK ,Ferguson MM, Danzmann RG (2008) Determination of Quantitative Trait Loci (QTL)

for Early Maturation in Rainbow Trout (Oncorhynchus mykiss). *Marine Biotechnology*, **10**:579–592.

Hamilton WD (1964) The Genetical Evolution of Social Behaviour. *Journal of Theoretical Biology*, **7**:1-16.

Hancock JM, Santibanez-Koref MF (1998) Trinucleotide Expansion Diseases in the Context of Micro- and Minisatellite Evolution Hammersmith Hospital, April 1–3, 1998.

Handley LJL, Estoup A, Evans DM, Thomas CE, Lombaert E, Facon B, Aebi A, Roy HE (2011) Ecological genetics of invasive alien species. *BioControl*, **56**:409–428.

Hänfling B, Brandl R (1998) Genetic differentiation of the bullhead Cottus gobio L across watersheds in central Europe evidence for two taxa. *Heredity*, **80**:110–117.

Hänfling B, Kollmann J (2002) An evolutionary perspective of biological invasions. *Trends in Ecology & Evolution*, **17**:545–546.

Hänfling B, Weetman D (2006) Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the River Sculpin Cottus gobio. *Genetics*, **173**:1487–1501.

Hänfling B (2007) Understanding the establishment success of non-indigenous fishes: lessons from population genetics. *Journal of Fish Biology*, **71**:115–135.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, **83**, 145-154.

Hassan, M., Harmelin-Vivien, M., Bonhomme, F., 2003. Lessepsian invasion without bottleneck: example of two rabbitfish species (Siganus rivulatus and Siganus luridus). *Journal of Experimental Marine Biology and Ecology*, **291**, 219–232.

Haynes GD, Gilligan DM, Grewe P, Nicholas FW (2009) Population genetics and management units of invasive common carp Cyprinus carpio in the Murray–Darling Basin, Australia. *Journal of Fish Biology*, **75**:295–320.

Heath DD (2003) Rapid Evolution of Egg Size in Captive Salmon. *Science* **299**:1738.

Hedrick RP, McDowell TS, Gay M, Marty GD, Georgiadisl MP, MacConnell E (1999) Comparative susceptibility of rainbow trout *Oncorhynchus mykiss* and brown trout *Salmo trutta* to *Myxobolus* cerebralis, the cause of salmonid whirling disease. *Diseases of Aquatic Organisms*, **37**:173-183.

Hedrick RP, McDowell TS, Marty GD, Fosgate GT, Mukkatira K, Myklebust K, Matbouli M (2003) Susceptibility of two strains of rainbow trout (one with a suspected resistance to whirling disease) to *Myxobolus cerebralis* infection. *Diseases of Aquatic Organisms* **55**:37-44.

Heger T, Trepl L (2003) Predicting Biological Invasions. *Biological Invasions*, **5**(4):313-321.

Herborg LM, Weetman D, van Oosterhout C, Hänfling B (2007) Genetic population structure and contemporary dispersal patterns of a recent European invader the Chinese mitten crab Eriocheir sinensis. *Molecular Ecology,* **16**:231–242.

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.

Hickey GI, Flecther HL, Winter P (2007) Genetics, New York : Taylor & Francis

Hoarau G, Holla S, Lescasse R, Stam WT, Olsen JL (2002) Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish Platichthys flesus. *Molecular Biology and Evolution,* **19**:2261–2264.

Hoarau G, Coyer JA, Olsen JL (2009) Paternal leakage of mitochondrial DNA in a Fucus (Phaeophyceae) hybrid zone, *Journal of Phycology*, **45**(3):621-624.

Hoffman GL (1990) *Myxobolus cerebralis*, a worldwide cause of salmonid whirling disease. *Journal of Aquatic Animal Health,* **2**:30–37.

Holland BS (2001) Invasion Without a Bottleneck: Microsatellite Variation in Natural and Invasive Populations of the Brown Mussel Perna perna (L). *Marine Biotechnology* **3**:407–415.

Holway DA (1998) Factors governing rate of invasion: a natural experiment using Argentine ants. *Oecologia,* **115**:206-212.

Howard LO (1897) The spread of land species by the agency of Man: with especial reference to insects. *Proceedings of the American Association for the Advancement of Science*, 46: 3-26.

Hua X, Wang W, Yin W, He Q, Jin B, et al. (2009) Phylogeographical analysis of an estuarine fish Salanx ariakensis (Osmeridae Salanginae) in the northwestern Pacific. *Journal of Fish Biology,* **75**:354–367.

Huang A, Wang G, He L, Niu J, Zhang B (2011) Characterization of small RNAs from Ulva prolifera by high-throughput sequencing and *Bioinformatics*analysis. Chinese Science Bulletin, 56, (27): 2916-2921.

Hudson RR (1990) Gene genealogies and the coalescent process. 1–42. in Oxford Surveyin Evolutionary Biology Vol 7 Futuyama D, Antonovics J, eds. Oxford University Press Oxford.

Huey RB, Gilchrist GW, Hendry AP (2005) Using Invasive Species to Study Evolution. Case studies with Drosophila and Salmon. Species Invasions: Insights into Ecology, Evolution, and Biogeography, Edited by Dov F. Sax, John J. Stachowicz, and Steven D. Gaines

Hughes C, Queller D (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology,***2**:131-138.

Hughes JM, Mather PB, Sheldon AL, Allendorf FW (1999) Genetic structure of the stonefly, Yoraperla brevis, populations: the extent of gene flow among adjacent montane streams. *Freshwater Biology,* **41**:63-72.

Huxel GR (1999) Rapid displacement of native species by invasive species: effects of hybridization. *Biological Conservation* **89**:143-152.

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature,* 431: 931-945

Jamandre BW, Durand JD, Tzeng WN (2009): Phylogeography of the flathead mullet Mugil cephalus in the north-west Pacific as inferred from the mtDNA control region. *Journal of Fish Biology,* **75**:393–407.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance web service. *BMC Genetics,* **6**:13. v316 http //ibdws sdsu edu/.

Jeschke JM, Strayer DL (2005) Invasion success of vertebrates in Europe and North America. *Proceedings of the National Academy of Sciences USA*, **17**(20): 7198–7202.

Jin P, Warren ST (2000) Understanding the molecular basis of fragile X syndrome, *Molecular Genetics*, **9**(6):901-908.

Kalinowski, ST, Taper, ML & Marshall, TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**: 1099-1006.

Kelly DW, Muirhead JR, Heath DD, Macisaac HJ (2006) Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish water. *Molecular Ecology,* 15: 3641–3653.

Kang M, Buckley YM, Lowe AJ (2007) Testing the role of genetic factors across multiple independent invasions of the shrub Scotch broom (*Cytisus scoparius*). *Molecular Ecology*,**16**(22): 4662-4673.

Keane RM, Crawley MJ (2002) Exotic plant invsions and the enemy release hypothesis. *Trends in Ecology & Evolution*, **17**(4):164–170.

Kelly DW, Muirhead JR, Heath DD Macisaac HJ (2006) Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish waters. *Molecular Ecology*, **15**: 3641–3653.

Kimura M. (1968) Evolutionary rate at the molecular level. *Nature*, **217** (5129):624–626.

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

Kirk H, Paul J, Straka J, Freeland JR (2011) Long-distance dispersal and high genetic diversity are implicated in the invasive spread of the common reed, *Phragmites australis* (Poaceae), in northeastern North America *American Journal of Botany*, **98**:1180-1190.

Kirkpatrick M, Barton NH (1997) Evolution of a Species' Range. *The American Naturalist*, **150**(1):1-23

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, **16**(4):199-204.

Kolar CS, Lodge DM (2002) Ecological Predictions and Risk Assessment for Alien Fishes in North America. *Science* **8**(298):1233-1236**.**

Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, et al. (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature,* **431**:177–181.

Konishi M, Takata K (2004) Impact of asymmetrical hybridization followed by sterile F1 hybrids on species replacement in Pseudorasbora. *Conservation Genetics*, **5**:463–474.

Kreitman M, Akashi H (1995) Molecular evidence for natural selection. *Annual Review of Ecology and Systematics,* **26***:*403-422.

Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonák J, Lind K, Sindelka R, Sjöback R, Sjögreen B, Strömbom L, Ståhlberg A, Zoric N (2006) The real-time polymerase chain reaction. *Molecular Aspects of Medicine*, **27**:95-125.

Lau, JA, (2008) Beyond the ecological: Biological invasions alter natural selection on a native plant species. *Ecology*, **89**(4):1023–1031.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the Natural Academy of Sciences USA*, **104**:3883–3888.

Le Roux J, Wieczorek AM (2009) Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, **154**:1-17.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution,* **17**:386–391.

Lee JS, Hanford MG, Genova JL, Farber R (1999) Relative Stabilities of Dinucleotide and Tetranucleotide Repeats in Cultured Mammalian Cells. *Human Molecular Genetics*, **8**(13):2567-2572.

Levine JM, Adler PB, Yelenik SG (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters*, **7**:975–989.

Levine, JM (2008) Biological Invasions. *Current Biology*, **18**:57-60.

Lewontin, R. C.; J. L. Hubby (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura". *Genetics* **54**:595-609.

Lin C, Lin H, Wang J, Chao S, Chiang T (2010) Phylogeography of Hemibarbus labeo (Cyprinidae) secondary contact of ancient lineages of mtDNA. *Zoologica Scripta*, **39**:23–35.

Liu J, Brown CL, Yang T (2009) Population genetic structure and historical demography of grey mullet Mugil cephalus along the coast of China inferred by analysis of the mitochondrial control region. *Biochemical Systematics and Ecology*, **37**:556–566.

Liu HZ, Yang JQ, Tang QY (2010) Estimated evolutionary tempo of East Asian gobionid fishes (Teleostei Cyprinidae) from mitochondrial DNA sequence data. *Chinese Science Bulletin*, **55**:1501–1510.

Lockwood JL, Hoopes MF, Marchetti MP (2007) Invasion ecology. Blackwell Publishing.

Lodge, D.M. (1993) Biological invasions: lessons for Ecology. *Trends in Ecology and Evolution*, **8**:133–137.

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.

Lucas, J. S. and Southgate P. C. 2003. Aquaculture farming aquatic animals and plants. UK. Fishing News Books, a Blackwell Publishing Company. 502 p

Lucek K, Roy D, Bezault E, Sivasundar A (2010) Hybridization between distant lineages increases adaptive variation during a biological invasion:stickleback in Switzerland. *Molecular Ecology*,**19**: 3995–4011.

Lunt HD, Hyman CB (1997) Animal mitochondrial DNA recombination. *Nature*, **387**:247.

Lynch M, Crease TM (1990) The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution*, **7**:377–394.

MacCrimmon HR (1971) World distribution of rainbow trout, Salmo gairdneri. Journal of Fisheries Research Board Canada, 28:663-704.

MacDougall AS, Turkington R(2005) Are invasive species the drivers or passengers of change in degraded ecosystems? *Ecology*, **86**:42–55.

Madeira M, Gomez-Moliner B, Barbe A (2005) Genetic introgression on freshwater fish populations caused by restocking programmes. *Biological Invasions*, **7**:117–125.

Maron JL, Vila M, Bommarco R, Elmendorf S, Beardsley P (2004) Rapid evolution of an invasive plant. *Ecological Monographs*, **74**, 261–280.

Matthews KR, Berg, NH (1997) Rainbow trout responses to water temperature and dissolved oxygen stress in two southern California stream pools, *Journal of Fish Biology*, **50**:50-67.

McCauley DE, Smith RA, Lisenby JD, Hsieh C (2003) The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Molecular Ecology*, **12**:3227–3235.

McCusker MR, Parkinson E, Taylor EB (2000) Mitochondrial DNA variation in rainbow trout (*Oncorhynchus mykiss*) across its native range: testing biogeographical hypotheses and their relevance to conservation. *Molecular Ecology*, **9**(12):2089-2108.

McDowall RM (1997) The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Reviews in Fish Biology and Fisheries*, **7**:443-462.

McDowall RM (2001) Anadromy and homing: two life-history traits with adaptive synergies in salmonid fishes? *Fish and fisheries* **2**, 78-85.

McDowell RM (2002) The origin of the salmonid fishes: marine, freshwater … or neither? *Reviews in Fish Biology and Fisheries* **11**, 171-179.

Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, **98**(10):5446-5451.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, **49**, 718–726.

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.

Moyle PB, Light T (1996) Fish Invasions in California: Do Abiotic Factors Determine Success? *Ecology*, **77**(6):1666-1670.

Muirhead JR, Gray DK, Kelly DW, Ellis SM, Heath D, et al. (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology*, **17**:1020–1035.

Musil J, Adámek Z (2007) Piscivorous fishes diet dominated by the Asian cyprinid invader,topmouth gudgeon (Pseudorasbora parva). *Biologia*, **62**(4):488-490.

Nabholz B, Glemin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals – the longevity hypothesis. *Molecular Biology and Evolution*, **25**:120–130.

Nabholz B, Glemin S, Galtier N (2009) The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evolutionary Biology*, **9**:54.

Naylor RL, Goldburg RJ, Primavera JH, Kautsky N, Beveridge MCM, et al. (2000) Effect of aquaculture on world fish supplies. *Nature*, **405**:1017–1024.

Naylor RL, Hindar K, Fleming IA, Goldberg R, Williams S, Volpe J, Whoriskey F, Eagle J, Kelso D, Mangel M (2005) Fugitive Salmon: Assessing the Risks of Escaped Fish from Net-Pen Aquaculture. *BioScience*, **55** (5): 427-437.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**:1–10.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**:583-590.

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. *Molecular Biology and Evolution*, **6**:290–300.

Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Evolutionary Process and Theory (eds Nevo E, Karlin S), pp. 515-534. Academic, New York.

Newsome AE, Noble IR (1986) Ecological and physiological characters of invading species. In Ecology of biological invasions (ed.R. H. Groves & J. J. Burdon), pp. 1–20. Cambridge University Press.

Nielsen JL (1997) McCloud River trout: molecular genetics and stock identification using mtDNA and microsatellites. Technical report submitted 1/29/1997 to California Department of Fish and Game, Wild Trout Program, Rancho Cordova, CA.

Nielsen JL, Sage GK (2001) Microsatellite analyses of the trout of northwest Mexico. Genetica, **111**(1-3):269-278.

Ohno S (1970) *Evolution by gene duplication*. Springer-Verlag. ISBN 0-04-575015-7.

Okada M, Lyle M, Jasieniuk M (2009) Inferring the introduction history of the invasive apomictic grass Cortaderia jubata using microsatellite markers. *Diversity and Distributions*, **15**:148–157.

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R(K), Hedrick PW (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics* **26**: 177-187.

Palmer TF (1996) Naturalised Rainbow Trout of the Rivers Wye and Derwent. Report, Submitted To: The Committee of the Darley Dale Fly Fishing Club

Palti Y, Luo MC, Hu Y, Genet C, You FM, Vallejo RL, Thorgaard GH, Wheeler PA, Rexroad CE (2009) a first generation BAC-based physical map of the rainbow trout genome. *BMC Genomics*, **8**:10:462.

Panov VE, Rodionova NV, Bolshagin PV, Bychek EA (2007) Invasion biology of Ponto-Caspian onychopod cladocerans (Crustacea: Cladocera: Onychopoda). *Hydrobiologia,* **590**:3–14.

Pearse DE, Hayes SA, Bond MH, Hanson CV, Anderson EC, MacFarlane RB, Gazra JC (2009) Over the Falls? Rapid Evolution of Ecotypic Differentiation in Steelhead/Rainbow Trout

(Oncorhynchus mykiss). *Journal of Heredity*, **100**(5):515–525.

Pearsons TN, Phelos SR, Martin SW, Bartrand EL, McMichael GA (2007) Gene Flow between Resident and Anadromous Rainbow Trout in the Yakima Basin: Ecological and Genetic Evidence. Redband Trout: Resilience and Challenge in a Changing Landscape Oregon Chapter, *American Fisheries Society,**56.***

Phillips BL, Shine R (2004) Adapting to an invasive species: Toxic cane toads induce morphological change in Australian snakes. *PNAS*, **101**(49):17150-17155.

Pielou, E. C., 1991, After the Ice Age-the return to life to glaciated North America: The University of Chicago Press, 366 p.

Pillay TVR, Kutty MN (2005) Aquaculture: principles and practices, Wiley-Blackwell

Pimental D, Lach L, Zuniga R, Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *BioScience,* **50**:53–65.

Pinder AC, Gozlan RE (2003) Sunbleak and Topmouth Gudgeon- two new additions to Britain's freshwater fishes. *British Wildlife*, **19**(3):77-83.

Pinder AC, Gozlan RE (2005) Management and Ecological Note Dispersal of the invasive topmouth gudgeon, Pseudorasbora parva in the UK: a vector for an emergent infectious disease. *Fisheries Management and Ecology*, **12**:411–414.

Poulin J, Weller SG, Sakai AK (2005) Genetic diversity does not affect the invasiveness of fountain grass (Pennisetum setaceum) in Arizona, California and Hawaii. *Diversity and Distributions* **11**:241–247.

Prentis PJ, Wilson JRU, Domontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Plant Science,* **13**(6):288-294.

Prentis PJ, Woolfit M, Thomas-Hall SR, Ortiz-Barrientos D, Pavasovic A, Lowe AJ, Schenk PM (2010) Massively parallel sequencing and analysis of expressed sequence tags in a successful invasive plant. Annals of Botany, 06:1009-1017.

Preziosi RF, Fairbairn DJ (1992) Genetic Population Structure and Levels of Gene Flow in the Stream Dwelling Waterstrider, Aquarius

(=Gerris) remigis (Hemiptera: Gerridae). *Evolution*, **46**(2):430-444.

Pritchard JK, Stephens M, Donnelly P (2000): Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959.

Prospero S, Rigling D (2011) Invasion genetics of the chestnut blight fungus *Cryphonectria parasitica* in Switzerland. *Phytopathology* **101**. Accepted for publication.

Pyšek P, Jarošík V, Hulme PE, Kühn I, Wild J, Arianoutsoue M, Bacherf S, Chirong F, Didžiulish V, Essli F, Genovesij P, Gherardik F, Hejdaa M, Karkl S, Lambdonm PW, Desprez-Loustaun M, Nentwigo W, Pergla J, Poboljšajp K, Rabitschi W, Roquesq A, Royr DB, Shirleys S, Solarzt W, Vilàu M, Winter M (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. *Proceedings of the National Academy of Sciences of the United States of America*,**107**:12157–12162.

Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, et al. (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of National Academy Science USA*, **102**:15942–15947.

Rand DA, Kann LM (1996) Excess Amino Acid Polymorphism in Mitochondrial Among Genes from *Drosophila*, Mice, and Humans. *Molecular Biology and Evolution*, **13**(6):735-748.

Raymond M. & Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**:248-249.

Reise K, Olenin S, Thieltges DW (2006) Are aliens threatening aquatic ecosystems? *Helgoland Marine Research*, **60**:77-83.

Richardson DM, Pysek P, Rejmanek M, Barbour MG, Dane Panetta F, West CJ (2000): Naturalization and Invasion of Alien Plants: Concepts and Definitions, *Diversity and Distributions*, **6**(2):93-107.

Roberge C, Blanchet S, Dodson JJ, Guderley H, Bernatchez L (2008) Disturbance of Social Hierarchy by an Invasive Species: A Gene Transcription Study. *Plos One* **3**(6):2408.

Roberts S (2001): Life history of rainbow and steelhead trout (Oncorhyncus mykiss) and subspecies. [Online] Available: http://lochsa.lcsc.edu/~msatterw/BI403/papers/Roberts/rainbow.html

Robinson MD (2005) The ecological consequences of hybridization between native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and introduced rainbow trout (*Oncorhynchus mykiss)* in South Western Alberta. Masters Thesis, University of Lethbridge

Roman J (2006) Diluting the founder effect cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society B,* **273**:2453–2459.

Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution,* **22**:454–464.

Rosecchi E, Thomas F, Crivelli AJ (2001) Can life-history traits predict the fate of introduced species? A case study on two cyprinids in southern France. *Freshwater Biology,* **46**:845–853.

Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of invasion and intraspecific hybridization in Brachypodium sylvaticum (Hudson) Beauv. in North America. *Molecular Ecology*, **17**:4657–4669.

Rousset F (1997) Genetic Differentiation and Estimation of Gene Flow from FStatistics Under Isolation by Distance. *Genetics,* **145**, 1219-1228.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics,***19**:2496–2497.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**:365-386.

Saillant E, Patton JC, Gold JR (2005) Genetic Variation Kinship and Effective Population Size in a Captive Population of the Endangered Cape Fear Shiner Notropis mekistocholas. *Copeia*, **1**:20–28.

Sakai A, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA (2001) The population biology of invasive species. *Annual Review of Ecology, Evolution and Systematics,* **32**:305–32.

Sakai H, Goto A, Jeon S (2002) Speciation and Dispersal of Tribolodon Species (Pisces, Cyprinidae) around the Sea of Japan. *Zoological Science,* **19**(11):1291-1303.

Sala, O E, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF,Jackson R, Kinzig A, Leemans R, Lodge D, Mooney HA, Oesterheld M, Poff L, Sykes MT, Walker BH, Walker M, Wall D (2000) Global biodiversity scenarios for the year 2100. *Science,* **287**:1770-1774.

Sánchez CC, Smith TPL, Wiedmann RT, Vallejo RL, Salem M, Yao J, Rexroad CE (2009) Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. BMC Genomics, 10:559.

Sax DF, Brown JH, White EP, Gaines SD (2005) The Dynamics of Species Invasions: Insights into the Mechanisms That Limit Species Diversity, Pp. 447-465 in D. Sax, J. Stachowicz, and S. D. Gaines, eds. Species Invasions: Insights into Ecology, Evolution and Biogeography. Sinauer Publishing.

Schisler GJ, Myklebust KA, Hedrick, RP (2006) Inheritance of *Myxobolus cerebralis* resistance among F1-generation crosses of whirling disease resistant and susceptible rainbow trout strains. *Journal of Aquatic Animal Health,* **18**:109-115.

Schisler GJ (2009): Resistant Rainbow Trout in Colorado: Current Status and Uses. Colorado Division of Wildlife Paper.

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma*, **109**:365-371.

Schlötterer C, Tauz D (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Research,* **20**:211–215.

Selkoe KA, Toonen, RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers, *Ecology Letters*, ***9***:615–629.

Shea K, Chesson P (2002): Community ecology theory as a framework for biological invasions. *Trends in Ecology & Evolution*, **17**(4):170-176.

Sheehan, R.J., Shasteen, S.P., Suresh, A.V., Kapuscinski, A.R. and Seeb, J.E. (1999) Better growth in all-female diploid and triploid rainbow trout. *Transactions of the American Fisheries Society*, **128**: 491-498.

Shigesada N., Kawasaki K., 1997. Biological Invasions: Theory and Practice.Oxford Series in Ecology and Evolution, Oxford University Press, Oxford.

Shoubridge EA, Wai T (2007) Mitochondrial DNA and the mammalian oocyte. *Current Topics in Developmental Biology*, **77**:87-111.

Sigler WF, Sigler JW (1990) Recreational fisheries:management, theory, and application. University of Nevada Press, 1990 - 418 pages.

http://silurus.acnatsci.org/ACSI/field/Mongolia2006/fishes/Cyprinidae/other/1index.html

Simberloff D (2009) The Role of Propagule Pressure in Biological Invasions. *Annual Review of Ecology, Evolution and Systematics*, **40**:81–102.

Simon A, Britton R, Gozlan R, van Oosterhout C, Volckaert FAM, Hänfling B (2011) Invasive Cyprinid Fish in Europe Originate from the Single Introduction of an Admixed Source Population Followed by a Complex Pattern of Spread. *PLoS ONE*, **6**(6)

Smith D (2009) The Dictionary of Early Shasta County History. ShastaCountyHistory.com

Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. Trends in Ecology & Evolution, **14**(9):348-352.

Sota T, Vogler AP (2001) Incongruence of mitochondrial and nuclear gene trees in the Carabid beetles Ohomopterus. *Systematic Biology*, **50**(1): 39-59.

Springate JRC, Bromage NR, Elliott JAK, Hudson DL (1984) The timing of ovulation and stripping and their effects on the rates of fertilization and survival to eying, hatch and swim-up in the rainbow trout (Salmo gairdneri R.) *Aquaculture*, **43**:313-322.

Staley K, Mueller J (2000) Rainbow Trout(*Oncorhynchus mykiss*). *Fish and Wildlife Habitat Management Leaflet*,**13.**

Stamatakis A, Hoover P, Rougemont J (2008) A Fast Bootstrapping Algorithm for the RAxML Web-Servers. *Systematic Biology*, **57**:758–771.

Stenberg D (1988): Trout. The Hunting & Fishing Library

Strayer DL, Eviner VT, Jeschke JM, Pace ML (2006) Understanding the long-term effects of species invasions. *Trends in Ecology and Evolution,* **21**(11):645-651.

Su GS, Liljedahl LE, Gall GAE (1999) Estimates of phenotypic and genetic parameters for within-season date and age at spawning of female rainbow trout. *Aquaculture*, **171**:209–220.

Suarez AV, Tsutsui ND (2008): The evolutionary consequences of biological invasions. *Molecular Ecology*, **17**:351–360.

Sumida M, Kaneda H, Kato Y, Kanamori Y, Yonekawa H, Nishioka M (2000) Sequence variation and structural conservation in the D-loop region and flanking genes of mitochondrial DNA from Japanese pond frogs. *Genes and Genetic Systems,* **75**(2):79-92.

Sutterby R, Greenhalgh M (2005): Atlantic salmon: an illustrated natural history, Stackpole Books

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**(21): 2498–2504.

Tamkee P, Parkinson E, Taylor EB (2010) The influence of Wisconsinan glaciation and contemporary stream hydrology on microsatellite DNA variation in rainbow trout (Oncorhynchus mykiss). *Canadian Journal of Fisheries and Aquatic Sciences*, **67**(6): 919-935.

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

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

Tang KL, Agnew MK, Hirt MV, et al. (2010) Systematics of the subfamily Danioninae (Teleostei Cypriniformes Cyprinidae) *Molecular Phylogenetics and Evolution*, **57**:189–214.

Temminck CJ, Schlegel H (1843–1850) Pisces. In: Siebold's fauna japonica, parts 1–16. Luguduni, Batavorum, Leiden

Tessier, N., Bernatchez, L., Presa, P. and Angers, B. (1995). Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic. salmon. *Journal of Fish Biology,* **47**(Supplement A):156-163.

Therriault TW, Orlova MI, Docker MF, MacIsaac, HJ, Heath DD (2005) Invasion genetics of a freshwater mussel (Dreissena rostriformis bugensis) in eastern Europe: high gene flow and multiple introductions. *Heredity,* **95:**16–23.

Thibault I, Bernatchez L, Dodson JJ (2009) The contribution of newly established populations to the dynamics of range expansion in a one-dimensional fluvial-estuarine system: rainbow trout (Oncorhynchus mykiss) in Eastern Quebec. *Diversity and Distributions*, **15**:1060–1072.

Thorgaard GH (1983) Chromosome set manipulation and sex control in fish. In: Hoar W.S., Randall D.J., Donaldson E.M. (eds.): Fish Physiology. Vol. 9, Part B. Academic Press, New York. 405–434.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research,* **25**:4876–4882.

Ting N, Tosi AJ, Li Y, Zhang Y, Disotell TR (2008) Phylogenetic incongruence between nuclear and mitochondrial markers in the Asian colobines and the evolution of the langurs and leaf monkeys. *Molecular Phylogenetics and Evolution,* **46**: 466–474.

Torchin ME, Lafferty KD, Dobson AP, McKenzie VJ, Kuris AM (2003) Introduced species and their missing parasites. Nature, 421(6):628-630.

Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, et al. (2000) Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (hotshot). *Biotechniques*, **29**:52–54.

Urban MC, Phillips BL, Skelly DK, Shine R(2007): The cane toad's (Chaunus [Bufo] marinus) increasing ability to invade Australia is revealed by a dynamically updated range model, *Proceedings of the Royal Society B,* **274**:1414-1419.

van Oosterhout C, Hutchinson WF, Wills, DP, Shipley P (2004): MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**:535–538.

Vander Zanden, MJ (2005) The success of animal invaders (commentary). *Proceedings of the National Academy of Sciences (U.S.A.)* **102**: 7055-7056.

Verhoeven KJF, Macel M, Wolfe LM, Biere A (2011) Population admixture, plant invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B - Biological Sciences*, **278**: 2-8

Verspoor E, Stradmeyer L, Nielsen JL (2007) The Atlantic Salmon. Genetics, Conservation and Management. Blackwell Publishing.

Vidal O, Garcia-Berthou E, Tedesco PA, Garcia-Marin J (2010) Origin and genetic diversity of mosquitofish (Gambusia holbrooki) introduced to Europe. *Biological Invasions,* **12**:841–851.

Villablanca FX, Roderick GK, Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology*, **7**, 547–560.

de Visser JAGM, Cooper T, Elena SF The causes of epistasis. *Proceedings of the Royal Society,* **278**(1725):3617-3624.

Walbaum JJ. 1792. Petri Artedi Sueci Genera Piscium. In quibus systema totum ichthyologiae proponitur cum classibus, ordinibus, generum characteribus, specierum differentiis, observationibus plurimis. Emendata et aucta a Iohanne Iulio Walbaum. Pars III. *In* JJ Walbaum, ed. Petri Artedi renovati. Greifswald (Germany): Ant. Ferdin. Rose, pp. 1-723.

Walker AF (2003) Status of rainbow trout in Scotland: the results from a questionaire survey. Scottish Fisheries Information Pamphlet No. 23, Fisheries Research Services, Pitlochry.

Walters DM, Blum MJ, Rashleigh B, Freeman BJ, Porter BA, Burkhead NM (2008) Red shiner invasion and hybridization with blacktail shiner in the upper Coosa River, USA. *Biological Invasions*, **10**:1229–1242.

Wang X, Luan J, Li J, Su Y, Xia J, Liu S (2011) Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics*, 12:458

Waples RS, Pess GR, Beechie T (2008) Evolutionary history of Pacific salmon in dynamic environments. *Evolutionary Applications*, **1**:189-206.

Waples RS, Do C *(*2009*)* Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications, **3**(3):* 244-262.

Watanabe K, Iguchi K, Hosoya K, Nishida M (2000) Phylogenetic relationships of the Japanese minnows Pseudorasbora (Cyprinidae) as inferred from mitochondrial 16S rRNA gene sequences. *Ichthyological Research*, **47**:43–50.

Weigel DE, Peterson JT, Spruell P (2003) Introgressive hybridization between native cutthroat trout and introduced rainbow trout. *Ecological Applications*, **13**:38–50.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* , **38**(6), 1358-1370.

Williamson, M. 1996. Biological Invasions. London: Chapman and Hall.

Williamson M, Fitter A (1996) The characteristics of successful invaders. *Biological Conservation*, **78**:163–170.

Wright S (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the 6th international congress of genetics. 356-366

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

Wright S (1977) Evolution and the Genetics of Populations. Vol. 3: Experimental Results and Evolutionary Deductions (Chicago, University of Chicago Press, 1977)

Wright S (1978) Evolution and the Genetics of Populations. Vol.4: Variability Within and Among Natural Populations (Chicago, University of Chicago Press, 1978)

Wright JM, Bentzen P (1994) Microsatellites: genetic markers for the future. *Reviews in Fish Biology and Fisheries* **4**:384-388.

Xiao-Gu Z, Jin-Gou T, Bang-Xi X (2006) Applications of microsatellite markers in studies of genetics and breeding of fish. *Chinese Journal of Agricultural Biotechnology*, **3**:83-87.

Yang J, He S, Freyhof J, Witte K, Liu, H (2006) The phylogenetic relationships of the Gobioninae (Teleostei: Cyprinidae) inferred from mitochondrial cytochrome b gene sequences. *Hydrobiologia*, **553**:255–266.

Yang L, He S (2008) Phylogeography of the freshwater catfish Hemibagrus guttatus (Siluriformes Bagridae) Implications for South China biogeography and influence of sea-level changes. *Molecular Phylogenetics and Evolution*, **49**:393–398.

Yoshida T, Goka K, Ishihama F, Ishihara M, Kudo S (2007) Biological invasion as a natural experiment of the evolutionary processes: introduction of the special feature. *Ecological Research*, **22**:849–854.

Zhong N, Ju W, Xu W, Ye L, Shen, Y, Wu G, Chen S, Jin R, Hu X, Yang A, Liu X, Poon P, Pang C, Zheng Y, song L, Zhao P, Fu B, Gu H, Brown WT (1999) Frequency of the fragile X syndrome in Chinese mentally retarded populations is similar to that in Caucasians. *American Journal of Medical Genetics Special Issue: X-Linked Mental Retardation, Part II*, **84**(3)191-194

Zidana H, Turner GF, van Oosterhout C, Hänfling B (2009) Elevated mtDNA diversity in introduced populations of Cynotilapia afra (Gunther 1894) in Lake Malawi National Park is evidence for multiple source populations and hybridization. *Molecular Ecology*, **18**:4380–4389.

Appendix I. Matrix of pairwise genetic comparisons between all *P. parva* populations based on mtDNA F_{ST} values

	BS	CG	CH	CK	CRH	CY	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS	T	TI
BS																				
CG	0.74769																			
CH	0.46104	0.20089																		
CK	0.74769	0	0.20089																	
CRH	0.08395	0.57273	0.25644	0.57273																
CY	0.69144	0.22222	0.17457	0.22222	0.51802															
EB	0.43448	1	0.76963	1	0.45977	0.92929														
FG	0.26745	0.29333	0.04509	0.29333	0.12943	0.20149	0.59542													
G	-0.00956	0.58211	0.27341	0.58211	0.01664	0.52768	0.34244	0.10784												
HA	0.39093	0.16944	-0.03407	0.16944	0.2346	0.12638	0.67532	0.01341	0.21045											
HE	0.08759	0.65064	0.3361	0.65064	0.11266	0.60012	0.23292	0.21001	-0.0084	0.2542										
HG	0.06499	0.82478	0.5539	0.82478	0.19767	0.76729	0.16931	0.38197	0.06312	0.46982	0.01958									
HS	0.249	0.89909	0.6379	0.89909	0.31833	0.83359	0.02198	0.47257	0.18503	0.54341	0.06856	0.00539								
IN	0.5088	0.11358	-0.02556	0.11358	0.33751	0.15137	0.8022	0.08955	0.32553	-0.03727	0.37554	0.5942	0.67395							
PU	0.01281	0.48167	0.30202	0.48167	0.08471	0.46436	0.0553	0.20769	0.02221	0.25941	0.01701	-0.01827	0.01574	0.32874						
SC	0.06726	0.69312	0.3789	0.69312	0.09204	0.6408	0.29483	0.22814	-0.00938	0.31097	-0.03159	0.02847	0.12263	0.43104	0.01387					
SE	0.14694	0.55026	0.21386	0.55026	0.11309	0.5034	0.39286	0.13614	0.02001	0.13979	-0.01674	0.15028	0.2183	0.24952	0.07487	0.02743				
SWS	0.76717	0.85714	0.5267	0.85714	0.61589	0.44772	0.96798	0.34942	0.62916	0.42338	0.70816	0.83234	0.88918	0.56449	0.53079	0.73664	-0.65			
T	0.17265	0.53714	0.17867	0.53714	0.09408	0.4891	0.42143	0.12594	0.02661	0.11522	-0.01156	0.17662	0.24236	0.2244	0.0863	0.04427	0.04622	0.64662		
TI	0.43366	0.95763	0.68199	0.95763	0.31937	0.88213	0.90196	0.51416	0.35324	0.61997	0.43153	0.5227	0.68063	0.73977	0.18824	0.40772	0.45727	0.93267	0.44688	
TT	0.78701	1	0.84408	1	0.69872	0.94574	1	0.72821	0.71152	0.79592	0.76318	0.8398	0.90787	0.86922	0.48622	0.76984	0.7619	0.97456	0.76385	0.96032

Appendix II. Table S2. Matrix of pairwise genetic comparisons between all *P. parva* populations based on mtDNA D_{XY} values

	BS	CG	CH	CK	CRH	CY	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS	T	TI
BS																				
CG	0.01336																			
CH	0.01146	0.00351																		
CK	0.01336	0	0.00351																	
CRH	0.00931	0.01206	0.0107	0.01206																
CY	0.01466	0.00148	0.00479	0.00148	0.01308															
EB	0.00596	0.0148	0.01217	0.0148	0.00954	0.01628														
FG	0.01204	0.00771	0.00864	0.00771	0.01218	0.00826	0.01347													
G	0.0082	0.01173	0.01061	0.01173	0.01023	0.01282	0.00746	0.0116												
HA	0.01197	0.00471	0.0065	0.00471	0.01185	0.0058	0.01206	0.00949	0.01117											
HE	0.00794	0.01107	0.01005	0.01107	0.01017	0.01255	0.00504	0.01179	0.0087	0.01044										
HG	0.00624	0.01404	0.0118	0.01404	0.00949	0.01552	0.00296	0.01279	0.00786	0.01202	0.00645									
HS	0.00635	0.01382	0.01159	0.01382	0.00961	0.0153	0.00143	0.01297	0.00773	0.01163	0.00565	0.00387								
IN	0.01187	0.00278	0.00513	0.00278	0.01149	0.00426	0.01244	0.00869	0.01092	0.00615	0.01014	0.01212	0.01182							
PU	0.01643	0.02478	0.02242	0.02478	0.01966	0.02613	0.0136	0.02309	0.01815	0.02263	0.017	0.01503	0.01447	0.0228						
SC	0.00751	0.01184	0.01037	0.01184	0.00968	0.01332	0.00515	0.01177	0.00846	0.01096	0.00727	0.00627	0.00573	0.01071	0.01671					
SE	0.00863	0.00888	0.00865	0.00888	0.01031	0.01036	0.00658	0.01093	0.00908	0.0092	0.00773	0.0076	0.00689	0.0086	0.0182	0.00784				
SWS	0.01711	0.00428	0.00721	0.00428	0.01501	0.00319	0.01908	0.00931	0.01487	0.00785	0.01535	0.01831	0.01809	0.00705	0.02868	0.01612	0.01316			
T	0.00868	0.00822	0.00805	0.00822	0.00989	0.0097	0.00658	0.01059	0.00895	0.00873	0.00759	0.00761	0.00686	0.00808	0.01822	0.00779	0.00746	0.0125		
TI	0.00692	0.01294	0.01054	0.01294	0.00838	0.01442	0.00559	0.01234	0.00843	0.01175	0.00777	0.0063	0.00608	0.01156	0.0165	0.00706	0.00837	0.01721	0.00787	
TT	0.01583	0.01974	0.01798	0.01974	0.01711	0.02122	0.0148	0.02005	0.017	0.01919	0.01634	0.01535	0.01513	0.01881	0.025	0.01579	0.01678	0.02401	0.01612	0.01382

Appendix III. Output file of the Direct approach, relative proportion of each scenario found in the selected 500 closest dataset; Posterior probabilities of scenarios obtained through a logistic regression computed every 10% of the number of selected datasets.

Direct approach closest		Scenario 1		Scenario 2		Scenario 3
50	0.24	[0.0000,0.6144]	0.3	[0.0000,0.7017]	0.46	[0.0231,0.8969]
100	0.27	[0.0000,0.6591]	0.28	[0.0000,0.6736]	0.45	[0.0139,0.8861]
150	0.2667	[0.0000,0.6543]	0.2533	[0.0000,0.6346]	0.48	[0.0421,0.9179]
200	0.26	[0.0000,0.6445]	0.26	[0.0000,0.6445]	0.48	[0.0421,0.9179]
250	0.24	[0.0000,0.6144]	0.268	[0.0000,0.6562]	0.492	[0.0538,0.9302]
300	0.2533	[0.0000,0.6346]	0.2533	[0.0000,0.6346]	0.4933	[0.0551,0.9316]
350	0.2514	[0.0000,0.6317]	0.2771	[0.0000,0.6695]	0.4714	[0.0339,0.9090]
400	0.2625	[0.0000,0.6482]	0.275	[0.0000,0.6664]	0.4625	[0.0255,0.8995]
450	0.2622	[0.0000,0.6478]	0.2733	[0.0000,0.6640]	0.4644	[0.0273,0.9016]
500	0.264	[0.0000,0.6504]	0.264	[0.0000,0.6504]	0.472	[0.0344,0.9096]

Logistic regression closest		Scenario 1		Scenario 2		Scenario 3
3000	0.0882	[0.0000,0.1868]	0.2861	[0.0806,0.4916]	0.6257	[0.3878,0.8636]
6000	0.058	[0.0138,0.1022]	0.314	[0.1675,0.4604]	0.6281	[0.4687,0.7874]
9000	0.0519	[0.0199,0.0840]	0.2917	[0.1797,0.4038]	0.6563	[0.5343,0.7784]
12000	0.0471	[0.0219,0.0724]	0.275	[0.1819,0.3680]	0.6779	[0.5767,0.7791]
15000	0.0431	[0.0224,0.0638]	0.2626	[0.1822,0.3430]	0.6943	[0.6070,0.7816]
18000	0.0398	[0.0224,0.0572]	0.2596	[0.1871,0.3322]	0.7005	[0.6222,0.7789]
21000	0.0373	[0.0222,0.0524]	0.2609	[0.1938,0.3280]	0.7018	[0.6296,0.7740]
24000	0.0363	[0.0226,0.0501]	0.2627	[0.1998,0.3255]	0.701	[0.6336,0.7684]
27000	0.0353	[0.0227,0.0478]	0.264	[0.2047,0.3234]	0.7007	[0.6372,0.7641]
30000	0.0341	[0.0226,0.0456]	0.2655	[0.2092,0.3219]	0.7003	[0.6402,0.7605]

Appendix IV. Genotypic data at five di-nucleotide microsatellite repeat loci for 378 individuals from 15 population samples of *P. parva*. Population abbreviations follow Table 1. Samples are in GENEPOP format.

Title line: microsatellite data of European P parva populations
PA02
PA05
PA06
PA22
PA24
pop
BK01 , 325329 094096 212212 116118 142164
BK02 , 325329 096096 212312 116118 164164
BK03 , 329329 096100 212282 116118 164164
BK07 , 327395 094100 240240 116118 142164
BK09 , 331331 094106 240282 118118 124142
BK10 , 325329 094096 240270 116118 136142
BK11 , 329000 094096 240000 116118 162164
BK12 , 325000 094094 270312 118118 164164
BK13 , 327333 094096 212288 116118 142182
BK14 , 329395 094100 270270 116118 136142
BK15 , 329345 094094 212288 116116 136164
BK18 , 395395 094094 240258 118118 164164
BK20 , 325327 094094 240270 116118 142142
BK27 , 327329 094096 258270 116118 142186
BK28 , 327395 094096 312312 116118 142164
BK29 , 325329 094094 258258 116118 142186
BK32 , 329000 094094 212000 116118 142164
BK33 , 325000 094106 212312 116116 142142
BK36 , 325325 094094 000000 116118 142142
BK37 , 331000 094094 312000 116118 142164
BK38 , 395395 096106 212240 116118 164186
BK39 , 325329 094094 212270 116118 142164
BK40 , 327331 094094 212258 116118 134142
BK41 , 331331 094106 270288 116116 164186
BK42 , 325327 094094 288312 116118 134164
BK43 , 327000 094094 288288 116118 182186
BK44 , 325325 094094 240288 116118 164212
BK45 , 325327 094096 000000 116118 142164
BK46 , 327331 094094 000000 116118 142142
BK47 , 325401 094094 312312 116118 134142
pop
BSa01 , 395395 094096 270288 116118 142186
BSa02 , 327395 094096 258258 118118 142162
BSa03 , 325000 094094 270312 116118 142142
BSa04 , 325327 094104 240270 118118 142186
BSa05 , 395000 094094 270312 118118 130164
BSa06 , 325325 094094 288312 116116 142164
BSa07 , 331395 094094 288288 118118 130142
BSa08 , 325329 094096 270270 118118 164186
BSa09 , 329000 106106 212288 118118 142140
BSa10 , 395395 094094 270288 118118 142186
BSa11 , 395000 094094 270312 116116 142180

BSa12 , 329331 094106 244288 116118 142186
BSa13 , 329329 094094 212270 116118 130130
BSa14 , 325327 094096 212240 116118 164180
BSa15 , 325327 094096 270312 116116 164186
BSa16 , 329329 094094 270288 116118 164186
BSa17 , 327331 094096 258312 118118 142164
BSa18 , 333395 094096 258258 116118 142186
BSa19 , 325000 094106 258312 116118 142186
BSa20 , 327329 094096 288312 116116 130142
BSa21 , 325329 096094 288288 116118 130164
BSa22 , 325327 094094 270288 116118 142186
BSa23 , 331333 094096 270288 118118 130142
BSa24 , 225327 094096 270288 116118 124142
BSa25 , 331331 094094 288312 116118 164164
BSa26 , 225329 094094 240240 116118 164186
BSa27 , 225327 094096 288312 118118 142164
BSa28 , 331000 094096 270312 116118 142164
BSa29 , 329000 094094 288312 116118 142164
BSa30 , 327333 094094 270270 116116 164186
pop
CRH01 , 000000 094096 000000 116138 158158
CRH02 , 313313 094094 262326 116138 130148
CRH03 , 323329 094100 292326 118118 130138
CRH04 , 327327 096098 212326 116116 134164
CRH05 , 325329 094096 272272 116138 134152
CRH06 , 000000 094096 240266 118118 130172
CRH07 , 331373 094100 232326 116138 150182
CRH08 , 000000 000000 000000 118140 150168
CRH09 , 327327 094094 272312 118134 130144
CRH10 , 323327 094094 326326 118118 164168
CRH11 , 317325 094104 240272 118134 130148
CRH12 , 317323 098098 212276 118140 136154
CRH13 , 323325 094100 238326 118134 134172
CRH14 , 323323 096098 212288 116118 130138
CRH15 , 325327 094098 204288 118118 128150
CRH16 , 000000 096098 000000 118118 150168
CRH17 , 325331 094094 312314 000000 134158
CRH18 , 323323 094096 272272 000000 132174
CRH19 , 327327 094104 272314 000000 136172
CRH20 , 327000 094104 212236 000000 124192
CRH21 , 327327 096098 288288 000000 130184
CRH22 , 325327 094096 240314 000000 134150
CRH23 , 313329 096098 312312 000000 160190
CRH24 , 327331 094096 000000 000000 130164
CRH25 , 325327 094094 000000 000000 132140
CRH26 , 317325 096096 000000 000000 140196
CRH27 , 327000 094094 000000 000000 132148
CRH28 , 325327 094094 000000 000000 124188
CRH29 , 323327 094096 000000 000000 134156
CRH30 , 313325 000000 312312 000000 134156
pop

EB01 , 327395 094094 258258 116118 178180
EB02 , 333333 094094 312312 116118 152196
EB03 , 333333 094094 258258 116116 124152
EB04 , 333395 094094 258258 116116 152152
EB05 , 329329 094094 258258 116118 130140
EB06 , 331333 094094 258312 118118 142180
EB07 , 309309 094094 258312 116118 130196
EB08 , 333000 094094 258258 116116 152180
EB09 , 329333 094094 258312 118118 142196
EB10 , 333395 094094 258312 116118 130152
EB11 , 333000 094094 258312 118118 142152
EB12 , 333395 094094 258258 116116 142152
EB13 , 327395 094094 258312 118118 130142
EB14 , 327395 094096 258312 118118 130140
EB15 , 333333 094000 258258 116118 140152
EB16 , 000000 094000 258258 118118 140142
EB17 , 327000 094000 258258 116116 140142
EB18 , 395000 094000 258258 118118 140140
EB19 , 327333 094000 258312 116116 164164
EB20 , 395395 096106 258312 116118 142180
EB21 , 325327 094000 258270 116118 164164
EB22 , 327329 094000 258258 116118 142164
EB23 , 329333 000000 312312 118118 130182
pop
FG01 , 327333 096096 312312 118118 142196
FG02 , 327331 094096 270288 116000 124136
FG03 , 333339 094094 270270 118118 164164
FG04 , 325325 094096 240270 118000 142186
FG05 , 327345 094094 240288 118000 152196
FG06 , 327345 096096 240240 118000 148206
FG07 , 327327 094094 240270 116116 148148
FG08 , 327345 094106 270370 116000 136142
FG09 , 331331 094094 264312 116118 124148
FG10 , 325333 094094 240282 116116 142148
FG11 , 339339 094094 270288 118118 164196
FG12 , 331345 094096 240288 118118 164164
FG13 , 325327 094096 270270 000000 196196
FG14 , 325325 096096 282288 116118 142142
FG15 , 325325 094096 240312 116118 148196
FG17 , 325345 094094 240288 118000 152196
FG18 , 327333 094094 240288 118000 142196
FG19 , 331333 094094 284312 116116 142196
FG20 , 325333 094094 270282 118118 142208
FG21 , 333345 094094 270288 118118 124124
FG22 , 333339 094094 282288 118118 142186
FG23 , 331343 094096 240312 118118 142196
FG24 , 313333 094106 270288 116118 142142
FG25 , 325345 094096 288288 116118 142142
FG26 , 325327 088088 270288 116118 136142
FG27 , 333333 088106 288288 118118 164164
FG28 , 327333 094106 288288 116000 142186

FG29 , 327345 094106 282288 116118 142164
FG30 , 325333 094106 288288 118118 136142
FG31 , 331333 094094 000000 000000 164196
pop
G01 , 327329 094094 258312 118000 142164
G02 , 333000 094094 212288 116118 142142
G03 , 325325 094096 244288 118000 142164
G04 , 323325 094096 212212 116118 142142
G05 , 329329 094094 240270 116118 142142
G06 , 327327 094094 312312 118000 130142
G07 , 327331 094094 244332 116116 142186
G08 , 327331 094106 240240 118000 142142
G09 , 327329 094094 288312 116116 124164
G10 , 325329 094096 288312 118118 142186
G11 , 319319 094094 288312 118118 142186
G12 , 331333 094094 244288 118118 142164
G13 , 325327 096100 288288 118122 142164
G14 , 321331 096096 312332 118118 142164
G15 , 325325 094094 288288 118118 142142
G16 , 327335 096094 286312 116118 142164
G17 , 329000 094094 212312 118118 130164
G18 , 329401 094094 288370 116116 136138
G19 , 329331 094094 244270 118118 164212
G20 , 327000 094094 240312 118118 136142
G21 , 333333 094096 240288 118118 142142
G22 , 327000 106106 212258 116000 164164
G23 , 325000 094094 240288 118000 142164
G24 , 327333 094096 262370 116000 130142
G25 , 331333 096096 288312 118000 142142
pop
HA01 , 327327 094096 282312 116116 124178
HA02 , 333333 094094 312312 114116 142164
HA03 , 327401 094094 240282 114116 124164
HA04 , 327331 096098 282312 118118 136152
HA05 , 327331 094096 282282 118118 164164
HA06 , 325333 094094 240312 116118 134160
HA07 , 333333 094094 282288 116118 124142
HA08 , 329337 088094 282312 118118 124142
HA09 , 325327 088094 270294 116116 136148
HA10 , 327327 088106 240270 118118 136182
HA11 , 327327 094094 240000 116118 142142
HA12 , 325327 094094 240240 118118 142162
HA13 , 327341 094096 240240 114118 124160
HA14 , 327327 094106 270312 118118 142142
HA15 , 325329 094094 240240 116118 160168
HA16 , 325327 094094 000000 118118 130182
HA17 , 325325 094094 312312 114116 148148
HA18 , 327341 094094 282312 116118 142164
HA19 , 327341 094094 240286 114118 162162
HA20 , 325333 094094 312000 118118 142164
HA21 , 333337 094094 000000 114116 124164

HA22 , 327331 094094 312000 114116 142162
HA23 , 325331 094094 240000 118118 162162
HA24 , 325333 096096 240288 116118 142000
HA25 , 327329 094096 270298 118118 124124
HA26 , 327331 094096 270270 118118 142158
HA27 , 327369 088096 222312 116118 142164
HA28 , 331331 094094 240312 116118 124152
HA29 , 333333 094094 240312 116118 142164
HA30 , 331333 094096 286312 116118 124124
pop
HE01 , 327327 094096 212270 116118 124130
HE02 , 333333 094094 270312 116118 142142
HE03 , 327333 094096 312000 116118 124182
HE04 , 327327 094094 288312 116116 186204
HE05 , 325327 094096 244244 116116 142196
HE06 , 327327 094106 244244 116116 164164
HE07 , 327333 094094 240288 114116 142162
HE08 , 327329 094094 240270 114116 142196
HE09 , 327331 094094 244312 116116 142164
HE10 , 325327 094094 258288 118118 142142
HE11 , 327327 094096 270270 116122 142164
HE12 , 327331 094094 270312 116122 142164
HE13 , 325327 094094 240240 116116 124130
HE14 , 325329 094096 244000 116116 144164
HE15 , 327337 094094 270270 116116 142142
HE16 , 325327 094094 270270 116116 144164
HE17 , 325327 094094 270270 116144 124186
HE18 , 325333 094106 270270 118144 142196
HE19 , 325331 094096 240000 116122 124142
HE20 , 327331 094094 256288 116122 130196
HE21 , 327329 096108 240312 116144 124142
HE22 , 327333 096096 240000 116116 124164
HE23 , 327333 094094 270000 116116 142142
HE24 , 327333 094096 240270 116144 162166
HE25 , 327329 094096 270000 118118 124182
HE26 , 329329 094094 270000 116118 142196
HE27 , 325327 094096 270000 116118 186186
HE28 , 327331 094094 270000 000000 142164
HE29 , 325325 094096 212240 116118 142164
HE30 , 325327 094094 248270 116118 142162
pop
HG01 , 331331 106106 240314 118118 196196
HG02 , 333333 094094 312312 116116 142152
HG03 , 331331 094094 312312 116116 164196
HG04 , 327327 094096 312000 118118 140152
HG05 , 325327 094096 212240 118118 140196
HG06 , 327333 094096 212212 116116 138164
HG07 , 327333 094094 270312 118000 164164
HG08 , 327337 094094 312000 116118 130164
HG09 , 331000 088106 312000 116118 142164
HG10 , 227331 094106 258258 118000 136160

```
HG11    ,       229331 094106 288312 118000 164194
HG12    ,       227333 094106 240258 116000 160160
HG13    ,       225225 094094 240000 116000 142164
HG14    ,       331333 094096 288288 116000 152164
HG15    ,       327327 094106 270284 118000 164164
HG16    ,       331000 096096 270312 118000 130164
HG17    ,       327000 094094 270312 116000 138180
HG18    ,       327000 094096 270000 118000 164000
HG19    ,       327327 094094 312312 118000 142164
HG20    ,       327000 094094 212312 118000 124164
HG21    ,       329331 094096 212000 118118 142142
HG22    ,       327331 094094 240270 118118 142164
HG23    ,       327329 094094 258312 118118 142142
HG24    ,       327395 094106 288288 118118 192192
pop
HS01    ,       325395 094096 240282 116116 138164
HS02    ,       325333 094096 270282 116134 138142
HS03    ,       395401 094096 240282 116118 138142
HS04    ,       325325 094094 212240 118118 142164
HS05    ,       325337 094094 288288 118118 164164
HS06    ,       327337 094094 240312 118118 142158
HS07    ,       333333 094094 282282 118118 142164
HS08    ,       325325 094106 270288 116118 142164
HS09    ,       327329 094094 212312 118118 138164
HS10    ,       327327 094096 282288 116118 142164
HS11    ,       327333 094094 312312 116134 164164
HS12    ,       327333 094094 212312 118134 142164
HS13    ,       327333 094096 240288 116118 168186
HS14    ,       313333 094096 212282 116116 142164
HS15    ,       327329 094094 282282 118118 142198
HS16    ,       325333 094096 212212 130130 142180
HS17    ,       329329 090108 270282 118134 142164
pop
IN01    ,       325329 094094 212312 116118 142142
IN02    ,       327329 094094 212312 116118 182196
IN03    ,       327327 094094 240312 116118 140140
IN04    ,       325329 094094 240322 000000 136196
IN05    ,       309309 094094 270312 000000 130142
IN06    ,       327329 096096 288312 000000 194196
IN07    ,       327327 000000 270288 000000 000000
IN08    ,       325327 094096 266266 000000 142164
IN09    ,       327327 094094 288312 118118 158158
IN10    ,       327329 094096 240240 116116 164168
IN11    ,       325325 094096 240312 118118 142196
IN12    ,       323325 094094 288312 116116 142168
IN13    ,       325329 094094 288312 118118 142196
IN14    ,       325327 094094 192248 118118 124142
IN15    ,       325329 000000 192192 118118 000000
IN16    ,       325327 094096 286312 118118 142000
IN17    ,       327327 094094 270274 000000 142164
IN18    ,       333333 094094 270270 000000 000000
```

IN19 , 325327 094094 266270 000000 142142
IN20 , 331333 094094 200200 000000 334334
pop
PU01 , 331333 094094 000000 000000 154174
PU02 , 325327 094096 212240 116116 136180
PU03 , 327000 094096 286286 116116 142180
PU04 , 325327 094096 282288 116116 124136
PU05 , 327333 094096 282312 116116 164164
PU06 , 329329 094096 312312 116116 180196
PU07 , 327333 094096 312312 116116 142186
PU08 , 327327 094094 212240 116116 164196
PU09 , 329329 094094 212240 116118 142164
PU10 , 327333 094094 288288 118136 164136
PU11 , 329329 094094 312312 116116 164180
PU12 , 327327 094094 240258 118136 164186
PU13 , 333000 094096 000000 116116 142162
PU14 , 325000 094094 312000 116118 136142
PU15 , 325327 094094 312000 116116 142196
PU16 , 327329 094116 212240 116118 164168
PU17 , 325333 094096 258258 000000 164186
PU18 , 329000 094096 000000 000000 164210
PU19 , 329331 094094 258312 000000 124164
PU20 , 327000 094094 288332 000000 136164
PU21 , 325333 094094 288312 000000 124164
PU22 , 327331 106106 288332 116116 164180
PU23 , 327333 094096 282312 000000 164180
PU24 , 325333 094096 288312 118118 136142
PU25 , 333000 094094 212212 118118 124142
PU26 , 329000 094096 282312 116116 142164
PU27 , 325325 094096 288288 116118 124164
PU28 , 329329 094096 312332 116118 142164
PU29 , 327329 094094 312332 118118 142164
PU30 , 329331 094094 258312 116118 124142
pop
SC01 , 327329 094094 270282 116118 142164
SC02 , 325325 094094 212288 118118 164210
SC03 , 325327 096106 312312 116130 182182
SC04 , 325325 096106 212288 118118 142142
SC05 , 309329 094106 244288 118118 142194
SC06 , 325325 094094 312330 116116 164182
SC07 , 325327 096118 282312 118126 134142
SC08 , 000000 096096 212212 000000 164164
SC09 , 313325 094096 312312 116118 142164
SC10 , 329329 094096 212282 118118 142164
SC11 , 325000 094094 000000 118134 142182
SC12 , 325000 094106 282286 118118 142142
SC13 , 325000 094096 000000 116118 134186
SC14 , 329000 094096 248262 118118 142182
SC15 , 325329 096096 000000 116118 142182
SC16 , 327000 096096 270270 116116 142182
SC17 , 327327 000000 244288 130130 142182

SC18 , 327000 000000 248248 000000 142164
pop
SE01 , 327327 094096 000000 164164 158174
SE02 , 327327 096096 288288 164186 164186
SE03 , 327327 094094 244244 164186 164186
SE04 , 327327 096096 244288 164186 164186
SE05 , 327327 094094 244312 164164 164164
SE06 , 327327 094096 244312 164164 164164
SE07 , 327327 094096 288288 186186 186186
SE08 , 327327 094094 278312 186186 164164
SE09 , 327327 094096 244248 000000 164186
SE10 , 329331 094096 240288 000000 164186
SE11 , 327327 104104 248270 000000 164210
SE12 , 327327 096096 000000 000000 164164
SE13 , 327327 000000 000000 000000 124164
pop
SWS01 , 331333 000000 312312 000000 130142
SWS02 , 327000 000000 288312 118118 138138
SWS03 , 333333 000000 270270 116118 134134
SWS04 , 333333 000000 288312 116118 130136
SWS05 , 327331 000000 270312 116118 130138
SWS06 , 327331 000000 258258 116118 130142
SWS07 , 331401 000000 240288 116118 130142
SWS08 , 333000 000000 000000 116118 142142
SWS09 , 327333 000000 212212 116118 130142
SWS10 , 333000 096096 240288 116118 130142
SWS11 , 329331 096000 240288 116118 130134
SWS12 , 331000 106000 286312 116118 130142
SWS13 , 329331 106000 240312 116118 142162
SWS14 , 329329 096000 000000 116118 142142
SWS15 , 329000 106000 286312 116118 136142
SWS16 , 000000 096000 212212 116118 130134
SWS17 , 327333 094096 270286 116118 118118
SWS18 , 329331 094096 270288 000000 130138
SWS19 , 331331 094000 248248 116118 118142
SWS20 , 331331 106106 312000 116118 130142
SWS21 , 331331 104104 000000 116118 142142
SWS22 , 327000 094000 282282 116118 000000
SWS23 , 331000 096000 270288 116118 130142
SWS24 , 331000 000000 248270 116118 134142
SWS25 , 327333 000000 288312 116116 142196
SWS26 , 331000 000000 248270 116118 130138
SWS27 , 327333 000000 000000 116118 130138
SWS28 , 333000 000000 000000 116118 142194

Appendix V. Table S2: Genotypic data at five di-nucleotide microsatellite repeat loci for 378 individuals from 15 population samples of *P. parva*. Population abbreviations follow Table 1. Samples are in FStat format.

```
378 5 401 3
PA02
PA05
PA06
PA22
PA24
1       325329 094096 212212 116118 142164
1       325329 096096 212312 116118 164164
1       329329 096100 212282 116118 164164
1       327395 094100 240240 116118 142164
1       331331 094106 240282 118118 124142
1       325329 094096 240270 116118 136142
1       329223 094096 240190 116118 162164
1       325223 094094 270312 118118 164164
1       327333 094096 212288 116118 142182
1       329395 094100 270270 116118 136142
1       329345 094094 212288 116116 136164
1       395395 094094 240258 118118 164164
1       325327 094094 240270 116118 142142
1       327329 094096 258270 116118 142186
1       327395 094096 312312 116118 142164
1       325329 094094 258258 116118 142186
1       329223 094094 212190 116118 142164
1       325223 094106 212312 116116 142142
1       325325 094094 190190 116118 142142
1       331223 094094 312190 116118 142164
1       395395 096106 212240 116118 164186
1       325329 094094 212270 116118 142164
1       327331 094094 212258 116118 134142
1       331331 094106 270288 116116 164186
1       325327 094094 288312 116118 134164
1       327223 094094 288288 116118 182186
1       325325 094094 240288 116118 164212
1       325327 094096 190190 116118 142164
1       327331 094094 190190 116118 142142
1       325401 094094 312312 116118 134142
2       395395 094096 270288 116118 142186
2       327395 094096 258258 118118 142162
2       325223 094094 270312 116118 142142
2       325327 094104 240270 118118 142186
2       395223 094094 270312 118118 130164
2       325325 094094 288312 116116 142164
2       331395 094094 288288 118118 130142
2       325329 094096 270270 118118 164186
2       329223 106106 212288 118118 142140
2       395395 094094 270288 118118 142186
2       395223 094094 270312 116116 142180
2       329331 094106 244288 116118 142186
```

2	329329	094094	212270	116118	130130
2	325327	094096	212240	116118	164180
2	325327	094096	270312	116116	164186
2	329329	094094	270288	116118	164186
2	327331	094096	258312	118118	142164
2	333395	094096	258258	116118	142186
2	325223	094106	258312	116118	142186
2	327329	094096	288312	116116	130142
2	325329	096094	288288	116118	130164
2	325327	094094	270288	116118	142186
2	331333	094096	270288	118118	130142
2	225327	094096	270288	116118	124142
2	331331	094094	288312	116118	164164
2	225329	094094	240240	116118	164186
2	225327	094096	288312	118118	142164
2	331223	094096	270312	116118	142164
2	329223	094094	288312	116118	142164
2	327333	094094	270270	116116	164186
3	223223	094096	190190	116138	158158
3	313313	094094	262326	116138	130148
3	323329	094100	292326	118118	130138
3	327327	096098	212326	116116	134164
3	325329	094096	272272	116138	134152
3	223223	094096	240266	118118	130172
3	331373	094100	232326	116138	150182
3	223223	086086	190190	118140	150168
3	327327	094094	272312	118134	130144
3	323327	094094	326326	118118	164168
3	317325	094104	240272	118134	130148
3	317323	098098	212276	118140	136154
3	323325	094100	238326	118134	134172
3	323323	096098	212288	116118	130138
3	325327	094098	204288	118118	128150
3	223223	096098	190190	118118	150168
3	325331	094094	312314	112112	134158
3	323323	094096	272272	112112	132174
3	327327	094104	272314	112112	136172
3	327223	094104	212236	112112	124192
3	327327	096098	288288	112112	130184
3	325327	094096	240314	112112	134150
3	313329	096098	312312	112112	160190
3	327331	094096	190190	112112	130164
3	325327	094094	190190	112112	132140
3	317325	096096	190190	112112	140196
3	327223	094094	190190	112112	132148
3	325327	094094	190190	112112	124188
3	323327	094096	190190	112112	134156
3	313325	086086	312312	112112	134156
4	327395	094094	258258	116118	178180
4	333333	094094	312312	116118	152196
4	333333	094094	258258	116116	124152

4	333395	094094	258258	116116	152152
4	329329	094094	258258	116118	130140
4	331333	094094	258312	118118	142180
4	309309	094094	258312	116118	130196
4	333223	094094	258258	116116	152180
4	329333	094094	258312	118118	142196
4	333395	094094	258312	116118	130152
4	333223	094094	258312	118118	142152
4	333395	094094	258258	116116	142152
4	327395	094094	258312	118118	130142
4	327395	094096	258312	118118	130140
4	333333	094086	258258	116118	140152
4	223223	094086	258258	118118	140142
4	327223	094086	258258	116116	140142
4	395223	094086	258258	118118	140140
4	327333	094086	258312	116116	164164
4	395395	096106	258312	116118	142180
4	325327	094086	258270	116118	164164
4	327329	094086	258258	116118	142164
4	329333	086086	312312	118118	130182
5	327333	096096	312312	118118	142196
5	327331	094096	270288	116112	124136
5	333339	094094	270270	118118	164164
5	325325	094096	240270	118112	142186
5	327345	094094	240288	118112	152196
5	327345	096096	240240	118112	148206
5	327327	094094	240270	116116	148148
5	327345	094106	270370	116112	136142
5	331331	094094	264312	116118	124148
5	325333	094094	240282	116116	142148
5	339339	094094	270288	118118	164196
5	331345	094096	240288	118118	164164
5	325327	094096	270270	112112	196196
5	325325	096096	282288	116118	142142
5	325325	094096	240312	116118	148196
5	325345	094094	240288	118112	152196
5	327333	094094	240288	118112	142196
5	331333	094094	284312	116116	142196
5	325333	094094	270282	118118	142208
5	333345	094094	270288	118118	124124
5	333339	094094	282288	118118	142186
5	331343	094096	240312	118118	142196
5	313333	094106	270288	116118	142142
5	325345	094096	288288	116118	142142
5	325327	088088	270288	116118	136142
5	333333	088106	288288	118118	164164
5	327333	094106	288288	116112	142186
5	327345	094106	282288	116118	142164
5	325333	094106	288288	118118	136142
5	331333	094094	190190	112112	164196
6	327329	094094	258312	118112	142164

6	333223	094094	212288	116118	142142
6	325325	094096	244288	118112	142164
6	323325	094096	212212	116118	142142
6	329329	094094	240270	116118	142142
6	327327	094094	312312	118112	130142
6	327331	094094	244332	116116	142186
6	327331	094106	240240	118112	142142
6	327329	094094	288312	116116	124164
6	325329	094096	288312	118118	142186
6	319319	094094	288312	118118	142186
6	331333	094094	244288	118118	142164
6	325327	096100	288288	118122	142164
6	321331	096096	312332	118118	142164
6	325325	094094	288288	118118	142142
6	327335	096094	286312	116118	142164
6	329223	094094	212312	118118	130164
6	329401	094094	288370	116116	136138
6	329331	094094	244270	118118	164212
6	327223	094094	240312	118118	136142
6	333333	094096	240288	118118	142142
6	327223	106106	212258	116112	164164
6	325223	094094	240288	118112	142164
6	327333	094096	262370	116112	130142
6	331333	096096	288312	118112	142142
7	327327	094096	282312	116116	124178
7	333333	094094	312312	114116	142164
7	327401	094094	240282	114116	124164
7	327331	096098	282312	118118	136152
7	327331	094096	282282	118118	164164
7	325333	094094	240312	116118	134160
7	333333	094094	282288	116118	124142
7	329337	088094	282312	118118	124142
7	325327	088094	270294	116116	136148
7	327327	088106	240270	118118	136182
7	327327	094094	240190	116118	142142
7	325327	094094	240240	118118	142162
7	327341	094096	240240	114118	124160
7	327327	094106	270312	118118	142142
7	325329	094094	240240	116118	160168
7	325327	094094	190190	118118	130182
7	325325	094094	312312	114116	148148
7	327341	094094	282312	116118	142164
7	327341	094094	240286	114118	162162
7	325333	094094	312190	118118	142164
7	333337	094094	190190	114116	124164
7	327331	094094	312190	114116	142162
7	325331	094094	240190	118118	162162
7	325333	096096	240288	116118	142116
7	327329	094096	270298	118118	124124
7	327331	094096	270270	118118	142158
7	327369	088096	222312	116118	142164

7	331331	094094	240312	116118	124152
7	333333	094094	240312	116118	142164
7	331333	094096	286312	116118	124124
8	327327	094096	212270	116118	124130
8	333333	094094	270312	116118	142142
8	327333	094096	312190	116118	124182
8	327327	094094	288312	116116	186204
8	325327	094096	244244	116116	142196
8	327327	094106	244244	116116	164164
8	327333	094094	240288	114116	142162
8	327329	094094	240270	114116	142196
8	327331	094094	244312	116116	142164
8	325327	094094	258288	118118	142142
8	327327	094096	270270	116122	142164
8	327331	094094	270312	116122	142164
8	325327	094094	240240	116116	124130
8	325329	094096	244190	116116	144164
8	327337	094094	270270	116116	142142
8	325327	094094	270270	116116	144164
8	325327	094094	270270	116144	124186
8	325333	094106	270270	118144	142196
8	325331	094096	240190	116122	124142
8	327331	094094	256288	116122	130196
8	327329	096108	240312	116144	124142
8	327333	096096	240190	116116	124164
8	327333	094094	270190	116116	142142
8	327333	094096	240270	116144	162166
8	327329	094096	270190	118118	124182
8	329329	094094	270190	116118	142196
8	325327	094096	270190	116118	186186
8	327331	094094	270190	112112	142164
8	325325	094096	212240	116118	142164
8	325327	094094	248270	116118	142162
9	331331	106106	240314	118118	196196
9	333333	094094	312312	116116	142152
9	331331	094094	312312	116116	164196
9	327327	094096	312190	118118	140152
9	325327	094096	212240	118118	140196
9	327333	094096	212212	116116	138164
9	327333	094094	270312	118112	164164
9	327337	094094	312190	116118	130164
9	331223	088106	312190	116118	142164
9	227331	094106	258258	118112	136160
9	229331	094106	288312	118112	164194
9	227333	094106	240258	116112	160160
9	225225	094094	240190	116112	142164
9	331333	094096	288288	116112	152164
9	327327	094106	270284	118112	164164
9	331223	096096	270312	118112	130164
9	327223	094094	270312	116112	138180
9	327223	094096	270190	118112	164116

9	327327	094094	312312	118112	142164
9	327223	094094	212312	118112	124164
9	329331	094096	212190	118118	142142
9	327331	094094	240270	118118	142164
9	327329	094094	258312	118118	142142
9	327395	094106	288288	118118	192192
10	325395	094096	240282	116116	138164
10	325333	094096	270282	116134	138142
10	395401	094096	240282	116118	138142
10	325325	094094	212240	118118	142164
10	325337	094094	288288	118118	164164
10	327337	094094	240312	118118	142158
10	333333	094094	282282	118118	142164
10	325325	094106	270288	116118	142164
10	327329	094094	212312	118118	138164
10	327327	094096	282288	116118	142164
10	327333	094094	312312	116134	164164
10	327333	094094	212312	118134	142164
10	327333	094096	240288	116118	168186
10	313333	094096	212282	116116	142164
10	327329	094094	282282	118118	142198
10	325333	094096	212212	130130	142180
10	329329	090108	270282	118134	142164
11	325329	094094	212312	116118	142142
11	327329	094094	212312	116118	182196
11	327327	094094	240312	116118	140140
11	325329	094094	240322	112112	136196
11	309309	094094	270312	112112	130142
11	327329	096096	288312	112112	194196
11	327327	086086	270288	112112	116116
11	325327	094096	266266	112112	142164
11	327327	094094	288312	118118	158158
11	327329	094096	240240	116116	164168
11	325325	094096	240312	118118	142196
11	323325	094094	288312	116116	142168
11	325329	094094	288312	118118	142196
11	325327	094094	192248	118118	124142
11	325329	086086	192192	118118	116116
11	325327	094096	286312	118118	142116
11	327327	094094	270274	112112	142164
11	333333	094094	270270	112112	116116
11	325327	094094	266270	112112	142142
11	331333	094094	200200	112112	334334
12	331333	094094	190190	112112	154174
12	325327	094096	212240	116116	136180
12	327223	094096	286286	116116	142180
12	325327	094096	282288	116116	124136
12	327333	094096	282312	116116	164164
12	329329	094096	312312	116116	180196
12	327333	094096	312312	116116	142186
12	327327	094094	212240	116116	164196

12	329329	094094	212240	116118	142164
12	327333	094094	288288	118136	164136
12	329329	094094	312312	116116	164180
12	327327	094094	240258	118136	164186
12	333223	094096	190190	116116	142162
12	325223	094094	312190	116118	136142
12	325327	094094	312190	116116	142196
12	327329	094116	212240	116118	164168
12	325333	094096	258258	112112	164186
12	329223	094096	190190	112112	164210
12	329331	094094	258312	112112	124164
12	327223	094094	288332	112112	136164
12	325333	094094	288312	112112	124164
12	327331	106106	288332	116116	164180
12	327333	094096	282312	112112	164180
12	325333	094096	288312	118118	136142
12	333223	094094	212212	118118	124142
12	329223	094096	282312	116116	142164
12	325325	094096	288288	116118	124164
12	329329	094096	312332	116118	142164
12	327329	094094	312332	118118	142164
12	329331	094094	258312	116118	124142
13	327329	094094	270282	116118	142164
13	325325	094094	212288	118118	164210
13	325327	096106	312312	116130	182182
13	325325	096106	212288	118118	142142
13	309329	094106	244288	118118	142194
13	325325	094094	312330	116116	164182
13	325327	096118	282312	118126	134142
13	223223	096096	212212	112112	164164
13	313325	094096	312312	116118	142164
13	329329	094096	212282	118118	142164
13	325223	094094	190190	118134	142182
13	325223	094106	282286	118118	142142
13	325223	094096	190190	116118	134186
13	329223	094096	248262	118118	142182
13	325329	096096	190190	116118	142182
13	327223	096096	270270	116116	142182
13	327327	086086	244288	130130	142182
13	327223	086086	248248	112112	142164
14	327327	094096	190190	164164	158174
14	327327	096096	288288	164186	164186
14	327327	094094	244244	164186	164186
14	327327	096096	244288	164186	164186
14	327327	094094	244312	164164	164164
14	327327	094096	244312	164164	164164
14	327327	094096	288288	186186	186186
14	327327	094094	278312	186186	164164
14	327327	094096	244248	112112	164186
14	329331	094096	240288	112112	164186
14	327327	104104	248270	112112	164210

14	327327	096096	190190	112112	164164
14	327327	086086	190190	112112	124164
15	331333	086086	312312	112112	130142
15	327223	086086	288312	118118	138138
15	333333	086086	270270	116118	134134
15	333333	086086	288312	116118	130136
15	327331	086086	270312	116118	130138
15	327331	086086	258258	116118	130142
15	331401	086086	240288	116118	130142
15	333223	086086	190190	116118	142142
15	327333	086086	212212	116118	130142
15	333223	096096	240288	116118	130142
15	329331	096086	240288	116118	130134
15	331223	106086	286312	116118	130142
15	329331	106086	240312	116118	142162
15	329329	096086	190190	116118	142142
15	329223	106086	286312	116118	136142
15	223223	096086	212212	116118	130134
15	327333	094096	270286	116118	118118
15	329331	094096	270288	112112	130138
15	331331	094086	248248	116118	118142
15	331331	106106	312190	116118	130142
15	331331	104104	190190	116118	142142
15	327223	094086	282282	116118	116116
15	331223	096086	270288	116118	130142
15	331223	086086	248270	116118	134142
15	327333	086086	288312	116116	142196
15	331223	086086	248270	116118	130138
15	327333	086086	190190	116118	130138
15	333223	086086	190190	116118	142194

Appendix VI. Hardy-Weinberg probability test of European *P. parva* populations.

Genepop version 4.1.0: Hardy-Weinberg test
File: Pparva_genepop3.txt (Title line: microsatellite data of European P parva populations)

Number of populations detected: 15

Number of loci detected: 5

Hardy-Weinberg exact test for up to four alleles.

(complete enumeration)

Estimation of exact P-Values by the Markov chain method.

Markov chain parameters for all tests:

Dememorization: 10000

Batches: 20

Iterations per batch: 5000

Hardy Weinberg: Probability test

****

==========

Results by locus

==========

Locus "PA02"

Fis estimates

POP	P-val	S.E.	W&C	R&H	Steps
BK47	0.0263	0.0062	0.1463	0.1306	9056 switches
BSa30	0.1641	0.0083	0.1373	0.1249	20030 switches
CRH30	0.1006	0.0081	0.1407	0.0915	10379 switches
EB23	0.1808	0.0150	0.1605	0.2201	8806 switches

FG31 0.3284 0.0124 0.0844 0.0982 12746 switches

G25 0.1859 0.0274 0.1964 0.1950 4948 switches

HA30 0.1403 0.0157 0.1325 0.0651 7561 switches

HE30 0.8514 0.0065 -0.0672 0.0245 17207 switches

HG24 0.4590 0.0342 0.2272 0.1693 3340 switches

HS17 0.1530 0.0121 0.1597 0.1049 7191 switches

IN20 0.0056 0.0018 0.1512 0.2627 8131 switches

PU30 0.0668 0.0038 0.1200 0.0745 32668 switches

SC18 0.5827 0.0118 0.2500 0.1550 10957 switches

SE13 0.0400 - 0.5000 0.0417 2 matrices

SWS28 0.1150 0.0068 0.1436 0.1101 18943 switches

All (Fisher's method):

Chi2: 61.6913

Df : 30.0000

Prob : 0.0006

Locus "PA05"

Fis estimates

POP P-val S.E. W&C R&H Steps

BK47 0.9333 - 0.0229 -0.0180 451 matrices

BSa30 0.0970 - -0.1471 0.0576 153 matrices

CRH30 0.1709 0.0083 -0.0079 -0.0088 21129 switches

EB23 0.1034 - 0.3086 0.0165 3 matrices

FG31 0.0153 - 0.2129 0.2927 977 matrices

G25 0.0322 - 0.3191 0.3034 88 matrices

HA30 0.1302 0.0111 0.1333 0.0149 8072 switches

HE30 0.4391 - -0.0301 -0.0231 67 matrices

HG24 0.2299 - 0.0723 0.0712 225 matrices

HS17 0.0943 0.0040 -0.1294 -0.0547 4745 switches

IN20 0.3900 - 0.2273 0.2353 4 matrices

PU30 0.0106 - -0.1447 0.2380 72 matrices

SC18 0.7018 - 0.1346 0.0447 147 matrices

SE13 0.0677 - 0.3125 0.6157 22 matrices

SWS28 0.0603 - 0.5294 0.6875 23 matrices

All (Fisher's method):

Chi2: 64.1824

Df : 30.0000

Prob : 0.0003

Locus "PA06"

Fis estimates

POP P-val S.E. W&C R&H Steps

--- --- --- --- --- ---

BK47 0.8087 0.0073 0.1768 0.1726 21004 switches

BSa30 0.0333 0.0050 0.0368 0.1253 14500 switches

CRH30 0.0187 0.0071 0.2218 0.1000 1845 switches

EB23 1.0000 - 0.0244 0.0203 14 matrices

FG31 0.2241 0.0193 0.0841 0.0503 8296 switches

G25 0.5637 0.0251 0.0625 0.0413 4660 switches

HA30 0.5321 0.0150 0.1282 0.0640 5857 switches

HE30 0.0296 0.0071 0.2353 0.1161 5495 switches

HG24 0.0240 0.0045 0.2738 0.2206 8404 switches

HS17 0.4686 0.0085 0.1579 0.1412 21053 switches

IN20 0.0054 0.0046 0.1581 0.2191 3204 switches

PU30 0.0006 0.0004 0.1992 0.2509 14157 switches

SC18 0.0094 0.0040 0.2651 0.1968 4489 switches

SE13 0.2551 0.0142 0.1544 0.0556 5573 switches

SWS28 0.0007 0.0005 0.2279 0.4042 10078 switches

All (Fisher's method):

Chi2: 88.3591

Df : 30.0000

Prob : 0.0000

Locus "PA22"

Fis estimates

POP P-val S.E. W&C R&H Steps

BK47 0.0027 - -0.5890 -0.5931 16 matrices

BSa30 1.0000 - -0.0116 -0.0118 13 matrices

CRH30 0.0225 0.0025 0.1045 0.0103 15954 switches

EB23 0.4007 - 0.2326 0.2390 11 matrices

FG31 0.3497 - 0.2267 0.2339 7 matrices

G25 0.1520 - 0.3798 0.2285 11 matrices

HA30 0.2149 - 0.0140 -0.0552 180 matrices

HE30 0.9058 0.0072 -0.0270 -0.0104 14371 switches

HG24 0.0437 - 0.6508 0.6989 5 matrices

HS17 0.1023 - 0.2644 0.3979 250 matrices

IN20 0.2322 - 0.4118 0.4400 4 matrices

PU30 0.0547 - 0.2960 0.1396 30 matrices

SC18 0.1762 0.0087 0.2929 0.2343 8564 switches

SE13 0.5301 - 0.3000 0.3265 4 matrices

SWS28 0.0000 - -0.8405 -0.8431 14 matrices

All (Fisher's method):

Chi2: 75.5695

Df : 30.0000

Prob : 0.0000

Locus "PA24"

Fis estimates

POP P-val S.E. W&C R&H Steps

--- --- --- --- --- ---

BK47 0.9740 0.0044 -0.0063 -0.0172 4240 switches

BSa30 0.5774 0.0205 -0.1687 -0.0533 7250 switches

CRH30 0.1550 0.0550 -0.0139 0.0037 576 switches (low!)

EB23 0.6244 0.0159 0.0596 0.0577 6640 switches

FG31	0.0278	0.0063	0.1635	0.1300	6877 switches
G25	0.4708	0.0279	-0.0408	-0.0251	3931 switches
HA30	0.0000	0.0000	0.1742	0.1122	2770 switches
HE30	0.0036	0.0019	0.0327	0.0255	5037 switches
HG24	0.0855	0.0269	0.1833	0.1828	2391 switches
HS17	0.1949	0.0172	-0.2121	-0.0707	3577 switches
IN20	0.0582	0.0191	0.2010	0.2554	2111 switches
PU30	0.9957	0.0018	-0.1576	-0.0655	3615 switches
SC18	0.5314	0.0149	-0.0485	-0.0106	6383 switches
SE13	0.3180	0.0161	-0.0052	0.0040	4137 switches
SWS28	0.0975	0.0103	-0.0055	0.1084	5236 switches

All (Fisher's method):

Chi2: Infinity

Df : 30.0000

Prob : High. sign.

==========

Appendix VII. Linkage disequilibrium test of European *P. parva* populations.

Genepop version 4.1.0, Genotypic linkage disequilibrium, significantly different p-values are shown in green.

File: Pparva_genepop2.txt (Title line: microsatellite data of European P parva populations)

Number of populations detected : 15
Number of loci detected : 5

Markov chain parameters
Dememorisation : 10000
Batches : 20
Iterations per batch : 5000

Pop	Locus#1	Locus#2	P-Value	S.E.	Switches
BK	PA02	PA05	0.35361	0.035111	3099
BK	PA02	PA06	1.000000	0.000000	485
BK	PA05	PA06	1.000000	0.000000	2077
BK	PA02	PA22	0.103900	0.011124	8274
BK	PA05	PA22	0.102670	0.007953	11753
BK	PA06	PA22	0.870150	0.011488	6984
BK	PA02	PA24	0.445970	0.051067	1242
BK	PA05	PA24	0.393940	0.045924	3559
BK	PA06	PA24	1.000000	0.000000	684
BK	PA22	PA24	0.132820	0.012763	8887
BSa	PA02	PA05	0.438020	0.031686	4292
BSa	PA02	PA06	0.426860	0.063328	1149
BSa	PA05	PA06	0.203370	0.018759	5811
BSa	PA02	PA22	0.898830	0.008742	14111
BSa	PA05	PA22	0.526970	0.009620	25086
BSa	PA06	PA22	0.947220	0.003769	18018
BSa	PA02	PA24	0.291330	0.046775	1123
BSa	PA05	PA24	0.842860	0.015085	5432
BSa	PA06	PA24	0.085310	0.044938	1919
BSa	PA22	PA24	0.647210	0.011246	17025
CRH	PA02	PA05	0.699490	0.035987	3125
CRH	PA02	PA06	1.000000	0.000000	556
CRH	PA05	PA06	0.102730	0.024600	2255
CRH	PA02	PA22	1.000000	0.000000	3357
CRH	PA05	PA22	0.948350	0.005448	7254
CRH	PA06	PA22	No information		
CRH	PA02	PA24	1.000000	0.000000	479
CRH	PA05	PA24	1.000000	0.000000	1707
CRH	PA06	PA24	1.000000	0.000000	338
CRH	PA22	PA24	1.000000	0.000000	3731
EB	PA02	PA05	0.221640	0.013107	6059
EB	PA02	PA06	0.842370	0.012355	7516
EB	PA05	PA06	0.306090	0.010124	12950
EB	PA02	PA22	0.357790	0.010937	16183
EB	PA05	PA22	1.000000	0.000000	28363

EB	PA06	PA22	0.390900	0.006905	33552
EB	PA02	PA24	1.000000	0.000000	742
EB	PA05	PA24	0.344580	0.022144	5481
EB	PA06	PA24	0.626240	0.019213	6252
EB	PA22	PA24	1.000000	0.000000	13842
FG	PA02	PA05	0.495350	0.030934	4645
FG	PA02	PA06	1.000000	0.000000	1273
FG	PA05	PA06	0.739170	0.017475	5474
FG	PA02	PA22	0.434190	0.022662	6821
FG	PA05	PA22	0.510170	0.016354	17287
FG	PA06	PA22	0.262260	0.016382	8250
FG	PA02	PA24	0.692830	0.066905	975
FG	PA05	PA24	0.754850	0.021465	4540
FG	PA06	PA24	0.780580	0.040438	1260
FG	PA22	PA24	0.024560	0.005795	7414
G	PA02	PA05	0.914180	0.015828	2629
G	PA02	PA06	1.000000	0.000000	421
G	PA05	PA06	0.977740	0.005658	2609
G	PA02	PA22	1.000000	0.000000	3579
G	PA05	PA22	0.217360	0.013880	8815
G	PA06	PA22	0.875430	0.020970	3669
G	PA02	PA24	1.000000	0.000000	1216
G	PA05	PA24	0.998930	0.000525	4231
G	PA06	PA24	0.410920	0.056181	1383
G	PA22	PA24	0.243650	0.023117	6524
HA	PA02	PA05	0.780360	0.041526	2205
HA	PA02	PA06	1.000000	0.000000	955
HA	PA05	PA06	0.114850	0.029554	1925
HA	PA02	PA22	0.184350	0.018127	6790
HA	PA05	PA22	0.383930	0.011183	8326
HA	PA06	PA22	0.346710	0.023388	5607
HA	PA02	PA24	0.209000	0.064660	736
HA	PA05	PA24	0.015520	0.006675	1765
HA	PA06	PA24	1.000000	0.000000	613
HA	PA22	PA24	0.591150	0.025772	5556
HE	PA02	PA05	0.339940	0.019072	5900
HE	PA02	PA06	0.973350	0.012022	1416
HE	PA05	PA06	0.719640	0.027713	4326
HE	PA02	PA22	0.649180	0.026609	5393
HE	PA05	PA22	0.622730	0.014332	9976
HE	PA06	PA22	0.728320	0.024075	3389
HE	PA02	PA24	0.598570	0.041984	1948
HE	PA05	PA24	0.168770	0.022785	5200
HE	PA06	PA24	1.000000	0.000000	1052
HE	PA22	PA24	0.854160	0.019287	4552
HG	PA02	PA05	0.951060	0.009915	3571
HG	PA02	PA06	1.000000	0.000000	717
HG	PA05	PA06	0.254150	0.025998	3442
HG	PA02	PA22	0.607790	0.022924	5787
HG	PA05	PA22	0.841020	0.008882	13905
HG	PA06	PA22	0.281340	0.022756	5732

HG	PA02	PA24	1.000000	0.000000	601
HG	PA05	PA24	0.409030	0.034441	3328
HG	PA06	PA24	1.000000	0.000000	612
HG	PA22	PA24	0.534090	0.023847	5272
HS	PA02	PA05	0.310320	0.011210	6972
HS	PA02	PA06	1.000000	0.000000	1223
HS	PA05	PA06	0.146830	0.009159	6456
HS	PA02	PA22	0.687760	0.020967	4397
HS	PA05	PA22	0.003300	0.001276	12199
HS	PA06	PA22	0.533180	0.025325	3900
HS	PA02	PA24	0.705220	0.033670	2418
HS	PA05	PA24	0.833250	0.009570	8433
HS	PA06	PA24	1.000000	0.000000	2156
HS	PA22	PA24	0.778910	0.014873	6333
IN	PA02	PA05	0.664840	0.013211	10977
IN	PA02	PA06	1.000000	0.000000	1895
IN	PA05	PA06	0.688120	0.015426	8413
IN	PA02	PA22	0.684700	0.007853	18557
IN	PA05	PA22	0.500220	0.002639	42408
IN	PA06	PA22	0.545910	0.008991	15254
IN	PA02	PA24	1.000000	0.000000	2132
IN	PA05	PA24	0.909680	0.008367	8461
IN	PA06	PA24	1.000000	0.000000	997
IN	PA22	PA24	0.302740	0.009697	12566
PU	PA02	PA05	0.467210	0.015294	8048
PU	PA02	PA06	0.001400	0.001096	1919
PU	PA05	PA06	0.592480	0.016828	7578
PU	PA02	PA22	0.347740	0.016624	9225
PU	PA05	PA22	0.376830	0.010099	17366
PU	PA06	PA22	0.155510	0.012377	9240
PU	PA02	PA24	0.716850	0.061504	1111
PU	PA05	PA24	0.818010	0.014097	6156
PU	PA06	PA24	0.611310	0.063493	1304
PU	PA22	PA24	0.889650	0.009902	6903
SC	PA02	PA05	0.681230	0.015662	4841
SC	PA02	PA06	0.191870	0.041242	1544
SC	PA05	PA06	1.000000	0.000000	3837
SC	PA02	PA22	0.780570	0.022615	3052
SC	PA05	PA22	0.589150	0.014896	8429
SC	PA06	PA22	0.577070	0.027024	3772
SC	PA02	PA24	1.000000	0.000000	2354
SC	PA05	PA24	0.409430	0.023093	5265
SC	PA06	PA24	1.000000	0.000000	1774
SC	PA22	PA24	0.744040	0.022102	3640
SE	PA02	PA05	1.000000	0.000000	16825
SE	PA02	PA06	0.600910	0.007625	10462
SE	PA05	PA06	1.000000	0.000000	8370
SE	PA02	PA22	No contingency table		
SE	PA05	PA22	0.619440	0.004869	33151
SE	PA06	PA22	0.415310	0.007783	17410
SE	PA02	PA24	1.000000	0.000000	9276

SE	PA05	PA24	0.787260	0.007776	13210
SE	PA06	PA24	0.494840	0.015290	7463
SE	PA22	PA24	0.067940	0.002890	22361
SWS	PA02	PA05	1.000000	0.000000	4408
SWS	PA02	PA06	0.518090	0.044425	1648
SWS	PA05	PA06	0.316500	0.023724	4511
SWS	PA02	PA22	0.711990	0.011759	4671
SWS	PA05	PA22	No contingency table		
SWS	PA06	PA22	0.967420	0.007107	3503
SWS	PA02	PA24	0.951390	0.017350	1640
SWS	PA05	PA24	0.708020	0.014935	4006
SWS	PA06	PA24	0.808760	0.037818	1242
SWS	PA22	PA24	0.174070	0.015148	3100

P-value for each locus pair across all populations
(Fisher's method)

Locus pair		Chi2	df	P-Value
PA02	& PA05	17.243461	30	0.969538
PA02	& PA06	20.877967	30	0.891665
PA05	& PA06	26.412199	30	0.653936
PA02	& PA22	18.501545	28	0.912808
PA05	& PA22	30.327365	28	0.347742
PA06	& PA22	19.610871	28	0.878232
PA02	& PA24	10.436742	30	0.999648
PA05	& PA24	22.488983	30	0.835671
PA06	& PA24	10.948970	30	0.999429
PA22	& PA24	30.355546	30	0.447557

Appendix VIII. Input file for Structure using European *P. parva* populations.

	PA02	PA05	PA06	PA22	PA24	
BK01	1	325	94	212	116	142
BK01	1	329	96	212	118	164
BK02	1	325	96	212	116	164
BK02	1	329	96	312	118	164
BK03	1	329	96	212	116	164
BK03	1	329	100	282	118	164
BK07	1	327	94	240	116	142
BK07	1	395	100	240	118	164
BK09	1	331	94	240	118	124
BK09	1	331	106	282	118	142
BK10	1	325	94	240	116	136
BK10	1	329	96	270	118	142
BK11	1	-9	94	-9	116	162
BK11	1	-9	96	-9	118	164
BK12	1	-9	94	270	118	164
BK12	1	-9	94	312	118	164
BK13	1	327	94	212	116	142
BK13	1	333	96	288	118	182
BK14	1	329	94	270	116	136
BK14	1	395	100	270	118	142
BK15	1	329	94	212	116	136
BK15	1	345	94	288	116	164
BK18	1	395	94	240	118	164
BK18	1	395	94	258	118	164
BK20	1	325	94	240	116	142
BK20	1	327	94	270	118	142
BK27	1	327	94	258	116	142
BK27	1	329	96	270	118	186
BK28	1	327	94	312	116	142
BK28	1	395	96	312	118	164
BK29	1	325	94	258	116	142
BK29	1	329	94	258	118	186
BK32	1	-9	94	-9	116	142
BK32	1	-9	94	-9	118	164
BK33	1	-9	94	212	116	142
BK33	1	-9	106	312	116	142
BK36	1	325	94	-9	116	142
BK36	1	325	94	-9	118	142
BK37	1	-9	94	-9	116	142
BK37	1	-9	94	-9	118	164
BK38	1	395	96	212	116	164
BK38	1	395	106	240	118	186
BK39	1	325	94	212	116	142
BK39	1	329	94	270	118	164
BK40	1	327	94	212	116	134
BK40	1	331	94	258	118	142
BK41	1	331	94	270	116	164
BK41	1	331	106	288	116	186
BK42	1	325	94	288	116	134

BK42	1	327	94	312	118	164
BK43	1	-9	94	288	116	182
BK43	1	-9	94	288	118	186
BK44	1	325	94	240	116	164
BK44	1	325	94	288	118	212
BK45	1	325	94	-9	116	142
BK45	1	327	96	-9	118	164
BK46	1	327	94	-9	116	142
BK46	1	331	94	-9	118	142
BK47	1	325	94	312	116	134
BK47	1	401	94	312	118	142
BSa01	2	395	94	270	116	142
BSa01	2	395	96	288	118	186
BSa02	2	327	94	258	118	142
BSa02	2	395	96	258	118	162
BSa03	2	-9	94	270	116	142
BSa03	2	-9	94	312	118	142
BSa04	2	325	94	240	118	142
BSa04	2	327	104	270	118	186
BSa05	2	-9	94	270	118	130
BSa05	2	-9	94	312	118	164
BSa06	2	325	94	288	116	142
BSa06	2	325	94	312	116	164
BSa07	2	331	94	288	118	130
BSa07	2	395	94	288	118	142
BSa08	2	325	94	270	118	164
BSa08	2	329	96	270	118	186
BSa09	2	-9	106	212	118	142
BSa09	2	-9	106	288	118	140
BSa10	2	395	94	270	118	142
BSa10	2	395	94	288	118	186
BSa11	2	-9	94	270	116	142
BSa11	2	-9	94	312	116	180
BSa12	2	329	94	244	116	142
BSa12	2	331	106	288	118	186
BSa13	2	329	94	212	116	130
BSa13	2	329	94	270	118	130
BSa14	2	325	94	212	116	164
BSa14	2	327	96	240	118	180
BSa15	2	325	94	270	116	164
BSa15	2	327	96	312	116	186
BSa16	2	329	94	270	116	164
BSa16	2	329	94	288	118	186
BSa17	2	327	94	258	118	142
BSa17	2	331	96	312	118	164
BSa18	2	333	94	258	116	142
BSa18	2	395	96	258	118	186
BSa19	2	-9	94	258	116	142
BSa19	2	-9	106	312	118	186
BSa20	2	327	94	288	116	130
BSa20	2	329	96	312	116	142

BSa21	2	325	96	288	116	130
BSa21	2	329	94	288	118	164
BSa22	2	325	94	270	116	142
BSa22	2	327	94	288	118	186
BSa23	2	331	94	270	118	130
BSa23	2	333	96	288	118	142
BSa24	2	225	94	270	116	124
BSa24	2	327	96	288	118	142
BSa25	2	331	94	288	116	164
BSa25	2	331	94	312	118	164
BSa26	2	225	94	240	116	164
BSa26	2	329	94	240	118	186
BSa27	2	225	94	288	118	142
BSa27	2	327	96	312	118	164
BSa28	2	-9	94	270	116	142
BSa28	2	-9	96	312	118	164
BSa29	2	-9	94	288	116	142
BSa29	2	-9	94	312	118	164
BSa30	2	327	94	270	116	164
BSa30	2	333	94	270	116	186
CRH01	3	-9	94	-9	116	158
CRH01	3	-9	96	-9	138	158
CRH02	3	313	94	262	116	130
CRH02	3	313	94	326	138	148
CRH03	3	323	94	292	118	130
CRH03	3	329	100	326	118	138
CRH04	3	327	96	212	116	134
CRH04	3	327	98	326	116	164
CRH05	3	325	94	272	116	134
CRH05	3	329	96	272	138	152
CRH06	3	-9	94	240	118	130
CRH06	3	-9	96	266	118	172
CRH07	3	331	94	232	116	150
CRH07	3	373	100	326	138	182
CRH08	3	-9	-9	-9	118	150
CRH08	3	-9	-9	-9	140	168
CRH09	3	327	94	272	118	130
CRH09	3	327	94	312	134	144
CRH10	3	323	94	326	118	164
CRH10	3	327	94	326	118	168
CRH11	3	317	94	240	118	130
CRH11	3	325	104	272	134	148
CRH12	3	317	98	212	118	136
CRH12	3	323	98	276	140	154
CRH13	3	323	94	238	118	134
CRH13	3	325	100	326	134	172
CRH14	3	323	96	212	116	130
CRH14	3	323	98	288	118	138
CRH15	3	325	94	204	118	128
CRH15	3	327	98	288	118	150
CRH16	3	-9	96	-9	118	150

CRH16	3	-9	98	-9	118	168
CRH17	3	325	94	312	-9	134
CRH17	3	331	94	314	-9	158
CRH18	3	323	94	272	-9	132
CRH18	3	323	96	272	-9	174
CRH19	3	327	94	272	-9	136
CRH19	3	327	104	314	-9	172
CRH20	3	-9	94	212	-9	124
CRH20	3	-9	104	236	-9	192
CRH21	3	327	96	288	-9	130
CRH21	3	327	98	288	-9	184
CRH22	3	325	94	240	-9	134
CRH22	3	327	96	314	-9	150
CRH23	3	313	96	312	-9	160
CRH23	3	329	98	312	-9	190
CRH24	3	327	94	-9	-9	130
CRH24	3	331	96	-9	-9	164
CRH25	3	325	94	-9	-9	132
CRH25	3	327	94	-9	-9	140
CRH26	3	317	96	-9	-9	140
CRH26	3	325	96	-9	-9	196
CRH27	3	-9	94	-9	-9	132
CRH27	3	-9	94	-9	-9	148
CRH28	3	325	94	-9	-9	124
CRH28	3	327	94	-9	-9	188
CRH29	3	323	94	-9	-9	134
CRH29	3	327	96	-9	-9	156
CRH30	3	313	-9	312	-9	134
CRH30	3	325	-9	312	-9	156
EB01	4	327	94	258	116	178
EB01	4	395	94	258	118	180
EB02	4	333	94	312	116	152
EB02	4	333	94	312	118	196
EB03	4	333	94	258	116	124
EB03	4	333	94	258	116	152
EB04	4	333	94	258	116	152
EB04	4	395	94	258	116	152
EB05	4	329	94	258	116	130
EB05	4	329	94	258	118	140
EB06	4	331	94	258	118	142
EB06	4	333	94	312	118	180
EB07	4	309	94	258	116	130
EB07	4	309	94	312	118	196
EB08	4	-9	94	258	116	152
EB08	4	-9	94	258	116	180
EB09	4	329	94	258	118	142
EB09	4	333	94	312	118	196
EB10	4	333	94	258	116	130
EB10	4	395	94	312	118	152
EB11	4	-9	94	258	118	142
EB11	4	-9	94	312	118	152

EB12	4	333	94	258	116	142
EB12	4	395	94	258	116	152
EB13	4	327	94	258	118	130
EB13	4	395	94	312	118	142
EB14	4	327	94	258	118	130
EB14	4	395	96	312	118	140
EB15	4	333	-9	258	116	140
EB15	4	333	-9	258	118	152
EB16	4	-9	-9	258	118	140
EB16	4	-9	-9	258	118	142
EB17	4	-9	-9	258	116	140
EB17	4	-9	-9	258	116	142
EB18	4	-9	-9	258	118	140
EB18	4	-9	-9	258	118	140
EB19	4	327	-9	258	116	164
EB19	4	333	-9	312	116	164
EB20	4	395	96	258	116	142
EB20	4	395	106	312	118	180
EB21	4	325	-9	258	116	164
EB21	4	327	-9	270	118	164
EB22	4	327	-9	258	116	142
EB22	4	329	-9	258	118	164
EB23	4	329	-9	312	118	130
EB23	4	333	-9	312	118	182
FG01	5	327	96	312	118	142
FG01	5	333	96	312	118	196
FG02	5	327	94	270	-9	124
FG02	5	331	96	288	-9	136
FG03	5	333	94	270	118	164
FG03	5	339	94	270	118	164
FG04	5	325	94	240	-9	142
FG04	5	325	96	270	-9	186
FG05	5	327	94	240	-9	152
FG05	5	345	94	288	-9	196
FG06	5	327	96	240	-9	148
FG06	5	345	96	240	-9	206
FG07	5	327	94	240	116	148
FG07	5	327	94	270	116	148
FG08	5	327	94	270	-9	136
FG08	5	345	106	370	-9	142
FG09	5	331	94	264	116	124
FG09	5	331	94	312	118	148
FG10	5	325	94	240	116	142
FG10	5	333	94	282	116	148
FG11	5	339	94	270	118	164
FG11	5	339	94	288	118	196
FG12	5	331	94	240	118	164
FG12	5	345	96	288	118	164
FG13	5	325	94	270	-9	196
FG13	5	327	96	270	-9	196
FG14	5	325	96	282	116	142

FG14	5	325	96	288	118	142
FG15	5	325	94	240	116	148
FG15	5	325	96	312	118	196
FG17	5	325	94	240	-9	152
FG17	5	345	94	288	-9	196
FG18	5	327	94	240	-9	142
FG18	5	333	94	288	-9	196
FG19	5	331	94	284	116	142
FG19	5	333	94	312	116	196
FG20	5	325	94	270	118	142
FG20	5	333	94	282	118	208
FG21	5	333	94	270	118	124
FG21	5	345	94	288	118	124
FG22	5	333	94	282	118	142
FG22	5	339	94	288	118	186
FG23	5	331	94	240	118	142
FG23	5	343	96	312	118	196
FG24	5	313	94	270	116	142
FG24	5	333	106	288	118	142
FG25	5	325	94	288	116	142
FG25	5	345	96	288	118	142
FG26	5	325	88	270	116	136
FG26	5	327	88	288	118	142
FG27	5	333	88	288	118	164
FG27	5	333	106	288	118	164
FG28	5	327	94	288	-9	142
FG28	5	333	106	288	-9	186
FG29	5	327	94	282	116	142
FG29	5	345	106	288	118	164
FG30	5	325	94	288	118	136
FG30	5	333	106	288	118	142
FG31	5	331	94	-9	-9	164
FG31	5	333	94	-9	-9	196
G01	6	327	94	258	-9	142
G01	6	329	94	312	-9	164
G02	6	-9	94	212	116	142
G02	6	-9	94	288	118	142
G03	6	325	94	244	-9	142
G03	6	325	96	288	-9	164
G04	6	323	94	212	116	142
G04	6	325	96	212	118	142
G05	6	329	94	240	116	142
G05	6	329	94	270	118	142
G06	6	327	94	312	-9	130
G06	6	327	94	312	-9	142
G07	6	327	94	244	116	142
G07	6	331	94	332	116	186
G08	6	327	94	240	-9	142
G08	6	331	106	240	-9	142
G09	6	327	94	288	116	124
G09	6	329	94	312	116	164

G10	6	325	94	288	118	142
G10	6	329	96	312	118	186
G11	6	319	94	288	118	142
G11	6	319	94	312	118	186
G12	6	331	94	244	118	142
G12	6	333	94	288	118	164
G13	6	325	96	288	118	142
G13	6	327	100	288	122	164
G14	6	321	96	312	118	142
G14	6	331	96	332	118	164
G15	6	325	94	288	118	142
G15	6	325	94	288	118	142
G16	6	327	96	286	116	142
G16	6	335	94	312	118	164
G17	6	-9	94	212	118	130
G17	6	-9	94	312	118	164
G18	6	329	94	288	116	136
G18	6	401	94	370	116	138
G19	6	329	94	244	118	164
G19	6	331	94	270	118	212
G20	6	-9	94	240	118	136
G20	6	-9	94	312	118	142
G21	6	333	94	240	118	142
G21	6	333	96	288	118	142
G22	6	-9	106	212	-9	164
G22	6	-9	106	258	-9	164
G23	6	-9	94	240	-9	142
G23	6	-9	94	288	-9	164
G24	6	327	94	262	-9	130
G24	6	333	96	370	-9	142
G25	6	331	96	288	-9	142
G25	6	333	96	312	-9	142
HA01	7	327	94	282	116	124
HA01	7	327	96	312	116	178
HA02	7	333	94	312	114	142
HA02	7	333	94	312	116	164
HA03	7	327	94	240	114	124
HA03	7	401	94	282	116	164
HA04	7	327	96	282	118	136
HA04	7	331	98	312	118	152
HA05	7	327	94	282	118	164
HA05	7	331	96	282	118	164
HA06	7	325	94	240	116	134
HA06	7	333	94	312	118	160
HA07	7	333	94	282	116	124
HA07	7	333	94	288	118	142
HA08	7	329	88	282	118	124
HA08	7	337	94	312	118	142
HA09	7	325	88	270	116	136
HA09	7	327	94	294	116	148
HA10	7	327	88	240	118	136

HA10	7	327	106	270	118	182
HA11	7	327	94	-9	116	142
HA11	7	327	94	-9	118	142
HA12	7	325	94	240	118	142
HA12	7	327	94	240	118	162
HA13	7	327	94	240	114	124
HA13	7	341	96	240	118	160
HA14	7	327	94	270	118	142
HA14	7	327	106	312	118	142
HA15	7	325	94	240	116	160
HA15	7	329	94	240	118	168
HA16	7	325	94	-9	118	130
HA16	7	327	94	-9	118	182
HA17	7	325	94	312	114	148
HA17	7	325	94	312	116	148
HA18	7	327	94	282	116	142
HA18	7	341	94	312	118	164
HA19	7	327	94	240	114	162
HA19	7	341	94	286	118	162
HA20	7	325	94	-9	118	142
HA20	7	333	94	-9	118	164
HA21	7	333	94	-9	114	124
HA21	7	337	94	-9	116	164
HA22	7	327	94	-9	114	142
HA22	7	331	94	-9	116	162
HA23	7	325	94	-9	118	162
HA23	7	331	94	-9	118	162
HA24	7	325	96	240	116	-9
HA24	7	333	96	288	118	-9
HA25	7	327	94	270	118	124
HA25	7	329	96	298	118	124
HA26	7	327	94	270	118	142
HA26	7	331	96	270	118	158
HA27	7	327	88	222	116	142
HA27	7	369	96	312	118	164
HA28	7	331	94	240	116	124
HA28	7	331	94	312	118	152
HA29	7	333	94	240	116	142
HA29	7	333	94	312	118	164
HA30	7	331	94	286	116	124
HA30	7	333	96	312	118	124
HE01	8	327	94	212	116	124
HE01	8	327	96	270	118	130
HE02	8	333	94	270	116	142
HE02	8	333	94	312	118	142
HE03	8	327	94	-9	116	124
HE03	8	333	96	-9	118	182
HE04	8	327	94	288	116	186
HE04	8	327	94	312	116	204
HE05	8	325	94	244	116	142
HE05	8	327	96	244	116	196

HE06	8	327	94	244	116	164
HE06	8	327	106	244	116	164
HE07	8	327	94	240	114	142
HE07	8	333	94	288	116	162
HE08	8	327	94	240	114	142
HE08	8	329	94	270	116	196
HE09	8	327	94	244	116	142
HE09	8	331	94	312	116	164
HE10	8	325	94	258	118	142
HE10	8	327	94	288	118	142
HE11	8	327	94	270	116	142
HE11	8	327	96	270	122	164
HE12	8	327	94	270	116	142
HE12	8	331	94	312	122	164
HE13	8	325	94	240	116	124
HE13	8	327	94	240	116	130
HE14	8	325	94	-9	116	144
HE14	8	329	96	-9	116	164
HE15	8	327	94	270	116	142
HE15	8	337	94	270	116	142
HE16	8	325	94	270	116	144
HE16	8	327	94	270	116	164
HE17	8	325	94	270	116	124
HE17	8	327	94	270	144	186
HE18	8	325	94	270	118	142
HE18	8	333	106	270	144	196
HE19	8	325	94	-9	116	124
HE19	8	331	96	-9	122	142
HE20	8	327	94	256	116	130
HE20	8	331	94	288	122	196
HE21	8	327	96	240	116	124
HE21	8	329	108	312	144	142
HE22	8	327	96	-9	116	124
HE22	8	333	96	-9	116	164
HE23	8	327	94	-9	116	142
HE23	8	333	94	-9	116	142
HE24	8	327	94	240	116	162
HE24	8	333	96	270	144	166
HE25	8	327	94	-9	118	124
HE25	8	329	96	-9	118	182
HE26	8	329	94	-9	116	142
HE26	8	329	94	-9	118	196
HE27	8	325	94	-9	116	186
HE27	8	327	96	-9	118	186
HE28	8	327	94	-9	-9	142
HE28	8	331	94	-9	-9	164
HE29	8	325	94	212	116	142
HE29	8	325	96	240	118	164
HE30	8	325	94	248	116	142
HE30	8	327	94	270	118	162
HG01	9	331	106	240	118	196

HG01	9	331	106	314	118	196
HG02	9	333	94	312	116	142
HG02	9	333	94	312	116	152
HG03	9	331	94	312	116	164
HG03	9	331	94	312	116	196
HG04	9	327	94	-9	118	140
HG04	9	327	96	-9	118	152
HG05	9	325	94	212	118	140
HG05	9	327	96	240	118	196
HG06	9	327	94	212	116	138
HG06	9	333	96	212	116	164
HG07	9	327	94	270	-9	164
HG07	9	333	94	312	-9	164
HG08	9	327	94	-9	116	130
HG08	9	337	94	-9	118	164
HG09	9	-9	88	-9	116	142
HG09	9	-9	106	-9	118	164
HG10	9	227	94	258	-9	136
HG10	9	331	106	258	-9	160
HG11	9	229	94	288	-9	164
HG11	9	331	106	312	-9	194
HG12	9	227	94	240	-9	160
HG12	9	333	106	258	-9	160
HG13	9	225	94	-9	-9	142
HG13	9	225	94	-9	-9	164
HG14	9	331	94	288	-9	152
HG14	9	333	96	288	-9	164
HG15	9	327	94	270	-9	164
HG15	9	327	106	284	-9	164
HG16	9	-9	96	270	-9	130
HG16	9	-9	96	312	-9	164
HG17	9	-9	94	270	-9	138
HG17	9	-9	94	312	-9	180
HG18	9	-9	94	-9	-9	-9
HG18	9	-9	96	-9	-9	-9
HG19	9	327	94	312	-9	142
HG19	9	327	94	312	-9	164
HG20	9	-9	94	212	-9	124
HG20	9	-9	94	312	-9	164
HG21	9	329	94	-9	118	142
HG21	9	331	96	-9	118	142
HG22	9	327	94	240	118	142
HG22	9	331	94	270	118	164
HG23	9	327	94	258	118	142
HG23	9	329	94	312	118	142
HG24	9	327	94	288	118	192
HG24	9	395	106	288	118	192
HS01	10	325	94	240	116	138
HS01	10	395	96	282	116	164
HS02	10	325	94	270	116	138
HS02	10	333	96	282	134	142

HS03	10	395	94	240	116	138
HS03	10	401	96	282	118	142
HS04	10	325	94	212	118	142
HS04	10	325	94	240	118	164
HS05	10	325	94	288	118	164
HS05	10	337	94	288	118	164
HS06	10	327	94	240	118	142
HS06	10	337	94	312	118	158
HS07	10	333	94	282	118	142
HS07	10	333	94	282	118	164
HS08	10	325	94	270	116	142
HS08	10	325	106	288	118	164
HS09	10	327	94	212	118	138
HS09	10	329	94	312	118	164
HS10	10	327	94	282	116	142
HS10	10	327	96	288	118	164
HS11	10	327	94	312	116	164
HS11	10	333	94	312	134	164
HS12	10	327	94	212	118	142
HS12	10	333	94	312	134	164
HS13	10	327	94	240	116	168
HS13	10	333	96	288	118	186
HS14	10	313	94	212	116	142
HS14	10	333	96	282	116	164
HS15	10	327	94	282	118	142
HS15	10	329	94	282	118	198
HS16	10	325	94	212	130	142
HS16	10	333	96	212	130	180
HS17	10	329	90	270	118	142
HS17	10	329	108	282	134	164
IN01	11	325	94	212	116	142
IN01	11	329	94	312	118	142
IN02	11	327	94	212	116	182
IN02	11	329	94	312	118	196
IN03	11	327	94	240	116	140
IN03	11	327	94	312	118	140
IN04	11	325	94	240	-9	136
IN04	11	329	94	322	-9	196
IN05	11	309	94	270	-9	130
IN05	11	309	94	312	-9	142
IN06	11	327	96	288	-9	194
IN06	11	329	96	312	-9	196
IN07	11	327	-9	270	-9	-9
IN07	11	327	-9	288	-9	-9
IN08	11	325	94	266	-9	142
IN08	11	327	96	266	-9	164
IN09	11	327	94	288	118	158
IN09	11	327	94	312	118	158
IN10	11	327	94	240	116	164
IN10	11	329	96	240	116	168
IN11	11	325	94	240	118	142

IN11	11	325	96	312	118	196
IN12	11	323	94	288	116	142
IN12	11	325	94	312	116	168
IN13	11	325	94	288	118	142
IN13	11	329	94	312	118	196
IN14	11	325	94	192	118	124
IN14	11	327	94	248	118	142
IN15	11	325	-9	192	118	-9
IN15	11	329	-9	192	118	-9
IN16	11	325	94	286	118	-9
IN16	11	327	96	312	118	-9
IN17	11	327	94	270	-9	142
IN17	11	327	94	274	-9	164
IN18	11	333	94	270	-9	-9
IN18	11	333	94	270	-9	-9
IN19	11	325	94	266	-9	142
IN19	11	327	94	270	-9	142
IN20	11	331	94	200	-9	334
IN20	11	333	94	200	-9	334
PU01	12	331	94	-9	-9	154
PU01	12	333	94	-9	-9	174
PU02	12	325	94	212	116	136
PU02	12	327	96	240	116	180
PU03	12	-9	94	286	116	142
PU03	12	-9	96	286	116	180
PU04	12	325	94	282	116	124
PU04	12	327	96	288	116	136
PU05	12	327	94	282	116	164
PU05	12	333	96	312	116	164
PU06	12	329	94	312	116	180
PU06	12	329	96	312	116	196
PU07	12	327	94	312	116	142
PU07	12	333	96	312	116	186
PU08	12	327	94	212	116	164
PU08	12	327	94	240	116	196
PU09	12	329	94	212	116	142
PU09	12	329	94	240	118	164
PU10	12	327	94	288	118	164
PU10	12	333	94	288	136	136
PU11	12	329	94	312	116	164
PU11	12	329	94	312	116	180
PU12	12	327	94	240	118	164
PU12	12	327	94	258	136	186
PU13	12	-9	94	-9	116	142
PU13	12	-9	96	-9	116	162
PU14	12	-9	94	-9	116	136
PU14	12	-9	94	-9	118	142
PU15	12	325	94	-9	116	142
PU15	12	327	94	-9	116	196
PU16	12	327	94	212	116	164
PU16	12	329	116	240	118	168

PU17	12	325	94	258	-9	164	
PU17	12	333	96	258	-9	186	
PU18	12	-9	94	-9	-9	164	
PU18	12	-9	96	-9	-9	210	
PU19	12	329	94	258	-9	124	
PU19	12	331	94	312	-9	164	
PU20	12	-9	94	288	-9	136	
PU20	12	-9	94	332	-9	164	
PU21	12	325	94	288	-9	124	
PU21	12	333	94	312	-9	164	
PU22	12	327	106	288	116	164	
PU22	12	331	106	332	116	180	
PU23	12	327	94	282	-9	164	
PU23	12	333	96	312	-9	180	
PU24	12	325	94	288	118	136	
PU24	12	333	96	312	118	142	
PU25	12	-9	94	212	118	124	
PU25	12	-9	94	212	118	142	
PU26	12	-9	94	282	116	142	
PU26	12	-9	96	312	116	164	
PU27	12	325	94	288	116	124	
PU27	12	325	96	288	118	164	
PU28	12	329	94	312	116	142	
PU28	12	329	96	332	118	164	
PU29	12	327	94	312	118	142	
PU29	12	329	94	332	118	164	
PU30	12	329	94	258	116	124	
PU30	12	331	94	312	118	142	
SC01	13	327	94	270	116	142	
SC01	13	329	94	282	118	164	
SC02	13	325	94	212	118	164	
SC02	13	325	94	288	118	210	
SC03	13	325	96	312	116	182	
SC03	13	327	106	312	130	182	
SC04	13	325	96	212	118	142	
SC04	13	325	106	288	118	142	
SC05	13	309	94	244	118	142	
SC05	13	329	106	288	118	194	
SC06	13	325	94	312	116	164	
SC06	13	325	94	330	116	182	
SC07	13	325	96	282	118	134	
SC07	13	327	118	312	126	142	
SC08	13	-9	96	212	-9	164	
SC08	13	-9	96	212	-9	164	
SC09	13	313	94	312	116	142	
SC09	13	325	96	312	118	164	
SC10	13	329	94	212	118	142	
SC10	13	329	96	282	118	164	
SC11	13	-9	94	-9	118	142	
SC11	13	-9	94	-9	134	182	
SC12	13	-9	94	282	118	142	

SC12	13	-9	106	286	118	142
SC13	13	-9	94	-9	116	134
SC13	13	-9	96	-9	118	186
SC14	13	-9	94	248	118	142
SC14	13	-9	96	262	118	182
SC15	13	325	96	-9	116	142
SC15	13	329	96	-9	118	182
SC16	13	-9	96	270	116	142
SC16	13	-9	96	270	116	182
SC17	13	327	-9	244	130	142
SC17	13	327	-9	288	130	182
SC18	13	-9	-9	248	-9	142
SC18	13	-9	-9	248	-9	164
SE01	14	327	94	-9	164	158
SE01	14	327	96	-9	164	174
SE02	14	327	96	288	164	164
SE02	14	327	96	288	186	186
SE03	14	327	94	244	164	164
SE03	14	327	94	244	186	186
SE04	14	327	96	244	164	164
SE04	14	327	96	288	186	186
SE05	14	327	94	244	164	164
SE05	14	327	94	312	164	164
SE06	14	327	94	244	164	164
SE06	14	327	96	312	164	164
SE07	14	327	94	288	186	186
SE07	14	327	96	288	186	186
SE08	14	327	94	278	186	164
SE08	14	327	94	312	186	164
SE09	14	327	94	244	-9	164
SE09	14	327	96	248	-9	186
SE10	14	329	94	240	-9	164
SE10	14	331	96	288	-9	186
SE11	14	327	104	248	-9	164
SE11	14	327	104	270	-9	210
SE12	14	327	96	-9	-9	164
SE12	14	327	96	-9	-9	164
SE13	14	327	-9	-9	-9	124
SE13	14	327	-9	-9	-9	164
SWS01	15	331	-9	312	-9	130
SWS01	15	333	-9	312	-9	142
SWS02	15	-9	-9	288	118	138
SWS02	15	-9	-9	312	118	138
SWS03	15	333	-9	270	116	134
SWS03	15	333	-9	270	118	134
SWS04	15	333	-9	288	116	130
SWS04	15	333	-9	312	118	136
SWS05	15	327	-9	270	116	130
SWS05	15	331	-9	312	118	138
SWS06	15	327	-9	258	116	130
SWS06	15	331	-9	258	118	142

SWS07	15	331	-9	240	116	130
SWS07	15	401	-9	288	118	142
SWS08	15	-9	-9	-9	116	142
SWS08	15	-9	-9	-9	118	142
SWS09	15	327	-9	212	116	130
SWS09	15	333	-9	212	118	142
SWS10	15	-9	96	240	116	130
SWS10	15	-9	96	288	118	142
SWS11	15	329	-9	240	116	130
SWS11	15	331	-9	288	118	134
SWS12	15	-9	-9	286	116	130
SWS12	15	-9	-9	312	118	142
SWS13	15	329	-9	240	116	142
SWS13	15	331	-9	312	118	162
SWS14	15	329	-9	-9	116	142
SWS14	15	329	-9	-9	118	142
SWS15	15	-9	-9	286	116	136
SWS15	15	-9	-9	312	118	142
SWS16	15	-9	-9	212	116	130
SWS16	15	-9	-9	212	118	134
SWS17	15	327	94	270	116	118
SWS17	15	333	96	286	118	118
SWS18	15	329	94	270	-9	130
SWS18	15	331	96	288	-9	138
SWS19	15	331	-9	248	116	118
SWS19	15	331	-9	248	118	142
SWS20	15	331	106	-9	116	130
SWS20	15	331	106	-9	118	142
SWS21	15	331	104	-9	116	142
SWS21	15	331	104	-9	118	142
SWS22	15	-9	-9	282	116	-9
SWS22	15	-9	-9	282	118	-9
SWS23	15	-9	-9	270	116	130
SWS23	15	-9	-9	288	118	142
SWS24	15	-9	-9	248	116	134
SWS24	15	-9	-9	270	118	142
SWS25	15	327	-9	288	116	142
SWS25	15	333	-9	312	116	196
SWS26	15	-9	-9	248	116	130
SWS26	15	-9	-9	270	118	138
SWS27	15	327	-9	-9	116	130
SWS27	15	333	-9	-9	118	138
SWS28	15	-9	-9	-9	116	142
SWS28	15	-9	-9	-9	118	194

Appendix IX. F_{ST} values and geographical distances between European *P. parva* populations used for Mantel-test calculations

Genetic

D	BK	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
BK	0	0.101	0.0322	0.0151	0.0447	0.054	0.0372	0.032	0.065	0.0466	0.0479	0.2253	0.1375
EB	0.101	0	0.1281	0.1194	0.1093	0.145	0.0964	0.1219	0.1366	0.1036	0.1456	0.3115	0.1703
FG	0.0322	0.1281	0	0.0189	0.0291	0.0574	0.0281	0.0278	0.0327	0.0487	0.0495	0.1952	0.1341
G	0.0151	0.1194	0.0189	0	0.0339	0.0691	0.0231	0.0179	0.0328	0.0373	0.0292	0.2041	0.1361
HA	0.0447	0.1093	0.0291	0.0339	0	0.0432	0.0231	0.0224	0.0465	0.0403	0.0568	0.1995	0.1416
HE	0.054	0.145	0.0574	0.0691	0.0432	0	0.061	0.0631	0.0627	0.0391	0.0794	0.1841	0.1591
HG	0.0372	0.0964	0.0281	0.0231	0.0231	0.061	0	0.0348	0.0298	0.0225	0.0483	0.1478	0.1243
HS	0.032	0.1219	0.0278	0.0179	0.0224	0.0631	0.0348	0	0.0537	0.0359	0.0338	0.2048	0.1551
IN	0.065	0.1366	0.0327	0.0328	0.0465	0.0627	0.0298	0.0537	0	0.038	0.0522	0.1638	0.1549
PU	0.0466	0.1036	0.0487	0.0373	0.0403	0.0391	0.0225	0.0359	0.038	0	0.0521	0.1577	0.146
SC	0.0479	0.1456	0.0495	0.0292	0.0568	0.0794	0.0483	0.0338	0.0522	0.0521	0	0.1814	0.0996
SE	0.2253	0.3115	0.1952	0.2041	0.1995	0.1841	0.1478	0.2048	0.1638	0.1577	0.1814	0	0.2731
SWS	0.1375	0.1703	0.1341	0.1361	0.1416	0.1591	0.1243	0.1551	0.1549	0.146	0.0996	0.2731	0

Geographic

D	BK	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
BK	0	528.51	604.92	181.98	1263.45	1049	1036.81	1216.08	1012	973.43	966.07	1137	599.99
EB	528.51	0	784.79	658.09	1723.98	1543.71	1521.3	1672.94	1551.55	1367.44	1437.29	1494.71	344
FG	604.92	784.79	0	762.02	1720	1448.67	1469.38	1675	1165.01	1504.02	1415.28	726.94	539.16
G	181.98	658.09	762.02	0	1093.12	888.86	875	1045	927.15	790	795.8	1219.39	774.28
HA	1263.45	1723.98	1720	1093.12	0	293.58	253.83	50.44	834	396	301	1782	1868.88
HE	1049	1543.71	1448.67	888.86	293.58	0	73.31	277.14	554	424	148	1474	1651
HG	1036.81	1521.3	1469.38	875	253.83	73.31	0	221	620	353	92	1538.78	1640
HS	1216.08	1672.94	1675	1045	50.44	277.14	221	0	828	346	262	1753	1824.23
IN	1012	1551.55	1165.01	927.15	834	554	620	828	0	919.5	649	983	1539.24
PU	973.43	1367.44	1504.02	790	396	424	353	346	919.5	0	281	1715	1555.7
SC	966.07	1437.29	1415.28	795.8	301	148	92	262	649	281	0	1517	1565
SE	1137	1494.71	726.94	1219.39	1782	1474	1538.78	1753	983	1715	1517	0	1272
SWS	599.99	344	539.16	774.28	1868.88	1651	1640	1824.23	1539.24	1555.7	1565	1272	0

Appendix X. Matrix of F_{ST} values of pairwise genetic comparisons between all *O. mykiss* populations based on mtDNA.

	A	AK	APH	B	BB	BC	C	CC	CF	CO	DR	E	EL	GL	H
A															
AK	0.54545														
APH	1	0.44444													
B	0.5	0.2766	0.5												
BB	0.65625	0.32595	0.725	-0.13194											
BC	0.53333	0	0.22222	0.12281	0.22475										
C	0	0.54545	1	0.5	0.65625	0.53333									
CC	0.33333	0	0.33333	0.10345	0.13542	-0.0303	0.33333								
CF	0	0.54545	1	0.5	0.65625	0.53333	0	0.33333							
CO	0	0.43902	0.85714	0.19048	0.25926	0.4	0	0.21739	0						
DR	0	0.2	0	-0.33333	-0.30556	0	0	0	0	-0.25					
E	0.25	0.39796	0.8125	0.0625	0.09167	0.34615	0.25	0.17857	0.25	-0.22222	-0.3				
EL	0.90909	0.56604	0.90909	0.6383	0.73333	0.56	0.90909	0.30769	0.90909	0.8	0.54545	0.76596			
GL	0	0.17391	0	-0.23529	-0.2037	0	0	-0.07692	0	-0.16667	-0.66667	-0.2	0.46154		
H	0.36364	0.35294	0.63158	0.01587	0.13119	0.26866	0.36364	0.09091	0.36364	0.13953	-0.06667	0.08491	0.53165	-0.05128	
HC	0.16327	0.33826	0.54945	-0.00539	0.0776	0.23166	0.16327	0.10787	0.16327	-0.04762	-0.38571	-0.05861	0.62585	-0.27551	0.00262
ID	0.19512	0.08824	0	0.03743	0.06118	0	0.19512	0.04444	0.19512	0.13836	-0.02703	0.11538	0.20809	-0.05195	0.1123
KA	0.5	0.26042	0.3	-0.12069	-0.05233	0.09483	0.5	0.03571	0.5	0.24	-0.41667	0.13793	0.55814	-0.29412	0.07639
LA	0	0.4	0.8	0.06667	0.09211	0.35	0	0.17647	0	-0.4	-0.33333	-0.35714	0.75862	-0.22222	0.0625
ML	1	0.44444	0	0.5	0.725	0.22222	1	0.33333	1	0.85714	0	0.8125	0.90909	0	0.63158
NV	0.88889	0.41026	0	0.41379	0.59184	0.19512	0.88889	0.28	0.88889	0.75	0	0.7027	0.83333	-0.16667	0.56471
OR	0.8	0.29091	0	0.35294	0.5	0.05	0.8	0.18919	0.8	0.66667	0	0.61905	0.76923	-0.09091	0.51064
PC	0.58333	0.25	0.16667	-0.02381	0.06818	0.06944	0.58333	0.05556	0.58333	0.38095	-0.22222	0.29167	0.42222	-0.16667	0.17143
PP	0.18333	0.28289	0.35526	0.03022	0.06897	0.18257	0.18333	0.05469	0.18333	0.04918	-0.25	0.02922	0.41414	-0.2561	0.04808
SC	0.27273	-0.02632	0.11111	-0.04255	-0.01582	-0.10909	0.27273	-0.27273	0.27273	0.14634	-0.1	0.09184	0.22222	-0.19048	0.03125
SH	1	0.44444	0	0.5	0.725	0.22222	1	0.33333	1	0.85714	0	0.8125	0.90909	0	0.63158
UW	0.33333	0.04	0.06667	0.08333	0.13636	-0.09804	0.33333	-0.05051	0.33333	0.23457	-0.11111	0.20202	0.40952	-0.17949	0.2
VH	1	0.375	1	0	0.08333	0.3	1	0.2	1	0.66667	0	0.5	0.85714	0	0.3
W	1	0.64286	1	0.72727	0.83824	0.65	1	0.5	1	0.8	0.5	0.78571	0.93333	0.4	0.65
WA	0.5	-0.04348	0.25	-0.09091	-0.03676	-0.2	0.5	-0.125	0.5	0.28571	-0.2	0.1875	0.54545	-0.15385	0.15789
WF	0.36364	0.40559	0.83254	0.08931	0.13188	0.35543	0.36364	0.1886	0.36364	-0.12397	-0.26515	-0.16928	0.77922	-0.17677	0.1131
YR	0.5	0.53947	0.8913	0.42754	0.59615	0.49667	0.5	0.35938	0.5	0.34483	0.17857	0.39535	0.84034	0.13514	0.25

	HC	ID	KA	LA	ML	NV	OR	PC	PP	SC	SH	UW	VH	W	WA	WF
YR	0.07481															
KA	-0.00501	0.00815														
LA	-0.1118	0.11864	0.13889													
ML	0.54945	0	0.3	0.8												
NV	0.47619	-0.01481	0.24	0.69565	0											
OR	0.42017	-0.01914	0.2	0.61538	0	-0.125										
PC	0.13629	-0.01826	-0.17949	0.3	0.16667	0.13333	0.11111									
PP	-0.05395	0.03954	-0.03056	0.00543	0.35526	0.3012	0.27698	0.0125								
SC	0.04034	-0.03731	-0.09783	0.1	0.11111	0.05405	0	-0.10526	-0.00694							
SH	0.54945	0	0.3	0.8	0	0	0	0.16667	0.35526	0.11111						
UW	0.10823	-0.02682	0.02151	0.2	0.06667	0.01587	-0.0625	-0.01333	0.03623	-0.11111	0.06667					
VH	0.34921	0.08333	0.125	0.5	1	0.8	0.66667	0.16667	0.23438	0	1	0.22222				
W	0.60952	0.25	0.70833	0.75	1	0.92308	0.85714	0.72222	0.48958	0.42857	1	0.48148	1			
WA	0.1155	-0.00885	-0.1087	0.2	0.25	0.21053	0.06897	-0.1	0.07609	-0.22222	0.25	-0.11111	0	0.66667		
WF	-0.01851	0.11797	0.16074	-0.26738	0.83254	0.71901	0.63273	0.30622	0.05061	0.0979	0.83254	0.21059	0.54545	0.81283	0.20096	
YR	0.21813	0.21447	0.5	0.3125	0.8913	0.81188	0.74545	0.56522	0.2625	0.30263	0.8913	0.35616	0.82143	0.82143	0.5	0.44156

Appendix XI. Matrix of F_{ST} values of pairwise genetic comparisons between all *O. mykiss* populations based on mtDNA.

	A	AK	APH	B	BB	BC	C	CC	CF	CO	DR	E	EL	GL	H
A															
AK	0.00886														
APH	0.00483	0.00725													
B	0.0029	0.00757	0.0029												
BB	0.00215	0.00707	0.00268	0.00193											
BC	0.00725	0.00741	0.00435	0.00551	0.00531										
C	0	0.00886	0.00483	0.0029	0.00215	0.00725									
CC	0.00725	0.00886	0.00725	0.007	0.00644	0.00797	0.00725								
CF	0	0.00886	0.00483	0.0029	0.00215	0.00725	0	0.00725							
CO	0.0006	0.00825	0.00423	0.00254	0.00181	0.00664	0.0006	0.00694	0.0006						
DR	0.00242	0.00805	0.00242	0.0029	0.00242	0.0058	0.00242	0.00725	0.00242	0.00242					
E	0.00097	0.00789	0.00386	0.00232	0.00161	0.00628	0.00097	0.00676	0.00097	0.00109	0.00242				
EL	0.00664	0.01067	0.00664	0.00568	0.00503	0.00906	0.00664	0.00785	0.00664	0.00604	0.00664	0.00568			
GL	0.00362	0.00926	0.00362	0.00411	0.00362	0.007	0.00362	0.00785	0.00362	0.00362	0.00362	0.00362	0.00785		
H	0.00295	0.00913	0.0051	0.00338	0.00301	0.00719	0.00295	0.00738	0.00295	0.00289	0.00403	0.00284	0.0053	0.00523	
HC	0.00211	0.00876	0.00393	0.0032	0.00272	0.0067	0.00211	0.0074	0.00211	0.00226	0.00302	0.00236	0.00634	0.00423	0.00366
ID	0.01981	0.0219	0.01594	0.01807	0.01777	0.01932	0.01981	0.02174	0.01981	0.0192	0.01787	0.01884	0.02089	0.0186	0.02008
KA	0.00338	0.00773	0.00242	0.0028	0.00231	0.0056	0.00338	0.00676	0.00338	0.00302	0.0029	0.0028	0.00519	0.00411	0.00386
LA	0.00081	0.00805	0.00403	0.00242	0.0017	0.00644	0.00081	0.00684	0.00081	0.00101	0.00242	0.00113	0.00584	0.00362	0.00286
ML	0.00483	0.00725	0	0.0029	0.00268	0.00435	0.00483	0.00725	0.00483	0.00423	0.00242	0.00386	0.00664	0.00362	0.0051
NV	0.00543	0.00785	0.0006	0.0035	0.00329	0.00495	0.00543	0.00755	0.00543	0.00483	0.00302	0.00447	0.00725	0.00362	0.0057
OR	0.00604	0.00738	0.00121	0.00411	0.00389	0.00483	0.00604	0.00745	0.00604	0.00543	0.00362	0.00507	0.00785	0.00443	0.00631
PC	0.00483	0.00805	0.00242	0.00338	0.00295	0.0058	0.00483	0.00725	0.00483	0.00423	0.00362	0.00386	0.00453	0.00483	0.0047
PP	0.00403	0.0102	0.0051	0.00488	0.00432	0.00816	0.00403	0.00859	0.00403	0.00409	0.00456	0.00413	0.00664	0.0055	0.00543
SC	0.00886	0.0102	0.00725	0.00757	0.00707	0.00886	0.00886	0.00886	0.00886	0.00825	0.00805	0.00789	0.00906	0.00845	0.00859
SH	0.00483	0.00725	0	0.0029	0.00268	0.00435	0.00483	0.00725	0.00483	0.00423	0.00242	0.00386	0.00664	0.00362	0.0051
UW	0.00845	0.01006	0.00604	0.00773	0.00738	0.00821	0.00845	0.00996	0.00845	0.00815	0.00725	0.00797	0.01057	0.00785	0.00939
VH	0.00242	0.00644	0.00242	0.00145	0.00081	0.00483	0.00242	0.00604	0.00242	0.00181	0.00242	0.00145	0.00423	0.00362	0.00268
W	0.00242	0.01127	0.00725	0.00531	0.00456	0.00966	0.00242	0.00966	0.00242	0.00302	0.00483	0.00338	0.00906	0.00604	0.00537

WA	0.00483	0.00617	0.00322	0.00354	0.00304	0.00483	0.00483	0.00644	0.00483	0.00423	0.00403	0.00386	0.00664	0.00523	0.0051
WF	0.00101	0.00785	0.00382	0.00229	0.00159	0.00624	0.00101	0.00674	0.00101	0.00111	0.00242	0.00117	0.00564	0.00362	0.00284
YR	0.00134	0.0102	0.00617	0.0037	0.00349	0.00805	0.00134	0.00859	0.00134	0.00195	0.00376	0.00231	0.00798	0.00497	0.0034

	ID	KA	LA	ML	NV	OR	PC	PP	SC	SH	UW	VH	W	WA	WF
ID															
KA	0.01778														
LA	0.019	0.0029													
ML	0.01594	0.00242	0.00403												
NV	0.0163	0.00302	0.00463	0.0006											
OR	0.01683	0.00362	0.00523	0.00121	0.00161										
PC	0.01763	0.00314	0.00403	0.00242	0.00302	0.00362									
PP	0.02002	0.00483	0.00412	0.0051	0.00557	0.00622	0.00537								
SC	0.02158	0.00741	0.00805	0.00725	0.00745	0.00765	0.00765	0.00966							
SH	0.01594	0.00242	0.00403	0	0.0006	0.00121	0.00242	0.0051	0.00725						
UW	0.02101	0.00749	0.00805	0.00604	0.00634	0.00644	0.00755	0.00926	0.01087	0.00604					
VH	0.01739	0.00193	0.00161	0.00242	0.00302	0.00362	0.00242	0.00429	0.00644	0.00242	0.00725				
W	0.02126	0.0058	0.00322	0.00725	0.00785	0.00845	0.00725	0.00644	0.01127	0.00725	0.01087	0.00483			
WA	0.0182	0.0037	0.00403	0.00322	0.00382	0.00389	0.00403	0.00617	0.00725	0.00322	0.00725	0.00242	0.00725		
WF	0.0188	0.00278	0.00114	0.00382	0.00443	0.00503	0.00382	0.00414	0.00785	0.00382	0.00795	0.00141	0.00342	0.00382	
YR	0.02115	0.00472	0.00215	0.00617	0.00678	0.00738	0.00617	0.00537	0.0102	0.00617	0.0098	0.00376	0.00376	0.00617	0.00235

Appendix XII. Haplotype frequency table of *O. mykiss* populations.

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	H26	H27	H28	H29	H30	H31
A01											3																				
AK01																							1							2	
AM01																															1
APH02																						5									
B03												1			3							1									
BB01											2				6							1									
BC01		1															1				1	1							1		
C02											4																				
CC01			1	1												1											1				
CF01											15																				
CO01											3				1																
DR03											1											1									
E01											3				2																
EL02														3												1					
GL01							1				1																				
H04											1	3			2											3					
HC01											4	1										2				1					
ID01	1						1															3									
KA01											1				2							1			1						
LA03											2				1																
ML03																						3									
MT01																									1						
NV01							1															3									
OR01					1		1															4									
PC01														1	1							2									
PP03							1		1		4									1		1		1							

RU01																			1												
RW01				1																											
SA01																						1									
SC01			1			1												1				3									
SH01																						3									
TC01																						1									
UW01								1		1						1												1			
VH03															3																
W01													4																		
WA01		1													1							1									
WF03											7				5																
YR31											4	5																			

Appendix XIII. Allele frequency tables of the 5 microsatellite loci of the European *P. parva* populations. Populations that were corrected for null alleles are indicated in blue.

Locus 1	BK	Bsa	CRH	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
225	0	0.0652	0	0	0	0	0	0	0.0526	0	0	0	0	0	0
227	0	0	0	0	0	0	0	0	0.0526	0	0	0	0	0	0
229	0	0	0	0	0	0	0	0	0.0263	0	0	0	0	0	0
309	0	0	0	0.0556	0	0	0	0	0	0	0.05	0	0.0455	0	0
313	0	0	0.0833	0	0.0167	0	0	0	0	0.0294	0	0	0.0455	0	0
317	0	0	0.0625	0	0	0	0	0	0	0	0	0	0	0	0
319	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0
321	0	0	0	0	0	0.025	0	0	0	0	0	0	0	0	0
323	0	0	0.1875	0	0	0.025	0	0	0	0	0.025	0	0	0	0
325	0.2708	0.1739	0.2083	0.0278	0.2167	0.175	0.1667	0.2	0.0263	0.2353	0.275	0.1739	0.4545	0	0
327	0.1875	0.2174	0.3125	0.1667	0.2	0.225	0.35	0.4667	0.3421	0.2353	0.375	0.3043	0.2273	0.9231	0.1765
329	0.2083	0.1957	0.0625	0.1389	0	0.175	0.05	0.1	0.0526	0.1176	0.175	0.2609	0.2273	0.0385	0.1471
331	0.125	0.1304	0.0625	0.0278	0.1167	0.15	0.1333	0.0833	0.2368	0	0.025	0.087	0	0.0385	0.3824
333	0.0208	0.0652	0	0.3611	0.2333	0.125	0.1833	0.1333	0.1579	0.2353	0.075	0.1739	0	0	0.2647
335	0	0	0	0	0	0.025	0	0	0	0	0	0	0	0	0
337	0	0	0	0	0	0	0.0333	0.0167	0.0263	0.0588	0	0	0	0	0
339	0	0	0	0	0.0667	0	0	0	0	0	0	0	0	0	0
341	0	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0
343	0	0	0	0	0.0167	0	0	0	0	0	0	0	0	0	0
345	0.0208	0	0	0	0.1333	0	0	0	0	0	0	0	0	0	0
369	0	0	0	0	0	0	0.0167	0	0	0	0	0	0	0	0
373	0	0	0.0208	0	0	0	0	0	0	0	0	0	0	0	0
395	0.1458	0.1522	0	0.2222	0	0	0	0	0.0263	0.0588	0	0	0	0	0
401	0.0208	0	0	0	0	0.025	0.0167	0	0	0.0294	0	0	0	0	0.0294

Locus 2	BK	Bsa	CRH	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
88	0	0	0	0	0.05	0	0.0667	0	0.0208	0	0	0	0	0	0
90	0	0	0	0	0	0	0	0	0	0.0294	0	0	0	0	0
94	0.7	0.7	0.5	0.9	0.6333	0.7	0.7167	0.7333	0.6458	0.7059	0.8333	0.7167	0.4375	0.4583	0.2
96	0.1833	0.2167	0.25	0.0667	0.2167	0.22	0.1667	0.2167	0.1667	0.2059	0.1667	0.2333	0.4063	0.4583	0.4
98	0	0	0.1429	0	0	0	0.0167	0	0	0	0	0	0	0	0
100	0.05	0	0.0536	0	0	0.02	0	0	0	0	0	0	0	0	0
104	0	0.0167	0.0536	0	0	0	0	0	0	0	0	0	0	0.0833	0.2
106	0.0667	0.0667	0	0.0333	0.1	0.06	0.0333	0.0333	0.1667	0.0294	0	0.0333	0.125	0	0.2
108	0	0	0	0	0	0	0	0.0167	0	0.0294	0	0	0	0	0
116	0	0	0	0	0	0	0	0	0	0	0	0.0167	0	0	0
118	0	0	0	0	0	0	0	0	0	0	0	0	0.0313	0	0

Locus 3	BK	Bsa	CRH	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
192	0	0	0	0	0	0	0	0	0	0	0.075	0	0	0	0
200	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0	0
204	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0
212	0.2083	0.05	0.0952	0	0	0.1	0	0.0476	0.1111	0.1765	0.05	0.12	0.1667	0	0.0909
222	0	0	0	0	0	0	0.0208	0	0	0	0	0	0	0	0
232	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0
236	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0
238	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0
240	0.1667	0.0667	0.0714	0	0.1897	0.12	0.2708	0.1667	0.1111	0.1471	0.125	0.1	0	0.05	0.0909
244	0	0.0167	0	0	0	0.08	0	0.119	0	0	0	0	0.0667	0.3	0
248	0	0	0	0	0	0	0	0.0238	0	0	0.025	0	0.1	0.1	0.0909
256	0	0	0	0	0	0	0	0.0238	0	0	0	0	0	0	0
258	0.1042	0.1	0	0.6957	0	0.04	0	0.0238	0.1111	0	0	0.1	0	0	0.0455
262	0	0	0.0238	0	0	0.02	0	0	0	0	0	0	0.0333	0	0
264	0	0	0	0	0.0172	0	0	0	0	0	0	0	0	0	0
266	0	0	0.0238	0	0	0	0	0	0	0	0.075	0	0	0	0
270	0.1667	0.2833	0	0.0217	0.2241	0.04	0.125	0.381	0.1389	0.0882	0.15	0	0.1	0.05	0.1818
272	0	0	0.1667	0	0	0	0	0	0	0	0	0	0	0	0
274	0	0	0	0	0	0	0	0	0	0	0.025	0	0	0	0
276	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0
278	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0
282	0.0417	0	0	0	0.0862	0	0.1667	0	0	0.2941	0	0.08	0.1333	0	0.0455
284	0	0	0	0	0.0172	0	0	0	0.0278	0	0	0	0	0	0
286	0	0	0	0	0	0.02	0.0417	0	0	0	0.025	0.04	0.0333	0	0.0682
288	0.1458	0.2833	0.0952	0	0.3448	0.28	0.0417	0.0952	0.1389	0.1471	0.125	0.18	0.1333	0.3	0.1818
292	0	0	0.0238	0	0	0	0	0	0	0	0	0	0	0	0

294	0	0	0	0	0	0	0.0208	0	0	0	0	0	0	0	0
298	0	0	0	0	0	0	0.0208	0	0	0	0	0	0	0	0
312	0.1667	0.2	0.1429	0.2826	0.1034	0.22	0.2917	0.119	0.3333	0.1471	0.25	0.3	0.2	0.15	0.2045
314	0	0	0.0714	0	0	0	0	0	0.0278	0	0	0	0	0	0
322	0	0	0	0	0	0	0	0	0	0	0.025	0	0	0	0
326	0	0	0.1667	0	0	0	0	0	0	0	0	0	0	0	0
330	0	0	0	0	0	0	0	0	0	0	0	0	0.0333	0	0
332	0	0	0	0	0	0.04	0	0	0	0	0	0.08	0	0	0
370	0	0	0	0	0.0172	0.04	0	0	0	0	0	0	0	0	0

Locus 4	BK	Bsa	CRH	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
114	0	0	0	0	0	0	0.1167	0.0345	0	0	0	0	0	0	0
116	0.5	0.4167	0.2188	0.4565	0.325	0.2941	0.3167	0.6207	0.3333	0.2941	0.3182	0.6522	0.2813	0	0.5
118	0.5	0.5833	0.5	0.5435	0.675	0.6765	0.5667	0.2069	0.6667	0.5294	0.6818	0.3043	0.5625	0	0.5
122	0	0	0	0	0	0.0294	0	0.069	0	0	0	0	0	0	0
126	0	0	0	0	0	0	0	0	0	0	0	0	0.0313	0	0
130	0	0	0	0	0	0	0	0	0	0.0588	0	0	0.0938	0	0
134	0	0	0.0938	0	0	0	0	0	0	0.1176	0	0	0.0313	0	0
136	0	0	0	0	0	0	0	0	0	0	0	0.0435	0	0	0
138	0	0	0.125	0	0	0	0	0	0	0	0	0	0	0	0
140	0	0	0.0625	0	0	0	0	0	0	0	0	0	0	0	0
144	0	0	0	0	0	0	0	0.069	0	0	0	0	0	0	0
164	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5625	0
186	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4375	0

Locus 5	BK	Bsa	CRH	EB	FG	G	HA	HE	HG	HS	IN	PU	SC	SE	SWS
118	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0556
124	0.0167	0.0167	0.0333	0.0217	0.0667	0.02	0.1897	0.1333	0.0217	0	0.0313	0.01	0	0.0385	0
128	0	0	0.0167	0	0	0	0	0	0	0	0	0	0	0	0
130	0	0.1167	0.1333	0.1304	0	0.06	0.0172	0.05	0.0435	0	0.0313	0	0	0	0.2778
132	0	0	0.05	0	0	0	0	0	0	0	0	0	0	0	0
134	0.05	0	0.1167	0	0	0	0.0172	0	0	0	0	0	0.0556	0	0.0926
136	0.05	0	0.0333	0	0.0667	0.04	0.0517	0	0.0217	0	0.0313	0.1	0	0	0.037
138	0	0	0.0333	0	0	0.02	0	0	0.0435	0.1176	0	0	0	0	0.1111
140	0	0.0167	0.0333	0.1522	0	0	0	0	0.0435	0	0.0625	0	0	0	0
142	0.3833	0.35	0	0.1957	0.3167	0.54	0.2414	0.35	0.1957	0.3529	0.3438	0.2	0.4167	0	0.3704
144	0	0	0.0167	0	0	0	0	0.0333	0	0	0	0	0	0	0
148	0	0	0.05	0	0.1	0	0.0517	0	0	0	0	0	0	0	0
150	0	0	0.0833	0	0	0	0	0	0	0	0	0	0	0	0
152	0	0	0.0167	0.1957	0.0333	0	0.0345	0	0.0652	0	0	0	0	0	0
154	0	0	0.0167	0	0	0	0	0	0	0	0	0.0167	0	0	0
156	0	0	0.0333	0	0	0	0	0	0	0	0	0	0	0	0
158	0	0	0.05	0	0	0	0.0172	0	0	0.0294	0.0625	0	0	0.0385	0
160	0	0	0.0167	0	0	0	0.0517	0	0.0652	0	0	0	0	0	0
162	0.0167	0.0167	0	0	0	0	0.1034	0.05	0	0	0	0.0167	0	0	0.0185
164	0.35	0.25	0.05	0.1087	0.15	0.24	0.1552	0.1667	0.3261	0.3824	0.0938	0.3167	0.2222	0.5769	0
166	0	0	0	0	0	0	0	0.0167	0	0	0	0	0	0	0
168	0	0	0.05	0	0	0	0.0172	0	0	0.0294	0.0625	0.0167	0	0	0
172	0	0	0.05	0	0	0	0	0	0	0	0	0	0	0	0
174	0	0	0.0167	0	0	0	0	0	0	0	0	0.0167	0	0.0385	0
178	0	0	0	0.0217	0	0	0.0172	0	0	0	0	0	0	0	0
180	0	0.0333	0	0.087	0	0	0	0	0.0217	0.0294	0	0.1	0	0	0

182	0.0333	0	0.0167	0.0217	0	0	0.0345	0.0333	0	0	0.0313	0	0.2222	0	0
184	0	0	0.0167	0	0	0	0	0	0	0	0	0	0	0	0
186	0.0833	0.2	0	0	0.05	0.06	0	0.0667	0	0.0294	0	0.05	0.0278	0.2692	0
188	0	0	0.0167	0	0	0	0	0	0	0	0	0	0	0	0
190	0	0	0.0167	0	0	0	0	0	0	0	0	0	0	0	0
192	0	0	0.0167	0	0	0	0	0	0.0435	0	0	0	0	0	0
194	0	0	0	0	0	0	0	0	0.0217	0	0.0313	0	0.0278	0	0.0185
196	0	0	0.0167	0.0652	0.1833	0	0	0.0833	0.087	0	0.1563	0.05	0	0	0.0185
198	0	0	0	0	0	0	0	0	0	0.0294	0	0	0	0	0
204	0	0	0	0	0	0	0	0.0167	0	0	0	0	0	0	0
206	0	0	0	0	0.0167	0	0	0	0	0	0	0	0	0	0
208	0	0	0	0	0.0167	0	0	0	0	0	0	0	0	0	0
210	0	0	0	0	0	0	0	0	0	0	0	0.0167	0.0278	0.0385	0
212	0.0167	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0
334	0	0	0	0	0	0	0	0	0	0	0.0625	0	0	0	0

Appendix XIV. Adjusted allele frequencies of amplified alleles based on the four methods of null allele estimation, in the analyses Brookfield 1 correction were used.

CRH locus 3					
Class	Obs. allele freq.	Oosterhout	Chakraborty	**Brookfield 1**	Brookfield 2
204	0.0238	0.0241	0.0212	**0.0216**	0.0124
212	0.0952	0.1003	0.0847	**0.0863**	0.0495
232	0.0238	0.0241	0.0212	**0.0216**	0.0124
236	0.0238	0.0241	0.0212	**0.0216**	0.0124
238	0.0238	0.0241	0.0212	**0.0216**	0.0124
240	0.0714	0.0742	0.0636	**0.0647**	0.0372
262	0.0238	0.0241	0.0212	**0.0216**	0.0124
266	0.0238	0.0241	0.0212	**0.0216**	0.0124
272	0.1667	0.1271	0.1483	**0.1511**	0.0867
276	0.0238	0.0241	0.0212	**0.0216**	0.0124
288	0.0952	0.0742	0.0847	**0.0863**	0.0495
292	0.0238	0.0241	0.0212	**0.0216**	0.0124
312	0.1429	0.1003	0.1271	**0.1295**	0.0743
314	0.0714	0.0742	0.0636	**0.0647**	0.0372
326	0.1667	0.1548	0.1483	**0.1511**	0.0867

FG locus 5					
124	0.069	0.0531	0.063	**0.0641**	0.0591
136	0.069	0.0715	0.063	**0.0641**	0.0591
142	0.3276	0.3305	0.2994	**0.3042**	0.2808
148	0.1034	0.0903	0.0946	**0.0961**	0.0887
152	0.0345	0.0351	0.0315	**0.032**	0.0296
164	0.1379	0.0903	0.1261	**0.1281**	0.1182
186	0.0517	0.0531	0.0473	**0.048**	0.0443
196	0.1724	0.1695	0.1576	**0.1601**	0.1478
206	0.0172	0.0174	0.0158	**0.016**	0.0148
208	0.0172	0.0174	0.0158	**0.016**	0.0148

HA locus 5					
124	0.1607	0.1548	0.1496	**0.1511**	0.1397
130	0.0179	0.018	0.0166	**0.0168**	0.0155
134	0.0179	0.018	0.0166	**0.0168**	0.0155
136	0.0536	0.0551	0.0499	**0.0504**	0.0466
142	0.25	0.2441	0.2327	**0.235**	0.2173
148	0.0536	0.0364	0.0499	**0.0504**	0.0466
152	0.0357	0.0364	0.0332	**0.0336**	0.031
158	0.0179	0.018	0.0166	**0.0168**	0.0155
160	0.0536	0.0551	0.0499	**0.0504**	0.0466
162	0.1071	0.0742	0.0997	**0.1007**	0.0931
164	0.1607	0.1548	0.1496	**0.1511**	0.1397
168	0.0179	0.018	0.0166	**0.0168**	0.0155

178	0.0179	0.018	0.0166	**0.0168**	0.0155
182	0.0357	0.0364	0.0332	**0.0336**	0.031

HG locus 3					
212	0.1111	0.0871	0.0954	**0.0988**	0.0988
240	0.1111	0.1181	0.0954	**0.0988**	0.0988
258	0.1111	0.0871	0.0954	**0.0988**	0.0988
270	0.1389	0.1502	0.1193	**0.1235**	0.1235
284	0.0278	0.0282	0.0239	**0.0247**	0.0247
288	0.1389	0.0871	0.1193	**0.1235**	0.1235
312	0.3333	0.2929	0.2863	**0.2964**	0.2964
314	0.0278	0.0282	0.0239	**0.0247**	0.0247

SC locus 3					
212	0.1667	0.1437	0.1447	**0.1486**	0.1086
244	0.0667	0.0691	0.0579	**0.0595**	0.0435
248	0.1	0.0691	0.0868	**0.0892**	0.0652
262	0.0333	0.0339	0.0289	**0.0297**	0.0217
270	0.1	0.0691	0.0868	**0.0892**	0.0652
282	0.1333	0.1437	0.1158	**0.1189**	0.0869
286	0.0333	0.0339	0.0289	**0.0297**	0.0217
288	0.1333	0.1437	0.1158	**0.1189**	0.0869
312	0.2	0.1437	0.1737	**0.1784**	0.1304
330	0.0333	0.0339	0.0289	**0.0297**	0.0217

SWS locus 3					
212	0.0909	0.0465	0.0805	**0.0823**	0.0581
240	0.0909	0.0955	0.0805	**0.0823**	0.0581
248	0.0909	0.0707	0.0805	**0.0823**	0.0581
258	0.0455	0.023	0.0402	**0.0411**	0.0291
270	0.1818	0.1743	0.161	**0.1645**	0.1163
282	0.0455	0.023	0.0402	**0.0411**	0.0291
286	0.0682	0.0707	0.0604	**0.0617**	0.0436
288	0.1818	0.2023	0.161	**0.1645**	0.1163
312	0.2045	0.2023	0.1811	**0.1851**	0.1308