Population Level Variation of Atlantic Salmon in the Chalk Streams of Southern England and Neighbouring Regions

Submitted by Charles Isioma Ikediashi, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences, on June 2015.

The thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University

Signature...

Abstract

In this thesis, population level variation is elucidated for Atlantic salmon living in the chalk streams of southern England – a unique and unusual habitat – as well as in immediately surrounding regions. Salmon in these chalk streams have yet to be robustly investigated, despite individual populations standing out from neighbouring populations in several previous studies. This thesis attempts to identify how different they are and the reasons for it. Then, this thesis also investigates the effect of this distinction on their internal population structure, as well as the current and future trajectory. A panel of microsatellite markers from the SALSEA-merge project were used to complete four studies of population structure in Atlantic salmon.

In the first study, which served primarily, as a training exercise, a multi-national baseline was used to identify the origins of salmon recolonising the river Mersey in northwest England. Fish entering the Mersey originated from multiple sources, with the greatest proportion (45–60%) assigning to rivers in the geographical region just north of the Mersey, including Northwest England and the Solway Firth. The number of fish originating from proximal rivers to the west of the Mersey was lower than expected. The results suggested that the recolonisers were straying in accordance with the predominantly clockwise gyre present in the eastern Irish Sea.

In the second study, the relationship of salmon in the chalk streams of southern England to salmon outside this region was elucidated. Salmon from all five chalk streams in southern England with major salmon populations were found to all be genetically distinct from these neighbours and statistically less genetically diverse than salmon in southwest England and France. The reasons for this were relatively low immigration and a history of low effective population size.

In the third study, the extent of population structure of salmon between the chalk streams and within one chalk stream, the river Frome, was explored. The results suggested these salmon were divided into three groups, i.e. 1) the Frome & Piddle, 2) the Avon and 3) the Test & Itchen. A significant pattern of isolation by distance between salmon in these five rivers was also identified. Historic samples from the Avon were assigned to the contemporary three groups. Surprisingly, most

of these fish assigned to the Frome and Piddle group. Within the river Frome, further sub-structure was identified over two separate years of sampling. Salmon from 2009 comprised three genetic groups, and salmon in 2011 comprised just two.

In the fourth study, historic scale samples were used to assess the current trajectory of genetic diversity and effective population size of salmon populations across Scotland, England, Wales and France. The majority of samples greater than 30 years old proved ineffective using the SALSEA panel. However, data was compiled from samples from eight rivers ranging from the Tweed in Scotland to the Scorff in France and from 1972 to 2012. Contrary to our hypothesis, most populations showed increases in allelic richness. Populations from one chalk stream show the steepest temporal decline in genetic diversity, which we speculate is partly due to the low immigration into the region. Effective population size proved difficult to determine using a number of methods and no robust pattern was identified.

Together these studies indicate that low immigration of salmon into the chalk streams appears to be key to their low genetic diversity and genetic distinction. Low immigration may also have enabled marked within-river population structure and the current negative trajectory of genetic diversity. The implications for general understanding of Atlantic salmon population structure across their range, and for the conservation of this species are discussed.

Table of contents

List of figures

List of tables

List of Appendices

Acknowledgements

Chapter 1: Introduction

Introduction

Conservation genetics is an interdisciplinary field of science that aims to use genetic methods to conserve biodiversity. The importance of conserving genetic factors was first illustrated by Frankel (1974). Using plant crops as an example, Frankel (1974) argued that the current increase in productivity of a few species had been gained at the expense of biodiversity in the wild. He also argued that this was detrimental to long term productivity, because within the lost gene pools there were likely to be further sources of productivity breakthroughs. Unfortunately his view did not prevent further demise of biodiversity, and our planet is undergoing what is often referred to as the Sixth Extinction (e.g. Brook *et al.* 2008; Collins 2010)

The Atlantic salmon (*Salmo salar L.*), the subject of this thesis, embodies Frankel's argument. The current global salmon population is bigger than ever before (Parrish *et al.* 1998). Annual per capita consumption of freshwater and diadromous fish species has increased from 1.5 kg in 1961 to 6.5 kg in 2010 and this has been driven largely by an increased consumption of salmonids (FAO 2014). However, the vast majority of available Atlantic salmon biomass (ca. 98%) is the product of artificial fish farming methods (Parrish *et al.* 1998). Wild salmon populations have greatly reduced in size and have become extinct in 15% of their native rivers (WWF 2001). In fact, of the 19 countries that still possess wild salmon, populations are regarded as healthy within only four: Scotland, Ireland, Iceland and Norway (WWF 2001). Unfortunately, there are many reasons for their decline, and their recovery will require interdisciplinary science and multi-national management. One increasingly proficient tool from science is population genetics. This thesis uses population genetics to further our understanding of Atlantic salmon population re-colonisation, structure and change over time.

Distribution

Historically, the natural range of Atlantic salmon spanned throughout both sides of the Atlantic Ocean (Figure 1.1). Currently, in the west, salmon can be found from the Hudson River, which drains New York state, to outer Ungava Bay in Quebec (MacCrimmon & Gots 1979). In the east, salmon span southward from

Iceland, the Barents Sea, and south western parts of the Kara Sea along the coastal drainage to northern Portugal and the Bay of Biscay (MacCrimmon & Gots 1979). They are also found in over 60 Icelandic rivers (MacCrimmon & Gots 1979). However, populations have completely disappeared from parts of North America and many parts of Europe including Belgium, the Netherlands and Germany (Parrish *et al.* 1998).

The range of the species has also been artificially extended. Atlantic salmon are found in hatchery facilities in Mexico, South America, South Africa, India, Indonesia, Australia and New Zealand (MacCrimmon & Gots 1979). The scale of these fisheries is large; for example, 44,000 tonnes of Atlantic salmon were produced in Australia between 2011 and 2012, making Atlantic salmon production the country's highest value fisheries product (http://www.sustainableseafood.org.au/fish.php/1/6/atlantic-salmon).

Life history

Wild Atlantic salmon are anadromous, which is an uncommon characteristic found only in ca. 110 fish species (McDowall 1997). Anadromous fish begin life in freshwater, mature at sea and return to freshwater to spawn before either returning to sea to repeat the cycle, or dying (McDowall 1997) (Figure 1.2). A fuller description for Atlantic salmon follows in order to appreciate fully current research on the species.

An Atlantic salmon life cycle might begin at the point of egg laying, which occurs in freshwater. Across their range, eggs are typically laid during the autumn and winter months (Fleming *et al.* 1996). Sexually mature females lay their eggs within specially dug nests, and males compete to fertilise them (Myers & Hutchings 1987). After fertilisation, the eggs are covered with gravel to protect them from predators, other females, desiccation during low water and freezing (Fleming *et al.* 1997). Females dig multiple nests in tandem and a string of nests is referred to as a "redd" (Myers & Hutchings 1987). The nest building and egg laying period lasts approximately five to six days for females, but males are usually sexually active for at least a month (Fleming *et al.* 1997). Unlike some species of salmonid (e.g. Pacific salmon), Atlantic salmon do not all die after mating (Saunders & Schom 1985); but

the proportion of multiple spawners varies between rivers and the probability of surviving to spawn again decreases with increasing size (Fleming 1998).

Eggs typically hatch between late March and early April. Emerging juveniles are known as alevins (Verspoor & Nielsen 2007) (Figure 1.2) and possess a yolk-sac attachment enabling them to remain in the gravel where they hatched for four to five weeks. As the yolk sacs approach full exploitation, they emerge from the gravel in order to find food (Mills 1971; Verspoor & Nielsen 2007). Now known as fry (Figure 1.2), they remain in close proximity to their redd site and siblings in areas with shallow riffles and low water velocity, for up to two weeks (Verspoor & Nielsen 2007). Fry (and subsequently parr), feed on chironomids, stoneflies, caddisflies and, if given the opportunity, many terrestrial insects (Verspoor & Nielsen 2007). On this diet fry grow fast and may be four times their initial fry length after a year, at which point they are known as parr (Figure 1.2).

The length of time that parr spend in-river, before going to sea, varies roughly from one to six years and is dependent on their growth rate (Klemetsen *et al.* 2003), which is itself a function of food availability and temperature (Klemetsen *et al.* 2003; Verspoor & Nielsen 2007). As these factors vary with latitude, the age at which most salmon head to sea also varies accordingly. At low latitudes (e.g. UK and Spain), where the water temperature and productivity are both relatively high, this freshwater stage is generally 1-2 years. In northern latitudes (e.g. Norway), where the water is colder and productivity is lower, the freshwater stage is roughly 3-6 years (Klemetsen *et al.* 2003; Verspoor & Nielsen 2007).

It is worth noting that in most rivers, a small proportion of juveniles will reach sexual maturation before going out to sea (Fleming 1998; Klemetsen *et al.* 2003). These individuals, known as precocious parr, are almost always male and fertilise up to 40% of eggs in a river catchment (Fleming 1998). However, they are usually reliant upon sneaky mating tactics, as they are unable to compete with the fully-grown adult males (Fleming 1998).

Smolt transformation

Before migrating to sea, salmon parr (both precocious and non-precocious) must undergo a process known as smoltification, where morphological and physiological transformations prepare them for growth and survival in the marine environment (Boeuf *et al.* 1994). Initiation is size-dependent; if a parr reaches a critical size by spring it will undergo smoltification, and if not, it will wait until the following year (McCormick *et al.* 1998). Morphological changes include body silvering – caused by the deposition of guanine and hypoxanthine in the skin and scales (McCormick *et al.* 1998) –, and body streamlining – caused by a greater increase in body length than in weight (McCormick *et al.* 1998). Physiological changes include a shift in visual pigments from porphyropsin to rhodopsin, and increased buoyancy (McCormick *et al.* 1998). Salmon that choose to undergo smoltification also grow faster than their non-smolting counterparts, fuelled by an increase of growth hormone (Boeuf *et al.* 1994). Most importantly for life at sea, they also gain an increased salinity tolerance, the mechanisms of which, have been widely studied (Boeuf *et al.* 1994; McCormick *et al.* 1998). An increase in gill Na^+, K^+ -ATPase activity, the number and size of gill chloride cells and intestinal water permeability have each been linked to increased salinity tolerance of smolts (McCormick *et al.* 1998). Interestingly, if necessary parr can gradually become acclimatised to seawater, however smolts can survive entering seawater directly with minimal ionic disturbance (Hoar 1988).

Just before migration downstream, smolts develop increased sensitivity to environmental factors and increased olfactory sensitivity (Boeuf 1993). Following an environmental trigger, such as heavy rainfall or an increase in water temperature (Solomon 1978), smolts migrate *en masse* down river to sea. Smolt mortality during the migration downstream can be exceptionally low with estimates of up to 90% survival in some rivers (Jepsen *et al.* 1998). However, survival rates do vary, with estimates for some rivers being as low as 17-51% (Aarestrup & Koed 2003). Evidence also suggests that precocious parr are less likely to survive the migration (Lundqvist *et al.* 1988). Smolts that survive this migration then remain within estuaries, where the brackish water enables them to adapt slowly to increasing salinity (Boeuf 1993) before eventually moving to open water.

The open ocean – where do salmon go?

Although Atlantic salmon potentially have the whole of the northern Atlantic to explore, salmon migrate to specific feeding grounds depending on the region where they were spawned. Salmon from rivers in North America largely remain within the west Atlantic in waters off western Greenland, the shelf off Newfoundland and in the Labrador Sea (Verspoor & Nielsen 2007). Salmon from Europe move far into the north and northeast Atlantic. Salmon from Canada and southern Europe are large contributors to stocks in west coast of Greenland (Guerin *et al.* 2014). Ocean feeding areas in the Faroe Islands and above are largely populated by fish from Norway, Scotland and Russia. Understanding of this migration is still improving, and some recent progress has be made using isotope analysis (MacKenzie *et al.* 2011; Mackenzie *et al.* 2012). This analysis has identified that salmon from different natal origins in the UK feed in different oceanic regions suggesting, for example, that salmon from northeast England most likely feed within the Norwegian Sea, while salmon from the river Frome in southern England feed around the Icelandic shelf (MacKenzie *et al.* 2011). Isotopes also suggest that feeding is age structured, so that one-sea-winter fish and multi-sea-winter fish tend to feed in separate oceanic regions (Mackenzie *et al.* 2012).

Natal homing and straying

Atlantic salmon endure journeys that can be thousands of kilometres long to return to their natal river to spawn. This return migration can be split into two main stages (Hansen *et al.* 1993): the first is the orientation from the feeding grounds to the home region, and the second is a more directed homing phase from within the coastal and estuarine areas (Hansen *et al.* 1993). The trigger for homing is not known for certain, but is thought to be related to the onset of sexual maturation, which is itself influenced by factors such as growth and photoperiod (Verspoor & Nielsen 2007).

How salmon orient their way home from feeding grounds is another area of avid interest and a number of mechanisms have been proposed. The discovery that salmon are sensitive to electromagnetic fields (Jr & McCleave 1973), the detection of magnetic material along the lateral line (Moore *et al.* 1990) and the fact that many other animals migrate using the earth's magnetic field – most notably birds (e.g.

Alerstam *et al.* 2001) – have led to unproven speculation that geomagnetism might be used by salmon at long range (Stabell 1984). Within Pacific salmon, there is evidence for magnetic imprinting at the point when juveniles enter the sea (Putman *et al.* 2013). Typically however, for Atlantic salmon, olfactory imprinting and sun-orientation are believed to be the primary homing mechanisms (Stabell 1984). Either way, it is clear that Atlantic salmon have the ability to return to the river in which they were spawned (Stabell 1984).

What is less clear is how accurate homing is. In a critical review of tagging studies, it was determined that 2-6% of salmon stray into rivers that were not their natal waters (Stabell 1984). However, some studies have estimated straying rates of up to 19% (Kuparinen *et al.* 2009). Straying serves many purposes for the species, including the maintenance of genetic diversity within small populations (Ardren & Kapuscinski 2003; Consuegra *et al.* 2005) and the colonisation of new habitats (Griffiths *et al.* 2011). Conversely, straying can also have negative impacts. If there is too much immigration for example, adaptations crucial to local populations may be lost (Jonsson *et al.* 2003). Understanding the causes and frequencies of straying would therefore be useful in salmon management.

However, obtaining accurate return and straying rates in Atlantic salmon is difficult for a number of reasons. First and foremost, Atlantic salmon are largely indistinguishable between rivers, and change considerably between the point at which they leave their natal rivers and the point at which they return. To combat this, salmon homing studies often involve tags, which are attached to juveniles and provide an indicator of the natal river if and when the salmon is caught again (Drenner *et al.* 2012). Tags often serve solely as visual indicators of the natal river, but can also have audio or radio frequencies attached for automatic detection, preventing the need to re-catch the fish (Jepsen *et al.* 1998; Drenner *et al.* 2012). This is the case on the river Frome where juvenile salmon are fitted with Passive Integrated Transponder (PIT) tags that are registered by PIT sensors near the river mouth recording exiting salmon smolts as well the returning adults (Ibbotson *pers. comm.*).

The second problem is survival, which is low in Atlantic salmon. For example, it has been estimated that only 1-3.4% of salmon parr survive to return as adults

(Stabell 1984). Thus, the number of salmon that need to be tagged in order to accurately determine the proportion that return is large (Ensing *et al.* 2013), making tagging studies costly in term of money, time and effort. It is even more difficult to determine how many salmon stray into a river or where they have strayed from. In this regard, population genetics has been increasingly useful (discussed henceforth).

Value of Atlantic salmon

The economic value of wild Atlantic salmon is huge. In Canada alone, the annual economic value of wild Atlantic salmon is calculated to be approximately $255 million (Pinfold 2011). In England and Wales, where a licence is legally required for salmon and trout fishing, there were nearly 20,000 full licences and almost 7,000 seasonal licences granted in 2007 (Mawle & Peirson 2001) generating an income of almost £1.5 million. In addition to licences, the purchase of rods and other materials provided further economic investment (Mawle & Peirson 2001).

Salmon also offer a social benefit often considered much greater than that of other fish, which has even been compared to that of charismatic megafauna such as tigers and leopards. They are said to have an "existence value" (NASCO 2008), which may exceed the economic values described previously. For example, from surveys it has been estimated that the average household in England and Wales would be willing to spend £15.80 to prevent a severe decline in salmon populations which amounts to £350 million (Mawle & Peirson 2001). They are also a valuable indicator species for the health of a river (ICES 2014).

The population decline

As indicated previously, despite their high value, salmon population sizes are falling (Figure 1.3), and have been for over the past 200 years (Parrish *et al.* 1998). This has partly been established from catch statistics from rod and net fisheries, which are often used to estimate salmon population size (Chaput *et al.* 2005). From a peak catch of approximately 12,000 tonnes in 1967 and 1973, catch has fallen to less than 2,000 tonnes since 2006 (Figure 1.3) (Parrish *et al.* 1998). Some of the decrease can be attributed to changes in fishing effort, and one way that fisheries account for that is by also measuring the catch-per-unit-effort (CPUE) alongside catch itself. CPUE is defined as a "derived quantity obtained from the independent

values of catch and effort," but is simply the amount caught divided by the amount of effort (ICES 2014). It is measured both in rod fisheries - where effort may be measured by the number of rod licences sold or the anglers themselves, for example - and in net fisheries - where effort may be measured by the number of boats and the number of hours spent at sea. Although useful, CPUE is also fraught with difficulties. This is because it is influenced by various factors, including fishing conditions and experience of the angler (ICES 2014). Over greater time scales, CPUE is also strongly affected by measures taken to reduce fishing effort, like the post 1980 closure and regulation of fisheries (ICES 2014).

Although some of the decrease in catch is due to the decrease in fishing effort, most of the decrease in catch is thought to be due to a real decline in the number of salmon in the sea (NASCO 2013). Unfortunately, there are many threats to salmon populations, described henceforth.

Threats to salmon

Marine survival and climate

Poor marine survival of post smolts is believed to be a key contributing factor to the current poor state of the species (WWF 2001; Potter *et al.* 2003). Survival is thought to have decreased during the late 1980s and late 1990s (WWF 2001) and, if real, this phenomenon does not appear to have ceased; smolts in England and Wales, for example, are still considered to be suffering very low marine survival (Cefas & Environment Agency 2013).

Changes in the sea environment are suspected to be a responsible (NASCO 2013). There is evidence that the decline in the early 1990s was preceded by changes across multiple levels of the sea ecosystem (Mills *et al.* 2013) and that the decline in 1997 was preceded by an exceptionally low North Atlantic Oscillation (NAO) event (Mills *et al.* 2003). The NAO is a climatic phenomenon of air pressure in the North Atlantic Ocean which controls the strength and direction of winds. It is also the dominant force for between-year variability in atmospheric circulation and is highly correlated with sea surface temperature in many regions (Mills *et al.* 2013), which influences algal growth and marine productivity, the declines and preceding changes in climate indices (e.g. the North Atlantic Oscillation Index), physical

conditions (e.g. temperature and salinity) and biological characteristics (e.g. phytoplankton abundance and zooplankton community composition) (NASCO 2013; Mills *et al.* 2013). These changes affect salmon both directly, as temperature is proportional to growth rate, and indirectly, as any changes to prey availability will affect the survival of salmon. Research in this area is still ongoing, for example it has recently been concluded that the NAO is relatively unimportant in salmon productivity, in favour of the Atlantic Multidecadal Oscillation (AMO), which is a measure of sea surface temperature (Friedland *et al.* 2014).

River pollution

Atlantic salmon are particularly sensitive to changes in river conditions (ICES 2014). Within river pollution takes many forms, has many avenues of effect and is considered the most significant factor for the decrease in Atlantic salmon population numbers (WWF 2001). Pollution from industry and agriculture has damaged many salmon rivers and removed salmon in some rivers completely, e.g. the river Mersey in northern England (Ikediashi *et al.* 2012) and the Thames in London (Griffiths *et al.* 2011). In fact, within the UK salmon were plentiful in both England and Wales until the Industrial Revolution towards the end of the 18th Century, when the amount of industrial and domestic pollution entering rivers greatly increased (MacCrimmon & Gots 1979).

Acid rain, and the resulting acidification of rivers, has had a significant effect on salmon populations in many regions. Salmon in 18 stocks in Norway were made extinct as a direct result of acidification, which was caused by pollution in Europe (Sandøy & Langåker 2001). This resulted in the loss of an estimated 100,000-300,000 salmon each year (WWF 2001). Rivers in Canada have also suffered the effects of acid rain (WWF 2001). Regulation of factory pollution has eliminated the possibility of acid rain in the future, yet the effects of past acidification are expected to last many years (WWF 2001). Eutrophication, caused by fertiliser run-off, also causes a problem: as although salmon can tolerate low concentrations of oxygen (5-6.5 mg/l), concentrations below 8 mg/l are considered detrimental to spawners (Binkley & Brown 1993). Other key pollution chemicals are pesticides and hormone disrupters (WWF 2001).

Habitat loss

Habitat loss within rivers is also a key factor. Over the past few decades, the freshwater range of salmon has been significantly reduced largely due to anthropogenic activities (Parrish *et al.* 1998) that result in habitat loss and degradation. The single greatest cause of salmon extirpations (MacCrimmon & Gots 1979) is believed to be river constructions such as the building of dams without fish passages.. This is particularly true of populations in Spain and France where salmon in many rivers have become extinct, or approach extinction, as a consequence of river structures (WWF 2001).

Fish farms

Over 2 million tonnes of farmed salmon were produced globally in 2012 (Figure 1.4; NASCO 2013). This was over 1,300 times the reported catch of wild salmon in the North Atlantic over the same year. The majority of salmon production in the North Atlantic is by Norway and Scotland (79% and 11% respectively); although there is considerable production outside of the North Atlantic (26% of the 2012 total) - largely dominated by Chile.

While, in theory, such heavy supply from aquaculture should relieve pressure on wild stocks and help salmon populations recover, in reality farmed salmon and the process of salmon farming pose significant threats to wild stocks for the following reasons. Typically, farmed salmon are kept in cages at a much higher density than they would be in the wild, and are prone to pathogen infection. These can be passed on to wild fish through infected escapees and when wild fish come too close to sea pens. Sea lice (*Lepeophtheirus salmonis*) are a particular problem (Naylor *et al.* 2003; Krkosek *et al.* 2005)

Marine cages are vulnerable and subject to breakages, which result in frequent occurrences of large-scale escapes (McGinnity *et al.* 2003). Farmed salmon escapees are relatively unfit compared to natural populations; for example, farmed females have been found to construct fewer nests and suffer greater egg mortality than wild females. They typically demonstrate less than a third of the reproductive success of wild females (Fleming *et al.* 1996). Farmed males suffer poorer competitiveness with less success in courting and spawning equating to only 1-3%

success compared to wild males. Farm salmon globally tend to be from Norwegian farm strains (McGinnity *et al.* 2009) and are therefore genetically different from most wild populations that they encounter. Intense farming methods have resulted in further differentiation, through founder effect, inadvertent selection and genetic drift during domestication (Crozier 1993; McGinnity *et al.* 2003). The effects of hybridisation of wild and genetically less-fit farmed salmon are often negative. Studies in Ireland (McGinnity *et al.* 2003) identified reduced survival in F1 crosses, despite the faster growth of hybrid juveniles. These larger hybrids also displaced fully wild parr, which then had lower survival than they otherwise would have (McGinnity *et al.* 1997, 2003).

Mitigation

In an attempt to restore salmon populations to pre-decline numbers a large number of steps have been taken both nationally and internationally. These can be divided into steps taken at sea and steps taken within river described henceforth.

Management at sea

For a species with such a wide distribution and, crucially, where stocks from one country can be exploited at sea by another country, it is necessary to have international conservative efforts. Thus, the North Atlantic Salmon Conservation Organization (NASCO) was formed by an inter-governmental convention in 1984 (WWF 2001). Their objective is to "conserve, restore, enhance and rationally manage Atlantic salmon" (http://www.nasco.int/about.html), and they aim to do this using the best available scientific information as well as international collaboration (ICES 2014). Member states are still responsible for the management of wild salmon in their own rivers, but distant water fisheries, which exploit salmon originating from other another member state (e.g. Greenland and Faroe Islands), are regulated by NASCO (ICES 2014). Management is filtered through three Commission areas, which are the North American Commission (Canada and USA), West Greenland Commission (Canada, Denmark, the European Union and USA), and the North-East Atlantic Commission (Denmark, the European Union, Norway and the Russian Federation).

The aim of their management is to maintain all member state stocks above their conservation limits. Conservation limits are defined as the number of spawning fish that achieve a long term average maximum sustainable yield (MSY) (ICES 2014). Enforced measures include the limiting of fishing for salmon beyond a member states' jurisdiction, and limiting the range over which a member state can fish for salmon. Major salmon fisheries have also closed. This includes all major commercial fisheries in Canada by 2012 after incremental closures since 1992 (NASCO 2013), and the closure of commercial fisheries for export in West Greenland in 1998 (NASCO 2013). In England and Wales the majority of salmon fisheries have also closed through government enabled phase-out schemes (Cefas & Environment Agency 2013). These measures have led to a gradual decrease in the exploitation of salmon. In 2013, a total nominal catch of 1,296 tonnes was recorded across their range, which was 115 tonnes lower than 2012 and the lowest ever recorded (ICES 2014).

Management in river

Apart from the reduction of fishing effort, options for improving survival of salmon at sea are considered limited (WWF 2001); however there is considered to be much greater potential to improve stocks within their natal rivers. One method is the policy of catch and release, where salmon caught in river, usually by rod and line, are released alive. This method has been practiced in USA since 1984, but has increased in Europe since the 1990s (ICES 2014). The proportion of catch and release is noted by NASCO and ranges between countries. In 2013, for example, this ranged 15% in Norway to 80% in Scotland (ICES 2014).

There is an increasing focus within wildlife conservation on preserving the natural habitat of an organism, and encouraging natural process of recovery. This is also the case for Atlantic salmon, where by improving the water quality towards to pre Industrial Revolution levels encourages the return of salmon. This occurs through the process of straying, which has been identified on the river Thames (Griffiths *et al.* 2011) in England, as well as the river Seine in France (Perrier *et al.* 2010).

Hatcheries and stocking

Hatcheries have played a significant role in the mitigation of salmon population declines across their range, with variable results. Up until the late 1990s, it was common across Europe to supplement local populations with the broodstock of salmon from rivers in countries with healthier populations. In many cases broodstock originated from Scotland, as it was for many rivers in France (Perrier *et al.* 2013), Spain (Saura *et al.* 2006) and England (Finnegan & Stevens 2008; Griffiths *et al.* 2010). However, the effectiveness of such stocking efforts – which was not easy to determine at the time – has in most cases been identified as poor (Saura *et al.* 2006; Finnegan & Stevens 2008; Griffiths *et al.* 2011). Thus, the process of supplemental stocking with exogenous fish has ceased.

Instead, supportive stocking with local fish is now widespread. In this case, local adults are caught in autumn months and farmed within hatcheries for broodstock. Artificial pairings are implemented and offspring are released to nearby rivers. This is the case in Spain (Saura *et al.* 2006), England and Wales (Cefas & Environment Agency 2013) and Norway (Saltveit 2006). In fact, many rivers in Norway, where dams prevent salmon from reaching spawning habitat, are entirely dependent upon such supportive stocking (Saltveit 2006) and successful females are kept within hatcheries for multiple years of eggs.

The effectiveness of even these stocking attempts is, however, facing increasing criticism, as incremental reports find evidence contradicting the success of this expensive process. One example is the river Tyne where, as mitigation for development work, 160,000 salmon were stocked annually from 1979 (Milner *et al.* 2004). Although initial increases were seen as the result of the stocking, a relatively recent report has identified that the majority of the recovery was likely natural (Milner *et al.* 2004).

Problem with population size estimates

Despite these significant measures to reduce global exploitation, natural population sizes remain at an all time low (NASCO 2013). However, at least some of this could be an artefact of the methods employed. Both population size and productivity of Atlantic salmon stocks are estimated by the International Council for

the Exploration of the Sea (ICES) using models based on catch statistics, CPUE and returns to natal rivers (Chaput *et al.* 2005; ICES 2014). These inherently have a number of flaws (WWF 2001). For example, fishing effort is never constant on a daily or yearly basis. This is particularly noticeable over greater time scales, where on one side fishing techniques for anglers and net fisheries have greatly improved (Lynch *et al.* 2012). On the other side, the imposed reductions and closure of fisheries, described previously, have significantly reduced effort.

Also, all stock estimates are based on estimates of adult returns, which are almost certain to be variable in accuracy. This makes comparisons between countries increasingly difficult. Finally, as with all models, no matter how complicated it is, it is only the best available estimate for that time. A recent study has attempted to improve upon the models that ICES use (Massiot-Granier *et al.* 2014). In doing so the author of the study (Massiot-Granier *et al.* 2014) indicates further problems with the initial models, which may "bias estimates of stock productivity." The study makes an important claim: that the previous model exaggerates the decrease in marine survival between 1971 and 2010. The study also claims that its new model – which improves upon the previous with additional parameters of density dependent egg-to-smolt survival – "dampens" the sharp decline between 1988 and 1990 (Massiot-Granier *et al.* 2014). This would have significant consequences for our current understanding of the population declines. It is worth noting that ICES recognises the new model and is moving to integrate it into future estimates (ICES 2014).

However, because there are significant problems with estimating salmon population sizes based on catch statistics, there is a clear need to supplement catch based estimates with other tools and techniques. One tool that has the potential to elucidate population size, and also to answer other difficult questions, is population genetics. Because population genetics is the primary tool used throughout this thesis, it is necessary to describe the following key concepts.

Basic Concepts for population genetics

As population genetics is the primary source of analysis within this thesis, it is important to understand some of its basic concepts. The field of population genetics is described by Verspoor & Nielsen (2007) as the "science of studying how genetic

variation is distributed among species, populations and individuals", and is concerned with how gene flow affects the distribution of genetic variation (Verspoor & Nielsen, 2007). Gene flow is defined by the following evolutionary forces of action: (natural) selection, mutation, migration and random genetic drift. Briefly, selection occurs when one iteration of a gene (an allele) offers an advantage over another allele, and subsequently has an increased chance of being passed on to the next generation. Mutations are changes in the DNA sequence of an organism caused either when a single base pair is substituted by an incorrect base pair during DNA replication, or when a base pair is added or deleted. Migration is when and in individual moves from one population into another and reproduces there. Finally, genetic drift describes the process where alleles are lost by chance. The smaller a population or subpopulation is, the greater the chance that an allele will be lost.

Within Atlantic salmon, population genetics can be used to understand many aspects of their biology, but two key aspects are population structure and changes in population size. In understanding the population structure of the species, the aim is to identify specific units to target management more effectively. As described by Frankel (1971), it is also important to conserve as much biodiversity as possible for the sake of the species itself. Population structure in salmon is prevented from complete panmixia (where all individuals have equal chance of breeding of with each other) by their ability to return their natal river. However, while it is apparent that the majority of salmon return to their natal river, the level of population structure within individual rivers, and the level of connectivity between different rivers – i.e. the amount of straying – is still unknown. The reasons for straying, which would be useful for management, are also unknown.

Changes in population size can, in theory, be detected through the observation that a population that is decreasing in size is increasingly likely to lose alleles (Garza & Williamson 2001). Thus, from the number of alleles a population has, population geneticists are able to view a population decline (Garza & Williamson 2001; Nikolic & Chevalet 2014). This information is useful to complement population size estimates based on catch data, as described previously (Chaput *et al.* 2005). It is also important to conserve alleles, or more broadly genetic diversity, as is recognised by the IUCN (NASCO 2009). Genetic diversity incorporates not only

the number of alleles, but also the proportion of individuals within a population that have multiple (heteroyzgote) alleles, rather than just one (homozygote). Reduced genetic diversity is harmful to populations for two reasons (Reed & Frankham 2003). Firstly, it has been found to reduce the fitness of populations (Reed & Frankham 2003) in a number of ways, including a reduced ability to respond to new pathogens. Secondly, it is thought to also limit the future evolutionary potential of the species (Frankham *et al.* 1999). Bottleneck events, where the number of individuals with a population is greatly reduced, are bad for populations because of the reduction in genetic diversity. Species like the cheetah (*Acinonyx jubatus*), are believed to have low genetic diversity presently due to past bottleneck events (Menotti-Raymond & O'Brien 1993), which is typically cited as part of the reason for their poor breeding success (Menotti-Raymond & O'Brien 1993).

Another crucial element of population genetics, and this thesis, is the concept of effective population size (N_E). Within real populations only a fraction of individuals will breed, which is caused by multiple factors, including unequal sex ratios, harem social structure, and poor survival to adulthood. The genetic variability within a population will reflect the (usually smaller) number of individuals that breed, rather than the size of the population, i.e. the N_E. Put briefly, N_E corresponds to the number of individuals within a population that would, within an idealised population, show the same amount of genetic variation as the real population under random genetic drift (Nikolic *et al.* 2009). N_E is important, and increasingly monitored, because it determines how quickly alleles are lost from a population, which as described previously, is bad for populations (Reed & Frankham 2003).

Tools of population genetics

The methods used in population genetics are constantly evolving, and it is important to understand the genetic markers used in this thesis. Presently mitchondrial DNA (mtDNA) sequences, microsatellites and single nucleotide polymorphisms (SNPs) are the most frequently used markers for population genetics, and will be described henceforth.

Mitochondrial DNA sequences

For inferring the evolutionary relationship between populations within a species, mitochondrial DNA sequences were once commonly used. Mitochondrial DNA typically has a faster mutation rate compared to coding nuclear DNA. Within Atlantic salmon the ND1 gene is the most variable region (Verspoor *et al.* 1999) and is subsequently often the focus of research (Verspoor *et al.* 1999; Consuegra *et al.* 2002). Mitochondrial DNA has been used to explore the phylogeographic colonisation of salmon throughout North America (King 2000) and Europe (Verspoor *et al.* 1999; Consuegra *et al.* 2002; Finnegan *et al.* 2013). The use of full mitochondrial DNA sequences in published studies has dwindled since the development of SNP markers, however they can still be a useful method for determining evolutionary relationships and will be used in this thesis.

Microsatellites

These are currently the most accessible markers for studying population genetics in salmon, although they are due to be superseded by SNPs. As they are the primary tool of this thesis, a detailed description follows.

Microsatellites are short (<6 base pairs) tandem repeats of short sequence motifs (ca. 100s bp) found randomly within the genome of eukaryotes (Jarne & Lagoda 1996). They were initially regarded with disinterest, but by the late 1980s, they were considered to be "the most powerful mendelian markers ever found (Jarne & Lagoda 1996)." They had several advantages over allozymes. Firstly they possess greater polymorphism – while allozymes contain 1-5 alleles per locus, microsatellites provide 1-50 alleles (Jarne & Lagoda 1996). Secondly, they are selectively neutral – while proteins have functions and are prevented from mutating randomly (i.e. are under balancing selection), microsatellites have no function, and thus in most cases are expected to be selectively neutral. Thirdly, they are also much easier to score, so that many more individuals can be scored at many more loci (positions) for microsatellites compared to allozymes.

As non-coding portions of DNA, microsatellites have a much faster rate of mutation than allozymes; microsatellite mutation rates are approximately 10^{-2}-10^{-6} per locus per generation (Ellegren 2000), which is two to three orders faster than

allozymes. This is also faster than the average rate of mutation for a single nucleotide within a DNA sequence of 10^{-9} (Ellegren 2000). Most frequently microsatellite mutations occur via the process of polymerase slippage (Jarne & Lagoda 1996), where an error is made during the DNA replication process causing an increase or decrease in the number of repeats in the daughter cells. However, larger mutational changes can also occur during other processes such as unequal crossing-over during meiosis.

Several mutation models have been proposed in order to predict how often the different types of mutations occur. The most popular model is the simplest i.e. the stepwise mutation model (SMM) by Ohta & Kimura (1973). In the SMM model, all mutations are single steps, meaning that only one repeat unit is either gained or lost. There is also the two phase model (TPM) (Rienzo & Peterson 1994), where some proportion of the mutations are SMM, while the remaining proportion have the chance to mutate with larger jumps. These models also make the following assumptions: that a microsatellite is equally likely to expand or contract, and that the mutation rate stays constant no matter the size of the microsatellite. Another popular model is the infinite allele mutation model (IAM) (Kimura & Crow 1964), where mutations can only lead to new allelic states and can involve any change in size. A key feature of mutation models is that they vary according to species (Ellegren 2000). For Atlantic salmon, the single mutation model is most often used (Ribeiro *et al.* 2008; Nikolic *et al.* 2009; Perrier *et al.* 2013; Olafsson *et al.* 2014), although increasingly frequently a TPM model is used (Grandjean *et al.* 2009; Ellis *et al.* 2010) with commonly 95% single-steps and 5% multi-step mutations.

Single Nucleotide Polymorphisms (SNPs)
SNPs are single nucleotide polymorphisms found within the nuclear genome and make up 90% of all genetic variation with the human genome (Brumfield *et al.* 2003). It is only through the recent development of screening technologies (e.g. Buetow *et al.* 1999; Picoult-Newberg *et al.* 1999), that regular characterisation has become possible. SNPs have a number of advantages and disadvantages over microsatellite loci, which have been explained in full during a review by Brumfield *et al.* (2003). They have, on average, a slower mutation rate than microsatellites of 10^{-8}-10^{-9}. Subsequently, from a possible four states (from four nucleotides) most

SNP mutations are in fact only bi-allelic. While this would mean less differentiation using the same number of loci as microsatellites, the widespread presence of SNPs within a genome and ease of detection mean that a much larger number of SNPs can be amplified than microsatellites for the same effort or less.

At the beginning of this thesis, the resources were not available to begin a SNPs salmon study at Exeter University, whereas a panel of microsatellites had been produced for the species, which had widespread use (Ellis *et al.* 2011). Thus throughout this thesis microsatellites are the primary source of inference. This also enables direct comparison with previous studies (Ellis *et al.* 2010; Griffiths *et al.* 2010; Olafsson *et al.* 2014), and the SALSEA database in particular (Gilbey *et al.* unpublished). As microsatellites have been used for a much longer period of time than SNPs, many methods for analysing data are also robust. As the inclusion of SNPs also would have required more time, money and expertise than was available, they were not used within the studies.

Population structure

Population structure within regions

Microsatellites have improved our resolution of salmon population structure. Within Europe for example, using 12 microsatellites selected during the Atlantic Salmon Arc Project (ASAP) (Griffiths *et al.* 2010), salmon in the southern part of their European range have been divided into those from northern Scotland & Ireland, central Scotland & eastern Ireland, northern England & southern England, northern England & borders of Scotland, southwest England & Wales, northern France, northern Spain, and southern England (Figure 1.5) (Griffiths *et al.* 2010). Similar groupings have also been determined within salmon populations in Ireland (Dillane *et al.* 2007), France (Perrier *et al.* 2011), Norway (Tonteri *et al.* 2009), the Baltic Sea (Säisä *et al.* 2003), and America (McConnell *et al.* 1997; Dionne *et al.* 2008a).

The finding of regional structure has led to significant changes in the management of off shore fisheries. Upon genotyping salmon caught in fisheries, we are able to determine the proportion of fish that belong to each region, via a process of genetic assignment (e.g. Piry *et al.* 2004). ICES advises that all fisheries should be managed on the basis of individual stocks and microsatellites are now part of the

process of identifying what those stocks are. As such, ICES has integrated microsatellite analysis and the identification of regional populations into part of its management strategy (ICES 2014). This has been to significant effect; for example, with microsatellites it was determined that drift nets in northeast of Ireland caught fish from multiple stocks, making the Folye salmon fishery a mixed stock fishery (Ensing *et al.* 2013). Therefore from 2008, the managing body of the Foyle area, introduced measures to reduce the capture of mixed stock, which included a ban on fishing seaward of Lough Foyle and restricting the number and size of nets in the Foyle estuary (Ensing *et al.* 2013). Microsatellite analysis is also used to estimate the proportion of different stocks within fisheries in Greenland, the Faroes and many others (ICES 2014).

In this thesis the salmon are assigned into their regional groups (Chapter 2) and differences between them are investigated (Chapter 3). Using the knowledge of where one group ends and another one begins, investigation of population structure within them can be better focused and more consistent (Chapter 4).

Population structure within river

With the use of microsatellites, different populations have also been identified within large river catchments. This includes the river Teno in Norway, which contains the largest indigenous stock of Atlantic salmon, which is unsurprising given its large catchment area of 16,386 km^2. Using microsatellites, it was determined that different populations within different tributaries separated, they believe, by natal homing to the different tributaries (Vähä *et al.* 2007). Separate populations have also been identified within the river Foyle in Ireland (Ensing *et al.* 2011), which has a catchment area of 4,450 km^2.

Although population structure has been identified within smaller rivers, e.g. the rivers Tamar in southwest England (Ellis *et al.* 2010) and the river Varzuga in Russia (Primmer *et al.* 2006), it has not been in the form of distinct sub-populations. In these cases, rather than distinct sub-populations, the studies have identified patterns of isolation by distance. Isolation by distance is a simple model of population structure, where there is a correlation between geographic distance and genetic distance (Wright 1943). Thus, current evidence suggests that for population

sub-division to occur in Atlantic salmon, rivers are required to be large and have many tributaries.

Yet, the generalness of this rule requires investigation. For example, salmon in four large rivers in North America, ranging from ca. 10,000-19,000 km^2 showed varying levels of genetic differentiation (Dionne *et al.* 2008b). Although genetic differentiation between tributaries was sometimes as large as that of between rivers, in one large river, the Miramichi, panmixia could not be rejected (Dionne *et al.* 2008b). Therefore, there is likely that there are other factors responsible for population subdivision, rather than river size and dendricity alone.

Monitoring change over time

Population genetics theory dictates that a smaller population will lose genetic diversity faster than a larger one, largely from genetic drift (Garza & Williamson 2001). Therefore, a population with relatively low genetic diversity is more likely to have been through a population decline. With the use of Atlantic salmon scales collected by anglers in the past, studies have been able to compare genetic diversity and effective population size (N_E) of past salmon populations to contemporary populations (Fraser *et al.* 2007; Valiente *et al.* 2010), largely to investigate whether they can confirm the apparent declines in salmon population number during the last century or to investigate the effect of past stocking efforts.

Interestingly, in Atlantic salmon populations, there does not appear to be evidence of the declines in genetic diversity and N_E (e.g. Consuegra *et al.* 2005; Perrier *et al.* 2013), that should follow the widespread population declines (Parrish *et al.* 1998; NASCO 2013). Some of the temporal changes identified include increases admixture between groups, which lead some to the proposition that population structure between rivers is breaking down over time (Perrier *et al.* 2013). No studies of salmon populations in England, Scotland and Wales have examined temporal changes in genetic diversity or N_E. They will, however, be looked at within this thesis, along with some populations from France (Chapter 5).

The chalk streams of southern England

As described previously, population structure in salmon has been identified between broad regions using microsatellites. It may be worthwhile to look at population structure differences and similarities between clearly defined groups, as well as within them. There is one potential group within southern England that is likely to be particularly interesting.

Within southern England there are rivers known as chalk streams. This name describes the geological rock the rivers are formed on, which is calcareous chalk laid down during the Cretaceous Period (ca. 66-145 million years ago). This chalk is relatively rare and found only within England, France, Belgium and New Zealand (Hampshire Biodiversity Partnership 2000). 85% of chalk streams are in England and span from North Humberside, along the east coast and down to Dorset in southern England (Hampshire Biodiversity Partnership 2000; Environment Agency 2004).

The chalk streams have a number of unique properties, because of the calcareous chalk (Berrie 1992). As the rock is porous, with 40% of its volume taken up by pore spaces (Berrie 1992), the majority of rain water flows through the chalk, rather than over it (as occurs in most other rivers) and thus chalk streams are aquifer fed. Therefore, the rock also acts as a filter, causing the water in the streams to be relatively clear (Berrie 1992). This creates a unique fishing experience, where anglers are actually able to see and target the fish before casting. In fact, the act of fly fishing is anecdotally believed to have been invented on a chalk stream, the river Itchen. Although salmon are not unique to the chalk streams, in fact no species is (Smith *et al.* 2014), salmon here have shown signs of being distinct from salmon in neighbouring regions (Griffiths *et al.* 2010). In a broad geographical study of salmon across a large part of Europe, salmon from the Avon, Itchen and Test stood far apart from salmon in neighbouring rivers in England and France, as well as from populations in Wales, Scotland and Spain (Griffiths *et al.* 2010). In a subsequent study, salmon from the Avon and Itchen appeared to be less genetically diverse than neighbouring salmon populations (Finnegan *et al.* 2013). As both studies had their own aims, no attempts were made to explain the unusual differences of the chalk salmon. However, either indication could have significant implications on the sustainability of these salmon.

Within this thesis, an attempt will be made to fully identify the extent of their genetic differentiation from non-chalk salmon, and explain this and their possible low genetic diversity (Chapter 3). This could have significant implications for the conservation of salmon in this region, and increase scientific understanding of salmon population structure across their range. Population structure within the chalk streams will also be investigated (Chapter 4). Although there are many studies investigating population structure between rivers, and now between groups (Griffiths *et al.* 2010; Perrier *et al.* 2011, 2013), none have attempted to look specifically within known groups, and salmon in these chalk stream may well be the most abruptly defined group yet.

The investigation is also important for salmon in this region, as their population size is believed to have fallen drastically since the 1980s (Figure 1.6). There are many threats to salmon in the chalk streams, including physical modifications to the catchments, water abstraction and pollution from point and diffuse sources (Hampshire Biodiversity Partnership 2000; Smith *et al.* 2014).

Overall thesis objectives

The aim of this thesis was to elucidate population structure in the chalk streams of southern England. On the first level (Chapter 3), the first objective was to determine how distinct salmon populations in the chalk streams were, compared to populations in immediately surrounding regions, and to determine why they were distinct. We hypothesised that salmon from all of these chalk streams would be genetically distinct and genetically less diverse. We also hypothesised that the difference was either because they were subjected to a unique genetic bottleneck, because they had relatively little immigration from outside the region, or because they had been separated from the other populations since before the last glacial maximum.

At the next level (Chapter 4), the objective was to determine the level of population structure in salmon between the chalk streams themselves, and within one individual chalk stream. In the final analysis (Chapter 5), the objective was to determine the trajectory of salmon population genetic diversity in the chalk streams, as well as populations in Scotland, Wales, other rivers in England, and France. Before these studies began, a training exercise was completed (Chapter 2), where the objective was to determine the source of salmon recolonising the river Mersey in northwest England. More details follow:

The origins of Atlantic salmon (Salmo salar L.) recolonising the river Mersey in northwest England (Chapter 2)

The river Mersey in northern England, is one of many rivers that suffered the loss of Atlantic salmon as a result of the Industrial Revolution (Burton 2003). Since the 1970s, a significant effort has been made to improve its water quality (Jones 2000; Burton 2003). In a rare success, Atlantic salmon have begun returning to the river, and several adult fish have been sampled. However, it was unknown where these salmon were coming from. Therefore, the aim of this study was to identify the source rivers of these recolonising fish. This was to be attempted using methods of genetic assignment and the SALSEA microsatellite panel. We hypothesised that the 100+ adult salmon caught in the river strayed from nearby rivers, in particular the river Dee, with which it shares an estuary. This follows recent findings (e.g.

Vasemägi *et al.* 2001; Griffiths *et al.* 2011), which have identified salmon recolonising from the nearest possible rivers.

As well as the primary aim, this study also contributes to the thesis aim of understanding population structure in the species. Because the Mersey no longer had salmon of its own, all salmon entering the river were strays, so we could begin to answer questions regarding straying in salmon. Is it always from a near river, or do some rivers, maybe with large endemic populations, produce more strays than others? This study was published in Ecology & Evolution by Ikediashi, Billington and Stevens (Ikediashi *et al.* 2012).

The distinction of Atlantic salmon in the chalk streams of southern England (Chapter 3)

It has been shown that salmon form groups of genetically more similar individuals covering broad geographical regions (e.g. Griffiths *et al.* 2010). Many factors had been proposed to explain what confines each genetic group to the region and one promising factor is geology (Perrier *et al.* 2011). Previous studies have hinted at an unusually large distinction between salmon in several chalk streams of southern England and all other salmon they have been compared with (Griffiths *et al.* 2010). However, until now, the chalk stream salmon had never been the focus of their own population genetic research and whether salmon in all these chalk streams formed a group together, distinct from the remainder, was unknown.

The aims of this study were two-fold, the first was to identify the extent of the distinction of chalk stream salmon, and to determine if they were less genetically diverse. The second aim was to identify why they were different. This second aim was explored from three avenues. 1) Had the chalk populations been through a bottleneck event? 2) Did they have low immigration? and 3) Had they diverged from the other populations at a much earlier date? For the first time, salmon were included from all chalk streams in southern England with significant salmon populations, i.e. the Frome, Piddle, Avon, Test and Itchen. With this data, comparisons were made against populations from non-chalk rivers in southwest England, France and Spain. These populations were chosen because of their close proximity to the southern English chalk streams and because they suffer similar

pressures from climate and human exploitation. Also, as they are each below the lower limit of the last glacial maximum, they were likely to have shared recent phylogeographic history. Based on previous work (Griffiths *et al.* 2010; Finnegan *et al.* 2013), which suggested that salmon in some of these chalk rivers are less diverse than their neighbours and that they stand apart, we hypothesised that salmon from all of these chalk streams would be genetically distinct and genetically less diverse. If this proved correct, the second objective was to determine the reason(s) for these differences. We hypothesised that the salmon in the chalk streams were different because they were subjected to a unique genetic bottleneck, had relatively little immigration from outside the region, or had been separated from the other populations since before the last glacial maximum.

Genetic analysis indicates marked population structure of Atlantic salmon (Salmo salar L.) among the chalk streams of southern England (Chapter 4)

Following investigation of population structure between salmon in the chalk streams of southern England and salmon in non-chalk rivers in England, France and Spain (Chapter 3), the population structure of Atlantic salmon within these chalk streams was investigated. Geological boundaries had only recently been observed as a force of population differentiation (Perrier *et al.* 2011). Previous studies investigating population structure between rivers had chosen rivers on an ad hoc basis. This may explain why there have been mixed conclusions regarding the extent of population structure at this level. This study is almost unique in its goal to elucidate population structure within a tightly defined geological region.

Within rivers, the identification of marked in-catchment population structure for this species appeared to be almost exclusive to large and dendritic rivers, including for example, the river Foyle (Ensing *et al.* 2011). As chalk streams are characteristically short and linear, we hypothesised population structure to be minimal. However findings from the previous chapter gave reason to expect otherwise.

This investigation thus took place on two levels. First the extent of population structure in salmon between the five chalk streams with major salmon populations was investigated. Historic samples from one chalk stream, the river Avon, were used

to explore the temporal stability of between river population structure. Then the extent of population structure with a single chalk stream, the river Frome, was investigated. While many studies have investigated population structure between and within rivers (e.g. Primmer *et al.* 2006; Ellis *et al.* 2010; Ensing *et al.* 2013), no study had previously focused on the chalk streams. We hypothesised that there is population structure in salmon between the chalk streams. This is because population structure, in the form of isolation by distance, has been identified between rivers in regions previously. This includes, for example, in Spain (Campos *et al.* 2007), Norway (Glover *et al.* 2012) and France (Perrier *et al.* 2011). However, within rivers population structure appears to be exclusive to larger dendritic rivers (e.g. Ensing *et al.* 2011). The chalk stream are characteristically small and linear (Berrie 1992); therefore we hypothesised that there would be no population structure within an individual chalk stream. This study has been submitted to the Journal of Fish Biology with edits and is currently in revision.

Temporal stability of genetic diversity and effective population size in Atlantic salmon across Great Britain and France (Chapter 5)

By using historic material and microsatellites, it is theoretically possible to detect temporal changes in population size and population structure. Using salmon scales, such investigations have occurred across Europe (e.g. Consuegra *et al.* 2005; Horreo *et al.* 2011; Perrier *et al.* 2013). However, no studies had investigated salmon populations within England, Scotland or Wales in this way.

This study aimed to identify changes in genetic diversity and effective population size across England, Scotland, Wales and France. This was done by obtaining salmon scales collected by anglers from rivers within Great Britain and France. Following this, attempts were made to obtain DNA from the material and to amplify the SALSEA panel of 16 microsatellites. With the data, investigations for changes in genetic diversity and effective population size were made. Following recent research on the current global state of Atlantic salmon (WWF 2001; ICES 2014), it was hypothesised that salmon in Scotland would show temporal stability in genetic diversity and effective population size - based on apparent stability in population size based on catch statistics -, while salmon in England, Wales and France would show temporal decreases in both measures -based on apparent

temporal decreases in population size. Several studies have made use of historic scale samples to investigate temporal change in genetic diversity and effective population size (e.g. Horreo *et al.* 2011; Perrier *et al.* 2013), however there are no similar published studies populations in Scotland, England, and Wales.

Marker choice

Throughout this thesis, microsatellites were the primary molecular marker used. In order to compare the results with previous studies, the study endeavoured to use the same set of 16 microsatellite markers as those agreed by the international SALSEA panel (Ellis *et al.* 2011a; Gillbey unpublished).
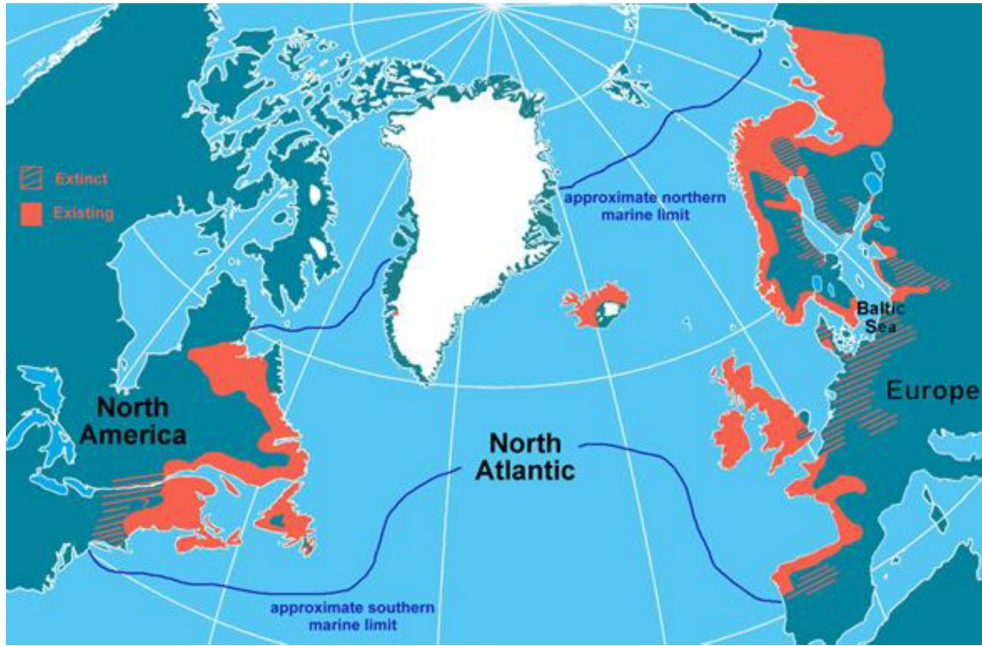
Figure 1.1- Natural range of Atlantic salmon (Image from Webb *et al*. 2007).
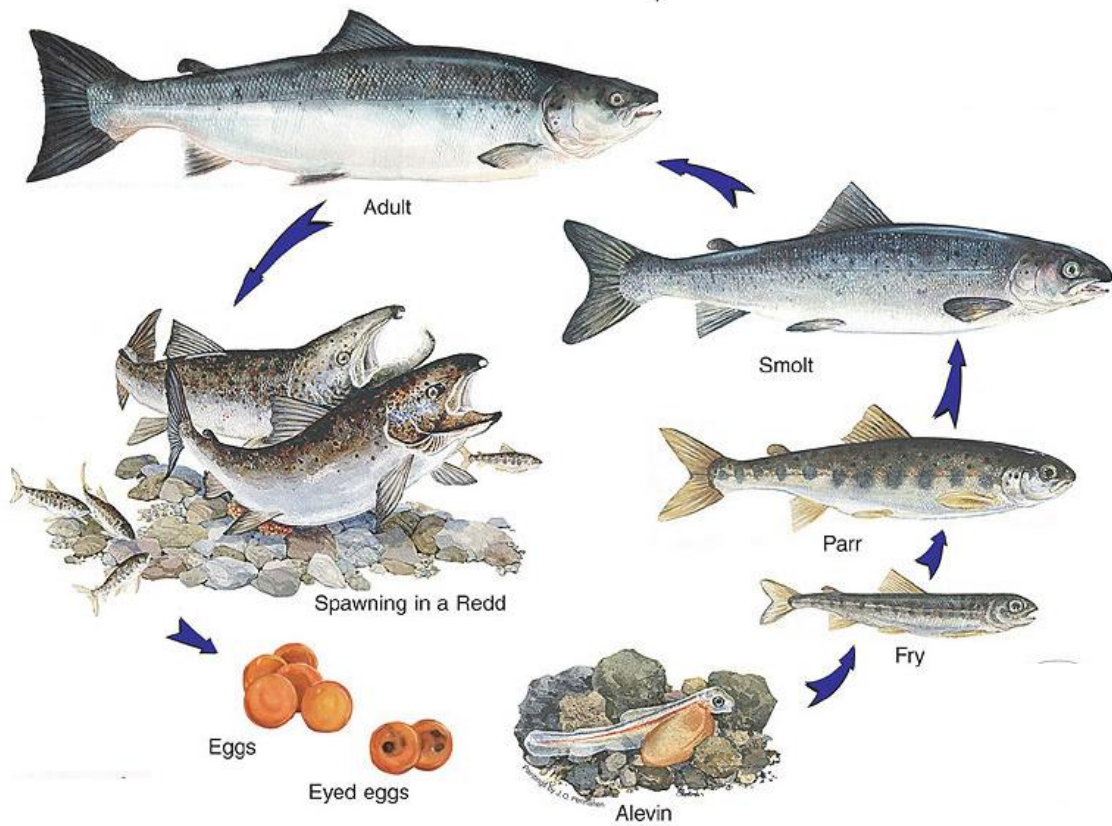
Figure 1.2- Diagram of Atlantic salmon life cycle. Image edited from the Atlantic Salmon Federation. Pictures not to scale.
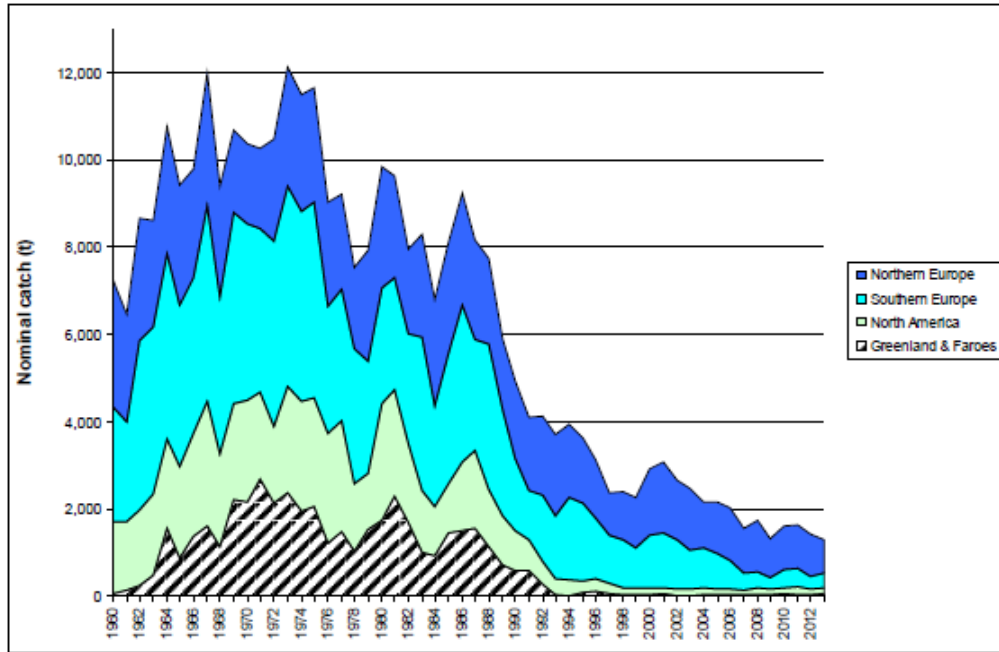
Figure 1.3- Reported total nominal catch of salmon. Catch in tonnes round fresh weight in the four North Atlantic regions between 1960 and 2012 (Image from ICES 2014).
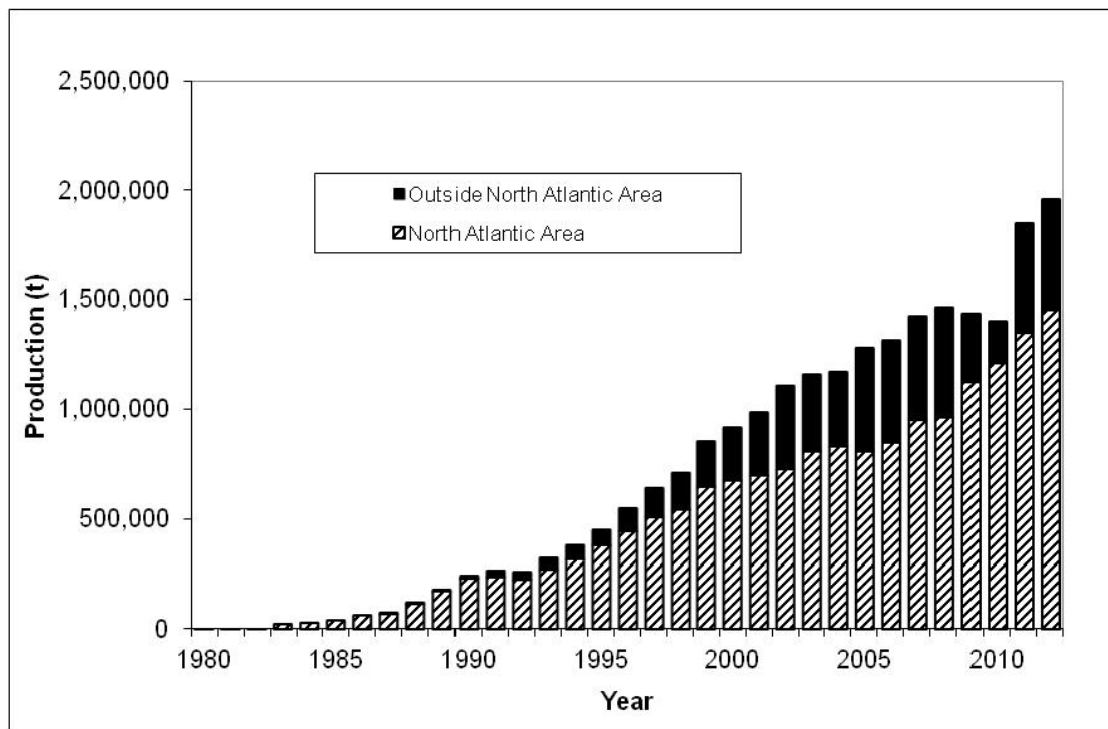
Figure 1.4- Worldwide production of farmed salmon. Production between 1980 - 2012 measure in tonnes. (Image from ICES 2014).
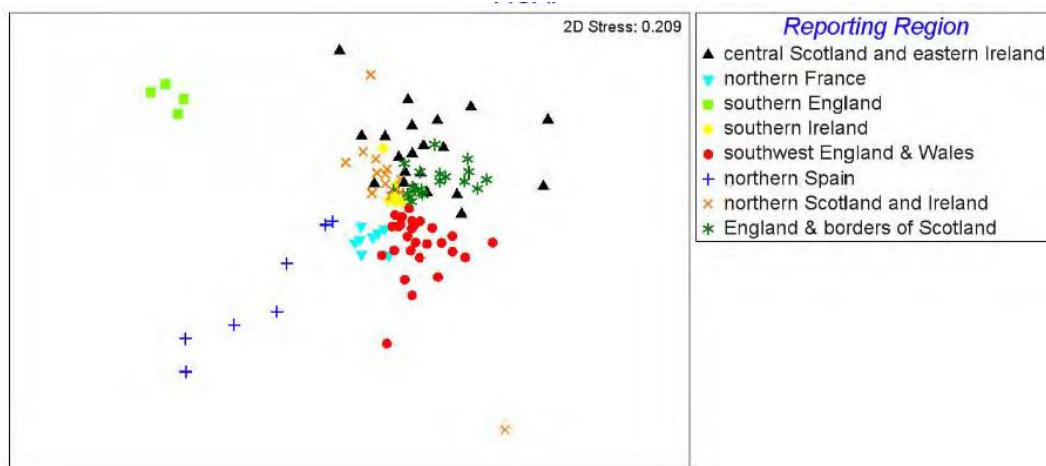
Figure 1.5- A multi dimensional scaling plot of salmon populations. The plot shows the genetic distance between salmon from rivers in England, Wales, Scotland, Ireland, France and Spain. Image from Griffiths *et al.* (2010).
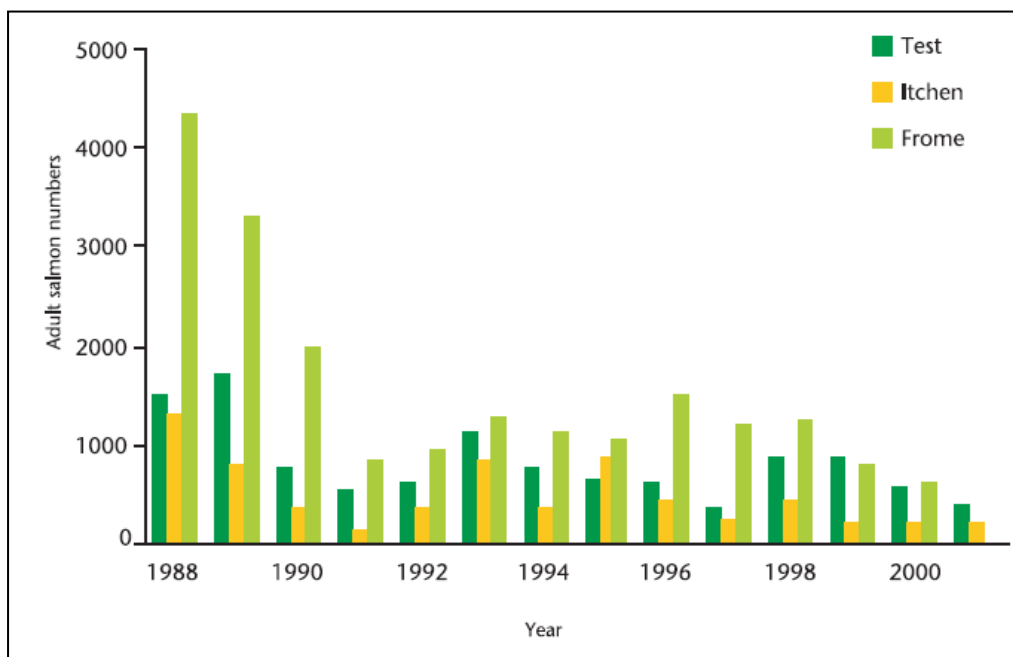
Figure 1.6- Estimated number of adult salmon in the rivers Test, Itchen and Frome. Numbers estimated from fish counter. Image modified from Environment Agency (2004).

Chapter 2: The origins of Atlantic salmon (Salmo salar L.) recolonising the river Mersey in northwest England

Charles Ikediashi, Sam Billington and Jamie R. Stevens (2012)

The original idea for this paper arose when the Environment Agency (Northwest Region) began to detect numbers of adult Atlantic salmon entering the river Mersey. Sam Billington (Environment Agency) approached us with a view to getting samples of these adult salmon genotyped and identified. Subsequently, Charles Ikediashi, Sam Billington and Jamie Stevens designed the research programme. Charles Ikediashi carried out all laboratory work and analysed the data. Charles Ikediashi, Sam Billington and Jamie Stevens interpreted analyses and wrote the paper.

Introduction

Global catch data shows that Atlantic salmon, *Salmo salar* L. (Figure 2.1), like many other fish have been in steep decline since the 1970s (e.g. Parrish *et al.*, 1998). The reasons appear to be multi-factorial, but include pollution and related reductions in water quality (Thorstad *et al.* 2007), reduction of access to waterways in which salmon spawn (Ugedal *et al.* 2008; Lin 2011), and an uncertain degree of marine mortality (Friedland 1998; Friedland *et al.* 2000).

Largely because of their iconic status and commercial value, huge amounts of money have been spent on reversing this downward trend, and a large proportion of this funding has been channelled through the controversial measure of stocking with hatchery-bred fish (Milner *et al.* 2004; Fraser 2008). Despite a clear lack of evidence regarding the success of stocking practices (e.g. Fraser 2008; Finnegan & Stevens 2008; McGinnity *et al.* 2009), it continues to be seen as a rapid solution to declining fish numbers by a significant number of fishery managers. Yet, in the light of genetic advances, stocking has come under further scrutiny as the limitations and, in many cases, negative impacts of the practice on the genetic diversity and population structure of endemic populations are revealed (Ayllon *et al.* 2006; Hutchings & Fraser 2008; Griffiths *et al.* 2009).

At the same time, the value of river restoration (in terms of both improved water quality and river access) is being recognised as a viable alternative, which can subsequently facilitate natural recolonisation. Examples have been reported for trout

(*Salmo trutta*) in Norway (Knutsen *et al.* 2001) and Germany (Schreiber & Diefenbach 2005), Atlantic salmon (*Salmo salar*) in France (Perrier *et al.* 2010) and coho salmon (*Oncorhynchus kisutch*) in Pacific northwestern USA (Kiffney *et al.* 2009). Most recently, a study by Griffiths *et al.* (2011), using microsatellite analysis, demonstrated no trace of the hundreds of thousands of Scottish and Irish-origin hatchery salmon stocked into the river Thames since 1975 (Griffiths *et al.* 2011); instead, all of the salmon caught in the Thames since 2003, were identified (by assignment analysis) as having originated in other proximal rivers in southern England. Genetic assignment methods have also been used to identify the origins of Atlantic salmon in the river Selja in Estonia (Vasemägi *et al.* 2001), the river Tambre in Spain (Saura *et al.* 2008) and the river Seine in France (Perrier *et al.* 2010), all of which have shown recent returns of Atlantic salmon after years of absence. In the cases of the Seine and the Thames, salmon appear to have returned naturally via straying after attempts at restocking were unsuccessful. If restoration is to be considered as a viable alternative to restocking for restoring Atlantic salmon, then more documented cases of natural recolonisation are required. In this regard, the river Mersey in northwest England presents an excellent case study.

The Mersey, which passes through the major urban areas of Liverpool and Manchester, suffered greatly as a result of the Industrial Revolution (see Jones, 2000, 2006 for full review). The 1820s saw the expansion of several industries (Gregory *et al.* 1953; Burton 2003) and industrial prosperity attracted huge numbers of people to the area (Handley & Wood 1999). Subsequent pollution had serious effects on fish stocks and by the 1850s fish were reportedly absent from the river Irwell, a major tributary of the river Mersey (Bracegirdle 1973; Holland & Harding 1984). Growth continued until the 1960s, particularly around the Mersey estuary and anecdotal evidence suggests that by the 1950s there were no fish in the river (Wilson *et al.* 1988; Jones 2000).

Water quality only began to improve in the 1970s, when a range of new legislation related to water quality was introduced. In 1983 a conference focusing on the Mersey was convened which led to the creation of the Mersey basin campaign (Jones 2000, 2006; Burton 2003). This heralded many changes that led to the Mersey becoming one of Britain's most high profile environmental success stories,

earning the inaugural International Thiess River prize in 1999 for best river system clean-up.

Although there is anecdotal evidence that salmonids began entering the Mersey estuary as early as the 1980s (Wilson *et al.* 1988), it is likely salmon began entering the river Mersey in the early 1990s as a result of the improving water quality (Jones 2000; Burton 2003). Video evidence of salmonids attempting to negotiate weirs on the river Bollin, a tributary of the Manchester Ship Canal, was taken in 1999 and 2000 (Jones, 2006; Environment Agency, unpublished data), and in 2001 the first salmon in several decades was caught by the Environment Agency (Jones, 2006).

Between 2001 and 2011, 158 untagged adult Atlantic salmon were caught within the river Mersey by the Environment Agency (England & Wales). A recent study found a proportion of tagged salmon (8/30) successfully ascended into the upper reaches of the river Mersey (Environment Agency 2012). During this period sampling effort and surveillance has been extensive and, although neither a 'run' of smolts to sea or a defined 'run' of returning adult salmon has been detected, three juveniles have been sampled. Therefore, we argue that although the Mersey is not yet a self-sustaining population, the river is in the early stage of an on-going process of natural recolonisation, following substantial improvements in overall river health. Assuming such improvements can be maintained, one can envisage that the recolonisation process could be actively encouraged once the source of recolonising adults has been identified. Moreover, if recolonising fish are shown to originate from similar (generally local) rivers, in which resident salmon are locally adapted, it seems probable that these fish may also exhibit some preadaptation to any proximal un-colonised river. Certainly, in studies of the Selja, Estonia (Vasemägi *et al.* 2001), and the Thames, UK (Griffiths *et al.* 2011), recolonisation appears to be predominantly by salmon from proximal rivers in the face of massive stocking with exogenous fish. Now that a comprehensive microsatellite baseline which includes fish from throughout their European range (Griffiths *et al.*, 2010; Gilbey *et al.,* unpublished) is available for Atlantic salmon, such identification is finally feasible.

The objective of this study was to identify the origin of adult and juvenile salmon sampled from the river Mersey between 2001 - 2011. To do this we

genotyped a sample of 149 Mersey salmon, three of which were juveniles, with a suite of 14 microsatellite loci used previously to assemble a baseline of genetic data from populations of salmon from across the southern part of their European range (Griffiths *et al.*, 2010). The Mersey genotypes were then assigned to a compiled baseline of probable source populations, which were taken from a previous study by Griffiths *et al.* (2010), and supplemented with additional populations from the SALSEA-Merge database (Gilbey *et al.*, unpublished).

Materials and Methods

Fish sampling

Ascending adult salmon were caught in a fish trap fitted to a Larinier fish pass built into Woolston weir on the river Mersey, 6.2 km upstream of the tidal limit. Fish were captured during August – October in the years 2001, 2002 and 2005 – 2010, with fishing effort being *ad hoc* over this period. The adult salmon were carefully removed from the trap, measured, weighed and scales removed for aging and genotyping. The total sample for genetic analysis was 149 Atlantic salmon (146 adults and three juveniles; (See Appendix I)

DNA extraction

Genomic DNA was extracted from individual scales using a chelex protocol (Estoup *et al.*, 1996). DNA from individual fish was genotyped using a panel of 14 apparently neutral loci: Ssa14 (McConnell *et al.* 1995); Ssa202, SSsp3016, Ssa197 (O'Reilly *et al.*, 1996); SsaF43 (Sánchez *et al.* 1996); SSspG7, SSsp1605, SSsp2210, SSsp2201 and SSsp2216 (Paterson *et al.* 2004); Ssa171, Ssa289, Ssa157 and SsaD144 (King *et al.* 2005). The loci were amplified within three multiplexed polymerase chain reactions (PCR), comprising: 1) SSspG7, Ssa14, Ssa202, SSsp3016; 2) Ssa197, SsaF43, SSsp1605, SSsp2210, SSsp2216; 3) SsaD157, Ssa171, Ssa289, SsaD144, SSsp2201. Loci were multiplexed on the basis of size using the Beckman Coulter three dye system (see Table 2.1 for dye details).

PCR reactions were carried out in 10μl reactions containing approximately 50ng of extracted Atlantic salmon template DNA, 3μl water, 5μl of Qiagen Taq PCR Mastermix and 1μl of primer mixture (details in Table 2.1). PCR conditions were as

follows: an initial denaturation step of 5 min at 95°C, followed by a touchdown PCR consisting of eight cycles with a 30s denaturation step at 95°C, a 90s annealing step starting at 62°C and decreasing the temperature 2°C every two steps until the touchdown temperature of 47°C was reached, with 3 minutes of extension at 72°C. The reaction ended with a final 10 minute extension at 72°C.

The size of the fluorescently labelled PCR products was determined using a Beckman-Coulter CEQ8000 automatic DNA sequencer and the associated fragment analysis software (Beckman Coulter). Data were checked for scoring errors due to stutter peaks, large allele dropout and null alleles using the program MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004).

Analysis

The genetic baseline used in this study represents a subset of the database developed by Griffiths *et al.* (2010), supplemented with genotypes from additional populations from rivers in Ireland, eastern Scotland and Norway from the SALSEA-Merge database (Gilbey *et al.* unpublished) to cover potential source rivers. Microsatellite genotypes acquired from the SALSEA Merge database required calibrating to match with Exeter genotypes, i.e. to correct for different scoring of alleles between laboratories. As part of the SALSEA database creation, a calibration study was completed, where the same samples were genotyped across 12 laboratories (Ellis *et al.* 2011). Using the results from the study by Ellis *et al.* (2011), the SALSEA data was transformed for the present study following specific rules for 11 loci (Table 2.2) The baseline comprised 5194 fish from 129 sampling sites within 60 rivers (Figure 2.2; Appendix II). All samples in the baseline were juveniles except those from all rivers in France, and the river Daleelva in Norway. It was the SALSEA Merge Consortium's decision to use juveniles in the baseline in order to reduce the risk that samples were strays from other rivers, which would reduce confidence in assignment. Assignment analyses were undertaken at the level of river and to reporting regions (see below). For assignment to river, where multiple samples were available from an individual baseline river, data from all sites were pooled prior to assignment. The effect of the adult samples from France and Norway on this baseline is expected to be small at river level assignment, because salmon are unlikely to stray far from their natal river. However it may affect confidence in

assignment between close rivers, for example between the Sée and Sélune. The effect is expected to be reduced at regional level as the majority of straying between rivers will be captured within the region.

In order to address the possibility that adult salmon sampled in the Mersey were salmon farm escapees, four populations from Norway were included in the baseline as surrogates for farmed fish. The vast majority of fish farmed in Britain are descended from Norwegian stock (Knox & Verspoor 1991), and recent research indicates a high degree of similarity between the genetic signatures of farmed fish and those of wild Norwegian salmon (Gilbey, pers. comm.).

Statistical treatment

FSTAT was used to calculate the number of alleles at each locus as well as each locus' allelic richness. Pair-wise F_{ST} values were calculated between rivers using the program FSTAT as previous studies have shown that for populations with very low F_{ST} (<0.1), assignment programs can be unreliable (Latch *et al.* 2006). Deviations from Hardy-Weinberg equilibrium were tested for using Arlequin v3.5 (Excoffier & Lischer 2010) and critical levels of significance were adjusted using the sequential Bonferroni procedure for multiple tests (Rice, 1989). To test the effectiveness of the baseline, the Leave-one out test, where each fish is systematically removed from its baseline population before having its own origins estimated using the rest of the baseline, was implemented in ONCOR (Kalinowski *et al.* 2007) and GeneClass 2 (Piry *et al.* 2004). Following these tests and the recommendations of Beacham *et al.* (2001), the rivers were grouped into broader, genetically based, reporting regions adapted from those proposed by Griffiths *et al.* (2010) for this part of the species' range.

Defining reporting regions

Reporting regions were created by pooling data from rivers based on their genetic similarity. Genetically similar groups were identified using the programs BAPS 5 (Corander *et al.* 2003) and STRUCTURE v. 2.3.3 (Pritchard *et al.* 2000). In BAPS this was done by using the 'clustering of groups of individuals' function and setting the maximum number of groups to 10, 20, 30, 50 and 60. STRUCTURE was run three times independently using the admixture ancestry model with 250,000

Markov Chain Monte Carlo (MCMC) replicates after a burn-in of 50,000 assuming 1-20 populations. STRUCTURE was run from $k = 1$ to 20 because due to the large sample size, running the program to $k = 60$ (as in BAPS) was computationally unfeasible. Following the results from BAPS, which was computationally much faster at determining the number of units, 20 was considered to be an appropriate maximum. The process was repeated ten times at different starting points along the MCMC chain. The most likely number of distinct genetic groups was inferred using the ΔK method of Evanno *et al.* (2005). However, because the Evanno method looks at the change in the likelihood score between runs, it is unable to include $k = 1$ in the comparison. As a solution to this, the Likelihood method, recommended by Pritchard *et al.* (2000), was also used to check the likelihood of one genetic unit ($k = 1$) and ensure it was not more parsimonious than k > 1. The reporting regions were then also tested for effectiveness for assignment by using the Leave-one out tests in ONCOR and GeneClass 2. Because there was uncertainty in STRUCTURE and BAPS regarding the placement of samples from the Southern Irish rivers - the Barrow, Boyne and Suir-, salmon from these three rivers were removed from the final assignment of Mersey samples.

Assignment

Genetic stock assignment of the Mersey salmon to the designated reporting regions was carried out using the programs ONCOR, which uses a maximum likelihood approach to assignment, and GeneClass 2, which uses a Bayesian approach. These methods have proven to be significantly more effective at assignment than previous distance-based methods (Cornuet *et al.* 1999). ONCOR was run under standard conditions and GeneClass 2 was run using the Rannala and Mountain (1997) algorithm.

A recognized flaw of assignment methods is the assumption that the source population is included within the baseline (Cornuet *et al.* 1999). In order to test this assumption, the exclusion method of assignment was performed according to (Vasemägi *et al.* 2001).

Testing assignment with samples from southern Ireland

In order to investigate the effect of removing salmon from southern Ireland, further analysis was completed. Firstly, in order to determine how best to group these salmon in a Reporting Region, a Leave-one out test was completed in ONCOR under two different scenarios. The first scenario placed the salmon from southern Ireland as a separate reporting region from any of the others, as might be expected from their geographical position. The second scenario placed the salmon in the same reporting region as salmon in Scotland, as was suggested by BAPS.

Secondly, following the results of the Leave-one out tests, assignment of Mersey salmon to the reporting region was repeated, whilst including the southern Irish populations. Following the results of BAPS, salmon from the three southern Irish rivers were combined with the salmon from Scotland. Under this scenario, assignment was completed with ONCOR and GeneClass 2 and the results were compared to the assignment without the southern Irish populations.

Results

Of the 149 Mersey salmon sampled, 134 adults and one juvenile were successfully amplified at 10 or more loci out of 14; unfortunately, due to the condition of the very limited amount of scale material collected, amplification was not successful from two of the three juveniles sampled. MICRO-CHECKER found no evidence of scoring errors due to stutter peaks or allele dropout. Evidence of null alleles was found at some loci. Of the 45 significant results, 10 were associated with locus SSspG7 and 8 with Ssa197. Previous work by Griffiths *et al.* (2010) showed the removal of loci with null alleles to be slightly detrimental towards the process of assignment. The issue has also been addressed by Carlsson (2008), who, from simulations, concluded that although null alleles can cause a slight overestimation of F_{ST} and a slight reduction in assignment power, their inclusion is not likely to alter the outcome of assignment; therefore, these loci were not removed from the analyses.

Genetic diversity within the baseline

The total number of alleles per locus ranged from eight in Ssa14 to 43 in SsaD157 and SsaD144 and allelic richness ranged from 2.27 in Ssa14 to 9.45 in

SsaD144 (Table 2.3). Heterozygosity was generally high but ranged from 0.934 in SSsp2201 to 0.366 in Ssa14.

F_{st} and Hardy Weinberg

The average inter-river F_{ST} for all rivers included in the baseline was 0.036 (Appendix III), which was less than the 0.05 recommended by Latch *et al.* (2006) for 97% accuracy of assignment. This was reduced to 0.0298 when looking within the UK alone, 0.027 after excluding populations from Ireland, and 0.019 after excluding the populations from Southern England. This confirmed the need to use reporting regions rather than individual rivers for subsequent assignment analysis. 27 alleles (1%) were found to be out of Hardy Weinberg equilibrium after Bonferroni correction (Appendix IV). As no allele or population was found to be consistently out of HW, no data was excluded due to this test.

Population structuring

There were several results from the STRUCTURE analysis. Firstly, the likelihood method - which was used to determine whether $k = 1$ was more parsimonious than dividing the baseline into two or more groups - indicated that $k = 1$ had a lower score than $k = 2$-20 (Figure 2.3a). Therefore it was appropriate to separate the baseline into several groups. While this method indicated that $k = 8$ had the highest score (followed by $k = 14$, and then $k = 9$), this method was not considered to be appropriate for determining the optimum number of genetic units (Evanno *et al.* 2005). Therefore, these results was not considered when determining the final number of reporting regions. The ΔK method - considered to be more reliable for determining the number of genetic units (Evanno *et al.* 2005) - identified the optimum number of genetic units from the STRUCTURE analyses to be $k = 6$, followed by $k = 7$ (Figure 2.3b).

The clustering of rivers function within BAPS identified seven groups (Figure 2.4). These seven groups agreed strongly with the seven groups identified in the STRUCTURE $k = 7$ run (Figure 2.4b, Figure 2.6), except for the following exception: BAPS placed the southern Irish rivers, the Barrow, Boyne and Suir, together with the rivers from Scotland, whereas STRUCTURE identified that each river contained a mosaic of genetic signatures matching those from Northern Ireland, Scotland and from around the Solway Firth (Figure 2.6).

Based on all the results, described henceforth, a decision was made to divide the baseline into seven groups (Figure 2.6). Firstly, the results from BAPS indicated seven clear groups. Secondly, there was general consistency between the results from the STRUCTURE $k = 7$ and results from BAPS. Thirdly, the ΔK method (Evanno et $al.$ 2005) also supported $k = 7$ after $k = 6$, as $k = 7$ was the next highest point (Figure 2.3b). Finally, although the Evanno method identified $k = 6$ as highest scoring, from the STRUCTURE plot, the same seven groups were identifiable in $k = 6$ and $k = 7$ (Figure 2.4a and 2.4b). Although the results indicated that salmon from southwest England and France were effectively in the same group (Figure 2.4a), their genetic makeup was visually very different.

However, due to the uncertainty regarding the position of genotypes from the Barrow, Boyne and Suir, salmon from these rivers were removed from the baseline used in the assignment analyses. This led to there being seven genetically based reporting regions, which were named as follows: Scotland, Solway & Northwest England, Southwest England & Wales, Southern England, Northern Ireland, France and Norway (a surrogate for Scottish farmed fish) (Figure 2.2 and 2.6).

Baseline test

The Leave-one out test found 46.5% in ONCOR (Table 2.4) and 47.5% in GeneClass 2 (results not shown due to the large size of the table) of fish correctly assigned back to the river from which they were sampled. After the formation of reporting regions, which excluded the fish from southern Ireland, the proportion of correctly self assigned individuals increased to 83% in GeneClass 2 and 84% in ONCOR (Table 2.5).

Assignment results

Exclusion analysis found that for 21 of the 135 salmon sampled from the Mersey, the probability of their assigning to any of the recognized reporting regions was less than 0.05 (Appendix IV). Therefore, the results of assignment analysis for these individuals are not considered further, but can be found in Appendix I.

Genetic assignment showed the remaining salmon from the Mersey to have a variety of different origins (Table 2.6). Both GeneClass 2 and ONCOR found the largest proportion of the Mersey salmon to be from the reporting region defined as

Solway & northwest England (44% and 59%, respectively). Both methods also found the next biggest contributing regions to be Scotland, followed by Wales & southwest England. Two fish were assigned to France by ONCOR, while the same two fish and one other were assigned to France by GeneClass 2. Three fish were assigned to Northern Ireland in both GeneClass 2 and ONCOR, and one other was also assigned to Northern Ireland in GeneClass 2. Four salmon were assigned to Norway in GeneClass 2, but none were assigned to Norway in ONCOR. The single juvenile that was sufficiently well genotyped to allow meaningful assignment was assigned to Solway & northwest England by both programs (see Appendix I for likelihood scores for assignment to each reporting region).

Effect of southern Ireland samples

There was a slight reduction in the average self-assignment results following the inclusion of samples from southern Ireland (Table 2.7). When the Scottish samples were grouped to form their own unique reporting region, the average self-assignment score (assessed only in ONCOR) fell to 78%, from the previous 84%. Crucially, the score for the Southern Ireland group was the lowest (58%), which was significantly lower than the next lowest, Scotland, which was now 67% and previously 72%. When the southern Irish samples were grouped together with salmon from Scotland, the average score fell approximately one percent, to 83%. However there was a consistent pattern of decreased confidence in each of the reporting regions.

Following the Leave-one out tests, the assignment with salmon from southern Ireland was completed with these salmon combined with the Scotland group to form a new group, "Scotland & Southern Ireland". In GeneClass 2, five of the 134 fish were assigned to different groups, compared to when southern Ireland was excluded (Table 2.8). Five fish that previously assigned to Solway & Northwest England now assigned to two different groups (Table 2.8). Three of these fish assigned to the new Scotland & Southern Ireland group, whilst the remaining two assigned to Northern Ireland.   In ONCOR, the inclusion of fish from southern Ireland changed the assignment of thirteen fish. One fish that previously assigned to Scotland was assigned to Solway & Northwest England. One fish that previously assigned to Northern Ireland was assigned to Scotland & Southern Ireland (Table 2.8). Of three

fish that previously assigned to southwest England & Wales, two assigned to Scotland & Southern Ireland and the last assigned to France. Finally, eight salmon that previously assigned to Solway & Northwest England now assigned to Scotland & Southern Ireland.

Discussion

Findings

This study aims to identify the origins of Atlantic salmon recolonising the river Mersey and in doing so, reveals some current limitations for genetic assignment within this region. Although most of the salmon now entering the Mersey could not be assigned to an exact river of origin, by identifying distinct genetic signatures of groups of salmon rivers, we are able to identify their region of origin with a high degree of probability. The reporting regions identified here match those identified by Griffiths *et al.* (2010), and they appear to be useful units for assignment, according to the results of the self-assignment test. For some reason, possibly an unidentified quantity of salmon translocation, the southern Irish rivers used in this study contain genotypes that fail to stand alone as a distinct reporting region. For this reason, these rivers were removed from the assignment analysis.

The genetic baseline used for assignment of Mersey fish was a subset of the populations used in the ASAP (Griffiths et al. 2010) and SALSEA (SALSEA consortium, unpublished) projects. Such a baseline was anticipated to provide comprehensive coverage of potential rivers of origin for those salmon now entering the Mersey. Nonetheless, even with such detailed coverage, the possibility remained that some fish might not assign to a population or region within the baseline. Accordingly, to address this possibility, we undertook exclusion analysis. This analysis found that 21 of the 135 salmon characterized did not assign to any of the reporting regions in our baseline; this may be because these fish really do originate from a population outside the area covered by our baseline, or may indicate that their genetic signatures are too general to assign to any reporting region with a sufficiently high score (above 0.05). This left 113 adults and one juvenile for assignment analysis, which identified multiple origins for salmon currently entering the river Mersey (Table 2.6). This finding is not unusual as previous studies also show

recolonisation from multiple source rivers (e.g., the River Seine, Perrier *et al.* 2010). Indeed, this should be beneficial for the long-term survival of any newly established population, as the potentially increased genetic variability should provide a broader basis for adaptation to local and possibly changing conditions

The Mersey is found to be on the border between two of the designated reporting regions. The majority of salmon in the Mersey clearly originate from rivers north of this border and, in particular, the Solway & Northwest England reporting region. Although this finding is not on its own surprising (the southernmost river of this reporting region being the Ribble, the mouth of which is approximately 40 km north of the Mersey), it was striking that so few (15/113 ONCOR; 18/113 GeneClass 2) appeared to have origins in the neighbouring Southwest England & Wales region (a trend reflected in assignment to river; Appendix I). In particular, this reporting region contains the river Dee, a highly productive salmon river that enters the sea in close proximity to the Mersey; the estuaries of these two major rivers are separated by the 11 km-wide Wirral peninsula. This finding may be due to the prevailing clockwise gyre in the eastern Irish Sea and an associated current, which for much of the year runs southwards down the northwest coast of England (Heaps & Jones 1977). Presumably, it is this current which carries some homing adult salmon past their natal rivers and southwards towards the Mersey, whilst simultaneously acting to move fish from the rivers of north Wales away from the Mersey.

This study finds evidence that, despite their well-known homing capabilities (Stabell 1984), Atlantic salmon can stray into distant rivers. Three fish were assigned to France by both programs. Previous work has shown that long distance colonisation does occur; for example, a study of recolonisers in the Séine (Perrier *et al.*, 2010) showed two out of seven fish assigned to a foreign baseline group better than any of the five French regions included in their analysis. A study by Griffiths *et al.* (2011), which found one of sixteen salmon sampled from the Thames to be from a French population, again demonstrates that salmon may stray relatively long distances to rivers in England.

An important caveat is that, despite evidence confirming that some of the stray adults caught in the weir do ascend into the Mersey's upper reaches (Environment Agency, 2012), within the limits of this study we cannot determine

which of the 135 genotyped adults would have ascended the river further and which would have left the catchment. However, the one juvenile, for which there was enough material to amplify the DNA reliably, assigned to the Solway & Northwest England region. We refrain from making major conclusions based on a single individual; however, the importance of this juvenile should not be overlooked. This result suggests that not only is the Solway & Northwest England the biggest source of strays, but also (because of their larger numerical contribution and their pre-adaptation to similar in-river conditions in their proximal rivers of origin) that salmon from this region are the most likely to successfully reproduce in the river Mersey at this time.

Farmed salmon

Four salmon populations from Norway were included in the baseline to represent the genetic signature of farmed fish of Norwegian-origin, which we considered might be a possible source of adult fish entering the Mersey. However, the results for this component of the analysis were inconclusive; four of the 135 Mersey fish assigned to Norway with GeneClass 2, while none were assigned to Norway with ONCOR. This discrepancy may indicate that the actual source population of these fish is not present within the baseline, as previous studies have concluded (e.g. Perrier *et al.*, 2010). However, additional evidence indicating a Norwegian genetic signature in the possible sources of Mersey fish comes from the STRUCTURE analysis (Figure 2.6). Some Scottish rivers such as the Clyde and Luce show clear evidence of resident salmon parr with Norwegian genetic signatures. These 'Norwegian' fish may be descendants of fish farm escapees but it is also possible that this reflects a shared common ancestry of northern salmon populations. Whatever their origins, one of our methodologies indicates that fish with at least a partial Norwegian signature are entering the Mersey. At this time, however, discrepancies in our assignment prevent us from making a firm conclusion, but improving the baseline, with the addition of hatchery stock, may resolve this issue.

Difficulty of assignment

To date, no study has made use of such an extensive baseline for the purpose of identifying the origin of unknown Atlantic salmon. While the epitome of genetic stock identification applications would be to identify any salmon to its river or

possibly tributary of origin, for this part of the species' range at least, that is beyond current means. The results of the Leave-one out test showed that less than one in two fish can be correctly assigned back to their river of origin; unfortunately, such a figure is insufficient for meaningful assignment. This was somewhat to be expected as previous research by Griffiths *et al.* (2010) also found lower accuracy of assignment in this region (Ireland, the west coast of Scotland, Northwest England and Wales), compared to that obtained when assigning to more southerly salmon populations. Inter-river F_{ST} values of 0.02 within each of the designated UK reporting regions analysed in this study, and many pair-wise inter-river F_{ST} values of less than 0.01 underline the inability to assign to individual rivers within this area; these values are far below the 0.05 suggested for 97% assignment accuracy (Latch *et al.*, 2006). Key to improving the accuracy of genetic assignment is improving the genetic distinction between populations within the baseline. One way of doing that is by increasing the number of markers used, either via the addition of more microsatellites, or with the use of Single Nucleotide Polymorphisms (SNPs) (Beacham *et al.* 2011). Currently however, the utility of SNPs for assignment purposes remains a topic of considerable discussion (e.g. Morin *et al.* 2004; Beacham *et al.* 2011). Another key approach is to reduce the sampling error, i.e. the difference between the estimated allele frequencies and the allele frequencies in the actual population (Beacham *et al.*, 2011). This would be achieved by increasing the sampling size of the baseline populations, which although not ideal, would be less effort than the cross calibration required with the addition of an extra microsatellite (e.g. Ellis *et al.*, 2011).

Samples from southern Ireland

The samples from southern Ireland proved difficult to place into a reporting region, and were consequently removed from the final assignment analysis (Ikediashi *et al.* 2012). However, to ensure that removing these samples was not leading to incorrect results, further analyses were completed with these samples included. First the results from the Leave-one out test indicated that if salmon from southern Ireland were to be included, it was better to include them together with Scotland, supporting the findings from the genetic analysis (i.e. BAPS) rather than their geographic placement (Table 2.7). Overall self assignment fell under both scenarios, but only 1% when these salmon were joined with Scotland, instead of 6%

when they were separate. This was largely driven by the low self assignment of salmon to Southern Ireland as a unique group (58%).

During assignment, when the southern Irish populations were placed in a group with Scotland as the genetic analyses indicated, the results did change slightly (Table 2.8). The two assignment methods were affected differently by the new scenario, but some changes were consistent. Most noticeably, in both programs, a significant number of salmon (3 in GeneClass and 8 in ONCOR) that previously assigned to Solway & Northwest England were assigned to Scotland & Southern Ireland instead. This highlights the difficulty of genetic assignment with this data, and the close similarity between salmon in Scotland and those in Solway & Northwest England. Including samples from southern Ireland, however, did not change the overall findings. The majority of salmon still assigned to Solway & Northwest England; where it was previously 44% and 59%, dependent on method, it became 40% and 55%. Besides greater numbers of salmon assigning to Scotland - now including salmon from southern Ireland - there were no further significant changes. This, we argue, indicates that the exclusion of southern Irish fish from the final analysis has not significantly affected the results.

It is worth discussing the similarities between salmon in Scotland and Ireland. Other studies have also identified several close genetic relationships between these countries (Gilbey & Coughlan unpublished), although at this stage there are no tested hypotheses. It is possible that the similarity is due to an historic common ancestor. For example, after the last glacial period, which terminated ca. 10,000 years ago (Finnegan *et al.* 2013), one population may have populated both regions. Alternatively, it is also possible that the similarities are due to recent movement of salmon from Scotland to Ireland. This could be due to artificial stocking, which was previously common practice across Europe (Martinez *et al.* 2001; Griffiths *et al.* 2010; Perrier *et al.* 2013). Or it could be due to natural migration between salmon in these regions, however this is perhaps the least likely possibility because of the large distance between these rivers - at least 200 km of the Irish Sea.

River restoration as a fisheries management tool

Overall, this study and others like it (Knutsen *et al.* 2001; Schreiber & Diefenbach 2005; Anderson & Quinn 2007; Kiffney *et al.* 2009; Perrier *et al.* 2010;

Griffiths *et al.* 2011) serve to underline the value of river restoration as an effective alternative to stocking to promote the recolonisation of rivers from which salmonids have been previously extirpated. Additionally, such an approach is likely to yield broader ecological benefits for a river ecosystem as a whole. For example, improvements in water quality have been shown to promote increased biodiversity of riverine invertebrate fauna (Chadwick & Canton 1986) and the return of larger animals, e.g. otters, in part due to improved water quality and partly due to increased availability of fish as food (Pountney *et al.* 2009; Crawford 2010).

Alternatively, in situations where the need for fish population restoration is urgent – for example, post-pollution mitigation – then assignment studies such as this offer, in combination with river restoration, robust insights as to which populations might best serve as donors for translocation, and thus more rapid recolonisation.

Conclusion

In conclusion, this study overcomes limitations in genetic assignment in order to ascertain the origins of Atlantic salmon recolonising the river Mersey. They appear to be from multiple regions primarily within England, Scotland and Wales, and in particular from the rivers in close proximity to the Solway Firth and the northwest of England. This key finding highlights an apparent clockwise direction of straying by Atlantic salmon in this region, which we speculate to be due to the clockwise gyre in the eastern Irish Sea. The one successfully analysed juvenile assigned consistently to this same region, which may indicate that not only is this region responsible for the greatest number of strays, but that these strays are also the most likely to successfully reproduce in this river. This study also finds that a small fraction of the recolonisers are from Northern Ireland, while a similar proportion appear to originate from France. The evidence suggests that salmon farm escapees, with a distinct Norwegian signature may be a fraction of the recolonisers, however, incongruence between the methods used prevented firm conclusions on this topic. While the information gained from this study increases our scientific understanding of the salmon life cycle, our findings are also especially useful for river management, as they demonstrate clearly the benefits of river restoration as a *bona fide* methodology for the re-establishment of salmonid populations in rivers from which they have been

previously extirpated; our results also serve to reconfirm the capacity for straying in this species otherwise famous for its homing ability.

Data archiving statement

Data for this study are available at: http://datadryad.org/ DOI: 10.5061/dryad.ck461

Figure 2.1- Sampling adult Atlantic salmon from the river Mersey. Courtesy of Sam Billington.
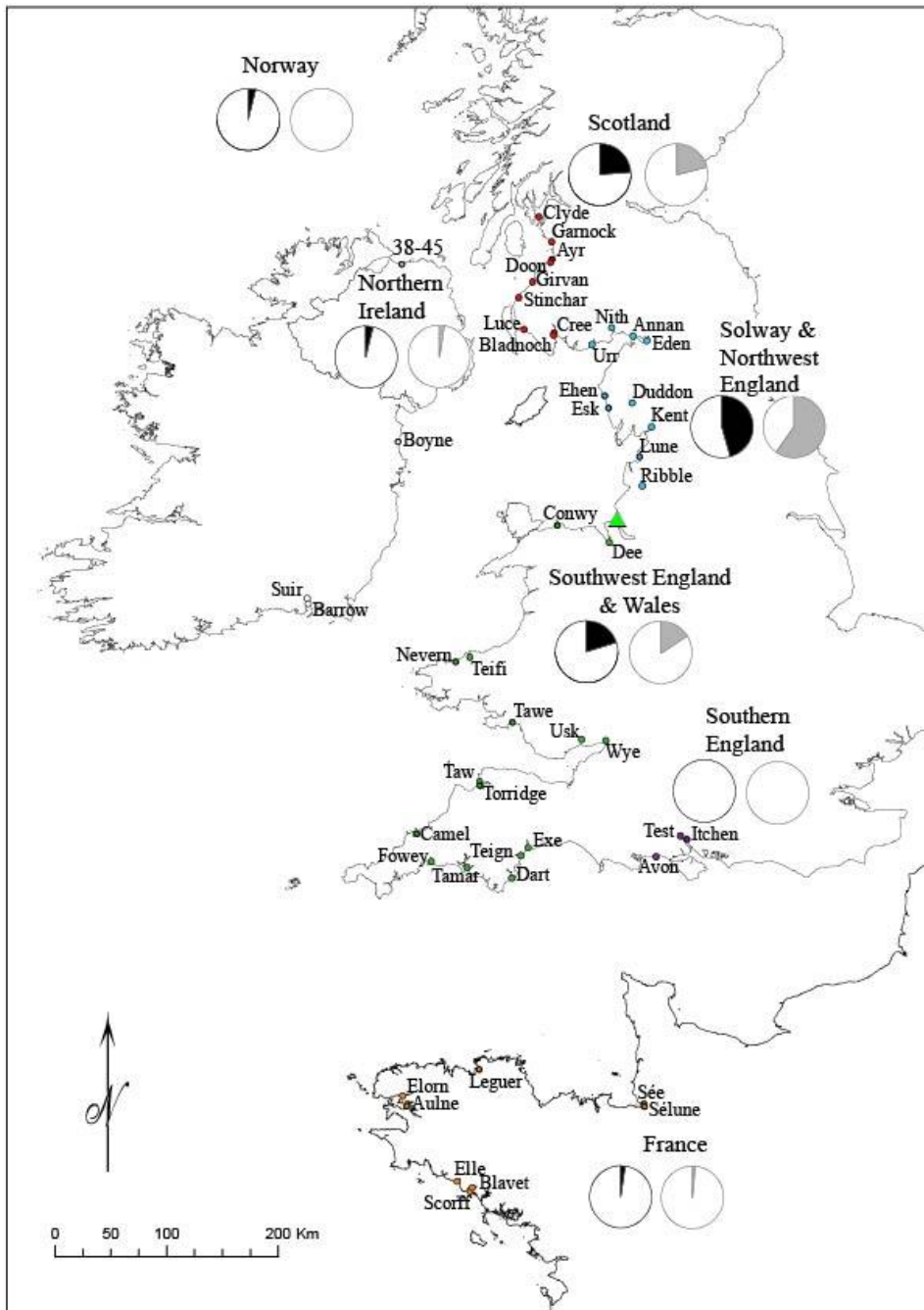
Figure 2.2- Map of genetic assignment to reporting regions. Points show the mouth locations of all rivers included within the baseline, excluding those in Norway. Rivers colour coded to show the designated reporting regions: Scotland: Red; Solway & Northwest England: Blue; Southwest England & Wales: Green; Southern England: Purple; France: Orange; Northern Ireland: Pink (N.b. Northern Ireland rivers enter Lough Neigh and share a common estuary – 38: Upper Bann; 39: Agivey; 40: Blackwater; 41: Clogh; 42: Grillagh; 43: Kells Water; 44: Moyola; 45: Six Mile). Pie charts show the proportion of Mersey samples assigned to each reporting region in GeneClass 2 (left) and ONCOR (right). The green triangle indicates mouth of the River Mersey.
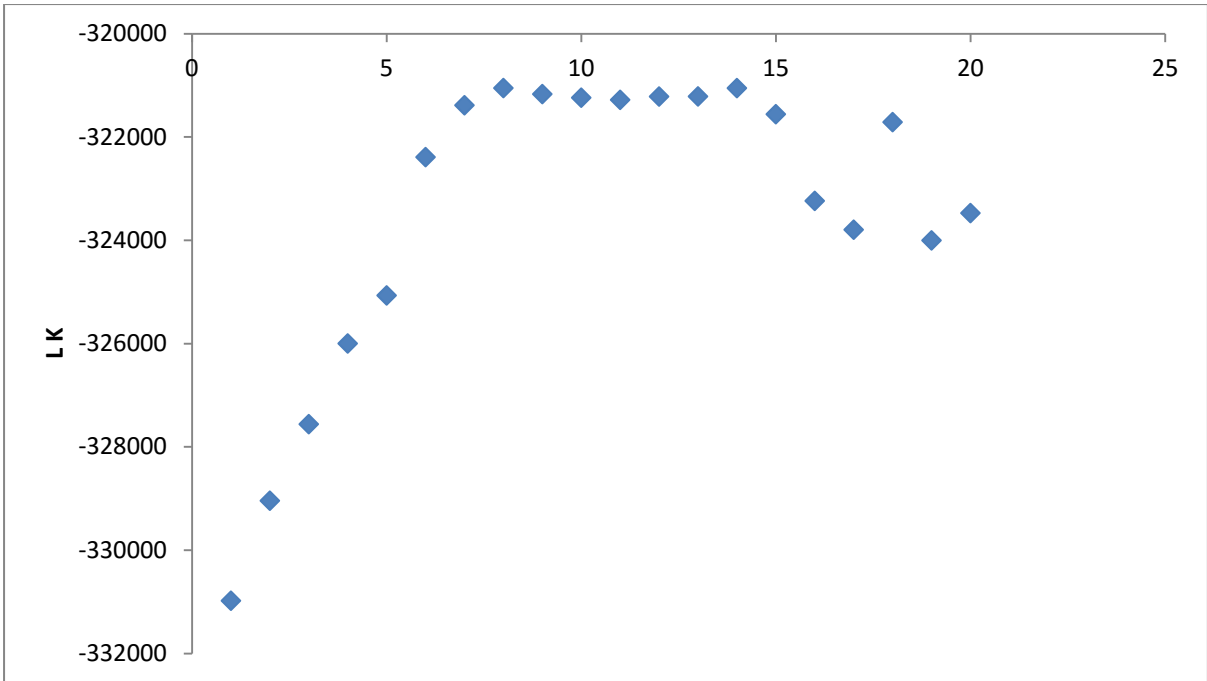
Figure 2.3a- STRUCTURE L K plot. The plot indicates the optimum number of genetic units calculated within the data in the program STRUCTURE using the Likelihood method.
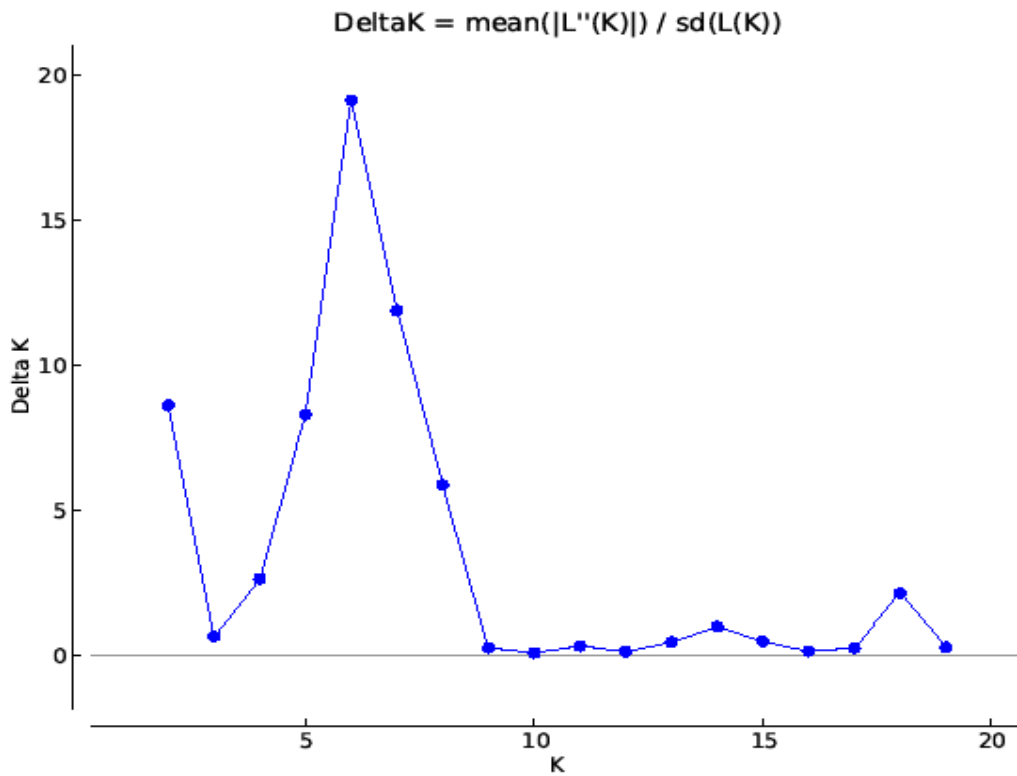
Figure 2.3b- STRUCTURE delta K plot. The plot indicates the optimum number of genetic units calculated within the data in the program STRUCTURE using the delta K method.
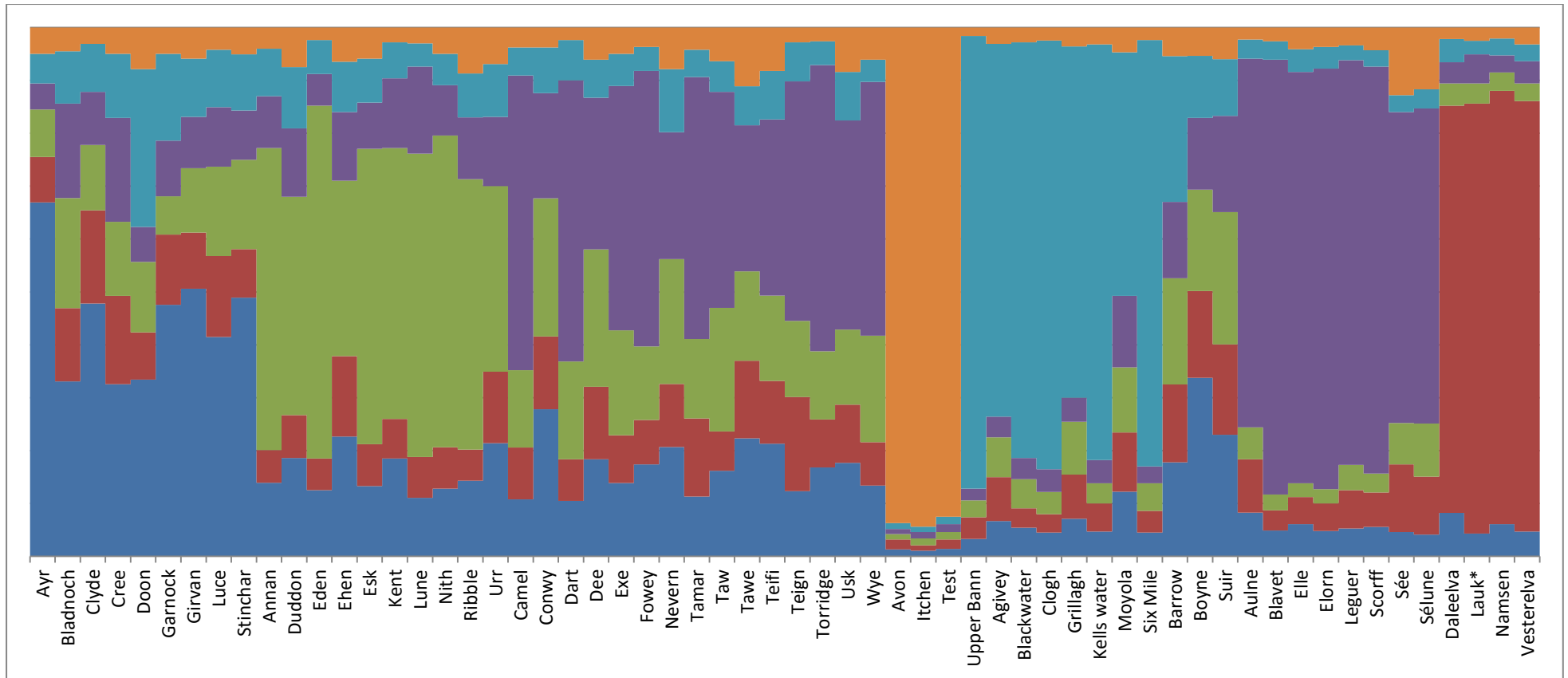
Figure 2.4a- Population STRUCTURE plot of salmon baseline for $k = 6$. Estimated proportions of the coefficient of admixture of each population's genome that originated from population k for $k = 6$. Each population is represented by a column. Thin black bars separate individual rivers, for which names are below the graphic.
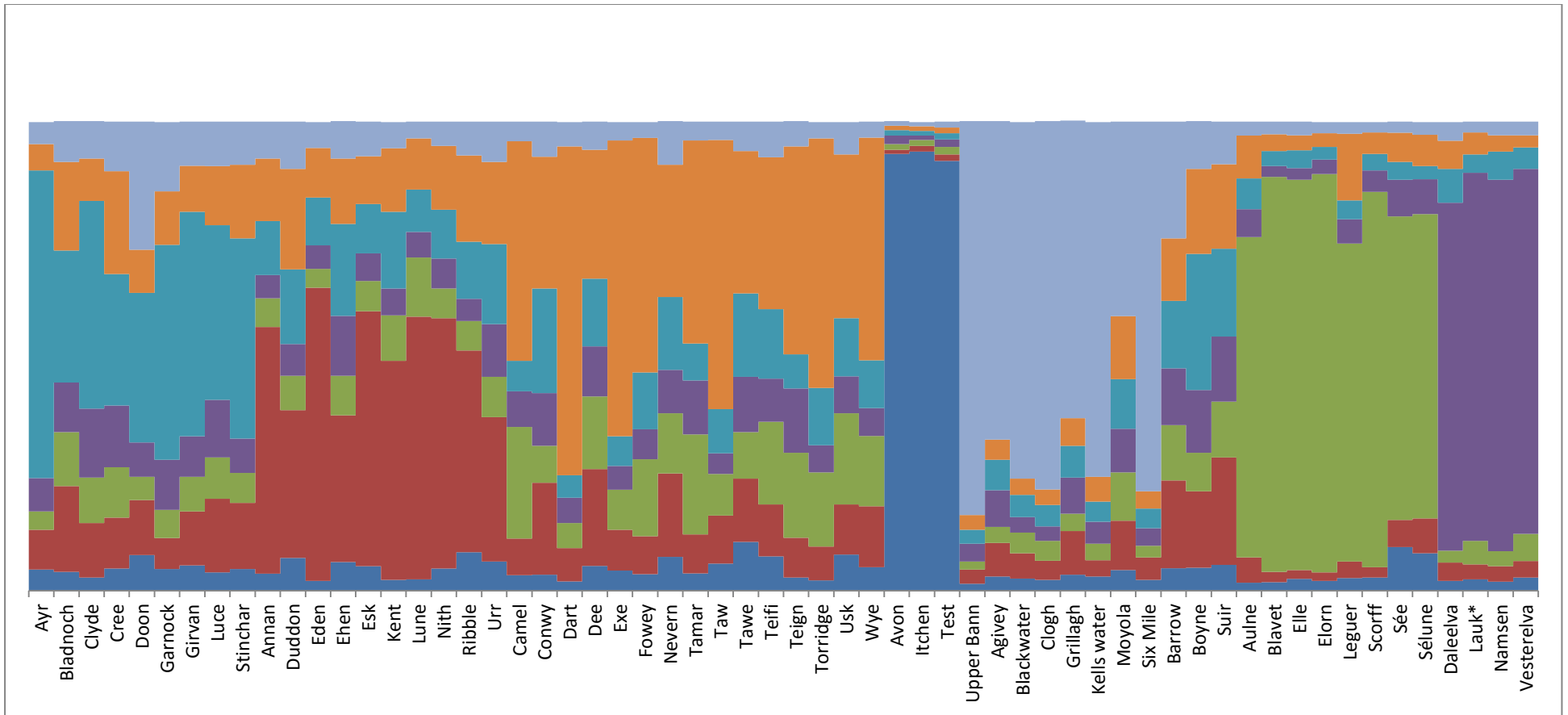
Figure 2.4b- Population STRUCTURE plot of salmon baseline for *k* = 7. Estimated proportions of the coefficient of admixture of each population's genome that originated from population *k* for *k* = 7. Each population is represented by a column. Thin black bars separate individual rivers, for which names are below the graphic.
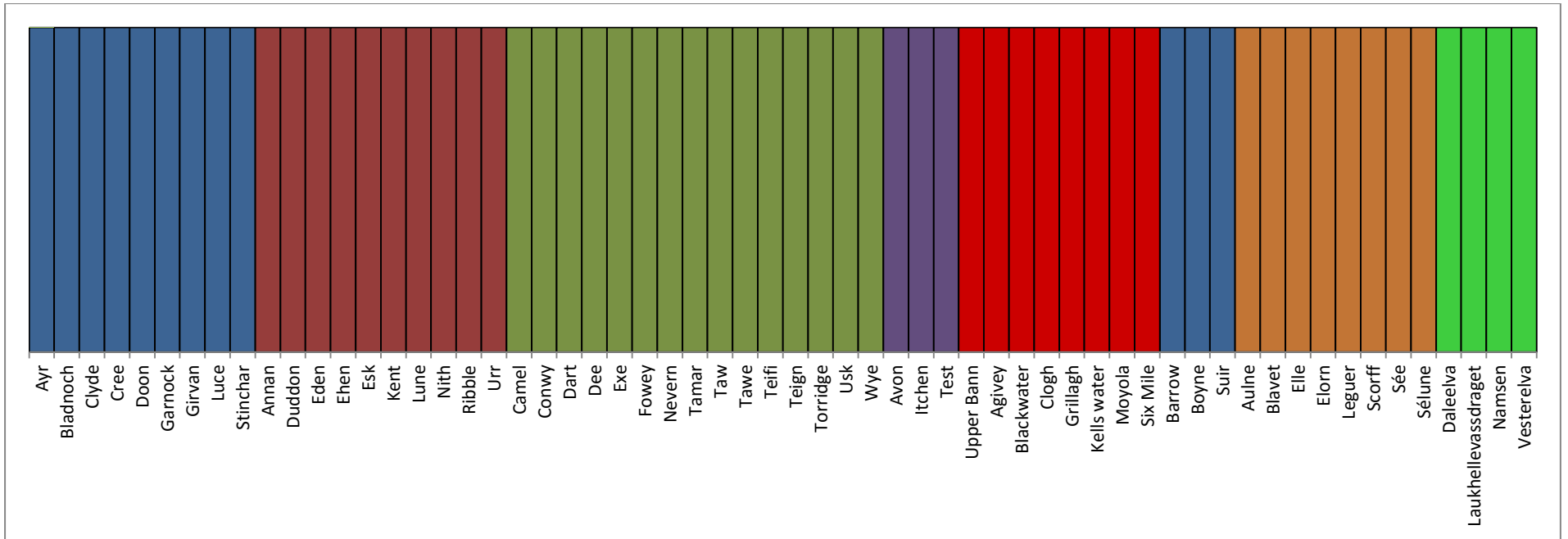
Figure 2.5- BAPS plot of salmon baseline. Estimated proportions of the coefficient of admixture of each populations genome that originated from population *k* for *k* = 7. Each population is represented by a column. Thin black bars separate individual rivers, for which names are below the graphic.
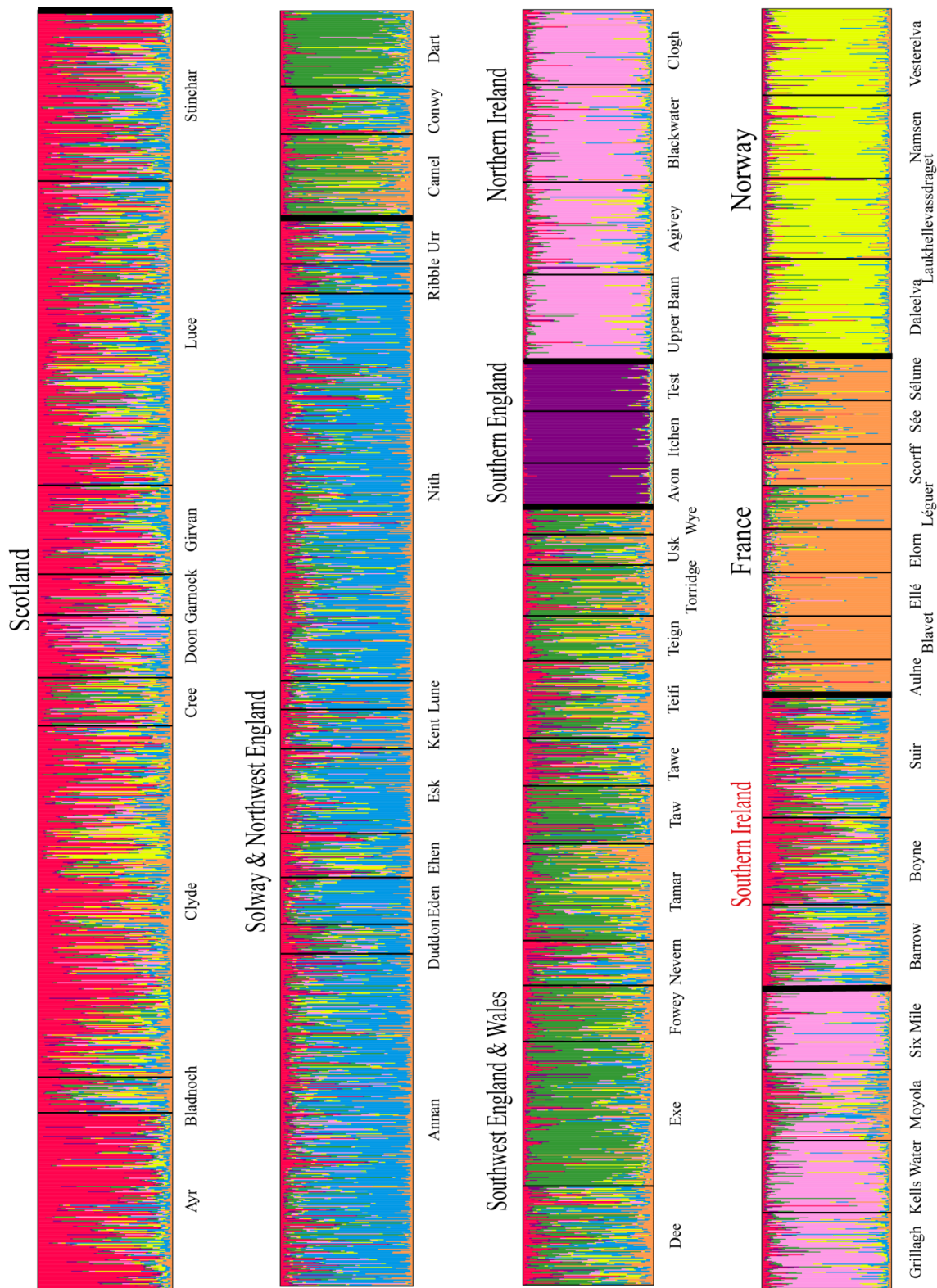
Figure 2.6- STRUCTURE plot of salmon baseline. Estimated proportions of the coefficient of admixture of each individual's genome that originated from population *k* for *k* = 7. Each individual is represented by a column. Thin black bars separate individual rivers, for which names are below the graphic. Thick black bars separate reporting regions, for which the names are above the graphic. The rivers from southern Ireland are in red because they are removed from assignment analysis. *

Table 2.1- Volume (µl) of primer within each 100 µl primer mixture. Primers added at 100 µM. Colours indicate the colour of the dye.

Multiplex A				Multiplex B		Multiplex C			
PCR I (55)		*PCRII (55)*		**PCR I**		**PCR I**		**PCR II**	
Ssosl417	1.6	Ssosl85	5	SSsp2216	4	SSaD144	5	Ssa157	8
Ssa202	4	Water	90	SsaF43	1.5	Water	90	Ssa171	3
Ssa14	4.5			SSsp2210	2.2			SSsp2201	3
SSsp3016	10			Ssa197	4			Ssa289	11
SSspG7	2.5			SSsp1605	4			Water	50
Water	54.8			Water	68.6				

Table 2.2- Table depicting the adjustment required to convert microsatellite data from the SALSEA Merge database to match the Exeter format.

Microsatellite marker	To transform Exeter data to SALSEA baseline	To transform SALSEA baseline to Exeter baseline
SSspG7	+2	-2
Ssosl417	-2	+2
Ssa202	-2	+2
Ssa197	+5	-5
SSsp2210	+2	-2
Ssa289	+5	-5
SSsp2201	+2	-2
Ssol85	-4	+4
Ssa14	-1	+1
SsaF43	+4	-4
SSsa2216	-1	+1

Table 2.3- Total number of alleles and allelic richness per locus for all baseline populations.

Locus	No. of alleles	Allelic richness	Observed heterozygosity
SSspG7	26	7.61	0.834
Ssa14	8	2.27	0.366
Ssa202	24	6.98	0.855
SSsp3016	20	7.32	0.861
Ssa197	33	8.37	0.871
SsaF43	13	4.05	0.666
SSsp1605	15	5.66	0.786
SSsp2210	18	5.98	0.755
SSsp2216	21	7.52	0.883
SsaD157	42	9.04	0.924
Ssa171	37	7.52	0.871
Ssa289	12	3.69	0.632
SsaD144	43	9.45	0.930
SSsp2201	37	9.4	0.934

Table 2.4- River baseline self-assignment scores from ONCOR. The percentage of individuals within the baseline that correctly assigned back to their own river and the river that contained the highest proportion of wrongly assigned individuals.

River	Correctly self assigned	Largest incorrect assignment	
Ayr	58.80%	Stinchar	8.10%
Bladnoch	23.50%	Cree	14.70%
Clyde	54.70%	Ayr	4.50%
Cree	50.00%	Bladnoch	50.00%
Doon	60.30%	Clyde	8.60%
Garnock	13.50%	Ayr	18.90%
Girvan	27.60%	Garnock	13.80%
Luce	52.90%	Stinchar	4.40%
Stinchar	35.30%	Girvan	8.00%
Annan	29.70%	Nith	18.40%
Duddon	36.70%	Nith	13.30%
Eden	43.80%	Annan	16.70%
Ehen	20.00%	Nith	11.10%
Esk	34.90%	Nith	16.30%
Kent	34.10%	Annan	19.50%
Lune	31.00%	Esk	13.80%
Nith	30.00%	Annan	17.10%
Ribble	41.90%	Annan	9.70%
Urr	30.00%	Clyde	10.00%
Camel	44.30%	Fowey	7.60%

Conwy	19.60%	Suir	15.20%
Dart	71.60%	Camel	5.40%
Dee	40.40%	Luce	4.50%
Exe	62.90%	Camel	7.60%
Fowey	33.30%	Teifi	11.10%
Nevern	35.90%	Dee	12.80%
Tamar	49.40%	Teifi	9.00%
Taw	32.70%	Exe	11.50%
Tawe	54.80%	Ehen	6.50%
Teifi	22.70%	Tamar	10.60%
Teign	26.30%	Fowey	10.50%
Torridge	44.90%	Nith	8.20%
Usk	16.70%	Wye	13.30%
Wye	34.60%	Dee	11.50%
Avon	81.40%	Itchen	11.60%
Itchen	76.00%	Avon	12.00%
Test	81.80%	Itchen	9.10%
Upper Bann	89.40%	Ayr	1.20%
Agivey	58.20%	Grillagh	17.60%
Blac	67.70%	Clogh	5.20%
Clogh	56.20%	Six Mile	15.10%
Grillagh	57.30%	Agivey	7.30%
Kells water	67.90%	Clogh	5.10%
Moyola	60.70%	Six Mile	6.60%
Six Mile	60.20%	Clogh	9.10%
Barrow	51.20%	Annan	6.00%
Suir	44.70%	Annan	3.90%
Aulne	36.10%	Leguer	8.30%
Blavet	41.50%	Elorn	17.10%
Elle	23.40%	Elorn	27.70%
Elorn	39.10%	Scorff	15.20%
Leguer	34.80%	Aulne	8.70%
Scorff	20.90%	Elle	23.30%
Sée	52.40%	Sélune	21.40%
Sélune	36.20%	Sée	27.70%
Daleelva	60.80%	Namsen	14.40%
Laukhellevassdraget	67.50%	Daleelva	13.30%
Namsen	70.60%	Daleelva	10.30%
Vesterelva	88.80%	Clyde	2.20%

Table 2.5- Reporting region self-assignment scores. The percentage of individuals within the baseline that correctly assign back to their own reporting region and the reporting region that contained the highest proportion of wrongly assigned individuals. *GeneClass 2 assigned one individual to Solway, southwest England & Wales, Northern Ireland and France.

Reporting Region	Correctly self assigned		Largest incorrect assignment		
	GeneClass 2	ONCOR		GeneClass 2	ONCOR
Scotland	70.0%	72.0%	Solway & Northwest England	10.7%	10.3%
Solway & Northwest England	76.6%	76.1%	Southwest England & Wales	9.3%	9.3%
Southwest England & Wales	70.8%	71.8%	Solway & Northwest England	10.6%	10.5%
Southern England	97.2%	97.1%	Solway & Northwest England *	0.7%	0.7%
Northern Ireland	89.2%	89.0%	Scotland	4.1%	4.1%
France	89.1%	88.8%	Southwest England & Wales	5.5%	5.5%
Norway	90.4%	90.2%	Scotland	4.3%	4.2%

Table 2.6- Results of assignment of adult Mersey fish to the seven reporting regions. Values show the exact number and percentage of individuals assigned to each reporting region in GeneClass 2 (left column) and ONCOR (right column).

Reporting region	GeneClass 2		ONCOR	
	n	%	n	%
Scotland	36	26.87	28	20.90
Solway & Northwest England	59	44.03	80	59.70
Wales & Southwest England	25	18.66	21	15.67
Southern England	1	0.75	0	0.00
Northern Ireland	5	3.73	3	2.24
France	3	2.24	2	1.49
Norway	5	3.73	0	0.00

Table 2.7- Baseline self-assignment scores including samples from southern Ireland. The percentage of individuals within the baseline that correctly assign back to their own reporting region and the reporting region that contained the highest proportion of wrongly assigned individuals. i) salmon from southern Ireland form a unique reporting region. ii) Salmon from southern Ireland are grouped with Scotland reporting region.

i)

Reporting Region	Correctly self assigned	Largest incorrect assignment	
Scotland	67.10%	Solway & Northwest England	9.10%
Solway & Northwest England	71.50%	Scotland	7.80%
Southwest England & Wales	68.00%	Solway & Northwest England	8.90%
Southern England	96.40%	Southern Ireland	1.50%
Northern Ireland	85.60%	Southern Ireland	5.70%
France	87.40%	Southwest England & Wales	5.20%
Norway	89.60%	Scotland	3.60%
Southern Ireland	57.80%	Scotland	12.40%

ii)

Reporting Region	Correctly self assigned	Largest incorrect assignment	
Southern Ireland and Scotland	68.40%	Solway & Northwest England	11.50%
Solway & Northwest England	74.90%	Southern Ireland and Scotland	10.30%
Southwest England & Wales	70.70%	Southern Ireland and Scotland	11.40%
Southern England	97.10%	Solway & Northwest England	0.70%
Northern Ireland	88.50%	Southern Ireland and Scotland	4.90%
France	88.50%	Southwest England & Wales	5.50%
Norway	90.20%	Southern Ireland and Scotland	4.50%

Table 2.8- Results of assignment of adult Mersey fish to the seven reporting regions both excluding and including salmon from southern Ireland. Values show the exact number and percentage of individuals assigned to each reporting region in GeneClass 2 (left column) and ONCOR (right column)..

Reporting region	GeneClass 2				ONCOR			
	Excluding S. Ire		Including S. Ire		Excluding S.Ire		Including S. Ire	
	n	%	n	%	n	%	n	%
Scotland (& S. Ireland)	36	26.87	39	29.10	28	20.90	38	28.36
Solway & Northwest England	59	44.03	54	40.30	80	59.70	73	54.48
Wales & Southwest England	25	18.66	25	18.66	21	15.67	18	13.43
Southern England	1	0.75	1	0.75	0	0.00	0	0.00
Northern Ireland	5	3.73	7	5.22	3	2.24	2	1.49
France	3	2.24	3	2.24	2	1.49	3	2.24
Norway	5	3.73	5	3.73	0	0.00	0	0.00

Chapter 3: How unique are Atlantic salmon in the chalk streams of southern England?

Introduction

It has become apparent that Atlantic salmon (*Salmo salar* L.) are clustered into groups of genetically similar fish spread over broad geographic regions (Griffiths *et al.* 2010; Perrier *et al.* 2011). This is likely to be because, after spending a year or more in the Atlantic Ocean feeding, most individuals return to the river in which they were spawned (Stabell 1984) via a combination of geomagnetic (Moore *et al.* 1990) and olfactory cues (e.g. Hansen & Jonsson 1994). Rivers vary in biotic and abiotic factors, including temperature, water chemistry, prey and predator availability. So it is possible that one or more of these factors could act as a selective agent for local adaptation (Fraser *et al.* 2011).

River geology has been identified as a key determinant of gene flow between these genetic groups (Perrier *et al.* 2011). Salmon in rivers on different geologies are much less likely to inter-breed, compared to those in rivers on the same geology (Perrier *et al.* 2011). Again, this could be due to local adaptation, or because freshwater within different geologies smells different. Either way, one geological region in southern England appears to harbour salmon that show disproportionately large differentiation for the species. This region is home to the aptly named "chalk streams", and endemic salmon appear to be particularly distinct from their neighbours in southwest England and southern Europe (Griffiths *et al.* 2010).

Chalk stream salmon samples have occasionally been included in broad population structure studies; however their uniqueness has only become apparent since the widespread application of microsatellites (Griffiths *et al.* 2010). In 1976, transferrin proteins could differentiate only two groups within the UK, and samples from two chalk streams – the rivers Avon and Stour – were grouped with the now de-bunked "Boreal race" rather than the "Celtic race" (Child *et al.* 1976). This Boreal race also contained their neighbours in southwest England. Protein electrophoresis later indicated a clear difference in allozyme frequency between the one included chalk stream – the river Itchen – and the remaining six rivers from across England and Wales (Hovey *et al.* 1989). Similar allozyme frequencies were identified within

proximate chalk streams, the Test (Thompson & C 1991) and Frome (Jordan *et al.* 2005). In 2010, microsatellites were used to investigate population structure within Europe, which included chalk salmon (Griffiths *et al.* 2010). Salmon from all three studied chalk rivers were distinct from salmon from the other rivers in the UK, France and Spain (Griffiths *et al.* 2010). However, as these studies each focused on the general population structure of salmon in a wide geographic area, very few chalk streams were sampled and the uniqueness of these salmon was either not observed or not discussed.

The evolutionary history of salmon in this region is also debatable. During the Pleistocene Period, major glaciers extended southward from the Arctic (Hewitt 1999), extirpating northern populations and leaving only refuges in warmer southern locations (Hewitt 1999). With the help of ca. 40,000 year old salmon remains found in Spain, a salmon refuge was identified within the Iberian Peninsula (Consuegra *et al.* 2002). It is suspected that after the ice melted, salmon emerged from this refuge to re-colonise large parts of Europe (Consuegra *et al.* 2002). More recently, a second refuge was proposed within northwest France (Finnegan *et al.* 2013). The authors found evidence that salmon currently along the eastern and southern coasts of England emigrated from this French refuge, as well as the previously identified Spanish refuge (Consuegra *et al.* 2002), after the glaciers retreated. The high genetic diversity of salmon within southwest England is proposed to be indicative of the meeting of emigrants from both of these proposed refuges (Finnegan *et al.* 2013). However, there is evidence that salmon in the southern English chalk streams have much lower genetic diversity (Finnegan *et al.* 2013), and no explanation for this has ever been given. When re-constructing the evolutionary history of populations in this region, the study by Finnegan *et al.* (2013) grouped chalk stream populations with those from throughout the UK based on mitochondrial DNA (mtDNA), despite the distinction they had identified using microsatellite data. It is therefore possible that these chalk stream salmon have an independent origin from salmon within the rest of the UK. During the last glacial maximum (LGM), which ended ca. 10,000-12,000 ybp (years before present), ice cover extended south over Britain to approximately 52°N (Hewitt 1999; Murton & Lautridou 2003). This would have left much of the chalk streams free from ice cover, therefore it is also possible that these salmon survived the LGM within them. Their distinction could therefore be the result

of having diverged from the remaining populations much earlier. Alternatively, if salmon re-colonised this region after the LGM, as suggested by Finnegan *et al.* (2013), strong selective pressure from the chalk stream environment may have forced rapid local adaptation. In effect, the relationship of the chalk stream salmon to their neighbours in southwest England, France and Spain is still unknown and this study aims to change that.

Chalk streams are a type of river formed on a rarely exposed substrate of calcareous chalk, which was laid down during the Cretaceous Period (Figure 3.1; 145.5-65.5 million years ago). Chalk is porous, which causes water to run through it rather than over it, and subsequently these rivers are relatively clear and of low stream order (Berrie 1992). The chalk gives these rivers further unique properties (Berrie 1992), any of which may be a selective agent of local adaptation. As the chalk acts as an aquifer, a relatively steady flow of water is released throughout the day, and the temperature rarely deviates from 10 ^{o}C. The chalk also causes the chalk streams to be alkaline (ca. pH 8; Mann 1989), rather than acidic like most rivers. It is worth noting that 85% of the world's chalk streams reside in southern and eastern England. Several are Sites of Special Scientific Interest (SSSIs) including the river Avon System and the river Itchen, which have also both been designated Special Areas of Conservation. The presence of Atlantic salmon is an important part of these designations.

This study has two aims, which build on previous research that indicated an unusually large difference between salmon in some of these chalk streams and their neighbours (Griffiths *et al.* 2010) and lower genetic diversity in the Avon and Itchen (Finnegan *et al.* 2013). The first is to identify the extent of genetic differentiation and genetic diversity reduction of the chalk stream populations relative to those in neighbouring southwest England, France and Spain, using samples from all UK chalk streams with major salmon populations. The second is to explain the current diversity and genetic differentiation of the chalk stream salmon. This is done by the analysis of migration rates, testing for evidence of a genetic bottleneck and inferring the evolutionary relationship of chalk stream salmon with neighbouring populations. In order to do this both microsatellites and mitochondrial DNA have been used.

Method

Sampling and DNA preparation

Salmon fin clips were obtained from juveniles (0+ parr) from three rivers in southwest England – the Camel, Dart and Exe; five chalk streams in southern England – the Frome, Piddle, Avon, Test and Itchen (Figure 3.2). The rivers in southwest England and the majority of chalk streams were sampled by the Environment Agency during routine national surveys and management programmes between 2004 and 2012. Sampling of the Frome and Piddle was carried out by the Game and Wildlife Conservation Trust in September 2009 and 2011 during routine juvenile abundance surveys. Samples were obtained using electrofishing to attract juveniles before cutting and collecting their adipose fins, which conformed to national agency ethical guidelines. Microsatellite genotypes were also obtained for four rivers in France – the Sée, Léguer, Ellé and Scorff and seven rivers in Spain – the Ason, Deva, Sella, Narcea, Eo, Miño and Ulla (Figure 3.2) from the SALSEA database (unpublished, J. Gilbey pers. comm.). In total, the microsatellite genotypes of 1,518 individual salmon were obtained from 19 rivers (Table 3.1).

All samples in the present study were juveniles, except those from France, where adults were sampled. The effect of this difference in life stages is likely to be minimal, except that the French samples may show upwards bias in genetic diversity indices. This is because adults have greater movement capabaility and are therefore likely to represent a larger spatial area. They also represent several different cohorts, while juveniles represent only one. This effect is likely to be very small and has been ignored in a previous study (Finnegan *et al*. 2013). The effect on any analysis of long-term effects should also be minimal as the small difference in sampling regime should not overshadow the effect of thousands of years of genetic drift and mutation. However the effect could be significant on the detection of recent migration for the two following reasons at least. Firstly as there are less likely to be strays in the juveniles and more likely to be strays in the adults, there may be a bias towards finding strays in the French samples. Secondly, the adult will serve to increase the time difference between samples as they will be from a previous generation.

DNA preparation

DNA was extracted from salmon fin clips using the HOTshot method (Truett *et al.* 2000) and 14 microsatellites were amplified according to Ikediashi *et al.* (2012). The region of the mitochondrial genome containing the ND1 gene was amplified and sequenced in chalk stream salmon according to Finnegan *et al.* (2013) using primers ND1-F and ND1-R (Nilsson *et al.* 2001). PCR reactions were carried out in a volume of 25 µl consisting of 1 x HotStar Taq Plus Master Mix (Qiagen, Manchester UK), 0.2 µM of each primer and approximately 50 ng of template DNA. PCR conditions were as follows. An initial denaturing step at 94 ^{o}C for 2 min 30 s, amplification proceeded for 35 cycles at 94 ^{o}C for 30s, 58 ^{o}C for 30 s, 72 ^{o}C for 60 s and a final extension at 72 ^{o}C for 10 min. PCR product (10 µl) was purified using 0.25 U of Exonuclease I (New England Biollabs, Hitchin, UK) and Antarctic Phosphotase (USB) with an initial incubation at 37 ^{o}C for 45 min followed by 80 ^{o}C for 15 min. Sequencing was carried out by Beckman Coulter Genomics (Essex, UK). These sequences were added to a ND1 dataset compiled by Finnegan *et al.* (2013). From southwest England, ND1 sequences were also obtained from salmon in the rivers Camel, Dart, Taw and Usk. Sequences from the Taw and Usk were included because they were available and would provide more genetic information for the analysis. Previous research based on microsatellites has shown salmon within these rivers to be part of the same genetic group as salmon in the Camel and Dart (Griffiths *et al.* 2010; Ikediashi *et al.* 2012; Finnegan *et al.* 2013), therefore they are likely to share the same evolutionary history. Microsatellite genotypes acquired from the SALSEA Merge database were calibrated to the Exeter genotypes to correct for different scoring of alleles between laboratories. Using the results from a previous calibration study (Ellis *et al.* 2011). the SALSEA data was transformed for the present study following specific rules for 11 loci (Table 3.2) The remaining loci did not require calibration.

Error checking

The microsatellite genotypes were checked for scoring errors due to stutter peaks, large allele dropout and null alleles using the program MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004). The confidence interval was set to 95% and 1000 iterations were performed. Genotypes were then checked for deviations from Hardy-Weinberg equilibrium, and linkage disequilibrium using Genepop on the web (Raymond & Rousset 1995). Both were calculated using 1000 de-memorisation

steps, 200 batches each with 10,000 iterations. The 95% significance level of each was adjusted using the False Discovery Rate method (Benjamini & Hochberg 1995). In order to prevent the false detection of population structure due to the presence of family groups, the program COLONY v2.0.4.1 (Jones & Wang 2010) was used to identify full siblings. The mating system was defined as polygamous for males and females and without inbreeding. Each run was of medium length, with high precision and using the Full-Likelihood method. Allele frequencies were not updated during the run and no prior sib-ship was assumed. An error rate of 0.02 was used for each locus based on testing by Ellis *et al.* (2011). COLONY was run twice independently, with different starting seeds to check consistency of the reconstruction. All members of each family group, except one, were removed if found to be in both runs with an average probability of greater than 0.5.

Descriptive Statistics

The number of alleles, expected heterozygosity (H_E) and observed heterozygosity (H_O) were calculated for salmon in each river using Genalex v6.5.0.1 (Peakall & Smouse 2012). In order to accommodate for differences in sample size between each river, which would bias the number of alleles, the allelic richness (A_R) within each river was calculated using Fstat v2.9.3 (Goudet 1995). In order to determine whether there were significant differences between the chalk streams and the other regions, two sided tests were performed within Fstat for A_R, H_O, H_E, heterozygote deficit (F_{IS}) and between-sample site F_{ST} between the chalk samples and the samples from southwest England, France and Spain independently.

Population structure

In order to determine whether the populations were significantly different from each other a test for significant differences in allelic frequencies was completed using GENEPOP on the web v4.2 (Raymond & Rousset 1995). Option 3, sub-option 2 was selected to infer whether alleles frequencies were equal between paired populations. Option 3, sub-option 4 was also selected to infer whether the distribution of genotypes was equal between paired populations. In both cases the program was run under default conditions. Also, pair-wise F_{ST} values between river samples were calculated in Genalex v6.5 (Peakall & Smouse 2012) and tested for significance difference via 999 permutations. A principal component analysis of the

individual genotypes was also performed using the package adegenet v1.4-1 (Jombart 2008) within R v.3.1.0 (Venables & Smith 2005). The optimal number of genetic units was determined using the programs BAPS 5 (Corander *et al.* 2003) and STRUCTURE v.2.3.4 (Pritchard *et al.* 2000). The "clustering of groups of individuals" function was performed within BAPS, with a maximum number of genetic units of 20. STRUCTURE was run using the admixture ancestry model with 250,000 Markov Chain Monte Carlo (MCMC) replicates after a burn-in of 50,000 assuming 1-20 populations. All other parameters were left as default. The process was repeated 20 times at different starting points along the MCMC chain. The most likely number of distinct genetic units was inferred using the ΔK method primarily (Evanno *et al.* 2005). However, because the method is unable to incorporate k=1 when deciding the optimum, the Likelihood method, recommended by Pritchard *et al.* (2000), was also used to check the likelihood of one genetic unit (*k*=1). A dendrogram was constructed to explore the relationship of salmon in each river, based on the microsatellite data. A neighbour joining dendrogram, based on Cavalli-Sforza and Edwards chord distance (D_{CE} - Cavalli-Sforza & Edwards 1967), was constructed using Populations v1.2.32 (Langella, 1999). The dendrogram was visualised in FigTree v1.4 (Rambaut & Drummond 2009).

Genetic assignment of samples was also used to investigate the population structure between rivers. A leave one out test was performed within the program ONCOR (Kalinowski *et al.* 2007), to determine the proportion of individuals that would assign back to the river and region they were sampled from.

Recent migration analysis

The recent rate of migration of salmon between regions was determined using the programs BIMr 1.0 (Faubet & Gaggiotti 2008) and BayesAss v1.3 (Wilson & Rannala 2003). BIMr uses a Bayesian approach and MCMC technique to detect the proportion of recent migrants, assuming sampling of individuals who have yet to migrate after spawning. Default parameters were used to run five independent runs with a burn-in and sampling size of 10^5 and a thinning of 50. The results from the five runs were checked for consistency between them and values were used from the run with the lowest deviation score as recommended by the program manual. BayesAss also uses Bayesian statistics to determine rates of recent migration, based on the mixed genotypes of second generations. Preliminary analyses were run to optimise

mixing parameters to obtain the target acceptance rate of between 20% and 60% for allele frequencies, inbreeding coefficients and migration rates, as recommended by the program manual. After this, analyses were run for 1×10^7 iterations after a burning of 1×10^6 iterations. The program was run five times with different starting seeds to check for convergence between runs. Migration results were selected from the run with the best model fit, determined from the highest likelihood score. From the neighbour-joining dendrogram, salmon from the river Sée were found to be equally related to salmon in southwest England and the remaining rivers in France (see results); therefore their effects on the calculation of migration rate were investigated by running both methods with the inclusion and exclusion of these samples.

Historic migration analysis

The programs IMa2 (Hey 2010) and MIGRATE (Beerli & Felsenstein 2001) were used to determine historic levels of migration between salmon in the four regions. IMa2 was the primary method used because it has been widely used (Wilson & Eigenmann Veraguth 2010; Leo *et al.* 2012; Finnegan *et al.* 2013), and would enable direct comparison with previous studies of Atlantic salmon (e.g. Finnegan *et al.* 2013). Although the program has the potential to use both genetic sequences and microsatellite data, the author warns that analysis including microsatellites takes a large amount of time (Hey 2011). As sequence data is considered to be more appropriate for historical processes - microsatellites evolve too quickly to be useful over larger time scales (Wang 2011)- the decision was made to run IMa2 with sequence data only. However, as a comparison, microsatellite data was also used to infer historic migration rates using the program, MIGRATE.

The program, IMa2, was initially used to calculate divergence dates from the mitochondrial ND1 gene (see details below). From the output of those analyses, the effective number of migrants per generation ($N_E m$) were calculated using the following equation: $N_E m = m \times (4 \times N_E \times \mu)/ 4$ where N_E (the effective number of individuals) was obtained from the output. The parameter, m, was calculated from the results, which gave the migration rate per generation (M), using the equation: $m = M / \mu$, *where μ* = the mutation rate of the gene (5.7×10^{-6} substitutions per year)

(Doiron *et al.* 2002). This mutation rate has been used in previous Atlantic salmon studies (Finnegan *et al.* 2013) but was calculated by Doiron *et al.* (2002) from the divergence date between two charr species and rainbow trout (*Oncorhynchus mykiss*). Doiron *et al.* (2013) investigated the mutation rate in 13 mitochondrial genes across a range of salmonids and identified that the mutation rates were similar. Following recommendation by Campagna *et al.* (2012) the effective number of migrants independent of mutation rate ($2N_Em$) was calculated by multiplying N_Em by two.

MIGRATE was run, using the microsatellite data, following the protocol of Barson *et al.* (2009). The program was run three times, with the first run using F_{ST}-based estimates as a searching start point in order to obtain posterior estimates of Θ and M. Θ is the effective population size multiplied by four times the mutation rate, which as is often used for salmonid microsatellites (Olafsson *et al.* 2014), was 5×10^{-4}. This rate is based on the date of origin for modern humans and an investigation of 15 di-nucleotide microsatellites (Goldstein *et al.* 1995). M is the migration parameter, which equals the migration rate (m) divided by the mutation rate (μ). M is also the number of new alleles introduced to a population by immigration relative to mutation. The subsequent runs used the results from the first run for prior values of Θ and M, so that $\Theta = 1.0$ and $M = 1000$ in all cases. Five thousands steps were taken within each long chain, and the burn-in was set to 10,000. A heating scheme was used with the following temperatures for all runs: 1.0, 1.2, 1.5, 3.0. From the posteriors of the final results, the effective number of migrants per generation (N_Em) was calculated from the following equation: $N_Em = (M \times \Theta) / 4$. The 95% confidence intervals of N_Em were compared in order to determine if migration into the chalk streams was significantly different from migration out of the chalk streams.

From the neighbour-joining dendrogram, salmon from the river Sée were found to be equally related to salmon in southwest England and the remaining rivers in France (see results). Further investigation informed us that the Sée was extensively stocked with exogenous fish during the last century, which may have had a signifcant effect on the genetic signature of these fish (Evanno, *pers. comm*), and may artificially influence the calculation of migration rate. Therefore their effects on

the calculation of migration rate were investigated by running MIGRATE with the inclusion and exclusion of these samples.

Bottleneck analysis

In order to test whether a bottleneck was the reason for apparently low levels of genetic diversity within the chalk stream salmon populations, the programs BOTTLENECK v1.2.02 (http://www.ensam.inra.fr/URLB), M-Ratio (Garza & Williamson 2001) and VarEff (Nikolic & Chevalet 2014) were used to test for significant population declines in all included rivers. BOTTLENECK is based on the principle that alleles are lost from a population faster than heterozygosity during a population decline, thus causing a heterozygote excess compared to what would be expected for the current allele frequencies, under mutation-selection equilibrium. The program was under both the single step mutation (SMM) and two-phase mutation (TPM) models, with the TPM set to 95% SMM with a variance of 12 as recommended (Piry *et al.* 1999; Tonteri *et al.* 2009). Significance was assessed using the Wilcoxon test (Piry *et al.* 1999). The program M-Ratio was used to calculate the ratio of number of alleles to the range in alleles (M ratio) and is based on the principle that the alleles are lost faster than the reduction in the range of alleles (Garza & Williamson 2001). The program, Critical M (Garza & Williamson 2001), was used to calculate the M ratio below which a population is likely to have been through a significant decline in population size. For this, Θ (= 4 x effective population size x mutation rate) was calculated. For the past effective population size a range of 100 to 10,000 was used and the microsatellite mutation rate used was 5 x 10^{-4} (Olafsson *et al.* 2014), therefore Θ ranged from 0.02 to 20. The mean step size was 3.5 following recommendation (Garza & Williamson 2001). VarEff uses a Bayesian coalescent approach to calculate posterior distributions of the past effective population size. Following Finnegan *et al.* (2013), a four year generation time was used, with a mutation rate of 5 x 10^{-4} (Olafsson *et al.* 2014). The prior effective population size for each sampled river was 20,000 and past effective population sizes were calculated up to 20,000 generations before the present. Following recommendation by Nikolic & Chevalet (2014), and the author of VarEff (Nikolic pers. comm.) only the SSM model was used as it has been identified as most accurate for these loci within this species (Nikolic pers. comm.).

Chalk divergence timing and phylogeography

In order to determine the age of the chalk stream lineage the date of their divergence from the southwest England, France and Spanish regions was estimated using the mitochondrial ND1 data and the isolation with migration model within the program IMa2 (Hey 2010), and also the microsatellite data within the program DIYABC (Cornuet *et al.* 2008). In both cases pair-wise comparisons were made between the chalk and either southwest England, French or Spanish populations. IMa2 uses a MCMC approach to estimate the posterior probability of the divergence date of populations (among other parameters) based on the differences between genes and the average mutation rate of that gene. It is also able to calculate and incorporate the rate of migration, which, if unaccounted for, would act to reduce estimated divergence dates by reducing the number of differences between sequences. DIYABC uses an approximate Bayesian computation (ABC) model to calculate the posterior distribution of time of divergence (among other parameters) (Antao *et al.* 2008). Crucially, unlike IMa2, rather than calculate and adopt the level of migration between populations, DIAYBC assumes that no migration occurs at all once the populations have diverged. Therefore, together these two programs are complementary.

IMa2 was run following the protocol of Finnegan *et al.* (2013) i.e. using a four year generation time and the following two mutations rates: A relatively slow mutation rate of 5.7×10^{-9} substitutions per site per year (Doiron *et al.* 2002), with upper and lower limits set to 1×10^{-8} and 1×10^{-9} respectively. This mutation rate is based on the mitochondrial genome divergence between two charr species (*Salvelinus fontinalis* and *Salvelinus alpinus*) and rainbow trout (*Oncorhynchus mykiss*). A relatively fast mutation rate of 1.537×10^{-8} substitutions per site per year (Jacobsen *et al.* 2012), which is based on recent mitochondrial divergence in whitefish species (*Coregonus* spp), was also used. No upper or lower limits were used in this case, because none were given (Jacobsen *et al.* 2012) The parameters were changed to match those recommended for small to medium sized data sets with medium heating, i.e the geometric model (-hfg) was used, the number of chains (-hn) was set to 40, the first heating pararameter (-ha) was set to 0.975 and the second heating parameter (-hb) was set to 0.75. Posterior probabilities of migration and divergence date parameters were at first calculated from shorter runs in order to

determine the appropriate priors. Once determined, two long runs were completed to check for consistency between the runs, each consisting of 25 x 10^6 generations of data, sampling every 100 generations, after a burn-in of 1 x 10^6.

Calculations of divergence dates using DIYABC require that the microsatellite loci are selectively neutral. Although these microsatellites are assumed to be selectively neutral, the programme LOSITAN (Antao *et al.* 2008) was used to be certain. For this, the dataset was divided into the four groups before comparing altogether via the generation of 100,000 simulated loci, providing an expected neutral distribution of F_{ST} values and an estimated p-value for each locus. After the removal of the one loci found to be under significant selection - Ssa171, see results - , DIYABC was run following the protocol with a SMM model. The summary statistics used were those recommended by Beaumont (2008), so that the single sample summary statistics selected were the number of alleles, heterozygosity and allele size variance, and the and the two population summary statistics were the mean number of alleles, heterozygosity and alleles size variance.

Although DIYABC has the ability to involve multiple groups at once, the program also makes the assumption of zero migration between the included groups. This was unlikely, therefore calculations with DIYABC were kept simple and limited to pair-wise comparisons between the English chalk streams and the other three groups. DIYABC is also able to determine whether salmon from one group (N1) colonised the other (N2), or if a mutual ancestor gave rise to them both. Therefore within each comparison, three scenarios were tested with two populations N1 and N2 (Figure 3.3), where N1 were the chalk stream salmon, and N2 was either salmon in southwest England, France or Spain. In Scenario 1, N1 was the ancestor population from which N2 diverged at time t, in scenario 2, N2 was the ancestor population from which N1 diverged at time t, and in scenario 3, both N1 and N2 diverged from a common ancestor (N3) of population of size N1+ N2 at time t. Parameter priors for effective population size (N) and t were determined from preliminary runs. In each case the Single Mutation Model was run following a previous study on Atlantic salmon (Olafsson *et al.* 2014), by setting the mean coefficient and individual locus coefficient to a minimum and maximum of 0. The mutation rates were left as default (mean = 5 x 10^{-4}). Following the preliminary runs,

the prior effective population size was set to a maximum of 300,000 for salmon in southwest England, 50,000 in the chalk streams, 200,000 in France and 50,000 in Spain (Figure 3.4). The prior time of divergence was a maximum of 100,000 generations in each case.

During each comparison, the minimum number of required datasets was simulated. The best scenario was then determined using the direct and logistic regression methods to determine each scenario's posterior probability. Model checking was then performed on the best scenario to ensure goodness of fit between the model and posterior parameters using a PCA, and posterior probabilities of each scenario were calculated using a logistic regression of 1% of the simulated data. In cases where the difference between the best scenarios was small the "confidence in scenario choice" function was initiated to identify the probability of type I (the probability of being rejected when in fact the true scenario) and type II (the chance of being accepted when in fact the wrong scenario) errors. Following this, the best scenario was run alone with the maximum number of datasets simulated, in order to obtain a more accurate estimate of divergence dates. The best scenario in each was also run with a different set of summary statistics, which were recommended by Guillemaud *et al.* (2010), in order to see how this would affect the estimated divergence dates. In this case, no single population summary statistics were selected, but the two sample statistics used were the mean number of alleles, heterozygosity, alleles size variance, F_{ST} and the classification index.

Results

The dataset

In total, the microsatellite genotypes of 1,518 individual salmon were obtained from 19 rivers (Table 3.1). Sibling analysis identified that the samples from the rivers Sella and Ason in Spain consisted of just four full-sibling families each, and that the Narcea consisted of 18. Therefore all samples from these rivers were removed from the dataset and all further analyses. After error checking and the removal of siblings the final microsatellite dataset contained 1,112 samples from 16 rivers (Table 3.1). This consisted of 500 fish from the Frome, Piddle, Avon, Test and Itchen; 265 from the Camel, Dart and Exe in southwest England; 184 from the Sée, Léguer, Ellé and Scorff in France, and 163 from the Deva, Eo, Miño and Ulla in Spain.

A total of 125 ND1 sequences were obtained (Table 3.3) each containing 1,150 base pairs. This included 39 from the rivers Camel, Dart, Taw and Usk in southwest England; 42 from the Frome, Piddle, Avon, Test and Itchen in southern English chalk region; 14 from the river Ulla and Sella in Spain; and 37 from the rivers Sée, Léguer, Ellé and Scorff in France (Table 3.3). The dataset contained 12 unique haplotypes, each of which had been identified previously (Finnegan *et al.* 2013). According to sibship analysis of microsatellite data, the majority of salmon from the Sella were full siblings – possibly explaining why only one haplotype was observed in the previous study (Finnegan *et al.* 2013) – therefore six out of eight sequences from this river were removed before further analysis.

Hardy-Weinberg and linkage disequilibrium

Initially, the genotype frequencies of 30 out of 224 comparisons (16%) deviated significantly from Hardy-Weinberg equilibrium. After false discovery rate (FDR) correction by locus this fell to 15. The largest number was found with Ssa197, which had four, and G7, which had six. Only nine loci were out of Hardy-Weinberg equilibrium when corrected using FDR by population, with a maximum of two cases per population. Out of 1,456 comparisons, there were a total of 130 cases of linkage disequilibrium between loci. This fell to 20 cases after FDR correction, with a maximum of four cases per population occurring in salmon from the Dart.

Genetic differentiation

The overall results for the allelic and genotypic pair-wise comparisons were "highly significant" in almost all comparisons (Chi^2 = infinity, df = 28, *P* = "Highly significant". This indicates that at for at least one of the microsatellites, a P value of 0 was inferred (Raymond & Rousset 1995). The exceptions were the same for both test, and were the Frome vs Piddle (Gene Chi^2 = 96.02532, df = 28, *p* = 0) (allelic Chi^2 = 90.94986, df = 28, *p* = 0), the Ellé vs Leguér (*p* = 0.000001, and 0.000002 respectively), Scorff vs Leguér (p = 0 and 0.000009 respectively), and Ellé vs Scorff (p = 0.003123 and 0.027649 respectively). The average F_{ST} value between each river was 0.042 and ranged from 0.006 between the Scorff and Ellé, to 0.083 between the Itchen and Ulla (Table 3.4). Salmon in all rivers were significantly different from each other (p < 0.05), except for in the rivers Scorff and Ellé (p = 0.569) (Table 3.4). Analysis with BAPS indicated four genetic groups, which consisted of salmon from 1) southwest England, 2) the chalk streams 3) France and 4) Spain (Figure 3.5). The likelihood analysis of the the STRUCTURE results indicated greater than 1 group (*K* = 11). The Δ*k* analysis of STRUCTURE results indicated two groups (*k* = 2; Figure 3.6), which consisted of salmon from all of the chalk rivers in one group, and salmon from all remaining rivers in the other (Figure 3.7a). Under *k* = 4, nine STRUCTURE runs identified the same genetic units as identified by BAPS (Figure 3.7b). For the other 11 runs, STRUCTURE grouped salmon in southwest England and France as one unit, and divided the chalk stream salmon into two groups (Figure 3.7c). The individual based principal component analyses (PCA; Figure 3.8) identified the four groups, with the chalk separation from the remaining rivers along axis 1 (Figure 3.8a and 8b). The separation of southwest England, France and Spain is visible along axis 3, albeit with overlap (Figure 3.8c).

The dendrogram indicated that salmon from the chalk streams are almost equally distant from salmon in southwest England, France and Spain (Figure 3.9). Salmon from the chalk, southwest England and Spanish rivers cluster into their respective groups. However, salmon in the French rivers do not cluster together as the river Sée was positioned just as close to the rivers in southwest England as the remaining rivers in France. For this reason, fish from the Sée were omitted from the analyses of historic divergence dates.

At river level the successful assignment of salmon to the rivers in which they were caught ranged from 34% for salmon from the Scorff to 93% for salmon on the river Deva (Table 3.5). The largest proportion of mis-assignment (i.e. the assignment of a fish to river that it was not sampled from) was, in 15 out of 16 cases, to another river within the same genetic group. Salmon from the Sée (France) were the exception, where four samples assigned to the Camel. At the regional level, successful assignment ranged from 94% in the French group to 99.8% in the chalk group (Table 3.5). Excluding samples from the Sée made little difference to regional assignment. The exception was for salmon in Spain, where with the Sée samples included, two Spanish samples assigned to southwest England and three assigned to France, and without the Sée samples, four Spanish samples assigned to Southwest England only. At river level, all salmon from the chalk streams assigned to the chalk rivers. At regional level, however, one chalk sample did assign to the French group.

Genetic diversity

The pair-wise statistical comparisons of genetic diversity revealed that salmon from the chalk streams had significantly lower allelic richness (A_R), observed heterozygosity (H_O) and expected heterozygosity (H_E) than salmon in southwest England ($P < 0.005$ for each) (Figure 3.10a) and significantly lower A_R, and He compared to salmon in France (Figure 3.10b). The observed heterozygosity between chalk and France was non-significant (2 sided test, $p = 0.124$; 1-sided test, $p = 0.124$). There were no significant differences between the chalk and Spanish populations. It is noteworthy that salmon in the Ulla and Miño of Spain have low heterozygosity (H_O and H_E) and allelic richness (Figure 3.10), which explains why there was no significant difference between the Spanish and chalk group. Between each group there was also no significant difference in between-river F_{ST} values.

Recent migration analysis

Due to the unusual placement of salmon from the Sée within the dendrogram (Figure 3.9), the migration analysis was performed with and without salmon from this river. With BayesAss, migration rates were not significantly affected by their removal (Table 3.6a). The average proportion of individuals found to be offspring of local (meaning from the river sampled) salmon was 0.98 with and without the Sée salmon, and ranged from 1.00 in the chalk streams to 0.96 in Spain. Straying rates ranged from 0.025 (or 0.021 without Sée salmon) from Spanish salmon straying into France to 0.0007 for chalk salmon straying into all other three regions. Contrastingly, BIMr was significantly affected by the inclusion or exclusion of samples from the Sée (Table 3.6b), although in both cases all five parallel runs were consistent, indicating convergence along the MCMC chain. Without salmon from the Sée, the average proportion of individuals found to be offspring of local salmon was 0.98 and ranged from 1 in southwest England, the chalk streams and France, to 0.92 in Spain.

Straying was limited to the Spanish salmon and ranged from 0 into the chalk region to 0.05 into France. When including fish from the Sée however, the average proportion of individuals found to be offspring of local salmon ranged from 1 in the chalk streams and France only, to 0.81 in southwest England. Straying into southwest England from France was particularly high (0.18). This was likely to be due to the salmon from the Sée containing genotypes matching those of southwest England. While this may indicate significant straying, this is unlikely because such high levels of straying would erode any genetic differentiation (Grandjean *et al.* 2009). However, because of the unusually high straying results, and the significant changes caused by the inclusion and exclusion of Sée fish, BIMr was considered to be unreliable for this dataset. The results from BayesAss indicate low migration between the four regions, but exceptionally low migration into and out of the chalk stream from the other three regions.

Historic migration analysis

The historic effective number of migrants calculated by IMa2 was low (Figure 3.11), ranging from 1.69×10^{-5} from the chalk group into southwest England to 1.60×10^{-8} from the Spain group to the chalk group. In each case, migration out of the chalk streams was larger than migration into the chalk stream (Figure 3.11). Calculations with MIGRATE yielded higher estimates (Figure 3.12). Again, effective migration out of the chalk streams was much higher than migration into the chalk stream. Using the 95% confidence intervals migration into the chalk streams from southwest England and France is significantly greater than the reverse migration of from the chalk streams into the southwest England and France (Figure 3.12). However, there was no significant difference between the migration of Spanish salmon into the chalk streams and the reverse.

Bottleneck analysis

The M ratio ranged from 0.89 for salmon in the river Exe to 0.57 for salmon in the river Miño (Figure 3.13). The average M ratio of samples from the chalk streams was 0.65. Salmon from the southwest English, French and Spanish rivers had average M ratios of 0.85, 0.79 and 0.68 respectively. Analysis indicated a critical M ratio of 0.70 and 0.77 when Θ was 0.2 and 2 respectively. It is also suggested that for any data set with seven loci or greater, an M ratio of less than 0.68 (Garza & Williamson 2001) is indicative of a population decline. Therefore salmon from all chalk streams, except the Test, and the rivers Eo and Miño had likely been through a bottleneck event in either case, while salmon in the chalk river Test, and Spanish rivers Deva and Ulla have been through a bottleneck assuming the upper threshold. All rivers in France and southwest England are at or above the upper threshold. The Wilcoxon analysis within the program BOTTLENECK identified no significant bottlenecks within any population, either under the SMM or TPM mutation model (Table 3.7).

Analysis with VarEff identified two major declines in the effective population size (N_E) of salmon populations during the last 80,000 years (Figure 3.14a and 3.14b). The oldest decline occurred 10,000-30,000 ybp (depending on the river) in all French rivers, two of three southwest English rivers, and the Spanish rivers Deva and Eo (albeit more gradually). N_E increased again in the majority of these

populations. A second decline was identified in all populations, except two, between 80-600 ybp (Figure 3.14b). The two exceptions are the salmon populations in the river Eo, which declined during the first bottleneck and the river Ulla, which declined even earlier. Both population sizes remained low before the start of this second bottleneck. The effective population sizes of all of the chalk stream salmon start relatively low, and four out of five stay low (although the Itchen and Test grew slightly) until the recent collapse in most populations. The exception is for salmon in the Avon, which grow in N_E from approximately 20,000, to a maximum of 120,000. However these also appear to have been through the recent decline.

Chalk divergence timing and phylogeography

Only one locus, Ssa171 was under significant divergent selection according to Lositan (Figure 3.15). This locus was removed during calculations of historic divergence time in order to prevent bias.

With IMa2, the posterior estimates were always consistent between the two long runs. Using the slower mutation rate (5.7 x 10^{-9} substitutions per site per year), the estimated modal divergence time (Table 3.8) between the chalk and French groups was 14,000 ybp (95% confidence interval, 1,000-62,000 ybp). The divergence date between the chalk salmon and Spanish salmon obtained a modal estimate of 15,000 ybp (95% CI, unknown), however the posterior probability failed to return to 0, indicating that the 95% estimates were unreliable. The date of divergence for the chalk and southwest groups was slightly earlier than both of the others at 21,000 ybp (95% CI, 1,100-100,000 ybp), but with overlapping 95% confidence intervals. With the faster mutation rate, the estimated divergence dates were 20-40% of the estimates from slower migration. The divergence dates were still largely overlapping and ranged from 3,200 ybp (95% CI, unknown) between the chalk and Spanish group and 6,500 (95% CI, 400-33,400 ybp) between the chalk and southwest England group. In summary, using the modal estimate of divergence date (Table 3.8) the results from IMa2 suggest that the chalk stream salmon separated from the other three groups at approximately the same time, most likely within the last 20,000 years, and therefore after the last glacial maximum (LGM). It has previously been determined that salmon in southern England arrived after the LGM, therefore this is considered to be a critical time period in their population

structure. Only the mean estimate with the slow mutation rate suggests much older divergence dates, but because this incorporates the entire range of values calculated in the posterior distribution, it is less precise. Conversely, the modal estimates indicate that most estimates were before 20,000 years.

Analysis of divergence dates using microsatellite data with DIYABC yielded a number of results (Table 3.9; also see Appendix VI for full results). Firstly, in pair-wise comparisons of the chalk salmon against southwest England and also against France, the scenario depicting the chalk streams as the source had the highest posterior probabilities of the three scenarios (Figures 3.16a and 3.16b). Using the summary statistics recommended by Beaumont (2008), under this scenario, the modal divergence time of the chalk group from the southwest England group was 15,600 ybp (95%, 6,000-92,800 ybp) and from the France group was 7,800 ybp (95%, 3,300-83,600 ybp; Table 3.9). In the comparison of Spanish versus chalk populations, the scenario where Spain was the source had the highest posterior probability (Figure 3.16c), however it was only slight and the logistic parameter indicated the same result. Therefore the type I (the probability of being rejected when in fact the true scenario) and type II (the chance of being accepted when in fact the wrong scenario) errors were investigated. The scenario where Spain was the source, of salmon recolonising the chalk region, had a higher type I error (0.59 compared to 0.52) and a lower type II error (0.54 compared to 0.61), therefore this scenario was accepted as the most likely scenario and the modal divergence time was 11,200 ybp (95%, 3,800-125,200 ybp) (Table 3.9). In each case, for the most likely scenario model checking indicated a good fit between the model and posterior parameters. These results indicate more time between the divergence of salmon in the chalk streams from the Spain, compared to the chalk divergence from the French group and then southwest England.

When run with the summary statistics recommended by Guillemaud *et al.* (2010), the modal divergence dates of the chosen scenarios were each increased between 8,000 and 13,000 years (Table 3.9). Thus, the divergence date of chalk from southwest England was now 28,320 ybp, from France was 18,400 ybp and from Spain was 25,720 ybp.

In the scenarios of Spain against France and Spain against southwest England – run for comparison and with only the Beaumont (2008) summary statistics – the scenario depicting Spain as the source had the greatest support. The modal divergence times of southwest England from Spain and France from Spain was estimated at 14,500 and 4,700 respectively (Appendix VI), while the mean divergence times were 30,400 and 17,000. In the scenarios comparing southwest England and France, the direct analysis indicated France as the source while the logistic comparison indicated southwest England as the source. The common ancestor scenario was the least likely in both sets of comparisons. The selected scenarios were in agreement with previous hypotheses of refugial zones in Spain (Consuegra *et al.* 2002) and France (Finnegan *et al.* 2013), which gave confidence in the present results.

Discussion

Genetic distinction

Before comparisons could be made between the included populations, it was important to identify how best to group the samples. Four units were identified, which were 1) southwest England, 2) the chalk streams of southern England, 3) France and 4) Spain. In the principal component analysis, the first principal component separated all individuals from the chalk streams from the remaining three regions (Figure 3.8a and 8b). The eigenvalues, which correspond to the ratio of the variance between groups over the variance within groups, indicate that the majority of the variation evident within these populations is captured from the separation of these chalk stream salmon form their non-chalk counterparts (Jombart 2008). The grouping of salmon with the Evanno method (Δk) – with all chalk salmon in one group, and the remaining salmon in the other – also supports this (Figure 3.7). Both these results complement the findings of Griffiths *et al.* (2010) and Finnegan *et al.* (2013), who also identified unusual separation of salmon from chalk streams.

However, the conclusion from the Evanno method (Δk), that the ideal number of genetic units is two, does stimulate questions. A large number of studies use STRUCTURE and Δk to identify the number of genetic units within their dataset; 4,534 studies have referenced the Evanno method at time of writing (data from Web of Science, last accessed 30/01/2015). Yet the present results show that it can produce simplistic results. Four groups are easily identified within BAPS and the PCA (although within the latter, there is overlap between the non-chalk groups). The use of STRUCTURE to define genetic clusters has been increasingly criticised (e.g. Kalinowski 2011), and in particular the Δk method. Perhaps most relevant to the present study is the criticism that within the repeated analysis generated to perform a reliable estimated second order of change, less than optimal STRUCTURE runs are included (Kalinowski 2011). For example, within the present study, some runs at $k = 4$ differentiate populations within the chalk stream while grouping all samples from southwest England and France (Figure 3.7c). STRUCTURE runs with these results have notably lower likelihood scores (-58986.9 compared to -58568.9 for the included) and thus bias the Δk calculation. Removal of runs with lower scores changed the result to include a peak at $k = 4$ where before there was none. Although

k = 2 was still the highest peak, similar effects would have also biased the results under each value of k and further promoted this simplistic finding. Thus this study indicates that the STRUCTURE method, and in particular Δ*k*, should not be used alone to determine the number of genetic units within a data set.

Genetic diversity

For the first time, salmon from all chalk streams containing major salmon populations have been sampled and in every case their genetic diversity – i.e. number of alleles, allelic richness, expected heterozygosity and observed heterozygosity – was significantly lower than their neighbours in southwest England and France (Figure 3.10a and 3.10b). This complements previous research, where relatively low genetic diversity was also identified in chalk stream salmon (Finnegan *et al.* 2013). Compared to the Spanish populations however, there was no significant difference. The Spanish populations, possibly due to their inhabiting the southern range limit of the species, have had exceptionally low sample sizes in recent years (WWF 2001; Saura *et al.* 2006; NASCO 2013). Therefore, it is not surprising that their genetic diversity is so low. A previous study of the Ulla identified that contemporary samples were as diverse as historic (pre-bottleneck) samples, so it is also possible that they have always had a low genetic diversity (Saura *et al.* 2006). Perhaps more surprising is the finding of relatively high diversity within the Deva and Eo. Previous research has also investigated why, despite having some of the lowest numbers of salmon in Europe, genetic diversity in the Eo, for example, is still high (Ribeiro *et al.* 2008). Between 1981 and 1991 these rivers were stocked with exogenous fish, which were mostly from Scotland (Ayllon *et al.* 2006). The sampling of salmon before and after stocking and subsequent identification of increased genetic diversity post stocking, along with evidence of foreign introgression from assignment analysis (Ayllon *et al.* 2006), supports the theory that stocking with foreign fish has increased local genetic diversity (Ayllon *et al.* 2006). Further evidence of foreign introgression into Spanish rivers has also been identified in previous studies (Martinez *et al.* 2001; Moran *et al.* 2005).

Admixture & Migration

In order to explain the evident distinction of chalk stream salmon and their significantly lower genetic diversity, compared to southwest England and France, this study identified migration rates between the four salmon groups. Results from the leave-one-out test, where only one out of 412 chalk stream salmon was assigned to a different group, as well as the STRUCTURE analysis suggest that migration is likely to be low.

The recent migration rates were indeed low (Table 3.6a and 6b). We favour the use of BayesAss over BIMr, as the result indicated less deviation caused by the somewhat unusual Sée population and the effect this had on the calculations, leading to a migration rate of 0.18 from France into southwest England. This level of migration is unlikely in reality, because this would prevent or erode differentiation between the regions (Grandjean *et al.* 2009). Therefore the high migration rate most likely signifies an existing closer genetic similarity between salmon in the Sée and salmon in southwest England. There are at least two possible explanations for why salmon in the Sée would closely match salmon in southwest England. The first possible explanation may be post glacial recolonisation. A study by Finnegan *et al.* (2013) found evidence to suggest that salmon from southwest England are colonisers from refugia in France and Spain after the last glacial maximum. However, this is unlikely to fully explain the similarity, as this would not explain why salmon in the Sée closer match, better than salmon in the other French rivers. An alternate scenario is that salmon in the Sée contain genotypes from southwest England due to artificial stocking from England or elsewhere in the United Kingdom. Although a recent study by Perrier *et* al. (2013) indicates that there was no official stocking of the Sée during 1950 - 1988, there has been supplementive stocking from the nearby river Aulne between 1989 and 2006 (Perrier *et al.* 2013). The Aulne was stocked with salmon from Scotland during 1950 - 1988 (Perrier *et al.* 2013), and therefore salmon from Scotland could yet be an artificial source of UK genotypes in the Sée.

Nonetheless, both recent migration methods indicate that exogenous migration into the chalk streams and migration of chalk stream salmon to non-chalk streams is currently relatively low, and possibly non-existent. However, the historic

migration rates indicated greater gene flow in the past between the chalk salmon and the non-chalk salmon (Figure 3.11 and 3.12). Both Migrate and IMa2 identify that migration out of the chalk streams was historically higher than migration in to the chalk streams. Both programs also suggest that historic migration from the chalk streams into southwest England was relatively high, in fact using MIGRATE, where all pair-wise comparisons were made, historic migration into southwest England from the chalk streams is the highest value, at 14 migrants per generation (Figure 3.12).

The number of effective migrants (N_Em) determined from both programs were considerably larger when calculated with microsatellite data using MIGRATE, compared to the calculations with mitochondrial ND1 data using IMa2. This can be explained however, as N_Em is proportional to the effective population size (N_E). The number of ND1 samples was far smaller than the number of microsatellite samples, so the effective population size estimates were likely to be smaller. Because mitochondrial genes are inherited maternally within salmonids (Allendorf & Seeb 2000), as with most, if not all eukaryotes, mtDNA have an effective size one quarter that of nuclear genes (Lynch *et al.* 2006). This leads to a further reduction in the effective population size and thus the calculation in number of effective migrants.

Nonetheless both of these results are in agreement with DIYABC, which identifies the scenarios where salmon from the chalk streams colonise southwest England and France is preferred over the alternative scenarios (Figure 3.16a and 3.16b).

A key finding is that at every stage, migration into the chalk streams has been lower than migration out. This may be a key factor for their low genetic diversity, as this means that the chalk stream salmon are relatively isolated, despite being surrounded by neighbouring salmon. Within the species there are numerous examples of inverse correlation between isolation and genetic diversity. For example, Atlantic salmon of the Baltic sea, are less genetically diverse than those in Eastern Europe (Säisä *et al.* 2005). These populations are unable to migrate into the Atlantic Ocean – instead migrating and growing within the Baltic Sea – and the Atlantic populations are unable to migrate in. Landlocked populations – salmon unable to migrate to sea that have lost their anadromous trait – have been identified with even lower genetic diversity, for example salmon locked within the river Namsen

in Norway for 9,500 years were found to have an expected heterozygosity of 0.31-0.38, which was much lower than a proximate anadromous population, which with the same markers were found to have an expected heterozygosity of 0.72 (Sandlund *et al.* 2014). Although salmon in the chalk streams are not landlocked, the results indicate that they are nonetheless isolated, and along with these examples it seems probable that this at least partly explains their lower genetic diversity (Säisä *et al.* 2005; Sandlund *et al.* 2014).

Conversely, some salmon populations, which are known to have declined more recently, have shown no reduction in genetic diversity and even some increases have been identified (Consuegra *et al.* 2005; Horreo *et al.* 2011; Perrier *et al.* 2013). This will be further investigated in the next chapter (Chapter 5).

Effective population size and bottlenecks

It was also possible that bottleneck events might have caused the low genetic diversity found in the English chalk stream populations. Considering the extent of population declines in many of these rivers, it was surprising that no bottleneck events were detected using the Wilcoxon test within BOTTLENECK. However, previous studies using this method have also failed to detect bottlenecks where recent population declines are known to have occurred. This includes salmon (e.g. Ribeiro *et al.* 2008) as well as other animals (e.g. Prairie chickens; Bellinger *et al.* 2003). Together these suggest that the Wilcoxon test may not be effective for the detection of bottlenecks. In an evaluation of heterozygote excess, Luikart & Cornuet (1998), determined that this method could only detect a recent bottleneck in 50-75% of cases. The authors suggest using at least 10 loci and greater than 30 individuals for a power >0.80 for detecting a 100-fold reduction. The number of individuals and loci within the present study surpass these criteria; however it is possible that any recent bottleneck was not long acting. It is also possible that too much time has passed since the bottleneck event and that populations are again in equilibrium (Cornuet & Luikart 1996). It might be a combination of both factors, or as the 50-75% detection rate indicates (Luikart & Cornuet 1998), none of these factors. However, the low M ratios for most of the chalk stream salmon indicated that they had been through a bottleneck. This discrepancy might be because the M ratio is able to infer a bottleneck greater than 100 generations ago (Garza & Williamson 2001). Together

these traditional analyses imply that chalk stream salmon have been through an historic bottleneck and not a recent one.

Results from VarEff indicate that what distinguishes the chalk stream salmon is not a recent bottleneck, which the method identifies within the majority of populations, but an historically and consistently low effective population size. This was found in all of the chalk stream populations, except one (the Avon). There are a number of possible reasons why the chalk streams would have an historically low N_E. Firstly, the chalk streams have typically smaller catchments than non-chalk streams (Berrie 1992), so would be expected to support fewer salmon if all other factors were constant. Secondly, after the LGM, it is possible that only a fraction of the salmon population was able to survive and spawn fertile offspring within these chalk streams – resulting in a founder effect (Mayr 1954). Thirdly, it is known that N_E is influenced by the effective size of the meta-population within which it lies (Kuparinen *et al.* 2009). As the migration analysis shows, salmon within the chalk streams are currently much less connected to the other regions. This inaccessibility to the larger meta-population may also have contributed to their historically lower effective population size. Salmon in the Avon appear to be an exception, which at their peak had an effective population size approximately four-times greater than the N_E of the remaining chalk streams (Figure 3.14a). This might also be explained by the meta-population theory (Kuparinen *et al.* 2009). The river Avon sits between the Frome and Piddle on one side, and the Test and Itchen on the other (Figure 3.2). Within this region, the salmon are structured in a pattern of isolation by distance between rivers (as identified in the following chapter); thus the Avon possibly receives a greater proportion of strays from the other chalk streams than they do from each other and is greater influenced by the meta-population that they form. Supportive stocking might also have been a factor, however there are no records of salmon stocking on the Avon (Russell *et al.* 1995; I Russell pers. comm.), so this is unlikely, although not impossible.

Salmon from the Eo, Ulla and Miño also have historically low N_E, as well as low M ratios. The Deva, which is closest to rivers in France (Figure 3.2), has a higher historic N_E and the highest M ratio of the Spanish populations. If the simple assumption is made that straying is inversely proportional to distance, then this result supports both the meta-population theory and the link between M ratio and historic

N_E calculations. The historically low effective population of salmon from the Dart is difficult to explain as they have a relatively high M ratio (0.83).

More studies are needed to verify the reliability of VarEff. It has the potential to clear up a number of inconsistencies identified by previous popular methods, such as when recent bottlenecks have not been detected when they are known to have occurred (e.g. Ribeiro *et al.* 2008). The results from VarEff in the present study also suggests that the M ratio might not just detect bottlenecks greater than 100 generations, but 5,000 generations, or whenever the southern English populations were generated. Also, although the program was used to investigate the presence of bottlenecks, it is worth discussing the surprisingly high historic N_E values calculated (Figure 3.14), which reach into the 10-100,000. Previous research by Nikolic & Chevalet (2014), who used the same method to investigate historic bottlenecks of Atlantic salmon in French and Scottish rivers, has identified similarly high historic values of N_E. In the study by Nikolic & Chevalet (2014), the authors hypothesise that the large historic N_E values are evidence of a larger common ancestor. This may be applicable to the present results but the results from DIYABC indicate that one common ancestor is unlikely. Alternatively, as discussed earlier, with the meta-population theory (Kuparinen *et al.* 2009), the effective population sizes of each river would have benefitted from the increased population sizes of all neighbouring rivers.

It is worth noting that bottleneck detection methods have come under increasing scrutiny (e.g. Chikhi *et al.* 2010; Broquet *et al.* 2010). A large part of this stems from the assumption that the sampled population is in isolation, which is unrealistic in many situations, not least in salmon. With BOTTLENECK in particular, a false bottleneck can be created by a reduction in immigration into the population in question (Broquet *et al.* 2010), thus it is also feasible that a bottleneck might not be detected if immigration into the population increases over the same time period. Even methods using a Bayesian framework, such as MSVAR, have been found to detect false population bottlenecks (or exaggerate them) e.g. in instances when multiple stationary but connected population are sampled but treated as one population (Chikhi *et al.* 2010). The authors of VarEff also note its failure to differentiate between changes in population size and changes in migration (Nikolic & Chevalet 2014). These caveats must be borne in mind when reaching conclusions based on bottleneck analyses.

Divergence

Lastly, it was possible that the distinction of chalk salmon was due to them having diverged from the other salmon at a much earlier date. By using ND1 sequences within IMa2, the dates calculated presently draw direct comparison with the study of Finnegan *et al.* (2013). Within that study, when the divergence dates between salmon in England and salmon in France and Spain (approximately 16,000 and 20,000 ybp respectively), salmon from these chalk streams were grouped with salmon in southwest and northwest England. When – within the present study – they are separated into a distinct group, the results indicate a similar divergence date for the chalk salmon from the French and Spanish salmon, as most of the estimates are younger than 20,000 ybp. But the results also indicate that these chalk salmon diverged from salmon in southwest England at approximately the same time, if you include 95% confidence intervals. Using DIYABC, the oldest divergence date estimated between the chalk and southwest English salmon (28,300 ybp; Table 3.8) predates recent estimates of the LGM (19,000-25,600 ybp; Clark *et al.* 2009). However, there may be bias; salmon in southwest England are thought to derived from multiple refugia (Consuegra *et al.* 2002; Finnegan *et al.* 2013), and the present study indicates that they are the sink of historic migrants (Figure 3.11 and 3.12). The exact effect this would have when calculating divergence date is unknown, but it is likely to be causing bias because the program DIYABC assumes no migration into the populations. Either way, it is also relevant that this is only one estimate out of four, and therefore should be considered sceptically. The remaining three estimates result in a greater probability that the divergence of southwest England from the chalk streams was also after the LGM. Divergence from southwest England was slightly earlier in all cases, but with overlapping confidence limits so the difference is unlikely to be significant. As expected, the use of two different mutation rates changed the estimated dates, but crucially both rates indicated divergence was complete pre-20,000 ybp, i.e. after the LGM.

Analysis with DIYABC enabled us to determine whether divergence events were from a common ancestor or if salmon branched from one group to colonise another. It appears more likely that salmon within these chalk streams colonised French rivers than the reverse, which was suggested by Finnegan *et al.* (2013). However, this is difficult to accept. The extent of ice sheet cover during the LGM

might have allowed salmon to survive in the chalk streams, as southern England extremes were not covered (Hewitt 1999). However had this been the case, salmon would probably have also survived in southwest England and France so there would have been no need for re-colonisation. Therefore it may be more likely that these chalk salmon found a refuge somewhere else during the glacial maximum. It is possible that analyses including salmon from elsewhere in Europe may further clarify this situation.

Summary

This study confirms that salmon in the chalk streams of southern England are distinct from their nearby relatives, and that their genetic diversity is relatively and statistically lower than salmon in southwest England and France. It is evident that the chalk salmon most likely shared a common ancestor with southwest English and French salmon until relatively recently (<20,000 ybp), and therefore their distinction is not the result of an older split. Migration appears to be a key factor as historically migration was higher out of the chalk streams to southwest England and France than migration into them, and in contemporary populations migration into and out of the chalk streams is much lower than the migration between the other groups. This factor alone would likely induce lower genetic diversity as has been identified within other studies.

These chalk salmon also appear to have been through a genetic bottleneck, although exactly when is unclear. While the traditional genetic methods indicate that salmon in the chalk streams have been through an historic bottleneck, the VarEff indicates that they have had a low effective population size for the past 80,000 years. Either scenario would also contribute to a reduced genetic diversity signature. Also, migration, effective population size and bottlenecks are not mutually exclusive, so it likely that the low migration into the chalk streams has at least contributed to the low effective population size. Consequently, both low migration – historic and contemporary – into the chalk streams and a lower effective population size or historic bottleneck event have resulted in chalk stream salmon in southern England that are distinct and possess lower genetic diversity.

Implications for conservation

Salmon in the English chalk streams possess a lower genetic diversity, are genetically distinct and have a reduced connectivity with salmon in neighbouring non-chalk regions; each of these factors is likely to affect their future conservation and sustainability. The low immigration into the chalk streams suggests that any reductions in population size are unlikely to be countered by an influx of neighbouring salmon. Their distinctiveness may make artificial attempts to replenish numbers an attractive proposition. Stocking with fish from neighbouring regions, for example, is likely to be difficult as fish from other regions are genetically dissimilar, and any hybrids may be less fit. Thus extra effort should be made to conserve the populations that are there. Although historically they may have had a high amount of emigration into the neighbouring regions, currently it is low compared to migration between the other groups. This inability to emigrate from the chalk streams suggests that these salmon may be poor at adapting to different environments, or at least that they are less fit than the endemic populations.

Further work

This study poses a number of questions. The relationship between these chalk salmon and non-chalk salmon should be further investigated as it appears more likely that salmon re-colonised southwest England and France than the other way round. This has significant implications for past studies, especially those indicating refugial zones in France (Finnegan *et al.* 2013) and Spain (Consuegra *et al.* 2002). Is it possible that there was also a refuge in the southern English chalk streams during the LGM, or that these fish are descendants from a refuge somewhere else in Europe that have possibly held their original signature better than salmon in France and southwest England.

Another avenue would be to investigate the cause of the low migration into the chalk streams. Do the properties of the water require a physiological or ecological adaptation in the species within one or more of its life stages, or do chalk stream salmon have such an advantage that they completely outcompete all other salmon to an extent not previously identified in salmon with open access to water? Single Nucleotide Polymorphism (SNP) analysis might be the best way to quickly identify regions under selection. Physiological experiments could also be done to

determine how well the gills of each fish function under the different ecological variables. Water pH might be a good place to initiate investigation because chalk streams are relatively alkaline, and past research has identified that pH affects the activity of gills in trout, in particular the activity of an ATPase enzyme within the gills (Nieminen *et al.* 1982). However, it could be a number of other factors. Understanding them could be critical in predicting how populations will change in the future, and importantly, might offer significant tools to improve the effectiveness of stocking and provide other mitigation steps.

The extent of population structure within the chalk streams is also another clear avenue for investigation, and is targeted in the following chapter. The final chapter sheds light on how the chalk streams and other regions have changed in the past and may change in the future.

Figure 3.1- Geological map of the United Kingdom, France and Spain. Map to indicate the range of calcareous rock (green), laid during the Cretaceous Period in England, Wales, France and Spain.
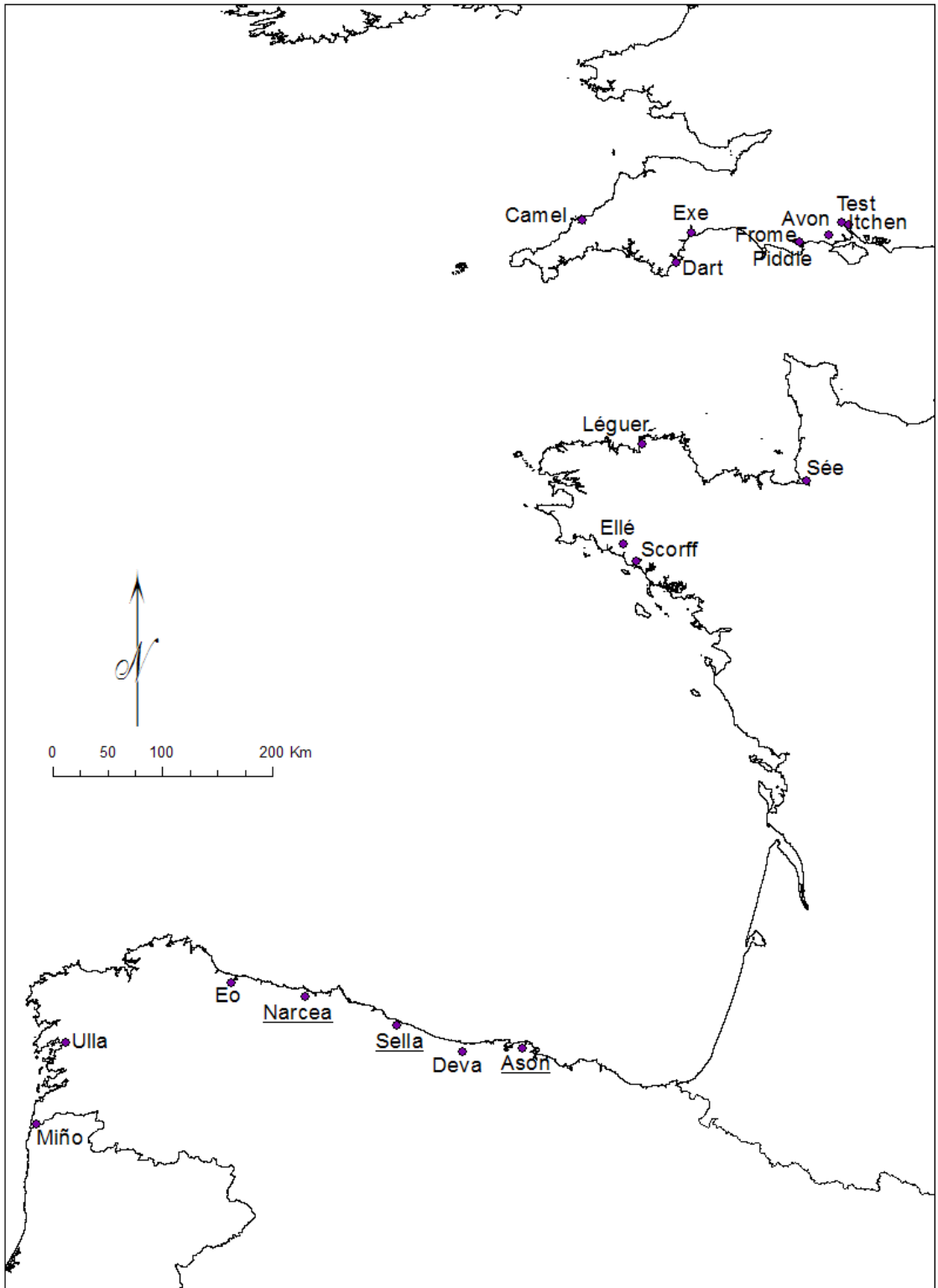
Figure 3.2- Map locations of sampled rivers. Map showing the location of rivers sampled for microsatellite loci in the present study. Underlined river names indicate the three rivers where all samples were removed after falling below 30 individuals after sib-ship analysis. The hatched circle indicates the location of the chalk streams. Salmon from the rivers Taw and Usk (blue circles) were sequenced only within the mitochondrial ND1 region.
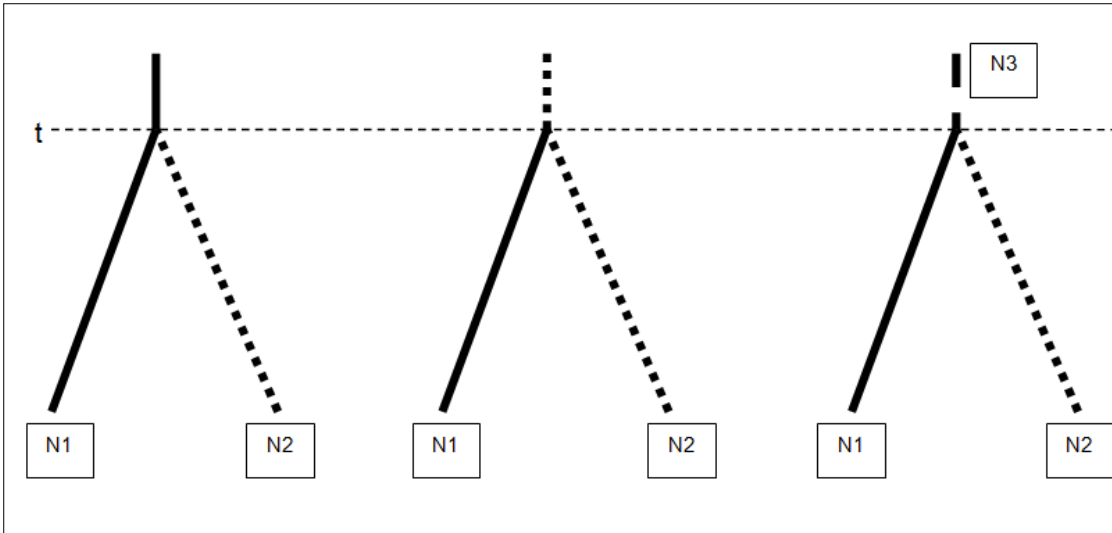
Figure 3.3– Illustration of the three scenarios compared within DIYABC. (From left to right) 1) Where N1 is the ancestral population, which at time t gives rise to N2, 2) where N2 is the ancestral population, which at time t gives rise to N1 and 3) where a common ancestor, N3 = N1 + N2, gives rise to N1 and N2.
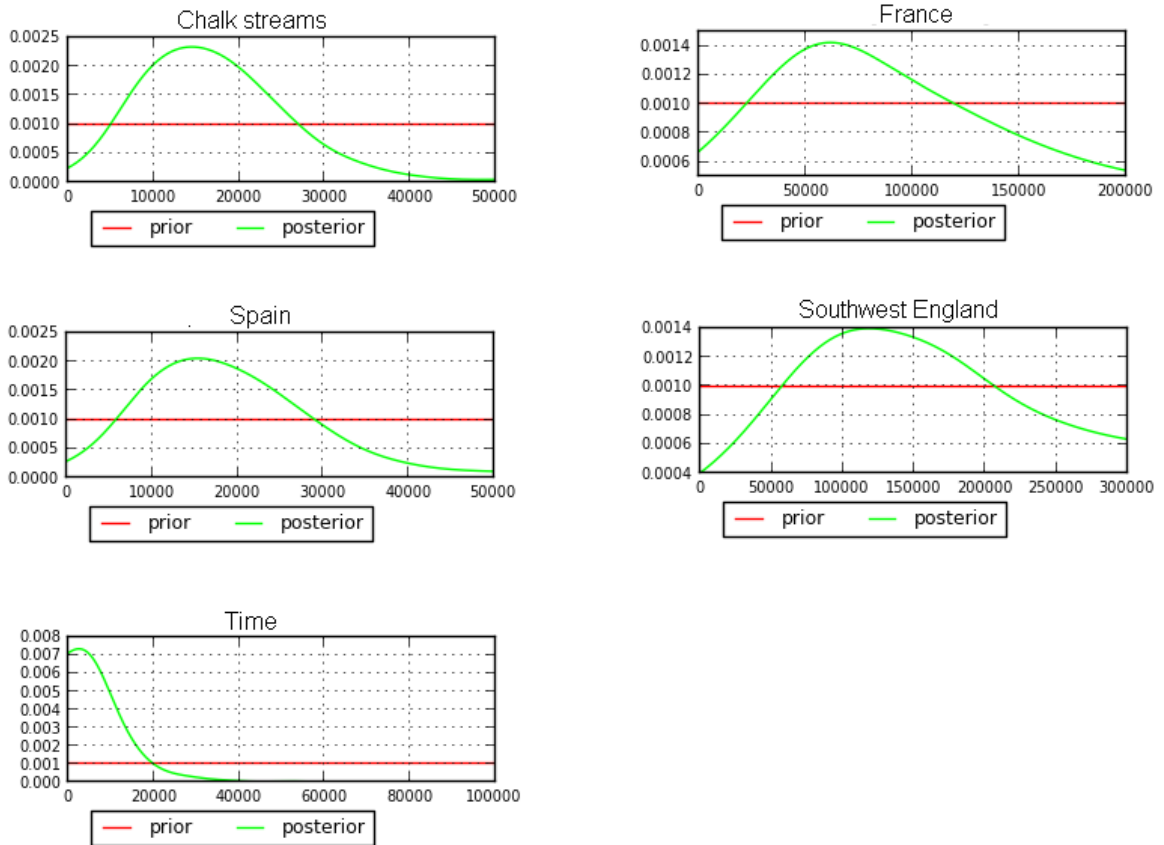
Figure 3.4- Posterior distribution graphs to show the effective population size (N_E) and time from preliminary DIYABC runs. Graphs indicate the posterior distribution of N_E of salmon from the chalk streams, Spain, southwest England and France from one run. The final graph indicates the posterior distribution of time from one run.
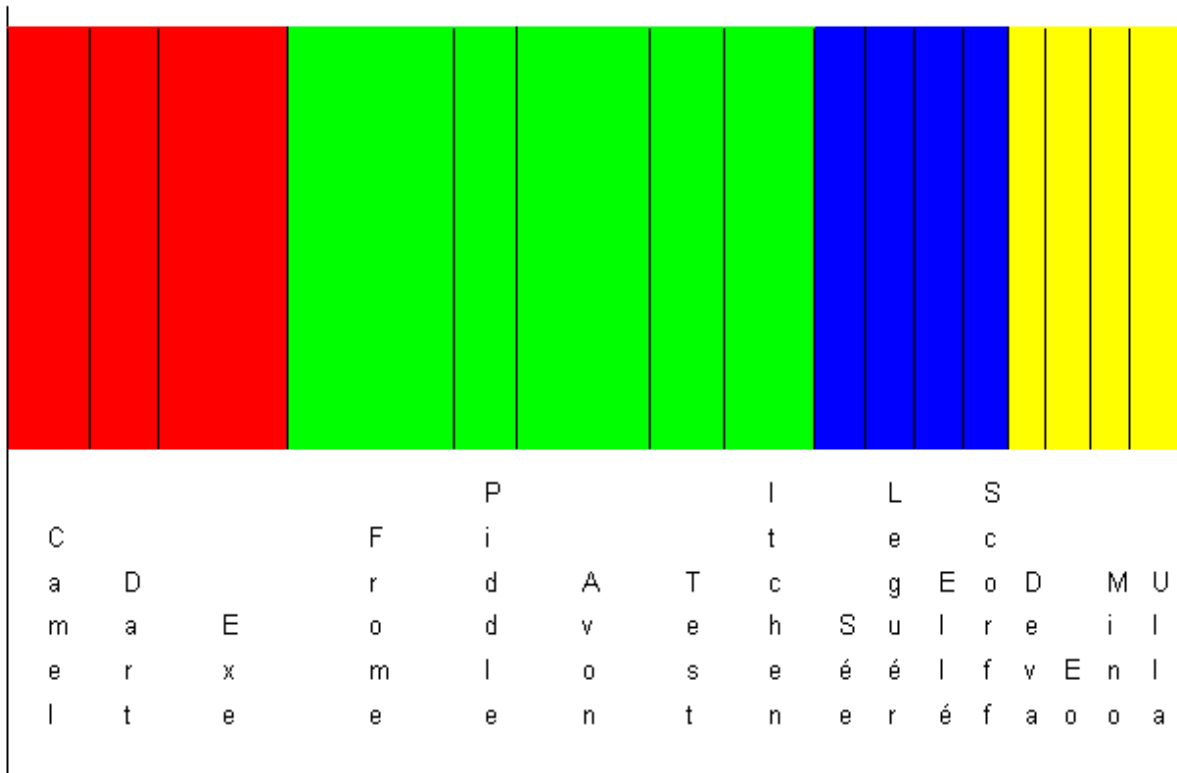
Figure 3.5- BAPS plot of salmon in southwest England, the chalk streams, France and Spain. Estimated assignment of each river population to a genetic group *k* where *k* = 4. Each river population is represented by a column separated by thin black bars, for which names are below the graphic.
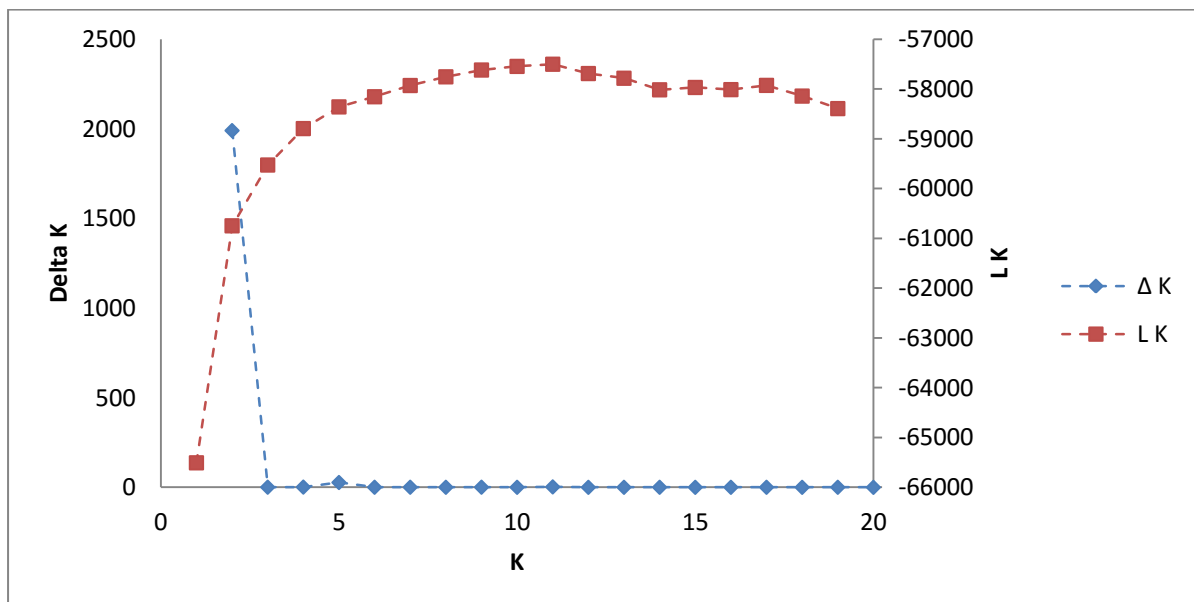
Figure 3.6- Graph of STRUCTURE Δ*k* and L*K* for the 16 salmon rivers. Mean Δ*k* and L*K* of over 20 runs for each *K* for 1 to 20 based on STRUCTURE analysis.

Figure 3.7- STRUCTURE plot of salmon in southwest England, the chalk streams, France and Spain. Each line is the estimated proportions of the coefficient of admixture of an individual's genome that originated from population k, for a) **top** k = 2, b) **middle** k = 4 alternative 1 and c) **bottom** k = 4 alternative 2. Each individual is represented by a column. Thin black bars separate individual rivers, for which names are given below the graphic. Hatched lines separate regions concluded from multiple analyses labelled within the box on
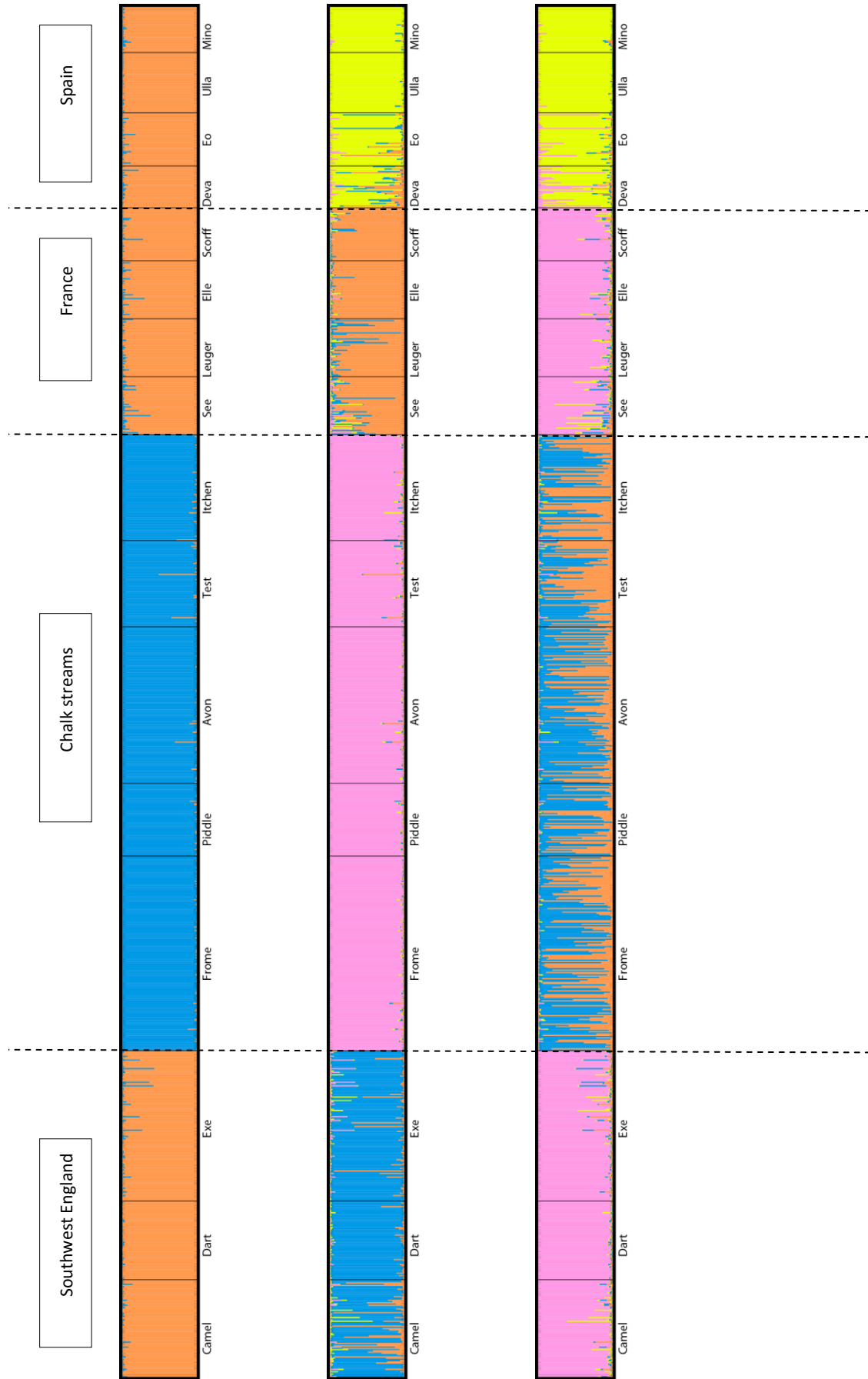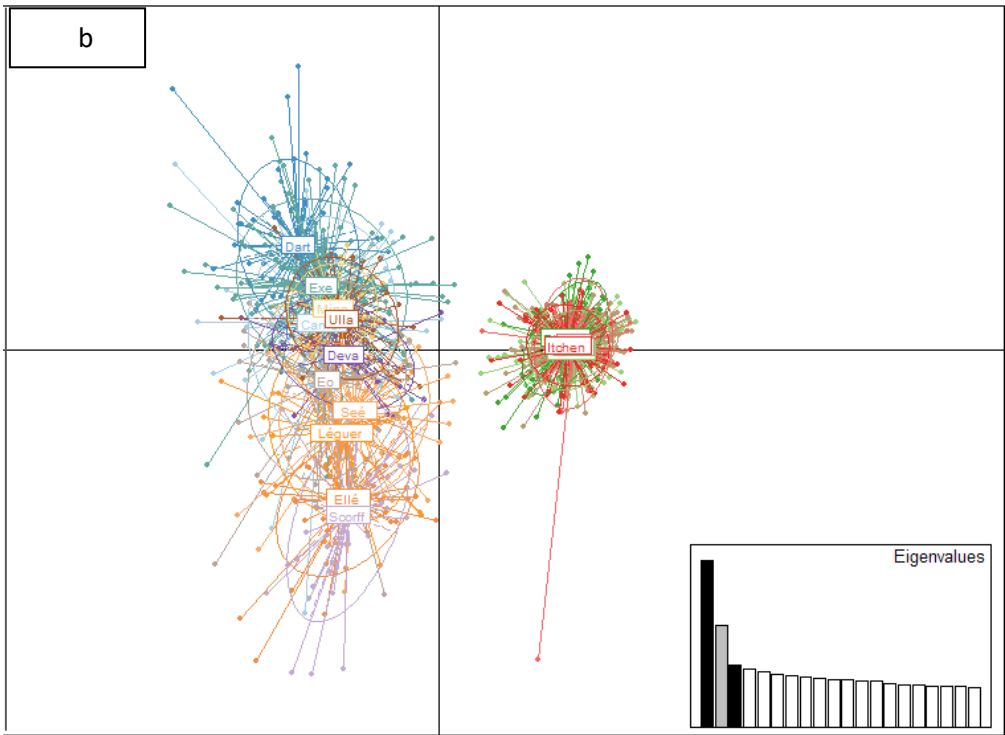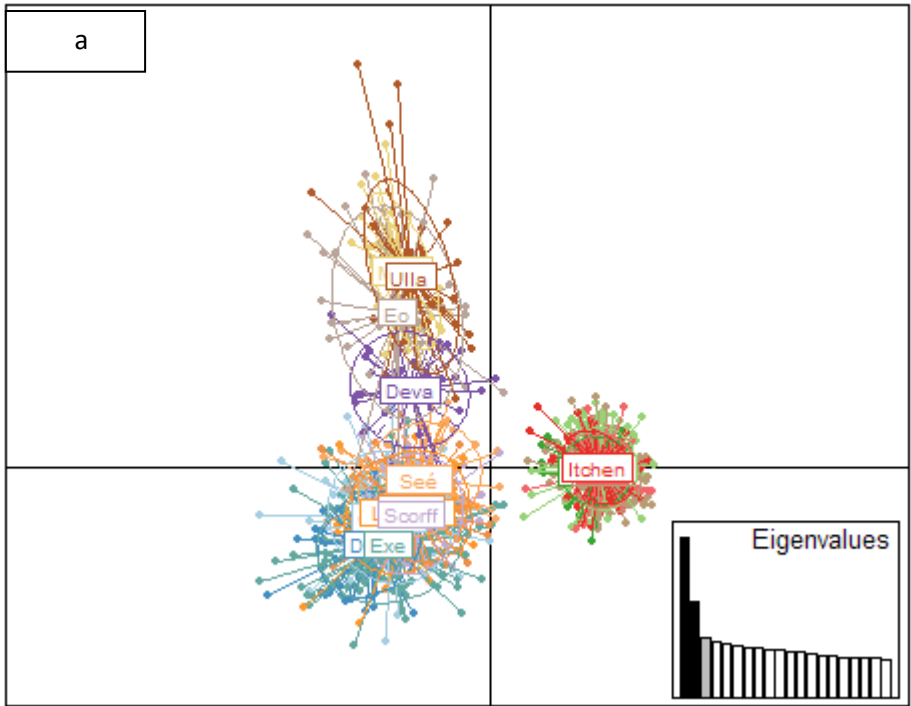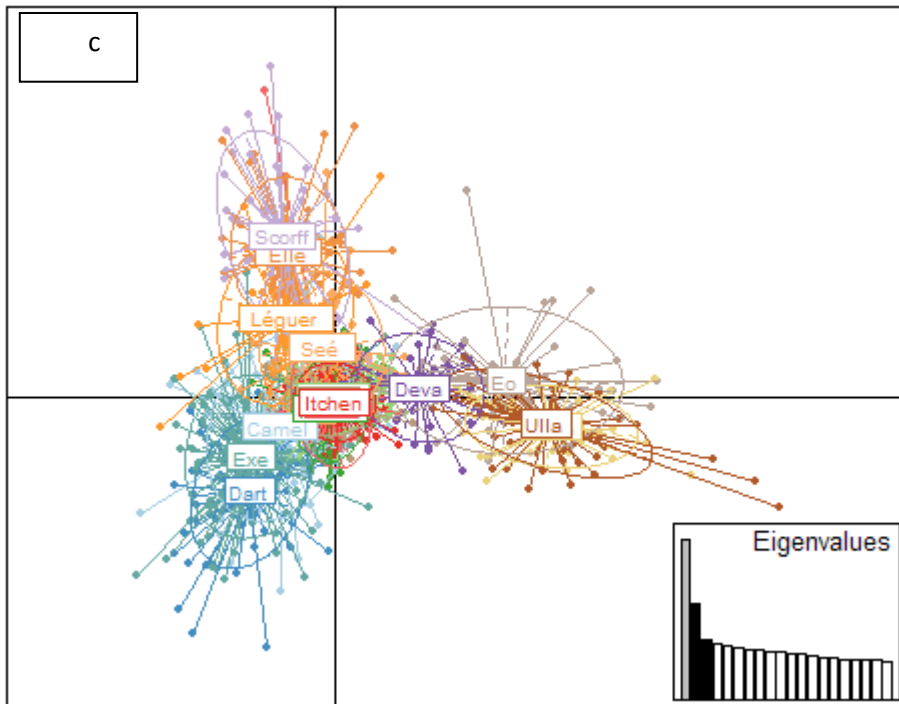
Figure 3.8- Principal component analysis of salmon microsatellite genotypes. a) Principal components 1 and 2. b) 1 and 3 and c) 2 and 3. Each dot represents an individual genotype and is coloured according to river sampled. Beneath the "Itchen" are the remaining chalk streams, the Frome, Piddle, Avon and Test and beneath the Ulla is the Miño (in yellow). The bar charts in the bottom right corner of each graph show the eigenvalue for the principal component analysis, and the black bars indicate which principal components are being displayed.
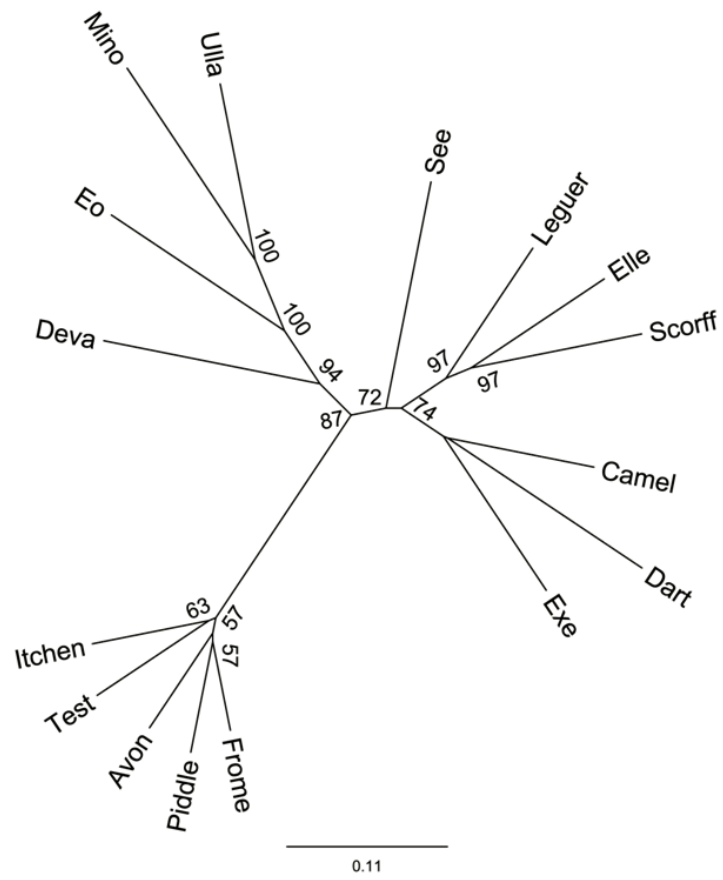
Figure 3.9- Neighbour joining dendrogram of salmon from 16 sampled rivers. Dendrogram is based on Cavalli & Edwards chord distance (D_{CE}) (Cavalli-Sforza & Edwards 1967).

a)
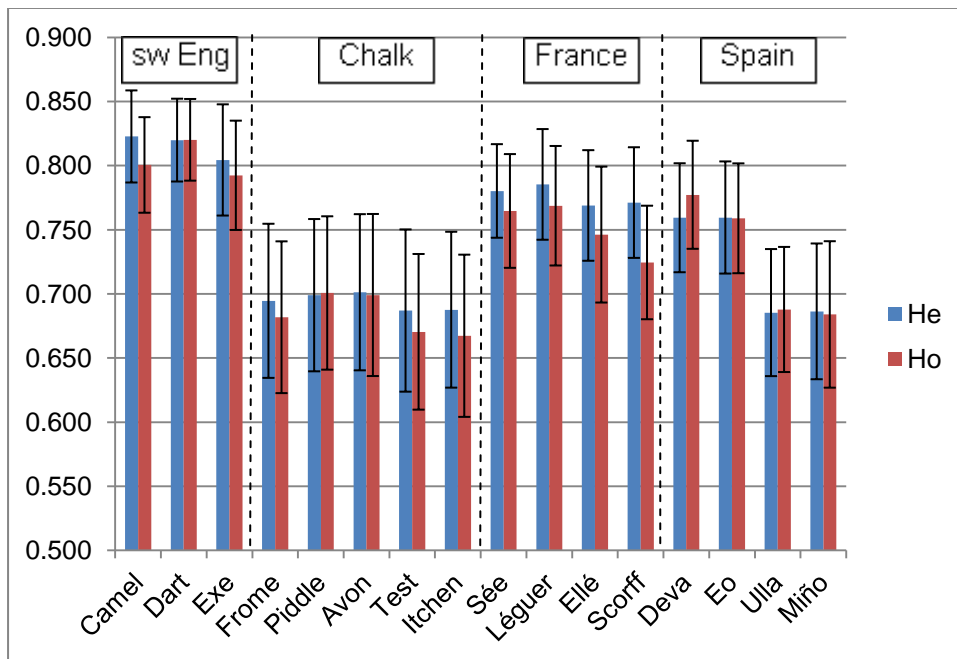
b)
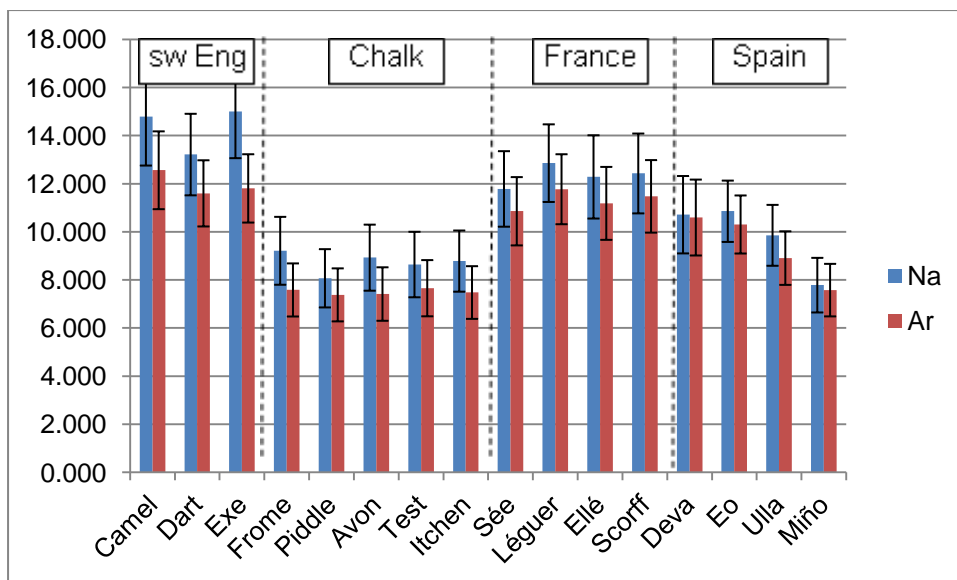
Figures 3.10- Graphs of a) expected heterozygosity (H_E) and observed heterozygosity (H_O) and b) number of alleles (N_A) and allelic richness (A_R) of salmon from the 16 rivers. Error bars indicate the standard error.
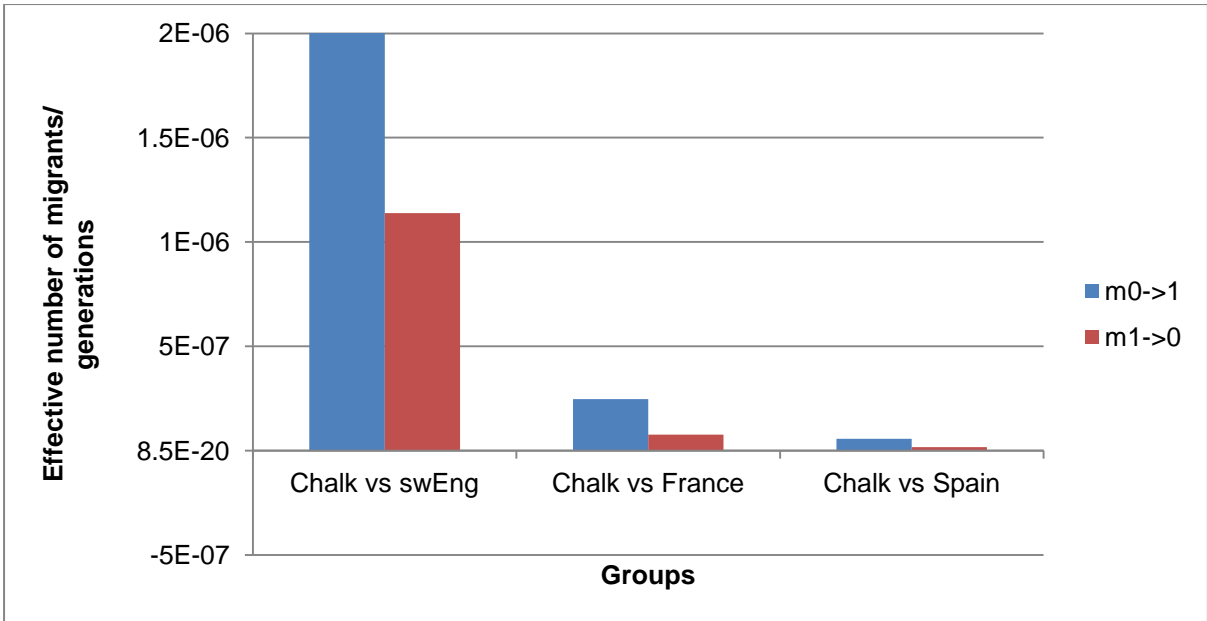
Figure 3.11- Historic number of effective migrants between groups per generation. Values calculated in IMa2 using mtDNA.   m0->1 indicates the number of migrants from the group on the left within each pair (0) to the group named on the right (1). m1 ->0 indicates the number of migrants from the right group to the left. Maximum value on the y axis has been limited for the purpose of clarity in the remaining calculations. * m Chalk -> swEng = 1.69 x 10^{-5}
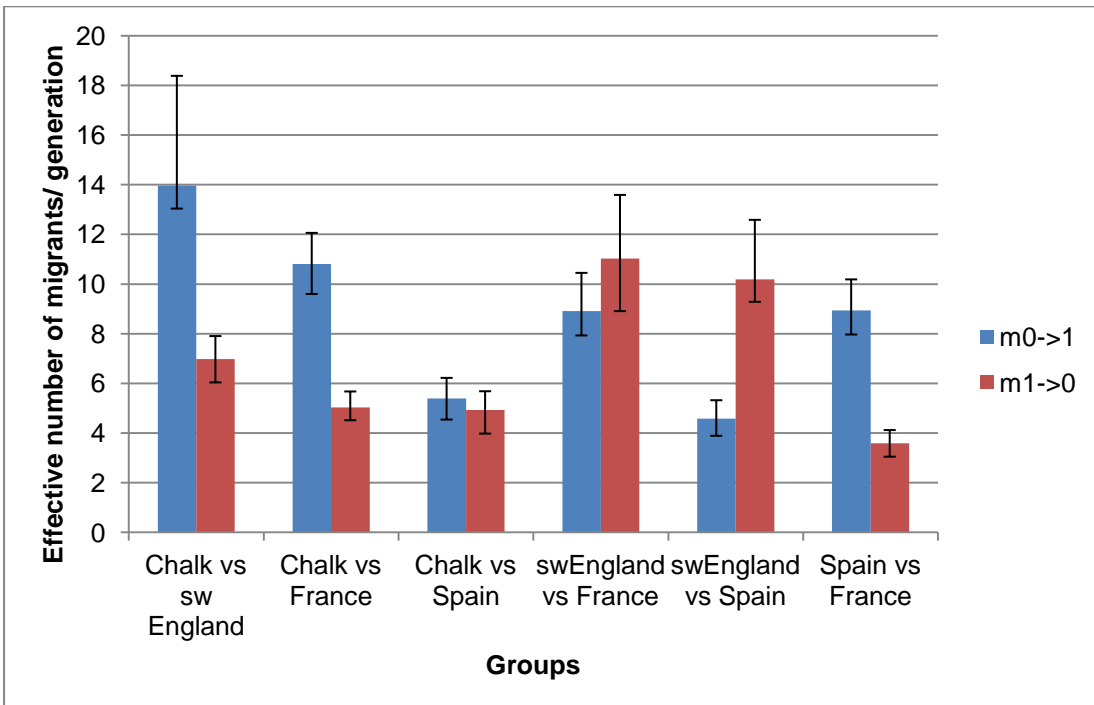
Figure 3.12- Number of effective migrants between groups per generation. Values calculated in MIGRATE using microsatellite DNA. m0->1 indicates the number of migrants from the first name of each pair (0) to the second name of each pair (1). m1->0 indicates the number of migrants from the second named group to the first.
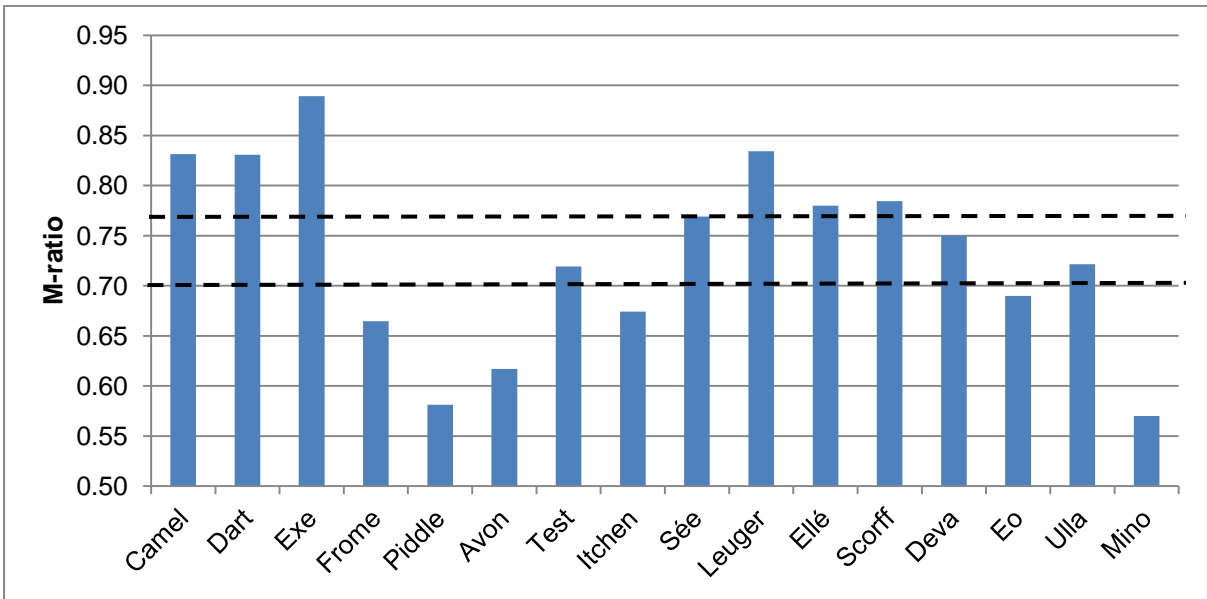
Figure 3.13- Graph to show the M ratio of salmon within each sampled river. Hatched horizontal line indicates the Critical M ratio at Θ = 0.2 (0.70) and Θ = 2 (0.77).
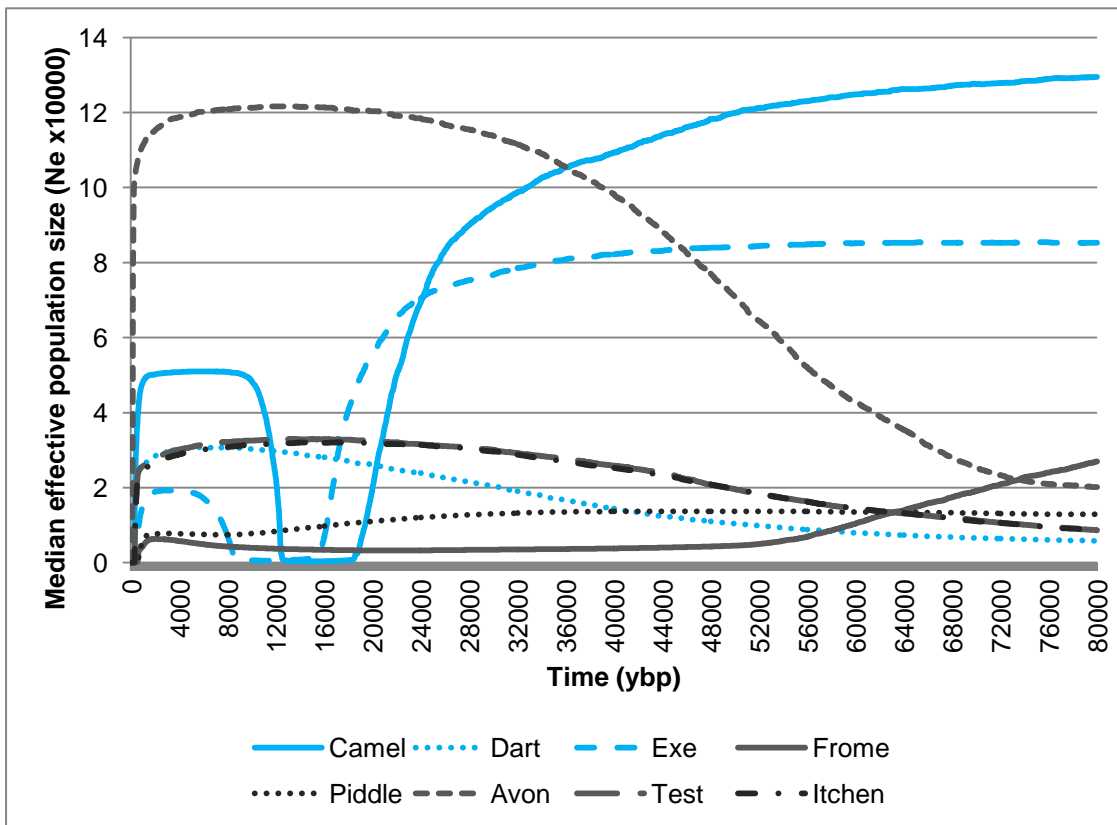
14ai)

14aii)
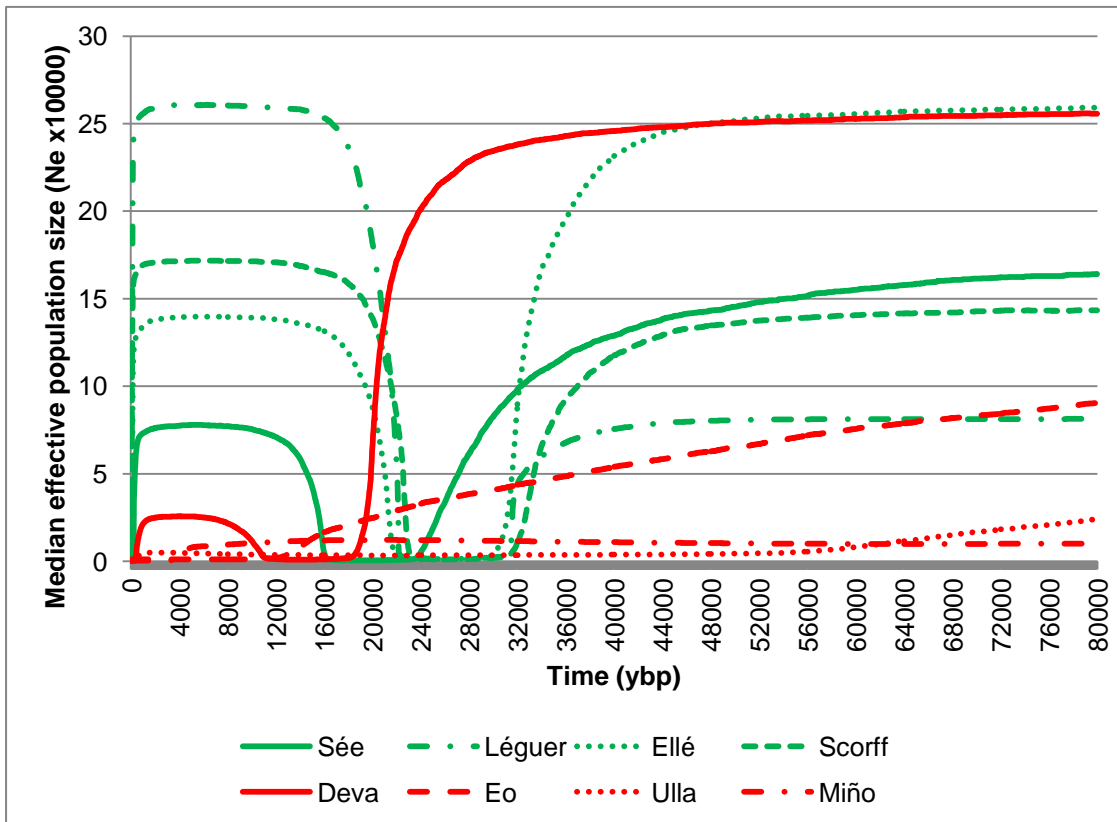
14bi)
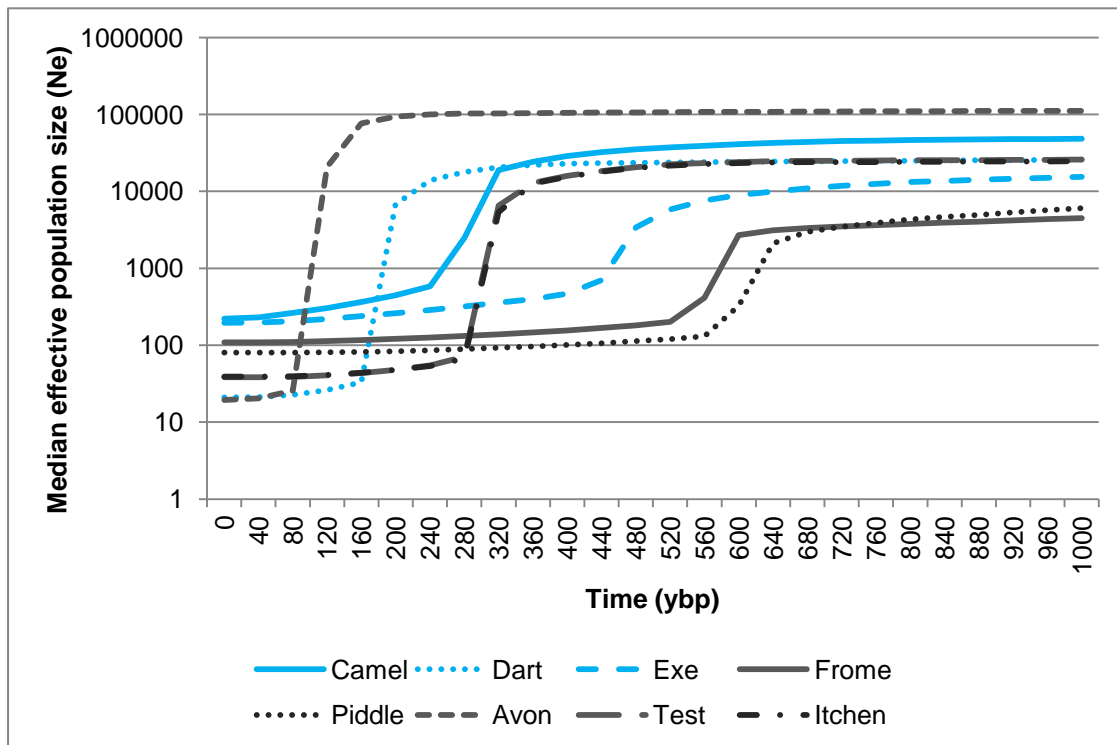
14bii)
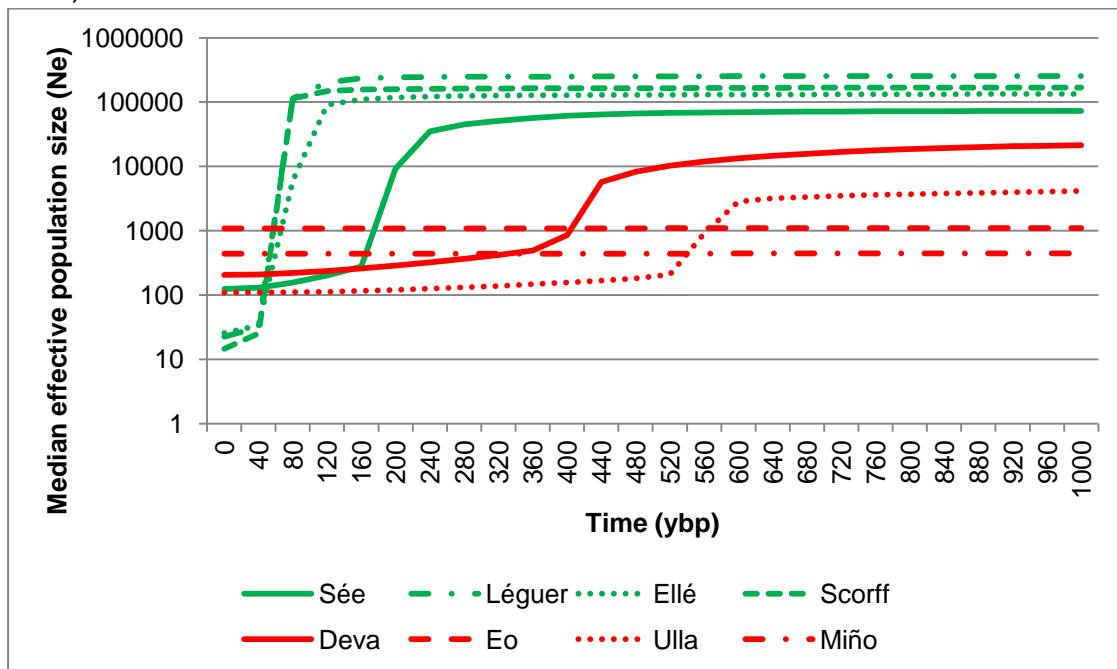
Figure 3.14- Past effective population size of salmon. Graphs indicate the median calculated effective population size of salmon populations during a) the past 80,000 years before present and b) the past 1,000 years before present, assuming a generation time of four years. Both a and b are split for clarity, into i) which contains rivers from southwest England and the chalk streams and ii) rivers from France and Spain.
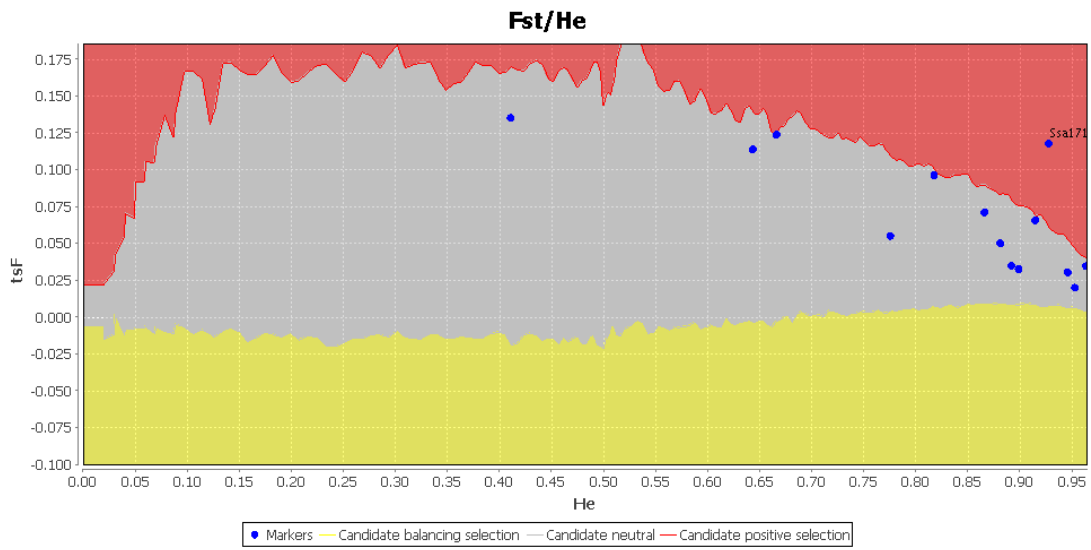
Figure 3.15- Graph of selective forces on each microsatellite locus. Result of analysis with Lositan showing whether or not each microsatellite locus is under selection positive selection (red area), balancing selection (yellow) or neutrality (grey). Only loci under significant selection are labelled.

a)

b)

c)
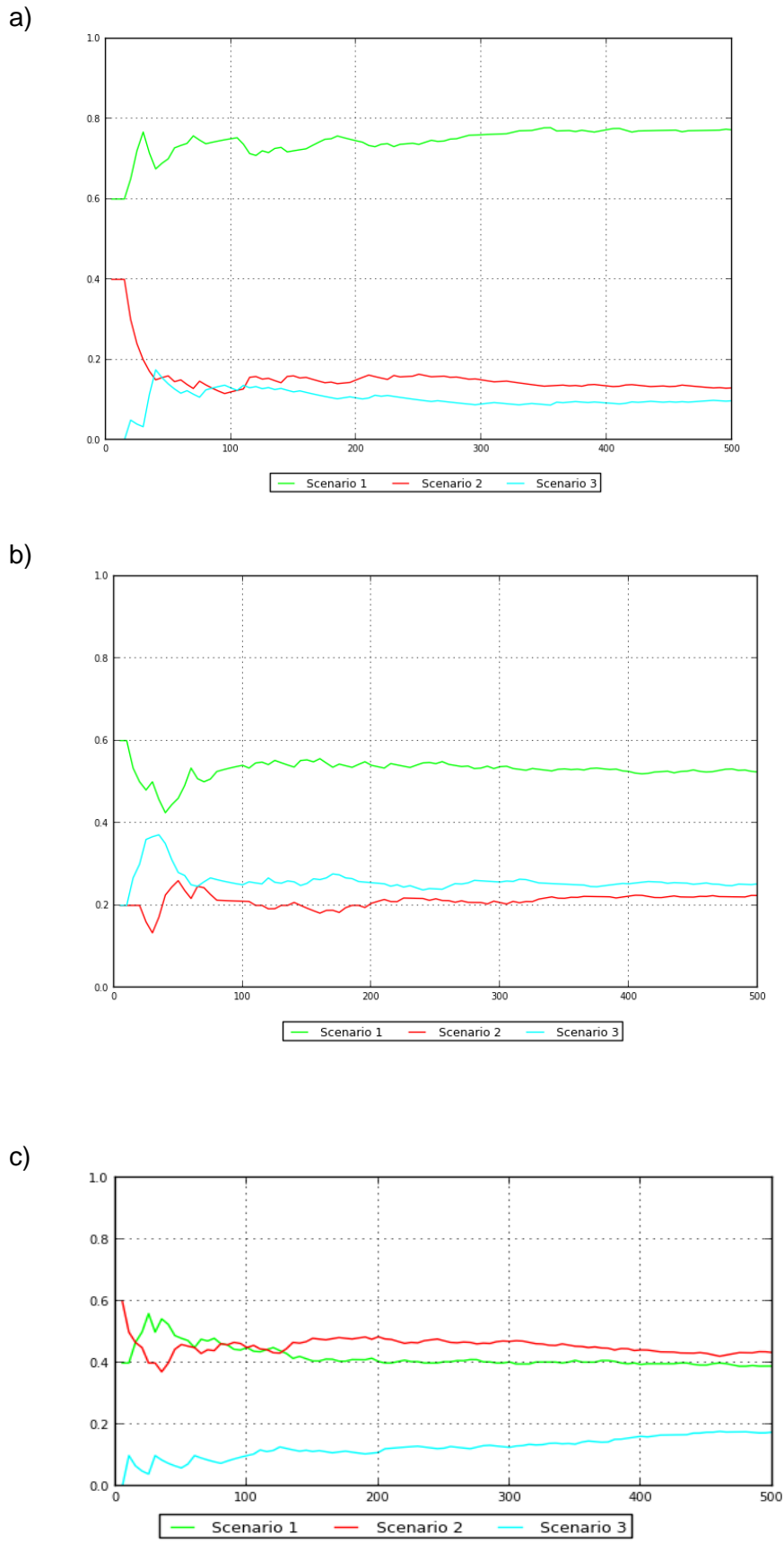
Figure 3.16- Graphs of direct likelihood of three scenarios in the pair-wise comparisons in DIYABC a) chalk vs southwest England, b) chalk vs France and c) chalk vs Spain.

Table 3.1- Details of samples from which microsatellite genotypes were obtained.

River	Year	No. Of sampling sites	No. of samples	After sibling removal	Life stage
Frome	2009	2	87	75	juvenile
	2011	2	104	83	juvenile
Piddle	2009	1	32	21	juvenile
	2011	2	89	38	juvenile
Avon	2004	2	42	39	juvenile
	2010	1	44	20	juvenile
	2012	3	100	68	juvenile
Test	2010	3	82	70	juvenile
Itchen	2004	1	27	26	juvenile
	2006	1	24	23	juvenile
	2010	1	46	37	juvenile
Camel	2005	3	88	79	juvenile
Dart	2005	1	44	35	juvenile
	2006	1	39	29	juvenile
Exe	2004	3	104	91	juvenile
	2005	1	38	31	juvenile
Sée	2005	1	47	47	adult
Leguér	2005	1	47	47	adult
Ellé	2005	1	47	47	adult
Scorff	2005	1	45	43	adult
Ason	2004	1	50	4	juvenile
Deva	2004	2	50	34	juvenile
Sella	2004	1	50	5	juvenile
Narcea	2004	1	50	18	juvenile
Eo	2004	1	46	43	juvenile
Mino	2004	1	47	37	juvenile
Ulla	2004	1	49	49	juvenile

Table 3.2- Table detailing the adjustment required to convert microsatellite data from the SALSEA Merge database to match the Exeter format.

Microsatellite marker	Exeter to SALSEA baseline	SALSEA to Exeter baseline
SSspG7	2	-2
Ssosl417	-2	2
Ssa202	-2	2
Ssa197	5	-5
SSsp2210	2	-2
Ssa289	5	-5
SSsp2201	2	-2
Ssol85	-4	4
Ssa14	-1	1
SsaF43	4	-4
SSsa2216	-1	1

Table 3.3- Number of each ND1 haplotype identified in salmon from each river. Each haplotype indicates a unique DNA sequence. Haplotype (H) numbers are defined in a previous study (Finnegan *et al.* 2013)

Region	River	H1	H2	H3	H4	H5	H6	H8	H14	H15	H16	H17	H20
Southwest England	Camel					5		2	2				
	Dart							7					1
	Taw			2		2		4	1		1	1	
	Usk		1			2			5	2			
English chalk streams	Frome	2		1				4		1			
	Piddle	2						3		2			
	Avon	5				1		1		3			
	Itchen					3		4		4			
	Test	3				2				1			
France	Ellé				3	5		1	2				
	Léguer			2	3	2		1	1				
	Scorff				4	3	1		2				
	Sée			4		1		1	1				
Spain	Ulla							6					
	Sella							2					

Table 3.4- Pair-wise F_{ST} values of salmon in 16 rivers. F_{ST} values (below diagonal) and significance values (above diagonal) determined from 999 permutations. Values in bold indicate no significant difference between the paired populations

	Ulla	Miño	Eo	Deva	Scorff	Ellé	Léguer	Sée	Itchen	Test	Avon	Piddle	Frome	Exe	Dart	Camel
Camel	0.055	0.053	0.032	0.027	0.019	0.019	0.016	0.018	0.044	0.045	0.041	0.041	0.042	0.011	0.012	-
Dart	0.057	0.054	0.035	0.032	0.028	0.029	0.023	0.027	0.052	0.054	0.052	0.052	0.054	0.020	-	0.001
Exe	0.067	0.064	0.044	0.035	0.028	0.028	0.022	0.032	0.051	0.053	0.049	0.047	0.047	-	0.001	0.001
Frome	0.078	0.074	0.058	0.053	0.042	0.043	0.041	0.051	0.011	0.009	0.008	0.008	-	0.001	0.001	0.001
Piddle	0.077	0.075	0.060	0.053	0.047	0.050	0.044	0.055	0.016	0.011	0.010	-	0.001	0.001	0.001	0.001
Avon	0.078	0.074	0.059	0.055	0.044	0.046	0.042	0.049	0.013	0.010	-	0.001	0.001	0.001	0.001	0.001
Test	0.079	0.073	0.059	0.055	0.046	0.048	0.045	0.056	0.008	-	0.001	0.001	0.001	0.001	0.001	0.001
Itchen	0.083	0.074	0.059	0.055	0.044	0.046	0.044	0.054	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sée	0.056	0.054	0.031	0.037	0.028	0.028	0.028	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Léguer	0.063	0.056	0.040	0.029	0.008	0.009	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Ellé	0.064	0.057	0.038	0.032	**0.006**	-	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Scorff	0.065	0.058	0.040	0.030	-	**0.569**	0.027	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Deva	0.028	0.027	0.024	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Eo	0.023	0.024	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Miño	0.019	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Ulla	-	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Table 3.5- 16 river self-assignment scores. The percentage of individuals that were correctly assigned to river and reporting region from which they were caught, and the river and reporting region that contained the highest proportion of wrongly assigned individuals. Region results display assignment including samples from the Sée (**above)**, and not excluding them (**below**).

River	Sample size	River Correct	Largest misidentification		Region correct	Largest misidentification	
Camel	70	67.10%	Exe	17.10%	96.70%	France	2.90%
Dart	58	81.00%	Camel	10.30%	98.30%	France	1.20%
Exe	113	80.50%	Camel	13.30%			
Frome	128	57.00%	Avon	15.60%	99.80%	France	0.20%
Piddle	44	50.00%	Frome	22.70%	99.80%	France	0.20%
Avon	107	57.90%	Frome	18.70%			
Test	66	47.00%	Frome	15.20%			
Itchen	67	71.60%	Avon	10.40%			
Sée	42	71.40%	Camel	9.50%	94.30%	SW Eng	5.10%
Léguer	46	56.50%	Ellé	15.20%	93.30%	SW Eng	6.70%
Ellé	47	53.20%	Scorff	19.10%			
Scorff	41	34.10%	Ellé	36.60%			
Deva	29	93.10%	Camel/Ulla	3.40%	96.60%	France	2.10%
Eo	40	92.50%	Ulla	5.00%	97.2%	SW Eng	2.80%
Ulla	46	78.30%	Miño	8.70%			
Miño	30	86.70%	Ulla	6.70%			

Table 3.6a and 6b- Calculated migration rates of salmon into rivers (across) and from which rivers (columns) using a) BayesAss and b) BIMr. Values on left indicate rates calculated without salmon from the river Sée, and values on right (in bold) indicate rates calculated without them.

a)

		Source							
		swEngland		Chalk		France		Spain	
Sink	swEngland	0.9795	**0.9872**	0.0026	**0.0026**	0.0137	**0.0059**	0.0041	**0.0042**
	Chalk	0.0007	**0.0007**	0.9979	**0.9979**	0.0007	**0.0007**	0.0007	**0.0007**
	France	0.0059	**0.014**	0.0022	**0.003**	0.9888	**0.9794**	0.0031	**0.0036**
	Spain	0.0112	**0.0125**	0.0032	**0.0034**	0.0251	**0.0214**	0.9605	**0.9627**

b)

		Source							
		swEngland		Chalk		France		Spain	
Sink	swEngland	1.00	**0.81**	0.00	**0.00**	0.00	**0.18**	0.00	**0.00**
	Chalk	0.00	**0.00**	1.00	**1.00**	0.00	**0.00**	0.00	**0.00**
	France	0.00	**0.00**	0.00	**0.00**	1.00	**1.00**	0.00	**0.00**
	Spain	0.03	**0.01**	0.00	**0.00**	0.05	**0.09**	0.92	**0.90**

Table 3.7- Table to show the significance of the Heterozygote excess test for population bottleneck calculated using the SMM and TPM mutation models.

River	SMM P	TPM P
Camel	0.9973	0.9824
Dart	0.9917	0.9877
Exe	0.9877	0.9406
Frome	0.8794	0.7292
Piddle	0.7684	0.4758
Avon	0.9031	0.8045
Test	0.9031	0.8521
Itchen	0.9137	0.8371
Sée	0.9877	0.9324
Léguer	0.9899	0.9852
Ellé	0.9979	0.9852
Scorff	0.9899	0.9791
Deva	0.8919	0.8302
Eo	0.9824	0.9547
Ulla	0.5151	0.311
Miño	0.9364	0.7929

Table 3.8- Estimated divergence dates (years before present) between four groups of salmon, as calculated by IMa2 using ND1 sequence data. Dates are calculated using two mutation rates: 5.7 x 10^{-9} (slow) and 1.537 x 10^{-8} (fast) substitutions per site per year. HPD95Lo indicates the lower bound of the estimated 95% higher posterior density (HPD) interval. HPD95Hi indicates the higher bound of the of the estimated 95% HPD interval. #? indicates an unreliable estimate.

Groups compared	Mutation rate	Mode	Mean	HPD95Lo	HPD95Hi
Chalk vs swEng	Slow	20614	36271	1140	90965
Chalk vs France	Slow	14474	26112	964.9	62193
Chalk vs Spain	Slow	14825	62665	789.5 #?	153947 #?
Chalk vs swEng	Fast	6539	13379	422.9	33409
Chalk vs France	Fast	5498	9717	711.6	36089
Chalk vs Spain	Fast	3155	22000	0#?	56441#?

Table 3.9- Estimated divergence dates (ybp) between paired groups, calculated by DIYABC from microsatellite data. Only the most likely of each scenario is shown.

Scenario	Summary statistics	Mean	Median	Mode	q050	q950
chalk colonise southwest England	Beaumont	34880	25600	15640	5960	92800
chalk colonise France	Beaumont	27120	17360	7840	3272	83600
Spain colonise chalk	Beaumont	36480	21600	11160	3768	125200
chalk colonise southwest England	Guillemaud	62400	52800	28320	16400	144000
chalk colonise France	Guillemaud	36240	29200	18400	8440	84400
Spain colonise chalk	Guillemaud	35240	26320	19160	6840	91600

Chapter 4: Genetic analysis indicates marked population structure of Atlantic salmon (Salmo salar L.) among the chalk streams of southern England

Introduction

The Atlantic salmon (*Salmo salar* L.) is a species of anadromous fish, which returns to its natal spawning grounds after reaching sexual maturity. Subsequently, the species shows marked sub-specific structuring into broad geographic groups, which is readily detectable using genetic methodologies (Stahl 1987; Verspoor *et al.* 2005) and in particular, analysis of microsatellite genotypes (e.g. King *et al.* 2001; Koljonen *et al.* 2005; Tonteri *et al.* 2009; Griffiths *et al.* 2010). Current research suggests that broad genetic groups are largely defined by a combination of geological substrate (Grandjean *et al.* 2009; Perrier *et al.* 2011), phylogeography (Finnegan *et al.* 2013) and environmental factors (Dillane *et al.* 2007), leading to the suggestion that salmon populations may be locally adapted to their in-river environment (Garcia de Leaniz *et al.* 2007; Griffiths *et al.* 2011; Perrier *et al.* 2011).

One sub-group of the species, within the chalk streams of southern England, were shown in a recent study (Griffiths *et al.* 2010) to be genetically distinct when compared to their geographical neighbours in non-chalk rivers in Britain, northwest France and Spain. In another study (Ikediashi *et al.* 2012), chalk stream salmon were also found to have a relatively low level of admixture with salmon in neighbouring regions. Admixture has for a long time been associated with a reduction in population differentiation; for example, Stahl (1987) calculated that in order to maintain genetic differences between two or more Atlantic salmon sub-populations of 2,500 to 10,000 fish, there had to be less than one migrant between them per year (Stabell 1984). Also, in more recent studies in Spain (Ayllon *et al.* 2006) and the Baltic Sea (Vasemägi *et al.* 2005), a reduction in the between-river population structure of salmon has been identified as a result of admixture with salmon farm escapees. Following this line of argument it is possible that chalk stream salmon, which have relatively little admixture with salmon in neighbouring regions (Ikediashi *et al.* 2012), may also have an increased level of population structure, compared to other populations in Europe. However, despite their

distinctiveness (Griffiths *et al.* 2010) and use in several studies (e.g. Child *et al.* 1976; Jordan & Cross 2005; Finnegan *et al.* 2013), the degree of population structure within and between the chalk streams had not yet been explored.

The reason for the distinction of these chalk stream salmon most likely stems from one or more of their unique abiotic factors, which are described in detail by Berrie (1992). The calcareous substrate, upon which chalk streams are formed, is porous, and thus chalk streams are aquifer fed. Thus, the water is relatively clear, stable in temperature throughout most of the year, and alkaline (ca. pH 8). Because of their unique assemblage of organisms (i.e. Bullheads, brook lamprey, sea lamprey and salmon), several chalk streams have been designated SSSIs (Site of Special Scientific Interest). However, despite spanning much of the east and southern coast of England (Environment Agency 2004), major salmon populations are present in just five. These are the rivers Frome, Piddle, Avon, Test and Itchen, which have each been sampled for the purpose of this study (Figure 4.1), and are henceforth referred to by their names only. Crucially, although the 161 chalk streams are located between Yorkshire in northeast England and Dorset in south England, the five with salmon span just ca. 70 km along the southern English coast. With so few chalk stream salmon populations, each of which has plummeted in recent decades (Environment Agency 2004), there is extra incentive to understand the extent of their local population genetic structure in this region. We anticipate that this information will be useful for the successful management and conservation of salmon within these rivers.

The aims of this study were two-fold; the first aim was to elucidate the population structure of Atlantic salmon among the five major chalk streams of southern England: the Frome, Piddle, Avon, Test and Itchen. This was assessed by testing for isolation by distance of salmon populations between the rivers, population differentiation and significant differences in the most commonly used indices of genetic diversity. We then explored the temporal stability of genetic profiles of salmon within the Avon via the assignment of salmon sampled from the Avon during 1986, 1987 and 1989. The second aim was to elucidate the population structure of salmon within the chalk stream catchments. To do this, samples collected throughout the salmon carrying part of the Frome catchment in 2009, and again in 2011, were

analysed. Spatial investigations were made of isolation by distance and population subdivision. Finally, the study tested for temporal divisions between the two years and differences between them in genetic diversity.

Methods

Sampling

Juvenile salmon (0+ parr) were sampled for fin clips from the five chalk streams of southern England still containing significant salmon populations - the Frome, Piddle, Avon, Test and Itchen (Figure 4.1). The Avon, Itchen and Test were sampled by the Environment Agency during routine national surveys and management programmes between 2004 and 2012. Sampling of the Frome and Piddle was carried out by the Game and Wildlife Conservation Trust in September 2009 and 2011 (Figure 4.1) during routine juvenile abundance surveys. In all cases, fin clip samples were obtained from salmon via electrofishing, which conformed to national agency ethical guidelines. A maximum of 50 parr samples were targeted at each sample site. Sample sites from the Frome included in this study were selected in order to maximize the coverage along the river and attempts were made to sample the same sites during both 2009 and 2011. Additionally, to test the temporal stability of salmon populations in the Avon, scale samples taken from adult salmon in 1985, 1986 and 1989 were also included in the analysis. In total, 1,261 juvenile salmon samples were genotyped at 16 microsatellites across 31 sampling sites in five rivers (Table 4.1) and an attempt was made to genotype 93 historical Avon samples.

Genetic data collection

DNA was extracted from fin clips using the HOTshot method (Truett *et al.* 2000) and DNA from scales was extracted using a Chelex method (Estoup *et al.* 1996). Sixteen microsatellite loci were genotyped. This followed the protocol by Ikediashi *et al.* (2012) to amplify the fourteen following loci: Ssa14 (McConnell *et al.* 1995); Ssa202, SSsp3016, Ssa197 (O'Reilly *et al.* 1996); SsaF43 (Sánchez *et al.* 1996); SSspG7, SSsp1605, SSsp2210, SSsp2201, and SSsp2216 (Paterson *et al.* 2004); Ssa171, Ssa289, Ssa157, and SsaD144 (King *et al.* 2005). Two additional loci, Ssosl85 and Ssosl417 (Slettan *et al.* 1995), were added to the first multiplex reaction described by Ikediashi *et al.* (2012). Potential salmon x trout hybrids were recognised by the presence of alleles larger than 350bp for locus SSsp1605, or alleles larger than 135bp for Ssa289. These fish were subsequently removed from the dataset.

Data checking

MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004) was used to check for scoring errors due to stutter peaks, large allele dropout or null alleles. In order to prevent the false detection of population structure due to the presence of family groups (Allendorf & Phelps 1981), the program COLONY v2.0.4.1 (Jones & Wang 2010) was used to identify full siblings. The mating system was defined as polygamous for males and females and without inbreeding; as the inbreeding setting is advised against unless inbreeding is considerably high (Jones & Wang 2010) Each run was of medium length, with high precision and using the full-likelihood method. Allele frequencies were not updated during the run and no prior sib-ship was assumed. An error rate of 0.02 was used for each locus based on analysis by Ellis *et al.* (2011a). COLONY was run twice independently, with different starting seeds to check consistency of the reconstruction. Full-sib families were reduced to one representative, if supported by an average likelihood of 50% or higher between the two runs. After this, all sample sites with fewer than 20 individuals were also removed from the dataset, as it is widely believed that small sample sizes (<20) bias estimates of population differentiation (Holsinger & Weir 2009; Willing *et al.* 2012). Linkage disequilibrium and deviations from Hardy-Weinberg equilibrium were detected using GENEPOP v4.2 (Raymond & Rousset 1995). The 95% significance level of each was adjusted using the False Discovery Rate (FDR) method (Benjamini & Hochberg 1995).

Descriptive statistics

The number of alleles (N_A), expected heterozygosity (H_E) and observed heterozygosity (H_O) were calculated using Genalex v6.5.0.1 (Peakall & Smouse 2012) for each sampling site. In order to account for differences in sample size between samples sites, the allelic richness (A_R) at each site was also calculated using the program Fstat v2.9.3 (Goudet 1995).

In order to determine whether there were any significant differences in genetic diversity between the five chalk streams, an analysis of variance (ANOVA) was implemented within R using the following equation (Hamilton *et al.* 2014):

```
x1<-glm(x~River+Locus)
x2<-update(x1,~.-year)
anova(x1,x2,test="F")
```

where glm = generalized linear model and x was either A_R, H_E, H_O, or F_{IS}. The results of each microsatellite loci were used -after obtaining them from Genalex-, rather than the average, in order to increase statistical power. If the ANOVA was significant, then the Tukey test was used to determine between which two years the significance lay. This was done using the following equation:

facriver<-factor(River)
anov1<-aov(x~facriver+Locus)
TukeyHSD(anov1,"facriver")

Population structure between chalk streams

In order to determine whether salmon in each chalk stream were significantly different from each other a test for significant differences in allelic frequencies was completed using GENEPOP on the web v4.2 (Raymond & Rousset 1995). Salmon were grouped by the river from which they were sampled. Option 3, sub-option 2 was selected to infer whether alleles frequencies were equal between paired populations. Option 3, sub-option 4 was also selected to infer whether the distribution of genotypes was equal between paired populations. Pair-wise F_{ST} values were calculated for salmon with each sample site and tested for significance with 999 permutations using Genalex. The FDR method was used to adjust the 95% confidence level (Benjamini & Hochberg 1995). In order to visualise the genetic distances between sample sites, a principal coordinate analysis (PCoA) of pair-wise F_{ST} was conducted in Genalex. To test whether the river populations were structured through a pattern of isolation by distance, Rousset (1997) genetic distance ($F_{ST}/1-F_{ST}$) was tested for significant correlation with geographic distance using a Mantel test (Mantel 1967) in Genalex. Geographic distances were determined between river mouths along the coastal line of southern England using arcGIS v10 (ESRI 2006).

The program STRUCTURE v2.3.4 (Pritchard *et al.* 2000) was used to identify the number of distinct genetic units (*k*) within the chalk-stream salmon populations. STRUCTURE attempts to cluster individuals into groups that are in Hardy-Weinberg and linkage equilibrium. STRUCTURE was run with the LOCPRIOR model, in order to detect weak population structure (Hubisz *et al.* 2009), from *k* = 1 to *k* = 10 with 250,000 Markov Chain Monte Carlo (MCMC) replicates, after a burn-in of 50,000

from ten independent starting points. The Evanno method (Δk: Evanno *et al.* 2005) was primarily used to determine the optimum number of genetic units (k) from the STRUCTURE results. However, because the Evanno method looks at the change in the likelihood score between runs, it is unable to include $k = 1$ in the comparisons. As a solution to this, the Likelihood method, recommended by Pritchard *et al.* (2000), was also used to check the likelihood of one genetic unit ($k=1$).

Spatial and temporal analysis of population structure in the Frome

In order to investigate population structure of Atlantic salmon within the chalk streams, sample sites from the Frome, a river that was sampled intensively during 2009 and 2011, were subject to the following additional analyses. The program STRUCTURE was used to identify the number of genetic units within each of these year classes, as described above. Because STRUCTURE attempts to cluster individuals (rather than groups of individuals), sampling sites comprising fewer than 20 fish were also included in the analysis. The number of breeders (N_B) and the effective population size (N_E) at each sample site along the Frome were also identified using the sibling method (Wang 2009) within the program COLONY, which could indicate possible barriers to migration. Statistical comparisons were made, where possible, using a two-sided t-test. The Mantel test was used to test for isolation by distance within each year, also as described above. Geographic waterway distances between sampling sites (Figure 4.1), which were calculated using arcGIS v10.

In order to investigate differentiation between temporal cohorts on the Frome, the following tests were performed. First a PCoA of F_{ST} values were calculated for all Frome 2009 and Frome 2011 sites. Secondly, differences in A_R, H_O, H_E, and F_{IS} between Frome 2009 samples and Frome 2011 samples were tested for significance using the ANOVA test described previously. As there were only two rivers, there was no need to implement a Tukey test following a significant finding.

Results

Number of individuals

In total, 1,261 juvenile salmon samples were genotyped at 16 microsatellites across 31 sampling sites in five rivers (Table 4.1). Two potential salmon x trout hybrids were detected within the Frome and five were detected within the Avon. After the removal of siblings and sampling sites containing fewer than 20 individual genotypes, the dataset was reduced to 724 samples from 25 sites (Table 4.1). Fifty-eight of 93 historical Avon samples amplified reliably with at least 12 of the 16 microsatellite loci and 55 individuals remained after removing full siblings. As the 1987 dataset contained only 10 samples, samples from 1986 and 1987 were combined to form a group of 32, while the 1989 group included 23 samples.

After applying the false discovery rate (FDR) correction, linkage disequilibrium was detected at seven out of a total of 3,000 comparisons. These were not consistent between sample sites, therefore no loci were removed. Across the 25 sample sites, after FDR correction only two cases of loci out of Hardy-Weinberg equilibrium were found, therefore no further sample sites were removed.

Genetic diversity

Between all sample sites (See Table 4.2 for details and key), the observed heterozygosity (H_O) of juvenile salmon ranged from 0.74 in AVNbrd04, to 0.64 in TSTbro10. Expected heterozygosity (H_E) ranged from 0.71 in FROgbc11 to 0.65 in TSTbro10. Allelic richness (A_R) ranged from 5.34 in FROlm09 to 4.77 in TSTbro10. Statistical comparisons made between all five rivers were non-significant in A_R, H_E or F_{ST} (Table 4.3). However, there were significant differences in H_O ($P < 0.047$). Further analysis indicated that the greatest statistical difference lay between salmon within the Piddle and Itchen (Tukey F, p = 0.061) , however no individual pair-wise comparison was significant (Tukey F, p > 0.05).

Structure between the chalk streams

The average pair-wise F_{ST} value between salmon in all chalk-stream sampling sites was 0.023, and ranged from 0.041 (between the salmon from the AVNbut12 and TSTbro10), and 0.010 (between salmon from FROcfmr11 and FROeb11 and also between FROcfmr11 and FROesg09) (Appendix VIII). All pair-wise F_{ST} comparisons were significant both before and after FDR correction, except between two pairs of salmon sites- sampling sites AVNbri12 and AVNbrd04, and ITCbis05 and ITCbis06.

The overall results for the allelic and genotypic pair-wise comparisons were "highly significant" in almost all comparisons (Chi^2 = infinity, df = 28, P = "Highly significant". This indicates that at for at least one of the microsatellite loci, a P value of 0 was inferred(Raymond & Rousset 1995). The exceptions were the same for both test, and were the Frome vs Piddle (Gene Chi^2 = 96.02532, df = 28, p = 0) (allelic Chi^2 = 90.94986, df = 28, p = 0), the Ellé vs Leguér (p = 0.000001, and 0.000002 respectively), Scorff vs Leguér (p = 0 and 0.000009 respectively), and the Ellé vs Scorff (p = 0.003123 and 0.027649 respectively), which were all still signficantly different.

Within the principal coordinate analysis (PCoA), the first and second axes indicated three possible groups (Figure 4.2). These groups consisted of salmon in 1) the Frome and Piddle, 2) the Avon and 3) the Test and Itchen. From the STRUCTURE analysis, the likelihood method identified that there was more than one genetic unit (Figure 4.3). The Evanno method (Δk) identified three genetic units as the optimum (k = 3; Figure 4.3), which were the same as those identified by the PCoA (See Figure 4.4a). The Mantel test between salmon within the five rivers proved significant (R^2 = 0.475, P = 0.03, Figure 4.5), indicating a significant pattern of isolation by distance.

Spatial and temporal analysis within the Frome

The likelihood scores from the STRUCTURE analyses indicated that salmon in the Frome constituted more than one genetic unit during both of the year groups analysed (Figure 4.6a and 4.6b). For the 2009 Frome samples, the Evanno method

(Δk) identified three genetic units ($k = 3$; Figure 4.6a) as optimum, while for the 2011 Frome samples, two units ($k = 2$; Figure 4.6b) were identified as optimum.

For the three units identified within the 2009 Frome salmon (Figure 4.4b), the mean number of breeders (N_B) was 25.5, 34, and 37 respectively, while the effective population size (N_E) was 15, 28, and 38 respectively. Comparison between the first unit and the other two combined were significant for both indices (two-sided t = -3.75 and -3.55, P = 0.01 and 0.02 respectively, df = 5). For the two units identified within the 2011 Frome salmon (Figure 4.4c), mean N_B was 24 and 34.4 respectively and the mean N_E was 10 and 25.6 respectively (excluding the sample sites with fewer than 20 samples). However, there were insufficient samples in the first unit, which consisted of just Bradford Peverell, to make the same statistical comparison.

In the test for isolation by distance, the Mantel test was non-significant for both the 2009 (R^2 = 0.084, P = 0.150) and the 2011 Frome samples (R^2 = 0.201, P = 0.079), indicating no evidence for isolation by distance within the river.

In the temporal comparison between cohorts, there was a significant difference in allelic richness between the cohort of Frome 2009 (A_R = 9.078) and Frome 2011 salmon (A_R = 8.280) (ANOVA, p = 0.048, Table 4.4). All other indices calculated were not significantly different between the two years (ANOVA p > 0.05). The principal coordinate analysis of salmon from all Frome sites from 2009 and 2011, suggested a segregation of Frome 2009 salmon away from the Frome 2011 populations (Figure 4.7) with a small degree of overlap. This suggests a slight genetic difference between the two cohorts.

Discussion

Overview

Populations of Atlantic salmon within the chalk streams of southern England have plummeted in recent decades, yet despite this and their distinction from other European populations (Griffiths *et al.* 2010), the genetic population structure within the chalk streams had previously not been investigated. This study revealed that for salmon in the five chalk streams, there is a significant pattern of isolation by distance and sub-division of fish into three sub-groups. These sub-groups consisted of salmon in 1) the Frome and Piddle, 2) the Avon, and 3) the Test and Itchen. Identification of these sub-groups significantly increases our understanding of the contemporary genetic structure within one of the key reporting regions identified by Griffiths *et al.* (2010) for Atlantic salmon in the southern part of the species' range. Further sub-division was also found within the Frome during two separate years, which to our knowledge, has not been identified previously in a river this small (ca. 48 km). This has significant implications for conservation and our understanding of salmon population structure.

Between-river population structure

Our analyses identified population subdivision of chalk stream salmon into three groups. These groups comprised salmon in 1) the Frome and Piddle, 2) the Avon, and 3) the Test and Itchen (Figure 4.2 and Figure 4.4a). These results also indicated a closer genetic relationship of salmon from the Avon to salmon from the Frome and Piddle, compared to their relationship with salmon from the Test and Itchen. As a significant pattern of isolation by distance was detected across the region, this similarity is most likely due to the closer proximity of the Avon to the Frome and Piddle (ca. 30 km coastal distance between estuaries), as compared to the distance between the Avon and the Test and Itchen (ca. 70 km). Isolation by distance has previously been reported between rivers in other locations, for example, Canada (Dionne *et al.* 2008a) and France (Perrier *et al.* 2011). These indicate that isolation by distance between rivers may be the norm under certain conditions.

Comparisons in genetic diversity showed that most indices were not significantly different between the rivers (Table 4.3). Only observed heteroyzgosity proved signifcantly different (ANOVA, p < 0.047), however comparisons between

each pair of rivers proved non-significant (Tukey, p > 0.05). Therefore genetic diversity in the five chalk streams is considered to be the same.

Within-river population structure

For salmon within the Frome, population sub-division was identified in both the 2009 and 2011 cohorts (Figures 4.4b and 4.4c). The difference in the number of subunits identified in each year group ($k = 3$ and 2, respectively) may seem unusual, but is readily explainable and accords with a metapopulation model of population substructure, as first identified in Atlantic salmon by Garant *et al.* (2000). This model suggests that although selective forces will drive the formation of subpopulations on a small geographical scale, the temporal stability of these subpopulations will vary depending on the temporal stability of suitable habitat. Thus, if on the Frome there was a different amount of suitable habitat between years, or indeed barriers, population structure could vary over time. The identification of different effective population sizes (N_E) and number of breeders (N_B) between salmon in different genetic units (Table 4.2), which were significant in 2009 (two-sided t-test, $P = 0.013$ and 0.016 for N_E and N_B respectively) and noticeable (but not statistically comparable) in 2011, support this. Some combination of factors has served to reduce the number of breeders between groups, which may include barriers, distance from the river mouth and density of the returning salmon.

Nonetheless, the identification of subdivision within the Frome contrasts with previous evidence for within-river population structure in Atlantic salmon, which has primarily been within larger, dendritic systems. This includes, for example, the rivers Teno and Näätämö in Norway (Vähä *et al.* 2008), which have catchment areas of 16,386 km^2 and 2,962 km^2 respectively, and the river Foyle in Ireland (Ensing *et al.* 2011), which has a catchment area of 4450 km^2. And while population structure has been previously identified in smaller rivers, for example within the river Tamar (length = 139 km, catchment area = 928km^2; Ellis *et al.* 2011b), and the Sainte-Marguerite river (length= 101 km, catchment area = 1000 km^2; Garant *et al.* 2000), these catchments are also noticeably dendritic. Thus, size and dendricity of the catchment areas appear to be key factors in the formation of intra-river population structure in this species, as identified by Perrier *et al.* (2011). In the present study, subdivision within the Frome, which is only ca. 48 km long and, with a catchment area of just 454

km^2 is relatively linear, suggests that the degree of admixture with neighbouring regions may be another key factor. Accordingly, we anticipate that further research will be needed to fully understand population structure within rivers known to have little admixture with neighbouring regions.

Finally, it should be noted, that recent improvements in methodology have also increased our ability to detect population structure. For example, increased number of microsatellite loci are now used in fisheries population genetic analyses (e.g. 16 in the current study compared to only five used by Garant *et al.,* 2000). There have also been significant improvements in statistical analyses – the LOCPRIOR model (within the program STRUCTURE) in particular has been used to detect population structure in rainbow trout (Heggenes & Beere 2011), sockeye salmon (Frazer & Russello 2013), brown trout and grayling (Junge *et al.* 2014) and Atlantic salmon (Olafsson *et al.* 2014), which may not have been detected otherwise (Hubisz *et al.* 2009).

Temporal structure within the frome

Atlantic salmon typically show considerable variation in the age at which they migrate to sea; therefore, the average smolt run contains fish of one – three years of age. However, the vast majority of chalk stream fish (98%) smolt after one year (Lauridsen *pers. comm.*). This may in part explain the apparent segregation of 2009 and 2011 fish (Figure 4.6). As a result, microsatellite analysis could be used to determine the number of years that each generation of chalk stream salmon spends between hatching and spawning, which varies considerably over their range (e.g. Johnson *et al. 1991;* Klemetsen *et al.* 2003). For example, if the spawning cycle is three years, as, for example, it is for salmon in Spain (Consuegra *et al.* 2005), then the 2012 cohort would have greater genetic similarity to the 2009 cohort than the 2011. This information would be extremely useful from a management perspective as selective pressures on particular cohorts may vary over time; such knowledge would aid in monitoring fish numbers and allow targeted conservation efforts to support particular year classes within a catchment.

Further implications for conservation

The five chalk streams studied are currently managed following county borders and Environment Agency regional borders, so that the Frome, Piddle and Avon are managed within the region of Wessex, while the Test and Itchen are managed within the Solent and South Downs region. This appears to be in keeping with their natural population structure, as this study reveals a higher degree of connectivity between salmon in the Frome, Piddle and the Avon (Figures 4.2 and 4.3), compared to their connectivity with salmon in the Test and Itchen. The identification of further substructure within the Frome, which changed between time points, supports the metapopulation model of population substructure described by Garant *et al.* (2000). It may also be an indication of the existence of temporally different barriers to upstream migration. Therefore, this method could be used to determine where barriers lay within the river at different times, and which could be removed or otherwise made passable to improve upstream salmon migration. The demonstration of robust sub-division, both between and within chalk stream salmon populations, reaffirms the need for bespoke management and conservation of these genetically distinctive fish.
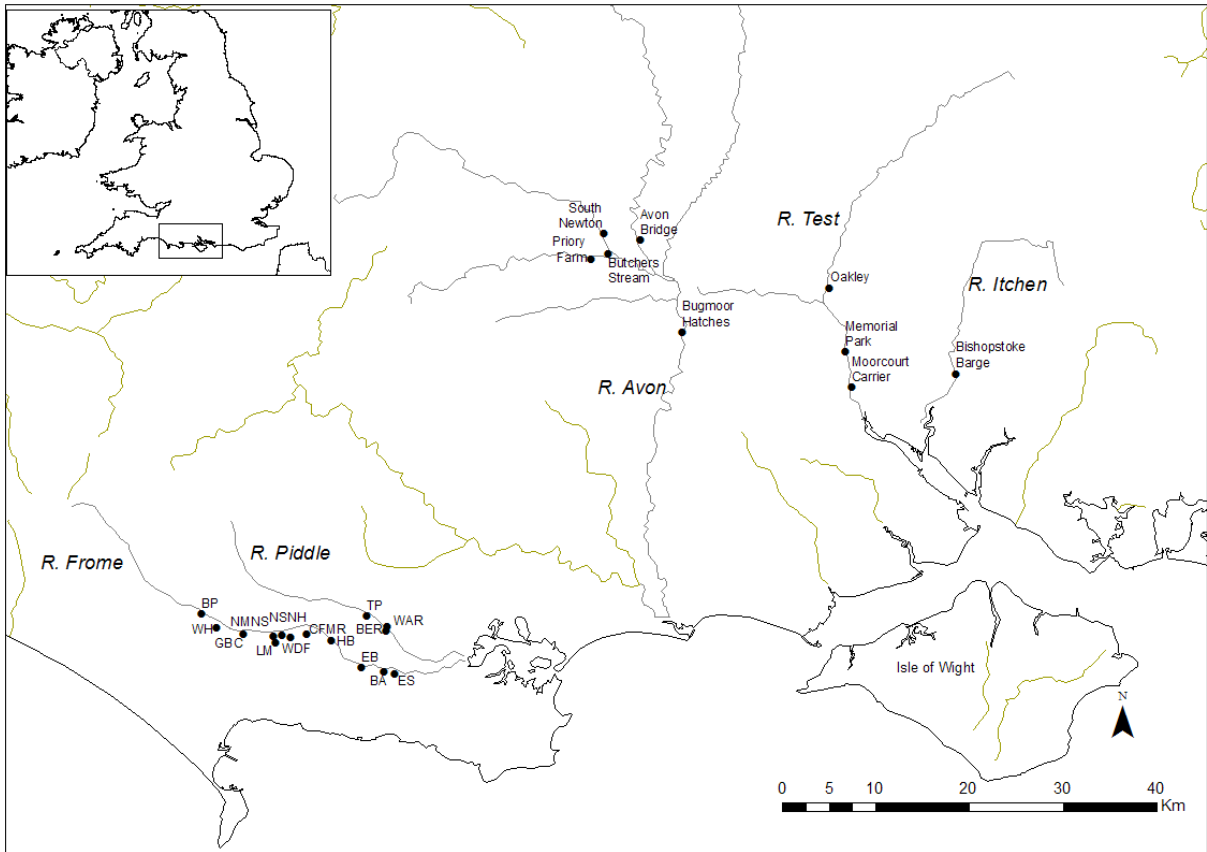
Figure 4.1- Sampling sites of chalk streams. Map depicting the location of all rivers included in this study and the sampling sites of juvenile fish. Sample codes are used for sample sites along the river Frome for clarity. BP- Bradford Peverell, GBC- Grey Bridge Carrier, NSNH- North Stream Nine Hatches, LM- Lewel Mill, WDF- Woodsford, CFMR- Clyffe Farm Main River, EB- East Burton, ES- East Stoke.
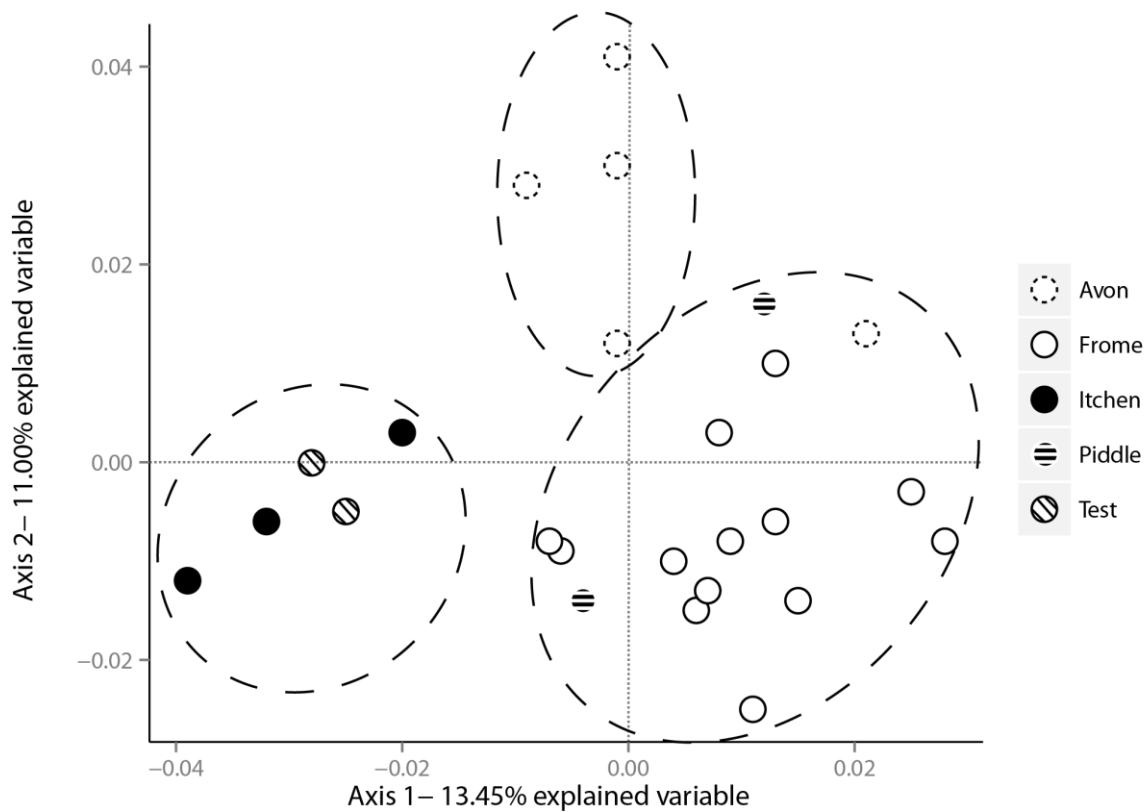
Figure 4.2- Principal coordinate analysis (PCA) of chalk stream salmon. PCA based on the pair-wise F_{ST} value between contemporary sample sites, which included at least 20 individuals. The hatched lines indicate the three groups identified by eye.
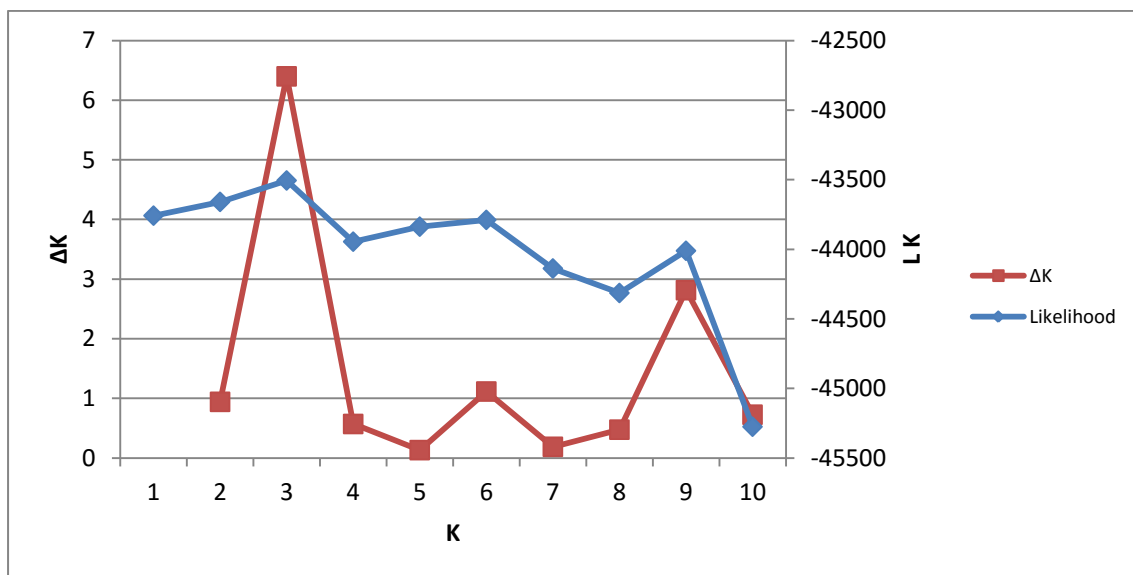
Figure 4.3- Graph of STRUCTURE Δ*k* and L*K* for chalk salmon. Mean Δ*k* and L*K* of over 10 runs for each *K* for 1 to 10 based on STRUCTURE analysis of all chalk streams with prior location information.
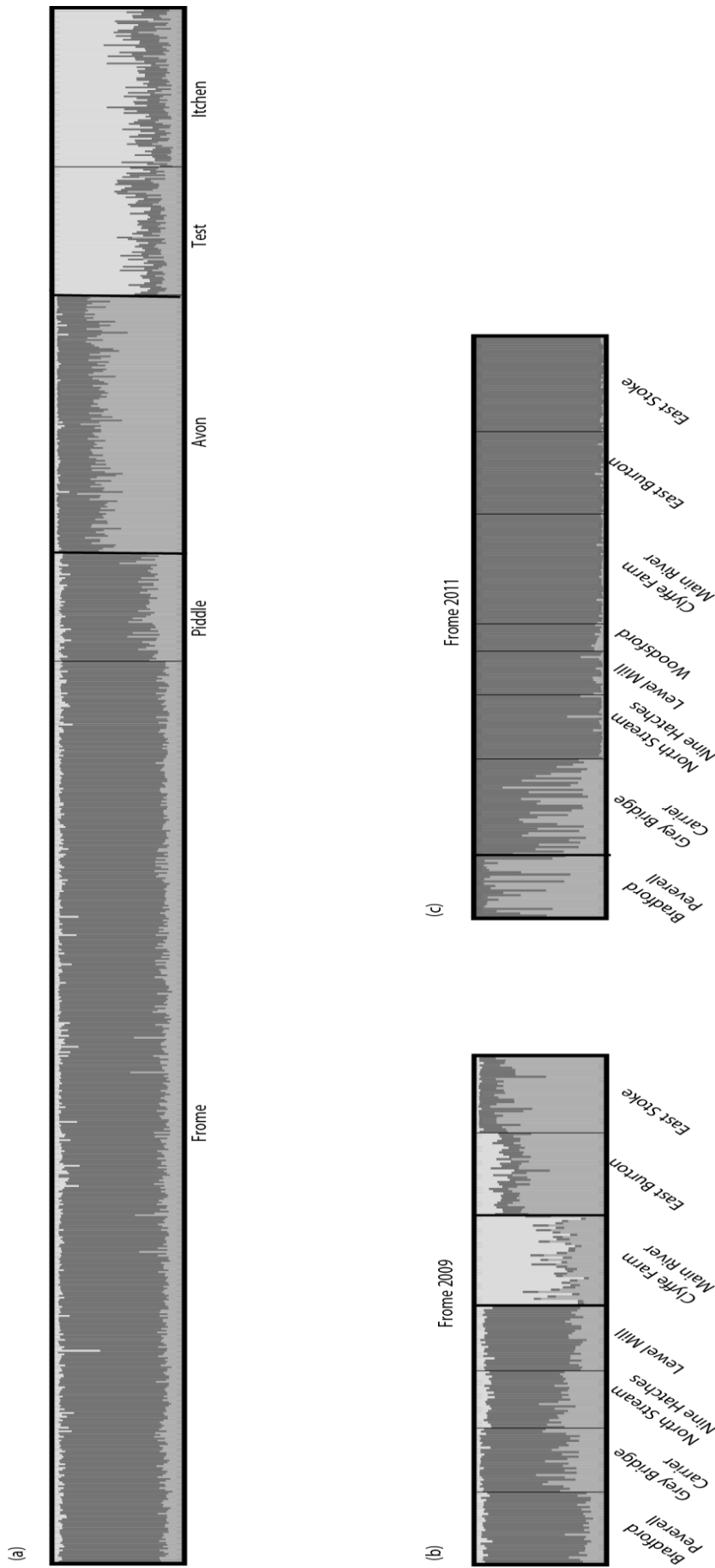
Figure 4.4- STRUCTURE plots of chalk stream salmon. Plots for optimal values of k are shown within: a) the five chalk streams ($k = 3$), b) Frome 2009 ($k = 3$), and c) Frome 2011 ($k = 2$). Each plot shows the estimated proportions of the coefficient of admixture of each individual's genome that originated from population k. The genetic profile of each individual fish is represented by a column. Thin black bars separate individual rivers (a), and sampling sites (b and c), for which names are given below the graphic. Thick black bars separate genetically distinct units.
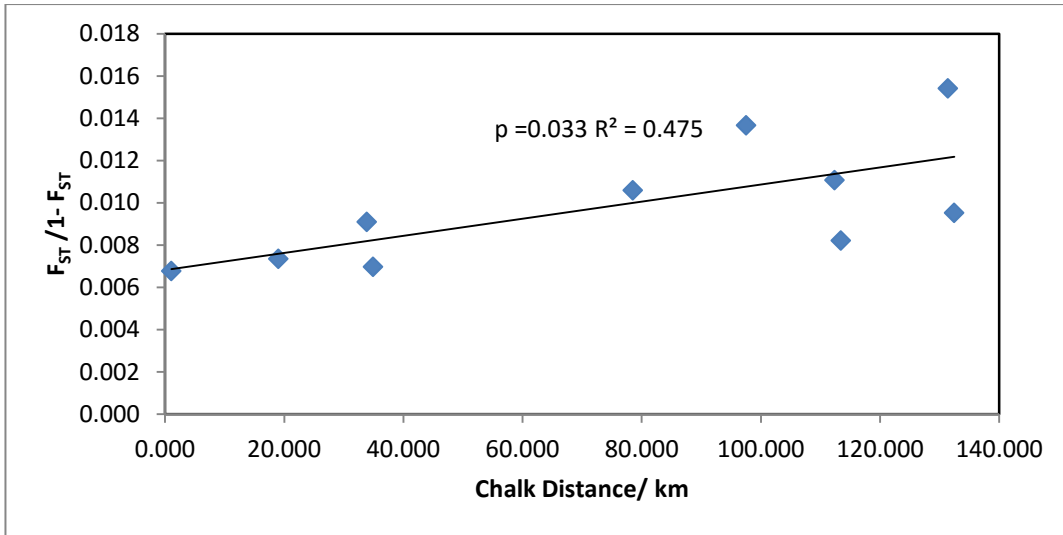
Figure 4.5- Graph of chalk streams Mantel test. Graph indicates geographic distance (km) between river mouths versus genetic distance (F_{ST} /1- F_{ST}) in salmon populations.

a)

b)
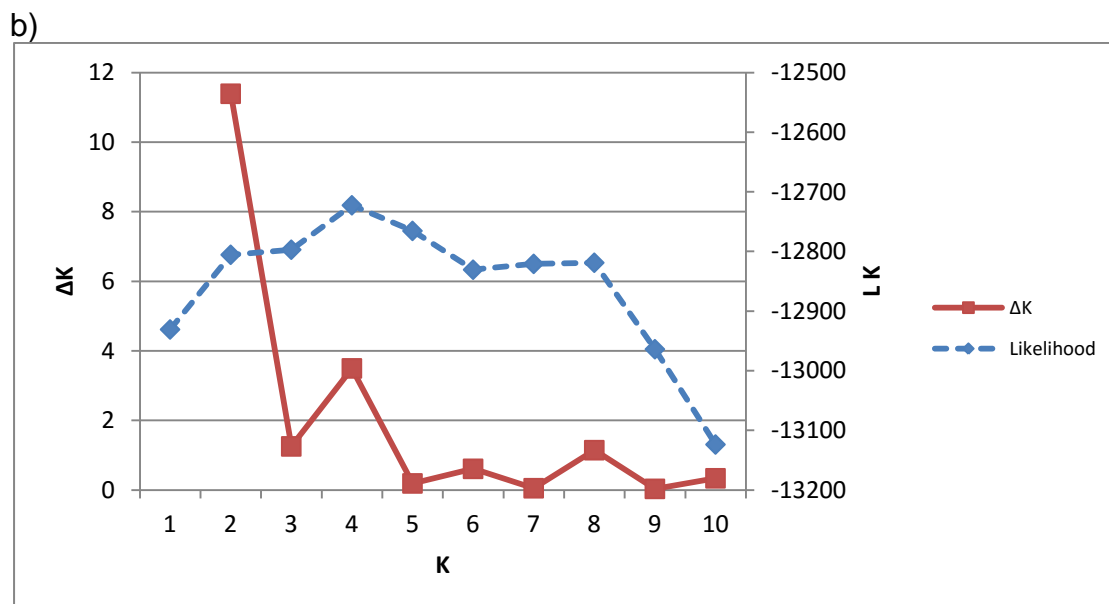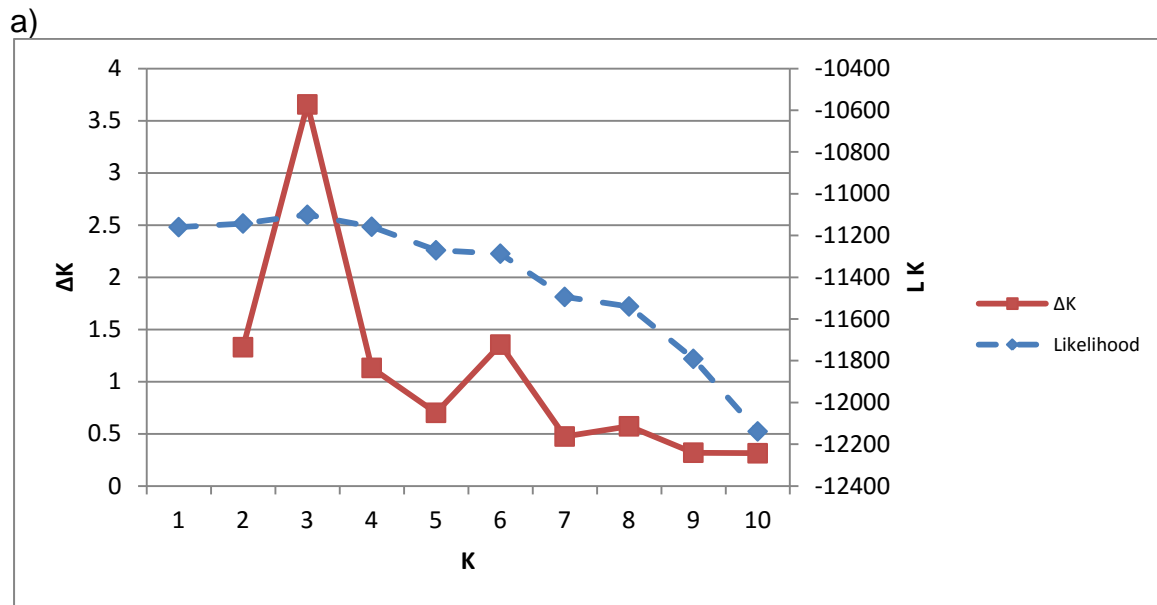
Figure 4.6- Graph of STRUCTURE Δ*k* and L*K* for Frome salmon. Mean Δ*k* and L*K* of over 10 runs for each *K* for 1 to 10 based on STRUCTURE analysis of all a) Frome 2009 and b) Frome 2011 samples with prior location information.
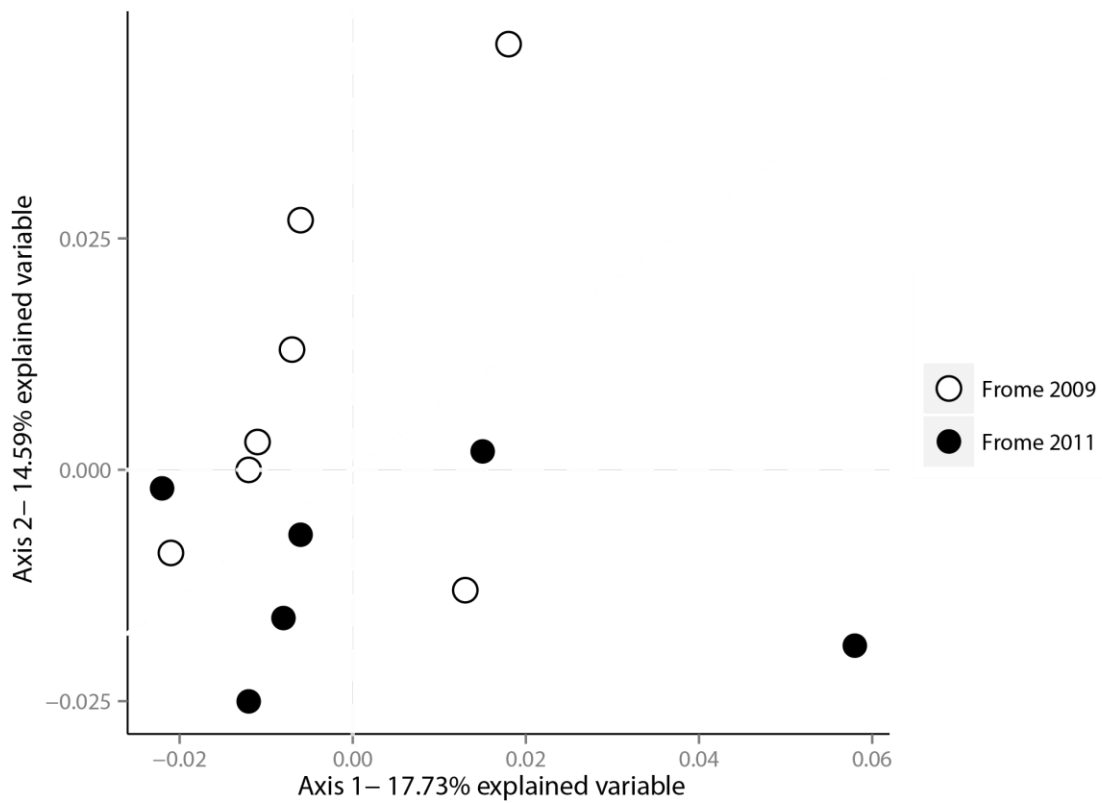
Figure 4.7- Principal coordinate analysis of Frome salmon. Graph is based on the pair-wise F_{ST} values of salmon from site sampled on the Frome in 2009 and 2011. Only sample sites containing at least 20 individuals are used.

Table 4.1- Details on the number of chalk salmon genotyped before and after sib-ship analysis. Sample in bold are adults.

Catchment-Year	No. of sampling sites	No. of samples	After sibling removal	Samples removed(%)
Frome 2009	7	302	221	26.8
Frome 2011	8	454	253	44.3
Piddle 2009	1	32	21	34.4
Piddle 2011	2	89	38	57.3
Avon 2004	2	42	39	7.1
Avon 2010	1	44	20	54.5
Avon 2012	4	117	82	29.9
Test 2010	3	89	70	21.3
Itchen 2005	1	27	26	3.7
Itchen 2006	1	24	23	4.2
Itchen 2010	1	46	37	19.6
Avon 1986+87	**1**	**35**	**33**	**5.7**
Avon 1989	**1**	**23**	**23**	**0**

Table 4.2- Key for each chalk sample site, including the full name of the sample site, the coordinates, the river and the year sampled. Also included are the salmon sample size, number of alleles (NA), allelic richness (AR), expected heterozygosity (HE), observed heterozygosity (HO), estimated effective population size (NE), estimated number of fathers, mothers and total number of breeders (NB) within all sample sites. Allelic richness was computed using a re-sample size of eight individuals. Sample sites in bold contain fewer than 20 individuals and were not included within any statistical comparisons. *FROnsnh11- amalgamation of two sites FROnsnh11 and nmns11 *FROcfmr11- amalgamation of two cites FROcfmr11 and hb11 *FROesg11- amalgamation of two sites FROesg11 and ba11.

Population	River	Sampling site	Year of sampling	X coordinate	Y coordinate	Original Sample size	Sample size post sib-ship	N_A	A_R	H_E	H_O	Ne	No. of Mothers	No. of Fathers	Parents (sum)
FRObp09	Frome	Bradford Peverell	2009	-2.482798	50.736346	42	31	7.563	5.1597	0.662	0.693	19	15	13	28
FROgbc09	Frome	Grey Bridge Carrier	2009	-2.419441	50.716733	49	28	6.875	5.1531	0.696	0.712	16	14	12	26
FROnsnh09	Frome	North Stream Nine Hatches	2009	-2.359681	50.71588	46	25	6.438	5.056	0.69	0.708	11	12	12	24
FROlm09	Frome	Lewel Mill	2009	-2.369884	50.70913	43	29	7.688	5.3484	0.698	0.704	14	12	12	24
FROcfmr09	Frome	Clyffe Farm Main River	2009	-2.322548	50.717482	46	39	6.938	4.7866	0.66	0.693	28	17	17	34
FROeb09	Frome	East Burton	2009	-2.240601	50.685788	41	36	7.125	4.9654	0.68	0.688	47	20	22	42
FROesg09	Frome	East Stoke	2009	-2.189794	50.679613	37	33	7.313	5.1314	0.682	0.704	29	17	15	32
FRObp11	Frome	Bradford Peverell	2011	-2.482798	50.736346	48	27	6.688	4.9965	0.675	0.704	10	15	9	24
FROgbc11	Frome	Grey Bridge Carrier	2011	-2.419441	50.716733	49	42	7.063	5.2326	0.707	0.691	42	20	20	40
FROnsnh11*	Frome	North Stream Nine Hatches	2011	-2.359681	50.71588	95	28	6.375	4.9203	0.663	0.707	20	10	16	26
FROlm11	**Frome**	**Lewel Mill**	**2011**	-2.369884	50.70913	**19**	**19**	**5.688**	**4.6712**	**0.656**	**0.662**	**21**	**9**	**9**	**18**
FROwdf11	**Frome**	**Woodford**	**2011**	-2.347811	50.714207	**48**	**12**	**5.188**	**4.6561**	**0.63**	**0.661**	**6**	**6**	**5**	**11**
FROcfmr11*	Frome	Clyffe Farm Main River	2011	-2.322548	50.717482	93	48	7.563	5.1029	0.691	0.685	22	17	21	38
FROeb11	Frome	East Burton	2011	-2.240601	50.685788	47	36	7	5.155	0.684	0.703	24	19	14	33
FROesg11*	Frome	East Stoke	2011	-2.189794	50.679613	55	41	7.813	4.9968	0.683	0.688	20	19	16	35
PIDber09	Piddle	Bere Stream	2009	-2.200775	50.725076	32	21	6.625	5.0329	0.682	0.728	12	-	-	-
PIDtp11	**Piddle**	**Turners Puddle**	**2011**	-2.232068	50.735405	**43**	**17**	**6.063**	**5.0551**	**0.677**	**0.642**	**11**	**-**	**-**	**-**
PIDwar11	Piddle	Warren	2011	-2.202387	50.721071	46	21	6.375	4.9488	0.681	0.719	16	-	-	-
AVNbrd04	Avon	Avon Bridge	2004	-1.816891	51.09558	23	20	6.375	5.1339	0.693	0.744	22	-	-	-
AVNbug04	**Avon**	**Bugmoor Hatches**	**2004**	-1.787263	51.007847	**19**	**19**	**6.625**	**5.2955**	**0.69**	**0.668**	**43**	**-**	**-**	**-**
AVNbrd10	Avon	Avon Bridge	2010	-1.816891	51.09558	44	20	6.438	5.0186	0.668	0.693	11	-	-	-
AVNbri12	Avon	Avon Bridge	2012	-1.816891	51.09558	21	21	6.625	5.2585	0.692	0.721	30	-	-	-
AVNbut12	Avon	Butchers Stream	2012	-1.866044	51.082822	45	21	6.188	4.9737	0.668	0.727	7	-	-	-
AVNprf12	Avon	Priory Farm	2012	-1.892028	51.077579	34	26	6.438	4.9603	0.682	0.684	22	-	-	-
AVNsnw12	**Avon**	**South Newton**	**2012**	-1.87176	51.101651	**17**	**14**	**5.313**	**4.7006**	**0.642**	**0.719**	**17**	**-**	**-**	**-**
TSTbro10	Test	Moorcourt Carrier	2010	-1.496397	50.953838	24	23	6.375	4.7656	0.646	0.643	30	-	-	-
TSToak10	Test	Oakley	2010	-1.528935	51.048915	25	22	6.938	5.2831	0.692	0.687	24	-	-	-
TSTmem10	**Test**	**Memorial Park**	**2010**	-1.505267	50.987364	**33**	**18**	**6.375**	**5.1998**	**0.685**	**0.724**	**22**	**-**	**-**	**-**
ITCbis05	Itchen	Bishopstoke Barge	2005	-1.337858	50.965754	27	26	6.75	4.9218	0.678	0.695	23	-	-	-
ITCbis06	Itchen	Bishopstoke Barge	3006	-1.337858	50.965754	24	23	7.188	5.2539	0.693	0.667	29	-	-	-
ITCbis10	Itchen	Bishopstoke Barge	2010	-1.337858	50.965754	46	37	7.313	5.028	0.685	0.667	24	-	-	-

Table 4.3- Genetic diversity and statistical significance of salmon in the chalk streams. Average observed heterozygosity (H_O), expected heterozygosity (H_E), allelic richness (A_R), inbreeding co-efficient (F_{IS}), and F_{ST} for sites within rivers. *P* indicates the significance of differences between all rivers. N/A indicates not attainable.

Statistic	Frome	Piddle	Avon	Test	Itchen	*P*
H_O	0.698	0.723	0.713	0.665	0.676	0.047
H_E	0.694	0.698	0.697	0.684	0.697	0.792
A_R	5.077	4.991	5.069	5.024	5.068	0.143
F_{IS}	-0.006	-0.036	-0.023	0.028	0.031	N/A
F_{ST}	0.019	0.018	0.02	0.031	0.009	0.537

Table 4.4- Genetic diversity and statistical significance of salmon in the river Frome in 2009 and 2011. Average observed heterozygosity (H_O), expected heterozygosity (H_E), allelic richness (A_R), inbreeding co-efficient (F_{IS}), and F_{ST} for sites within rivers. Allelic richness was calculated from 197 re-sampled individuals. *P* indicates the significance of differences between all rivers.

Statistic	Frome 2009	Frome 2011	P-value
A_R	9.078	8.280	0.048
H_O	0.674	0.662	0.528
H_E	0.676	0.678	0.865
F_{IS}	0.013	0.024	0.707
F_{ST}	0.020	0.019	0.847

Chapter 5: Temporal stability of genetic diversity and effective population size in Atlantic salmon across Great Britain and France

Introduction

Decline of fisheries and salmon

In 1900, a critical review (Garstang 1900) was presented to a sceptical scientific audience, which concluded that fisheries were not only exhaustible, but in a "rapid and continuous process of exhaustion." As this went beyond conventional belief, its warnings were mostly unheeded. A century later however, our fisheries are in dire states and, for example, it has been calculated that the biomass in the Atlantic Western Boundary Current fishery is 3-10% of historic levels (MacIntyre *et al.* 1995). So, it is not in isolation that Atlantic salmon (*Salmo salar* L.), have suffered global population decline (Parrish *et al.* 1998; WWF 2001). The species is threatened worldwide by multiple factors, which include, but are not limited to, overfishing, river engineer schemes, pollution and climate change (WWF 2001).

In 2001 the World Wildlife Federation (WWF) attempted to summarise the state of Atlantic salmon populations by collating information of salmon stocks within rivers across their entire natural range (WWF 2001). Using the best available local data they identified the number of rivers within each country that historically contained salmon. They then determined the proportion of rivers that were one of the following six categories: Healthy, Vulnerable, Endangered, Critically Endangered, Extinct or of Unknown status. These categories were defined by multiple parameters, including the current number of spawning adults and the speed of any decline in this number (WWF 2001). They identified that, at the turn of the 21st Century, salmon had become extinct from 15% of their historical rivers, and were considered healthy in only 43% (WWF 2001). Of these, 93% were found within just four countries: Norway, Ireland, Iceland and Scotland (WWF 2001). Approximately 50% of rivers analysed had suffered a greater than 70% reduction in rod-catch compared to 20 years earlier (WWF 2001).

Significant measures have been taken to tackle the decline. For example, within England and Wales, a rod licence is a legal requirement for rod-fishing of salmon and trout (Cefas & Environment Agency 2013). This enables the UK government to monitor rod-catch and enforce limits. On many rivers, a policy of catch and release – where salmon must be released alive after catching – is encouraged, and in a small number of rivers (the Wye, Taff and Ely) catch and release is mandatory (Cefas & Environment Agency 2013). This policy is also enforced in parts of Canada, USA and increasingly so in other parts of Europe (NASCO 2013). Major salmon fisheries at sea and on shore have been closed across their natural range. This includes all major commercial fisheries in Canada by 2012 after incremental closures since 1992 (NASCO 2013), and the closure of commercial fisheries for export in West Greenland in 1998 (NASCO 2013). In England and Wales the majority of fisheries have been closed through government enabled phase-out schemes (Cefas & Environment Agency 2013). The northeast coast fishery experienced the largest closure where phase-out began in 1993, and by 2013, 142 drift net licences had been reduced to just thirteen (Cefas & Environment Agency 2013).

Despite these measures, many salmon populations remain threatened and, as noted by the North Atlantic Salmon Conservation Organization (NASCO 2013), numbers do not appear to be improving as expected. However, it is possible that estimates are not accurate. The vast majority of salmon population size estimates are calculated from models, which are based on rod or net-catch data (Chaput *et al.* 2005; ICES 2014). Although these are common place, catch data is not always reliable (WWF 2001). Firstly, for both rod and net fisheries, the absence of fish being caught does not necessarily mean that fish are absent. This can be accommodated for by working out the proportion of the population caught per unit effort. However, measuring fishing effort is also rarely consistent: if, for example, Fisherman A tells Fisherman B that he has had a poor fishing session, then Fisherman B is increasingly unlikely to attempt to fish (Beaumont, GWCT pers. comm.). Further complications arise when temporally comparing population size. On one side, improvements are made in fishing methods, while on the other, fisheries are increasingly forced to reduce fishing effort.

Fish counters are expected to alleviate some uncertainty in catch data (ICES 2013). These are devices that autonomously count the passing of a fish through a specific point. Measurement is generally achieved via electrical resistivity, optically or hydroacoustically (Eatherley *et al.* 2005). Placed at or near a river mouth, in theory they are able to count every salmon that enters the river. However, they are rarely 100% effective as, for example, some fish will bypass the counter (Thorley *et al.* 2005) and some will pass counters multiple times (Dunkley & Shearer 1982; Thorley *et al.* 2005). There can also be false positives, where salmon are detected which did not pass and false negatives, where salmon that pass are not detected (Dunkley & Shearer 1982). Yet counters are considered to be more reliable than other methods (Eatherley *et al.* 2005), and are primarily limited by high installation and maintenance costs (Eatherley *et al.* 2005). Thus they have a limited current coverage, for example there are only thirteen used by the Environment Agency across the entirety of England and Wales (Cefas & Environment Agency 2013).

Where counters are installed, they can be used to estimate the effectiveness of rod-catch based calculations. Some studies have identified agreement between the two methods (e.g. Crozier 2001; Thorley *et al.* 2005; Jonsson *et al.* 2008). Crucially however, other studies have identified incongruence between the methods (O'Connell 2003; Eatherley *et al.* 2005; Thorley *et al.* 2005). For example, estimates based on rod-catch have been up to four times greater than estimates based on fish counters (O'Connell 2003). As neither measure is 100% accurate, it is clear that other methods for estimating the size and sustainability of populations are necessary.

Usefulness of population genetics

Population genetics is increasingly being used as an additional tool to estimate population size and sustainability. Population genetic theory dictates that a smaller population will lose genetic diversity, largely from genetic drift (Garza & Williamson 2001), indicating that a population with low genetic diversity may have been through a population decline. Genetic diversity is also a key indicator of the health of a population, as it depicts the evolutionary potential of the species (Laikre *et al.* 2009). Another parameter, the effective population size (N_E), is also increasingly monitored. N_E corresponds to the number of individuals within a population that would, within an idealised

population, show the same amount of genetic variation as the real population under random genetic drift (Nikolic *et al.* 2009). Crucially, N_E determines how quickly genetic diversity is lost. Although the effective population size can be calculated from demographic parameters, this is difficult, especially for species like salmon where viewing the majority of their natural life cycle is challenging.

The continual development of polymorphic markers has however, enabled easier calculations of N_E, with increasingly fewer samples (Nikolic *et al.* 2009). Monitoring population size using genetics has advantages over other methods as it can be employed without major financial investment in fish counters and also avoids reliance on inconsistent catch data from fishermen. A number of studies have used DNA derived from historic material – often scales in the case of salmonids – to compare the genetic diversity and effective population sizes (N_E) of historic populations against contemporary populations. These have included populations from Canada (Fraser *et al.* 2007; Palstra *et al.* 2009), America (Lage & Kornfield 2006), Denmark (Nielsen *et al.* 1997), Finland (Säisä *et al.* 2003), Spain (Consuegra *et al.* 2005; Ayllon *et al.* 2006), France (Valiente *et al.* 2010; Perrier *et al.* 2013), and Norway (Skaala *et al.* 2006; Glover *et al.* 2012).

Importantly, by using historic material, these studies have been able to obtain the trajectory of populations, rather than simply capturing a snapshot of the present. Decreasing genetic diversity and effective population size have been identified in some populations (Lage & Kornfield 2006; Horreo *et al.* 2011b), while in other countries the results have been mixed (Perrier *et al.* 2013). Proposed reasons for changes have included stocking with exogenous fish (Glover *et al.* 2012; Perrier *et al.* 2013), climate change (Valiente *et al.* 2010) and the reduction in size of endemic populations (Glover *et al.* 2012).

The gap in our knowledge

As yet, no published studies have used historic material in Great Britain, and the trajectory of genetic diversity and effective population size of salmon in this region is unknown. Although the majority of salmon rivers in Scotland (67%) were classified as Healthy (Table 5.1), only 33% were deemed Healthy in England and Wales (WWF 2001; NASCO 2013). Considering this poor state, there is a real need to identify the trajectory of genetic diversity, and the present

study aims to fill in that gap. Salmon from France were also included in the analysis as these salmon populations are also in poor health. Although fish in some French rivers have previously been subject to temporal analyses, at best only two time points have been included within a single study, i.e. before and after major exogenous stocking in the 1990s (Perrier *et al.* 2013). Critically, the results from that study (Perrier *et al.* 2013) could, in theory, reflect year to year variation, even if spread over decades. Thus, there is still a need to identify the trajectory of genetic diversity within this region.

The present study has two aims; the first is to determine the feasibility of generating sufficient microsatellite data from historic scales to identify long term genetic trends in salmon populations across Scotland, England, Wales and France. The second aim is to identify whether the trajectory of genetic diversity and the effective population size of Atlantic salmon populations in these countries reflect the consensus based on catch statistics (WWF 2001). Based on the findings by WWF (2001), our hypothesis is that the genetic diversity and effective population size of salmon populations within Scotland will be temporally stable, reflecting the majority of Healthy rivers. In England, Wales and France, where the majority of populations are not Healthy, we hypothesise that genetic diversity and effective population size will be decreasing over time.

Materials and Methods

Atlantic salmon scale samples were collected from Scottish (Conon between 1998 - 2012 and Tweed between 1992 - 2012), Welsh (Dee 1991 - 2011), English (Tyne 1991 - 2012, Frome 1954 - 2011, Avon 1951 -1986, Axe 1963 - 1975 and Exe 1966 - 2009) and French (Sée 1977 - 2000, Ellé 1968 - 2000 and Scorff 1972 - 2000) rivers (Figure 5.1, Table 5.2). The aim was to collect samples from rivers that had been sampled on at least three occasions, at least four years (one generation) apart. This was necessary in order to obtain trends in both the genetic diversity and effective population size. Primarily, scales from adults were targeted because they would most consistently contain the genetic signature across the entire river, and typically only adult scales were available for historic samples. Scale samples had either been collected from salmon caught by rod-and-line within a river catchment, or collected from salmon as they passed through a fish counter.

In order to determine which DNA extraction method provided the best quality DNA for the amplification of microsatellite DNA, preliminary tests were made with historic scales from the river Axe using the Chelex method (Estoup *et al.* 1996) and QIAGEN extraction kits. As the tests indicated no difference, the Chelex method was used for all of the samples because it was more cost and time effective. Genomic DNA was extracted from all salmon scales using the Chelex method of Estoup *et al.* (1996) with minor modifications. The volume of 5% Chelex solution was reduced to 50 ml and extracts were left for three hours, or more for the historic scales, to improve the concentration of the DNA. The intention was to amplify a total of 16 microsatellite loci. These were Ssa14 (McConnell *et al.* 1995); Ssa202, SSsp3016, Ssa197 (O'Reilly *et al.* 1996); Ssosl85 and Ssosl417 (Slettan *et al.* 1995); SsaF43 (Sánchez *et al.* 1996); SSspG7, SSsp1605, SSsp2210, SSsp2201, and SSsp2216 (Paterson *et al.* 2004); Ssa171, Ssa289, Ssa157, and SsaD144 (King *et al.* 2005). The loci were typically amplified within three multiplexed polymerase chain reactions (PCR), comprising: (Primer mix 1) SSspG7, Ssa14, Ssa202, SSsp3016, Ssosl85 and Ssosl417; (Primer mix 2) Ssa197, SsaF43, SSsp1605, SSsp2210, SSsp2216; and (Primer mix 3) SsaD157, Ssa171, Ssa289, SsaD144, SSsp2201. However, a number of changes were made in order to better amplify the microsatellites from the historic DNA. Locus SSspG7 was removed from Primer mix 1 and amplified in isolation. Similarly, SSsp1605 was removed from Primer mix 2 and amplified in isolation. PCR reactions were carried out in 10 µL reactions following Ikediashi *et al.* (2012), and the size of the fluorescently labelled PCR products was determined following Ikediashi *et al.* (2012) (Chapter 2).

Genotypes from adults from the Sée, Ellé and Scorff were also obtained from the SALSEA database (Gilbey *et al.* unpublished) for the year 2005 in order to extend the temporal range of the datasets. These required conversion to match the Exeter genotypes to correct for different scoring of alleles between laboratories. Using the calibration from the study (Ellis *et al.* 2011), the SALSEA data was transformed for the present study following specific rules for 11 loci (Table 5.3) The remaining microsatellites used did not require conversion.

Genotypes from juveniles were also obtained from previous research on the river Exe from 2004 (Griffiths *et al.* 2010) and 2009 (Counter thesis, unpublished) and the Avon from 2004 and 2012 (previous chapter), as adult

samples were not available (See Appendix IX and X for juvenile sampling locations). Juvenile and adult samples have previously been compared to infer genetic stability (e.g. Nielsen *et al.* 1997; Lage & Kornfield 2006). In order to better match the adult samples, which would represent salmon from across the catchment area, the following methods were employed. For the Exe 2004, Exe 2009 and Avon 2012 where multiple sites had been sampled (Appendix Appendix IX and Appendix X), 50 samples were randomly selected from each set of samples before further analysis. To minimise the chance of bias, this sub-sampling was repeated to make four additional replicates of Exe 2004 and 2009 and Avon 2012 samples. Genetic diversity and effective population size was calculated for each of these replicates using the methods described henceforth. For samples from the Avon from 2004, where only two sites had been sampled, all individuals were included for further analysis.

A total of 1918 adult scale samples were collected from 11 rivers (Table 5.2), 138 French adult genotypes were obtained from the Salsea database and 225 juvenile genotypes were obtained for salmon on the Exe and Avon (Table 5.2)

Data checking

As the aim of this study was to infer temporal changes within each river, and not to compare the genetic diversity and N_E between rivers, genotype data was analysed on an individual river basis. This enabled the maximum amount of genetic data to be retained. Microsatellite loci were removed if they failed to amplify in 20% or more individuals within a single sampling year, and individuals were removed if they contained 25% missing data or greater. In effect, samples from each river formed a unique temporal dataset. Each temporal dataset could differ in the makeup and number of microsatellites included.

In order to identify why some of the samples amplified poorly, a Bioanalyser was used to determine the concentration and size of DNA within the historic Axe samples. Six samples from 1963 and five samples from 1975 were included and one sample from the river Mersey from 2002 was included as a positive control.

MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004) was used to check for scoring errors due to stutter peaks, large allele dropout or null alleles within all samples. However, due to the degree of missing data – caused by low quality DNA – which may have reduced the efficacy of MICRO-CHECKER in a number of instances, another test was devised. Allelic dropout affects the largest alleles first (Takahashi *et al.* 1997), as the more base pairs there are within a sequence, the greater the chance of them being broken down into smaller fragments (Takahashi *et al.* 1997). Therefore, if allelic dropout were to occur, then longer microsatellites would show a disproportionate increase over time. Thus a graph of the rate of change in allelic richness ($d\text{A}_\text{R}/dt$) against the maximum known size of each microsatellite marker was plotted for each temporal dataset. If samples were subject to allelic dropout, we expected a significant positive correlation between $d\text{A}_\text{R}/dt$ and the maximum allele size.

In order to prevent the false detection of population structure due to the presence of family groups (Goldberg & Waits 2010), the program COLONY v2.0.4.1 (Jones & Wang 2010) was used to identify full siblings. The mating system was defined as polygamous for males and females and without inbreeding. Each run was of medium length, with high precision and using the full-likelihood method. Allele frequencies were not updated during the run and no prior sib-ship was assumed. An error rate of 0.02 was assumed for each locus based on analysis by Ellis *et al.* (2011a). COLONY was run twice independently, with different starting seeds to check the consistency of the reconstruction. Full-sib families were reduced to one representative, if supported by an average likelihood of 0.5 or higher between the two runs.

Linkage disequilibrium and deviations from Hardy-Weinberg equilibrium were detected using GENEPOP v4.2 (Raymond & Rousset 1995). Both were calculated using 1,000 dememorisation steps, 200 batches each with 10,000 iterations. The 95% significance level of each was adjusted using the False Discovery Rate (FDR) method (Benjamini & Hochberg 1995).

Data analysis

In order to investigate the extent of genetic change within temporal samples, a number of analyses were made within the temporal datasets. Firstly, pair-wise F_{ST} values were calculated within the datasets, and tested for significance with 999 permutations using Genalex v6.5 (Peakall & Smouse 2012). Secondly, significant temporal differences in allelic and genotypic frequencies were tested for using GENEPOP. To elucidate the trajectory of genetic diversity within each river, for each sample the average expected heterozygosity (H_E), observed heterozygosity (H_O), and inbreeding coefficient (F_{IS}) were calculated using Genalex, while the allelic richness (A_R) was calculated using Fstat v2.9.3 (Goudet 1995). Within each temporal dataset, the rate of change in each index over time was also determined by calculating the gradient of the change in the index against the change in time. In order to determine whether any changes between temporal samples were significant, an analysis of variance (ANOVA) was implemented within R using the following equation (Hamilton *et al.* 2014):

```
x1<-glm(x~year+Locus)
x2<-update(x1,~.-year)
anova(x1,x2,test="F")
```

where x was either A_R, H_E, H_O, or F_{IS}. The results of each microsatellite loci were used, rather than the average, in order to increase statistical power. If the ANOVA was significant, then the Tukey test was used to determine between which two years the significance lay. This was done using the following equation:

```
facyear<-factor(year)
anov1<-aov(x~facyear+Locus)
TukeyHSD(anov1,"facyear")
```

Calculating effective population size

Effective population sizes (N_E) were determined for salmon within each river and year using both single sample and temporal sample methods. The single sample methods implemented were the linkage disequilibrium (Do *et al.* 2014) and molecular co-ancestry (Nomura 2008) within the program NeEstimator (Do *et al.* 2014) and the heterozygote excess (Pudovkin *et al.* 1996) and sib-ship method within the program COLONY (Jones & Wang 2010). The first two methods were run following the manual (Do *et al.* 2014), while the methods run in COLONY were implemented following the protocol described earlier (Data checking). For temporal calculations of N_E, which are based on the temporal change in allele frequencies, multiple methods were used. These were the Nei & Tajima (1981), Pollak (1983) and Jorde & Ryman (2007) methods within the program NeEstimator. However, because the first two methods often calculated infinite 95% confidence intervals, only the Jorde & Ryman (2007) method was used for all temporal datasets. The generation time used was four years following previous examples (Nikolic *et al.* 2009; Finnegan *et al.* 2013). Preliminary analysis indicated that changing the generation time changed N_E by a consistent factor. Because the aim of this study was to identify trends in N_E and the numbers themselves were not significant, using one intermediate generation time was sufficient. 95% parametric confidence limits were also calculated. Because the COLONY method and the temporal methods both also provided 95% confidence limits, non-overlapping confidence limits were also used to indicate a significant difference in N_E calculations ($p < 0.05$).

Short term rod-catch

Although significant declines in population size have been found in many salmon populations over the past 30 years and more (Parrish *et al.* 1998; WWF 2001), it was possible that during the time period over which samples were collected population sizes may not have declined. This would be an important factor for detecting changes in genetic diversity and effective population size. Rod-catch data was collected for rivers from which samples had also been collected for over their sampling period (Appendix Table A5.4). The trend in rod-catch over time was determined by calculating Pearson's Correlation (r) between rod-catch and time within Microsoft Excel. The significance of the

correlation was determined by using the following equation to convert r into the t distribution value (t):

$$t = (r \times \text{sqrt} (n - 2)) / (\text{sqrt} (1 - r^2))$$

The significance of t was then determined within Excel using the t distribution 2-tailed test.

Further investigation- Nine loci comparisons

When calculating changes in effective population size and genetic diversity it was possible that some of the temporal datasets might have suffered bias from having fewer microsatellite loci (as low as nine, compared to a maximum of 16). To investigate this possibility, a new dataset was created where all samples used exactly the same loci. After the removal of one set of samples from the Sée from 1980-1, where the genotypes for five loci had amplified poorly (Ssa157, Ssa171, Ssa289, SsaD144 and SSsp2201), nine loci were common between the remaining samples. Thus, the datasets were each reduced to include these nine loci before running the sib-ship single sample and temporal calculation methods of effective population sizes again. The other single sample methods proved uninformative (see results), so these methods were not repeated with nine loci. Tests for significant differences in A_R, H_O, H_E and F_{IS} were also repeated, in order to identify whether significant differences had been missed due to the dataset possessing a reduced number of loci. This dataset also enabled the inference of genetic difference between every population at every time point, which could indicate whether differences between populations were stable, and possibly provide explanations for some of the identified changes in genetic diversity and N_E. This was achieved by calculating pair-wise F_{ST} values between all samples.

Past studies have found evidence of increased gene flow between salmon populations over time (Consuegra *et al.* 2005; Perrier *et al.* 2013), which would bias calculations of N_E (Wang 2009). To investigate that possibility, pair-wise F_{ST} values were calculated between salmon from the three French rivers (only these samples were used because only these were sampled at consistently matching time points). In order to determine whether differences between the years were significant, an ANOVA test was run within the program,

Genalex. Because there was a large range between the pair-wise Fst values, the data was further explored to ensure that statistical significance was not lost within the larger variation. This was achieved by systematically removing one of the pairwise comparisons and then repeating the ANOVA test.

Mutation and selection might also be considered factors likely to affect the analyses, however microsatellite mutation is unlikely to have had a noticeable effect within this short timescale (Hardy *et al.* 2003). Selection is unlikely to affect the analysis because these loci are believed to be selectively neutral (e.g. Olafsson *et al.* 2014), and have been tested in a previous chapter (Chapter 3) on a geographically broad collection of salmon.

Results

The dataset

Genotypes from historic samples from the rivers Axe, Frome and Tyne failed to amplify consistently within the timeframe and in particular with Primer mix 1 and Primer mix 3; therefore salmon from these three rivers were not included within the analyses. Analysis with the Bioanalyser identified that although the concentration of DNA between the contemporary and historic samples were on par with each other (Table 5.4), the average fragment size (base pairs) was significantly smaller in the historic samples. This would explain why microsatellite amplification within historic samples failed in many cases (Table 5.2). The Tyne samples are likely to have failed because they had been chemically cleaned for the purpose of scale reading, while the Axe samples are likely to have failed because of their age (39-51 years). The Frome samples were unsuccessful due to a combination of their age (primarily in Primer mix 1 and Primer mix 3) or because samples had been cleaned. For a number of rivers (Exe, Avon, Sée, Ellé and Scorff), the most historic samples were removed due to poor amplification of microsatellites.

When using as many loci as possible, the final dataset consisted of 1,386 salmon from eight rivers, which covered the years 1972-2012 (Table 5.2). The number of loci successfully amplified ranged from nine in samples from the Sée to 16 from those from the Tweed (Table 5.2). When using the common nine loci, where samples from the Sée 1980-81 were removed, the dataset consisted of 1,279 samples. The nine included loci were SSsp3016, Ssa197, SsaF43, SSsp1605, SSsp2210, SSsp2216, Ssa171, Ssa289 and SsaD144.

After false discovery rate (FDR) correction, there were 17 cases of linkage disequilibrium between pairs of loci, with Tweed 1992 salmon possessing the greatest number (10). After FDR correction, there were 15 pair-wise comparisons out of Hardy-Weinberg equilibrium, with salmon from the Ellé in 2005 possessing the greatest number (3). Therefore, as no loci were consistently out of Hardy-Weinberg equilibrium or linkage disequlibrium, no loci were removed as a result of these analyses. Analysis with MICRO-CHECKER identified no instances of allelic dropout, but there were 15 instances of null alleles out of 432 comparisons, as indicated by an excess of homozygotes.

In the plots of $d\text{A}_\text{R}/dt$ against the maximum allele size (Appendix XI), a positive correlation was found for salmon in six of the eight rivers. These were the Conon, Tweed, Exe, Avon, Sée and Ellé. However, only within one river, the Sée, was the correlation significant (Pearson's correlation = 0.94, p =0.0001). These samples were retained, because they would have been accepted if solely the traditional method (MICRO-CHECKER) was used, however they provide an added opportunity to investigate the effect of allelic dropout further.

Genetic differentiation

The average pair-wise F_{ST} values between temporal samples within rivers ranged from 0.005 in the River Tweed to 0.015 in the Avon (Figure 5.2). Significant F_{ST} values were found between temporal samples from the rivers Conon, Exe, Avon and Sée (See Appendix XII). Significant changes in genotype frequencies (p < 0.05) were found between at least one pair of years in the rivers Conon, Exe, Avon, Sée, Ellé and Scorff. The range of average pair-wise F_{ST} value using the nine loci was exactly the same as that of samples when using all available loci (see Table 5.9).

Genetic Diversity

As many microsatellites as possible (between 9 and 16) were used to investigate change in genetic diversity over time (see Table 5.2). Allelic richness was found to be decreasing over time in salmon within the Scottish rivers Conon and Tweed and the southern English river Avon, but increasing in the rivers Dee, Exe, Ellé, Scorff and Sée (Figure 5.3a). Significant differences were found only within the Sée (ANOVA, p = 0.019), between the years 2005 and 1980 (Tukey test, p = 0.047), and between 2005 and 1988 (Tukey test, p = 0.035) and also in the Ellé (ANOVA, p = 0.003), between the years 2000 and 1988 (Tukey test, p = 0.020), and between 2005 and 1988 (Tukey test, p = 0.032).

Expected heterozygosity (H_E) increased over time in salmon from all rivers except within the river Conon (Figure 5.3b). Significant differences in H_E were found for salmon within the Sée (ANOVA, p = 0.006), however no significant difference was detected between any two specific years. A significant difference was also detected within the Scorff (ANOVA, p = 0.006), between

samples from 1994 and 1988 (Tukey test, p = 0.030) and 2005 and 1988 (p = 0.002), 2005 and 2000 (p = 0.041).

Observed heterozygosity increased over time in salmon from all rivers except in those from the rivers Conon, Tweed and Scorff (Figure 5.3c). Significant differences in H_O were found in salmon from the river Sée only (ANOVA, p = 0.049), although no two specific years were significantly different. The inbreeding coefficient (F_{IS}) decreased in most populations but increased for salmon in the Conon and Scorff. The only significant difference in F_{IS} was found for salmon within the river Scorff (ANOVA, p = 0.003), between 2005 and 1988 (Tukey test, p = 0.037).

When using the nine consistent loci, significant differences were found only within the Sée and Scorff rivers. Within the Sée, differences were identified in H_O (ANOVA, p = 0.003) between salmon from 2000 and 1988 (Tukey test, p = 0.026) and in F_{IS} (ANOVA, p = 0.034) between the same years (Tukey test, p = 0.03). Within the river Scorff significant differences were identified in H_E (ANOVA, p = 0.006) between 2005 and 1988 (Tukey test, p = 0.0064) and F_{IS} (ANOVA, p = 0.002) between the same years (Tukey test, p = 0.026), both of which had been identified with the full set of loci. Therefore it is possible that more significant differences would have been detected if all temporal datasets had used 16 loci. This also indicates it is unlikely that significant differences were missed by using too many loci. The temporal trends in allelic richness remained the same, such that previous declines across time remained declines and increases remained increases (Table 5.5); however, the magnitude of the slope changed in many instances.

The results from Genepop indicated a change in the allelic and genotypic frequencies in many of the rivers between at least one pair of sampled years (Table 5.6). Only salmon in the Tweed and the Dee do not appear to have seen significant changes in these indices during the sampling period, while salmon from the Avon and Exe showed significant differences between each sampling year. Salmon in the three French rivers each showed a significant difference in both allelic and genotypic frequencies between 2005 and all of the other sampling points.

Effective population size

For the single sample N_E results, three methods contained a large proportion of infinite N_E estimates, which were thus uninformative (Appendix XIV). These were the heterozygote excess (where 100% of N_E estimates were infinite), linkage disequilibrium (30% with infinite N_E) and molecular co-ancestry (44% with infinite N_E). Using the single sample sib-ship method (Figure 5.4a), no calculations were infinite. With this method, the effective population size (N_{Esib}) ranged from 30 in the Sée in the year 2000, to 109 in the Dee in 1999 (Figure 5.4a). They also indicated stability in N_{Esib} within each temporal sample (Figure 5.4a), with largely overlapping 95% confidence intervals. Using the temporal method (Figure 5.4b), some calculations of N_{Etp} were negative. Following past studies (e.g. Palstra & Ruzzante 2008), these were interpreted as very large i.e. infinite. Subsequently, N_{Etp} ranged from 15.4 in the Sée between 2000 and 2005 and infinite in the Dee between 1991 and 1995 and between all Tweed samples. Significant changes were seen in some populations as indicated by non-overlapping 95% confidence limits.

A significant temporal decline was detected in the rivers Dee, Exe, Avon and Sée (Figure 5.4b). Although the River Conon showed a decrease, the overlapping 95% confidence limits indicate the difference was not significant. Salmon from the Ellé showed an initial decrease followed by an increase to the initial level. Salmon from the Scorff only showed a clear significant increase in N_E over the sample period. Repeated sub-sampling of salmon from the Exe 2004, Exe 2009 and Avon 2012 showed minimal deviations in calculated N_E by both the sib-ship method and the temporal method (Appendix XV).

Rod-catch

Over the different time periods of successfully amplified genotypes, increases in rod-catch were identified in the rivers Tweed, Conon, Dee, Exe and Sée (Table 5.7; see Appendix XI for rod-catch data), while decreases were identified in the Avon, Ellé and Scorff. However, there was only a significant correlation with time within the Tweed and Avon (Table 5.7). These results suggest significant declines in effective population size and genetic diversity should not be expected in any river but the Avon.

Further investigation

Nine loci F_{ST} comparisons

The mean pair-wise F_{ST} value between the three French rivers was 0.037 in 1988 (standard deviation (sd) = 0.022), 0.044 in 1994 (sd = 0.031), 0.032 in 2000 (sd = 0.020) and 0.026 in 2005 (sd = 0.017) (Table 5.8). This suggests a temporal decline in the genetic differentiation between the three rivers. The ANOVA test showed no significant difference when all three river samples were included (ANOVA, df = 11, p = 0.80). After omitting the Ellé vs Scorff F_{ST} values however, a significant difference was found (ANOVA, df = 7, p = 0.015). Differences were non-significant following the omission of either of the other pair-wise comparisons (p > 0.05). This suggests that there is a temporal decline in F_{ST} values between the rivers, but the variance caused by comparing these three rivers together is larger than the temporal difference.

Pair-wise F_{ST} values across all samples (Table 5.9) indicated that samples from the Scorff and Ellé were not significantly different during certain years (Scorff 1988, 2000 and 20005 against all Ellé samples). Samples from the Dee and from the Tweed were also not significantly different between certain years (Dee 2007 against all Tweed samples; and Dee 1991 against Tweed 2000).

Nine loci N$_E$

Using the nine loci dataset, calculations of N$_E$ with COLONY were lower in the majority of cases (Figure 5.5a). Samples from the Sée increased, but as the original dataset also included nine (different) loci, this was not truly an exception. All datasets still showed a pattern of temporal stability in effective population size, as they had done when using between 9-16 loci. In effect this result indicates that fewer loci would have reduced N$_E$ estimates for samples with fewer loci, e.g. Exe (11 loci) and Sée (nine loci), but that it is unlikely to have eliminated trends, because all time points were effected equally.

There were a number of changes to temporal effective population size estimates as a result of using the nine loci (Figure 5.5b). N$_E$ estimates were larger in 11 cases and smaller in four. This accentuated the temporal declines of N$_E$ seen in the Exe and Avon. The decline was also heightened within the Sée, but this was the result of different loci rather than fewer. The Conon samples now indicated a significant decline in N$_E$ over the two estimates, whereas previously the decline had been non-significant. The Ellé samples, which previously indicated a decrease in N$_E$ and then increase, now indicated a steep increase and then decrease. The Scorff samples appeared to be continually increasing, whereas before the calculation indicated an increase followed by stability. The effective population size of salmon on the Tweed remained infinite. The changes in some effective population sizes suggest that using fewer and different loci may have affected the outcome of the temporal analysis. However low estimates (N$_E$ < 100) are likely to remain low, while large estimates (N$_E$ > 100) can increase greatly.

Discussion

Genetic diversity

Of all the genetic diversity indices, the number of alleles is predicted to change the most rapidly in response to changes in population size (Garza & Williamson 2001; Hoban *et al.* 2014), while heterozygosity is predicted to change only when a population decline is particularly drastic or for an extended period of time (Garza & Williamson 2001). Only salmon from two rivers, the Sée and Ellé, showed significant changes in A_R over time, although surprisingly these were temporal increases. The oldest Sée samples were likely subject to allelic dropout (Appendix XIIfi) and without these samples, the difference was non-significant ($P > 0.05$). Allelic richness within the remaining populations is effectively stable over the time frame, which is what would be expected based on the recent rod-catch, except for within the Avon. A previous study by Perrier *et al.* (2013), which also investigated temporal change in allelic richness of salmon in French rivers identified no change in allelic richness for salmon in the Sée between 1977 and 2003 and the Scorff between the same years. So it is fitting that samples from the Sée (after removal of the most historic and subject to allelic drop out) and Scorff show no significant change. Although individually changes in A_R are non-significant in most cases, it is noteworthy that both salmon populations from Scotland showed temporal decreases in A_R, while all but one of the rivers in Wales, England and France showed temporal increases.

Biologically, the increases in allelic richness could be explained by an increased level of admixture with exogenous fish, as the alternative – mutation – is too slow to increase the allelic richness within this time frame. The temporal reduction of F_{ST} values between the French samples (Table 5.8) may be a useful indicator, as it suggests a reduction in between-river population structure, which could also be caused by increased admixture. Previous studies have identified similar increases in allelic richness over time (e.g. Horreo *et al.* 2011) and some studies have also proposed this to be the result of increased admixture, caused by changing water temperature, escapees from farms, stocking from non-local sources (Ayllon *et al.* 2006; Glover *et al.* 2012; Perrier *et al.* 2013), or climate change (Perrier *et al.* 2013). Much like in these examples (Ayllon *et al.* 2006; Glover *et al.* 2012; Perrier *et al.* 2013), stocking

has been attempted within most of these French rivers – the Sée and Scorff for definite (Perrier *et al.* 2013) – and English & Welsh rivers – Dee and Exe (Selly *et al.* 2014) – and historically, supplementary salmon have primarily been sourced from Scottish rivers (Russell *et al.* 1995). Stocking with foreign fish ceased in French rivers during the 1980s and where stocking continues it is from local sources. So the fact that the F_{ST} values continue to fall after 1994 (Table 5.8) is especially interesting, as it would suggest that stocking with local fish is also damaging, or that the fall in F_{ST} has nothing to do with stocking. Alternatively, the fall in F_{ST} may be further evidence of increased gene flow between natural populations, as proposed in salmon populations previously (Consuegra *et al.* 2005; Perrier *et al.* 2013). Falls in population declines do not result in a significant loss of genetic diversity, as identified in the Asón between 1950s and 1990s despite a significant fall in rod-catch (Consuegra *et al.* 2005). This has also been identified in other salmonids, for example Steelhead trout (Ardren & Kapuscinski 2003), but also in bluefin tuna (Riccioni *et al.* 2010) and toads (Beebee 2009).

The decreases identified in the Scottish populations could be the result of large endemic populations. Glover *et al.* (2012) identified a strong inverse relationship between the successful introgression of farmed escapees and the density of the native population. Although they identified this effect from farm escapees, straying does occur naturally between rivers, and may also be limited by the endemic population size. It is estimated that the Tweed supports 135,000 salmon (R Campbell pers. comm.), therefore the size of their populations could be preventing introgression from natural strays and the main sources of genetic change within these populations are mutation and genetic drift. However, the River Conon is also subject to significant restocking efforts, with over 2,500,000 ova stocked per annum, and this may also be a factor.

For salmon within the Avon, although it was non-significant, allelic richness decreased temporally more than in any of the other rivers. From the calculations of change in recent rod-catch data (Table 5.7), this is the only population where a decline would be expected. The decrease in genetic diversity could be the result of reduced population size and lack of immigration from neighbouring regions as identified by previous research (Chapter 3). In Chapter 3 it was determined that contemporary migration into the chalk streams

from other regions is minimal, if not completely absent. Thus, the results from this analysis lead us to suspect that while migration is increasing temporally within Great Britain and France, migration into the Avon, and possibly the other chalk streams, has not increased, and what we are viewing may be the un-mitigated effect of population decline. Interestingly, in the study of salmon within French rivers (Perrier *et* al. 2013), the greatest decrease was found within the river Bresle, which is also a chalk stream. These fish were also subject to the greatest change in F_{ST} between temporal samples. It might be unwise to draw too many conclusions from just two rivers; yet a pattern does appear to be emerging.

However, it is also possible that this decline could be an artefact of the different life stages of the samples. The historic samples were adults and so possibly captured more of the variability within the river for the following reasons. Salmon returning to rivers are different ages, due to the varying number of years they spend within river before smolting and the varying number of years they spend at sea. Therefore a sample of adults contains the genetic diversity of salmon spawned over several years. Also, only a fraction of returning adults is fortunate enough to spawn; thus a large amount of variability within the adults is not passed on. Adults are also able to move through the entire river, while the dispersal of juveniles is limited. If population sub-structure exists within river, then sampling a small region of river for juveniles may reduce the proportion of diversity captured even further. The Avon samples from 2004 samples might be particularly biased, because samples were from two sample sites only (Appendix X), whilst the 2012 samples were from a wider but not all encompassing four sample sites. There is however a precedent for sampling both life stages in both Atlantic salmon (Nielsen *et al.* 1997) and brown trout (Østergaard *et al.* 2008). The salmon samples from the Exe also include historic adult samples and contemporary juveniles. As they show a contrasting increase in allelic richness over time, it doesn't appear that any bias caused by the different sampling is repeatedly overwhelming. However, it is noteworthy that the Exe samples from 2009 were from a much larger portion of the catchment (Appendix IX), covering 19 sites over >30 km, while for the Avon in 2012 only 4 sites were sampled from within 5 km of each other (Appendix X). Nonetheless, increased confidence would be gained by obtaining more Avon samples in the future, or making another attempt to amplify microsatellites from the historic

Frome samples. It would be useful from a management perspective to see if this trend continues within the Avon.

No significant differences were detected in observed heterozygosity (H_O) between any temporal samples, although there was a slight temporal decrease in the Conon, Tweed and Scorff and a slight increase in the remainder. Expected heterozygosity (H_E) was only significantly different within the Scorff. A significant difference in the inbreeding coefficient (F_{IS}) is identified within the Scorff only, indicating that salmon in this river have gone from a somewhat expanding and out-breeding population to a somewhat shrinking and inbreeding population.

Change in allele frequencies

Although there were not many significant changes in genetic diversity, there were several changes in the allelic and genotypic frequencies of populations over time. Salmon in the Tweed and the Dee are the only ones that show no significant change in either index between any time points. This is fitting for the Tweed, at least, which has a large population size and is therefore less likely to suffer from the negative effects of smaller populations, such as genetic drift. The Dee population is much smaller, and subject to greater yearly variation. For this reason their temporal stability in allelic and genotypic frequency is more surprising but should be welcome because it shows stability.

Samples from the Exe and Avon show significant differences between each pair of time points. Unfortunately, this parameter is likely to suffer greatest from the sampling of different life stages between time-points. As they are the only samples with both juveniles and adults samples, and also the only samples where all comparisons are significant, it does appear more likely than not that this is an artefact, which may also effect calculations of effective population size, discussed henceforth.

The remaining French samples show one consistent pattern. The majority of the significance stems from the comparisons between the most recent samples -each from 2005- and all of the older samples. The only exception is that there is also a significant difference between the Sée samples from 1980-81 and 2000. There are two possible avenues. The first is that the results are genuine and correspond to a shift in allelic frequencies between

2000 and 2005 and the second is that this is an artefact of genotyping. As discussed previously, there is evidence for increased admixture for salmon in this region (Perrier *et* al. 2013). Although the study by Perrier *et al.* (2013) did not investigate allelic of genotypic frequencies, they did identify increasing admixture between many French populations over time. This may have created the change in frequencies evident in the present study. However, there is a possibility that this is an artefact of genotyping because the 2005 genotypes were the only ones obtained from the SALSEA database. Despite following the calibrations established by Ellis *et al.* (2011), it is possible that little differences in genotyping exist between laboratories. However the results could be genuine and perhaps an effect of the increasing admixture evident in rivers in France (Perrier *et al.* 2013).

Effective population size

These results highlight the importance of using multiple methods to calculate the effective population size of natural populations as there is some discrepancy between the two final methods settled upon. For the following reasons the temporal method is considered more reliable within this study. Firstly, this is the most commonly used genetic method of calculating effective population size (Serbezov *et al.* 2012). Simulations have identified it to be ten-times more effective than test based methods on the loss of heterozygosity (Luikart & Cornuet 1998). Secondly, these estimates more accurately reflect the large differences in population size within the targeted populations; for example rod-catch on the Tweed in 2012 was almost 13,000, while rod-catch on the Avon during the same year was 62, therefore one would expect differences in N_E more like N_{Etp} than N_{Esib}.

Starting with Scotland, the N_E of salmon on the Tweed is too large to be calculated using all available data,. Considering the large population size, this is the most likely population for such a result. More samples or more microsatellite markers would likely improve the ability to infer an estimate of N_E. On the Conon, a decline is detected, but with all the data it is not significant, however the nine loci analysis suggest it is significant. Although this seems to contradict the hypothesis of stability in Scotland, the Conon is one river in Scotland where salmon populations are struggling, relatively. Parr stocking is substantial as

described earlier, and rod-catch has fallen over the past 20 years (Figure 5.3), even if not over the sampling period (1998 – 2012).

In Wales, England and France, there is a complicated picture. In the Exe and Avon N_{Etp} is most likely downwardly biased and therefore unreliable (discussed under Caveats) because of the different life stages of the samples. Of the four remaining rivers, a significant temporal decline is found within the Sée only. These samples did initially show evidence of allelic dropout (Appendix XII Figure fi) however, after changing to the nine consistent loci (and removal of the samples from 1980-81), they did not (Appendix XII Figure fii). Yet a decline in N_E was still evident. This suggests that the N_E decline in the Sée population is real and not just an artefact. Also, even with both sets of nine loci, N_{Etp} of the salmon between 2000 and 2005 is exceptionally low (Figure 5.5b). As these samples were relatively recently sampled, and thus less likely to have suffered DNA degradation due to age (e.g. Table 5.4), this is unlikely to be bias caused by allelic dropout. It is also worth noting that a contrasting estimate of N_E was calculated for salmon in this river between 2002 and 2003 (Perrier *et al.* 2013). This study identified a much healthier N_E estimate of 189 over this period. However, different loci and sampling years were used. Also, the study by Perrier *et al.* (2013) used a novel program calculate N_E – VarEff (Nikolic & Chevalet 2014) – which has not yet undergone rigorous testing like the temporal method employed here and also does not provide confidence intervals, making determining confidence difficult.

The population in the Dee could also be said to have undergone a decline in N_E, if you consider that between 1995 and 1999, the effective population size was too large to calculate. For the remaining two populations in France, overall a slight temporal decrease is found within the Ellé, however this would not be considered significant. For the Scorff, there is a significant increase between the first two time points and the next two, which are stable. In summary, there was no consistent pattern in the N_{Etp} within all of these rivers. Considering the latitudes they span – from the Dee to the Scorff – perhaps this is not surprising.

Temporal genetic differentiation

The range of within-river temporal F_{ST} values (0.005 in the river Tweed to 0.015 in the Avon) is congruent with previous studies, for example (Perrier *et al.* 2013), which contained values between 0.001 and 0.04. Interestingly the rivers with the highest temporal F_{ST} within both studies are both for salmon on chalk geology. There was consistency between the detection of significant pair-wise F_{ST} and change in genotype frequencies, although salmon in the Ellé and Scorff showed significant differences in genotype frequencies only.

Caveats

Although we attempted to use consistent sampling techniques within each river, the rivers Exe and Avon consisted of historic adult samples, and contemporary juvenile samples. This is far from ideal, but is sometimes acceptable, provided that sampling has been sufficiently random to conform to Hardy-Weinberg equilibrium conditions (e.g. Spidle *et al.* 2003; Lage & Kornfield 2006). In these samples, siblings were removed and each of the samples was in Hardy-Weinberg equilibrium. However, simulations have shown that sampling of adults or siblings can introduce bias (Waples & Yokota 2007). Using barnacles as an example, Waples & Yokota (2007) identified that where juveniles are sampled, only the fraction of the population that reproduced the previous year is being represented and that this leads to a greater sampling variance than would be expected (and the difference in allele frequencies, F, between sampling years is greater), which leads to an underestimated N_E. The declines in the N_{Etp} calculations for the Exe and Avon populations are likely to be subject to this bias. Three possible solutions to this are 1) to sample adult salmon, which is difficult without a fish trap, but is possible with the help of fishermen, 2) to sample and pool together consecutive years of juveniles to represent a greater proportion of the salmon within the river, a technique often invoked for single sample N_E estimates (e.g. Lage & Kornfield 2006), or 3) sample a single cohort many generations henceforth. This last point relies on the fact that bias in N_E is largest for short time intervals (Waples & Yokota 2007), and often disappears after 5-10 generations (Waples & Yokota 2007). By allowing more time to pass between sampling points, more episodes of genetic drift are able to influence F (difference in allele frequencies), and subsequently increase the ratio of signal-to noise (Waples & Yokota 2007).

Although not critical to the aims of this study, it is worth noting that during the periods of increased N_{Etp} on the Dee, these salmon were not significantly different from at least one Tweed cohort (Table 5.9). It is possible that this is a case of homoplasy (Estoup *et al.* 2002), i.e. a similarity in allele size caused by independent mutation events and not shared ancestry. However, in a review involving simulations, it has been determined that, with multiple microsatellite loci, this is unlikely (Estoup *et al.* 2002). It is also possible that this is due to straying from the Tweed into the Dee, or vice versa, but because of the large distance between them (+500 km along the coastline) this is also unlikely. Stocking might therefore be the most likely explanation, especially considering that the fish from the Conon, which is between the Dee and the Tweed is less similar to both, than they are to each other (Table 5.9). Although, for decades it has been from native Dee brood-stock only (Ian Davidson, *pers. comm.*), there is a long history of salmon stocking on the Dee, which historically would have been from more successful exogenous rivers, like the Tweed in Scotland. More research is clearly required to determine whether this is a little noticed success, of which there is precedent; some success of stocking with Tweed fish has been found on the river Dart in southern England (Finnegan & Stevens 2008).

Data reliability

In order to determine whether or not samples were accurately genotyped, two methods of analyzing the microsatellite genotypes were employed. These were MICRO-CHECKER and the newly devised plots of dA$_R$/dt against maximum allele size. MICRO-CHECKER, in particular, is considered to be key for detecting technical artefacts and assessing data quality (Nielsen & Hansen 2008). However, there are several further methods that could have been taken in the laboratory to ensure that the data produced was reliable, and should be taken in future.

Taberlet *et al.* (1996) devised a series of steps - based on mathematical models and stochastic events associated with the laboratory steps involved in DNA amplification - that could be taken to ensure that reliable genotypes are obtained in the laboratory with 99% confidence. However as this involves the use of ten tubes per locus, it is impractical for larger studies, and would not have been possible in the present study without removing a large proportion of the samples.

The scientific literature contains many further recommendations for extracting and amplifying DNA reliably (Nielsen *et al.* 1997, 1999; Morin & Mccarthy 2007; Smith *et al.* 2011). Considering the volume of data that has already been obtained for the present study, the next logical step would be to determine the reliability of this data by repeating the PCRs for a proportion of samples (10%) from each river and time period completed. It would also be useful to repeat the DNA extraction process, for a smaller proportion of samples, followed by the PCR, following recommendations by Nielsen & Hansen (2008). However for many of these samples, only one scale was available for donation or donated.

Further work

The present study was limited by the availability of Atlantic salmon scales, the need to have three temporal samples from within a single river, and crucially the quality of the DNA extracted from the samples. Rather than attempting to amplify microsatellite data from past scale samples, it might prove effective to use contemporary and future samples. However, few regions have the resources to sample adults, so in order to continue building upon the collected database, it would be beneficial to identify a means of using and comparing juvenile and adult samples. This might be achieved by sampling and combining two or more consecutive juvenile cohorts. Because adults do not all return to spawn at the same time, sampling of adults within a single year will contain salmon born several years apart. By sampling juveniles of multiple years it might be possible to capture that range better. Alternatively, if the adult ages are known, then possibly the genetic diversity of just adults born during a single year could be compared to juveniles.

Many of the estimators of effective population size provided estimates with confidence limits too wide to be of use. This included the temporal estimator by Jorde & Ryman (1995) which we used here. This was unexpected as previous studies have used a comparable number of loci and samples to produce estimates of N_E, for example 11 loci and 11-65 individuals were used by Perrier *et al.* (2013), and nine loci and 18-60 individuals were used by Horreo *et al.* (2011). Increasing the number of samples or loci would likely increase precision. This would be an ideal opportunity to develop primer mixes involving

smaller loci, which would have a better chance of amplification, and a reduced chance of allelic dropout (Takahashi *et al.* 1997).

It is also clear from this study that the quality and quantity of DNA is inversely correlated with the age of historic material collected. If the wealth of available scales are to attain their potential usefulness, it seems prudent to identify a long term storage process for the genetic material within them. This may mean extracting and storing the DNA sooner rather than later.

Conclusion

This study sought to infer the genetic trajectory of Atlantic salmon within Scotland, England, Wales and France using historic material, and at the same time determine the feasibility. There were significant obstacles, and it's clear that salmon scales older than 30 years when stored ad hoc do not make effective sources of DNA for population genetics, at least not with these microsatellite markers. Despite the difficulties, genotypes were obtained from eight salmon populations ranging from the Conon in Scotland to the Sée in France. Within these samples no significant decrease in allelic richness were detected within any population, which was expected for populations in England Wales and France. This is partly because the time period of these samples is smaller than it could have been if genotypes from the oldest samples had successfully yielded DNA. In support, the analysis of rod-catch data (used as a surrogate for population size) also showed that over the sampling time periods it is likely that population size did not fall significantly. Another explanation for the lack of evident decline may be that genetic diversity is slow to respond to population declines; as possible evident by it scarcity within the literature.

Surprisingly, increases in allelic richness were detected in all English, Welsh and French population except one, and within two rivers the increases were significant. This may due to increased admixture, which may be an effect of past supplemental stocking with non-local fish, farm escapees, changes in climate or even a natural effect of population declines (Consuegra *et al.* 2005; Valiente *et al.* 2010; Horreo *et al.* 2011b). The present study does not have the power to determine which it might be.

There were significant changes in allelic and genotypic frequencies in all but two of the populations. However, for the Avon and Exe popuation, these

may be artefacts of sampling both juveniles and adults. For the French salmon it is possible that this is an artefact of genotyping discrepancy between the laboratories in France and Exeter. Despite the significant research previously done, to ensure calibration (Ellis *et al*. 2011), repeat genotyping of the 2005 samples within the Exeter laboratory would be needed to determine whether there is a laboratory based genotyping difference or not.

A single conclusion regarding the effective size of these populations is difficult. The sib-ship method indicates stability in all populations, with all datasets. The temporal method, which appears more responsive and is typically considered the most reliable estimate of N_E is likely to be the more reliable result. With this method, salmon in Scotland appear stable, as was hypothesised. The hypothesis of decline for the remaining populations, however, was perhaps too simple. In England and France, evident declines are likely to be downwardly biased, caused by differences in sample life stages and more work needs to be done to accommodate juvenile and adult samples. The remaining rivers do not show a consistent pattern, which might be a true reflection of their population size. To better elucidate the relationship between actual population size and effective population size, both better estimates of actual salmon population size, and more reliable methods of calculating N_E are needed.

Due to the difficulty of obtaining the microsatellite genotypes in historic samples, we were unable to take some essential laboratory steps, described in the discussion, to assess fully the reliability of the data. Although the obtained data was analyzed using two methods, it is still possible that the genotypes from some of the populations could be unreliable. Therefore the conclusions reached must be viewed with scepticism until repeats are completed.
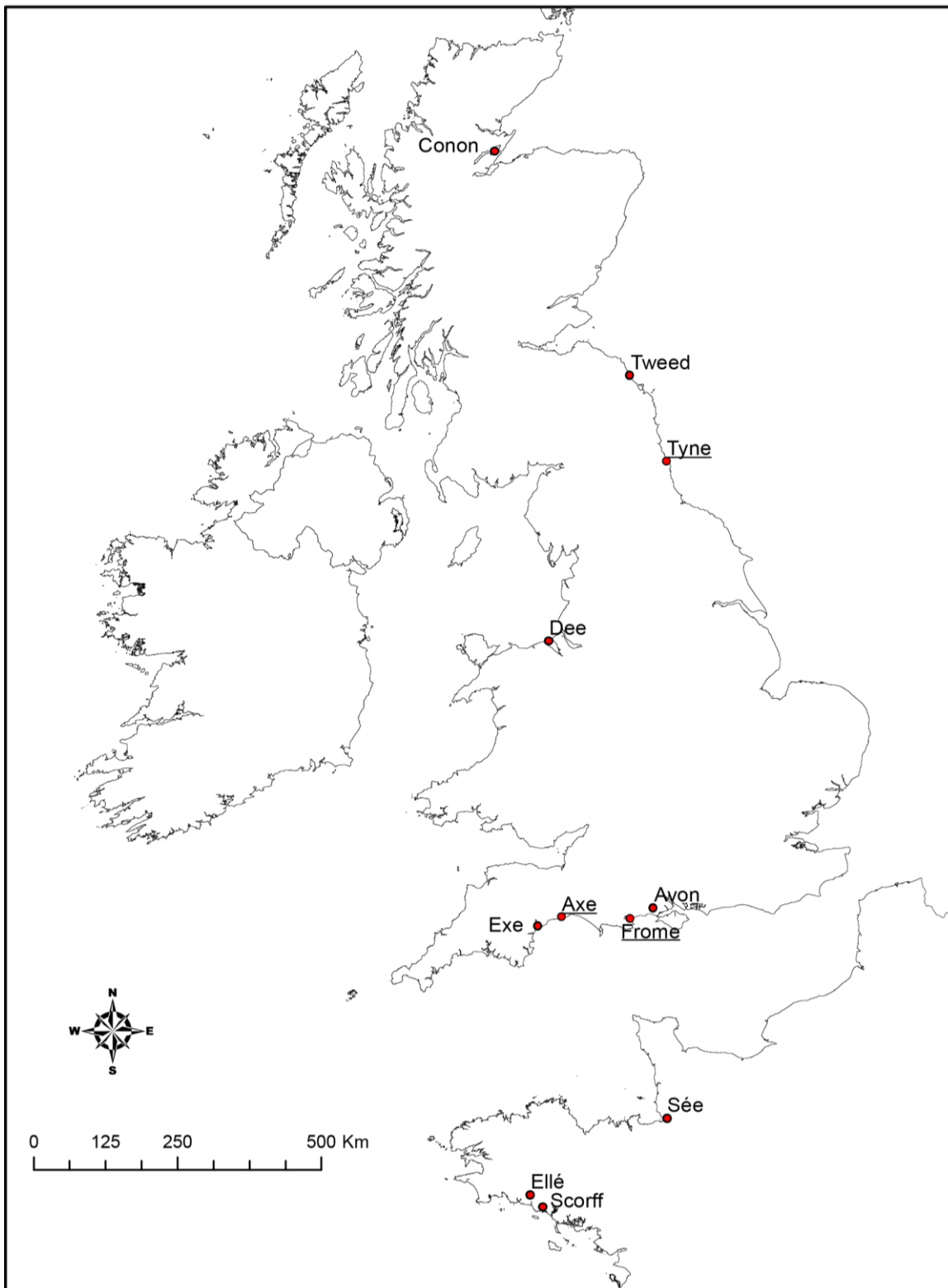
Figure 5.1- Map of historically sampled rivers. The map indicates the location of all rivers from which historic samples have been obtained during the present study. Underlined rivers indicate samples that failed to make the final dataset due to poor amplification of microsatellite DNA.
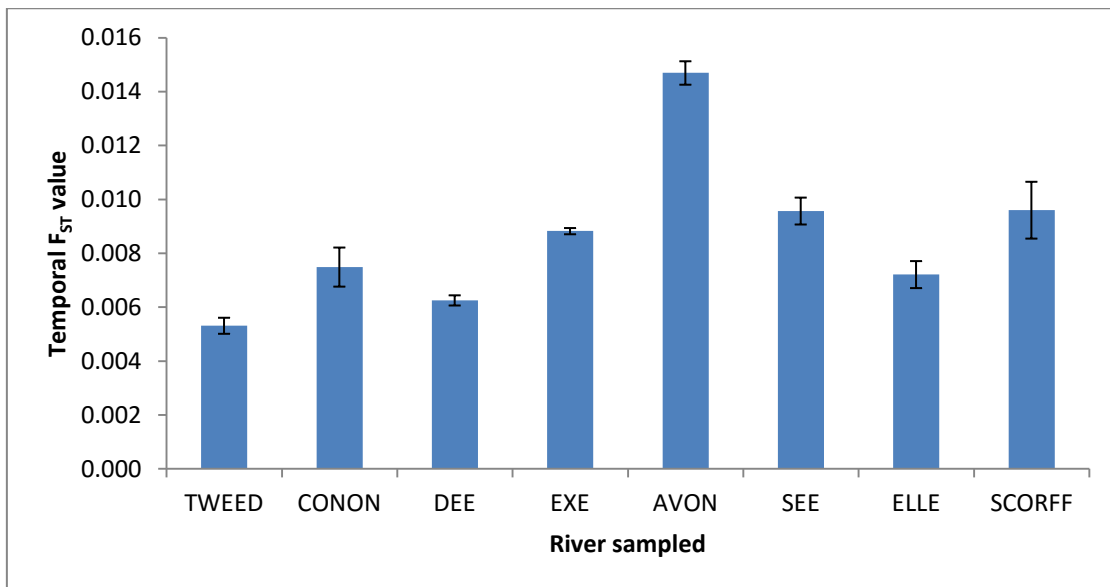
Figure 5.2- Temporal pair-wise F_{ST} values within sampled rivers. The average of pair-wise F_{ST} values between each sample within each temporal dataset, as calculated using all available loci. Error bars indicate the standard error surrounding the mean.

3a)
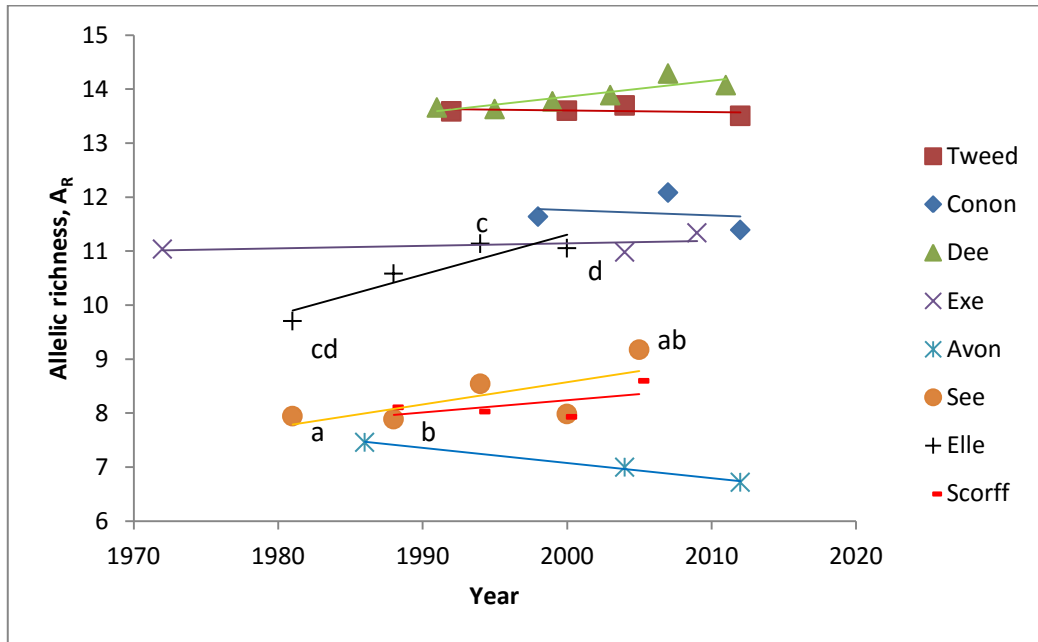
3b)
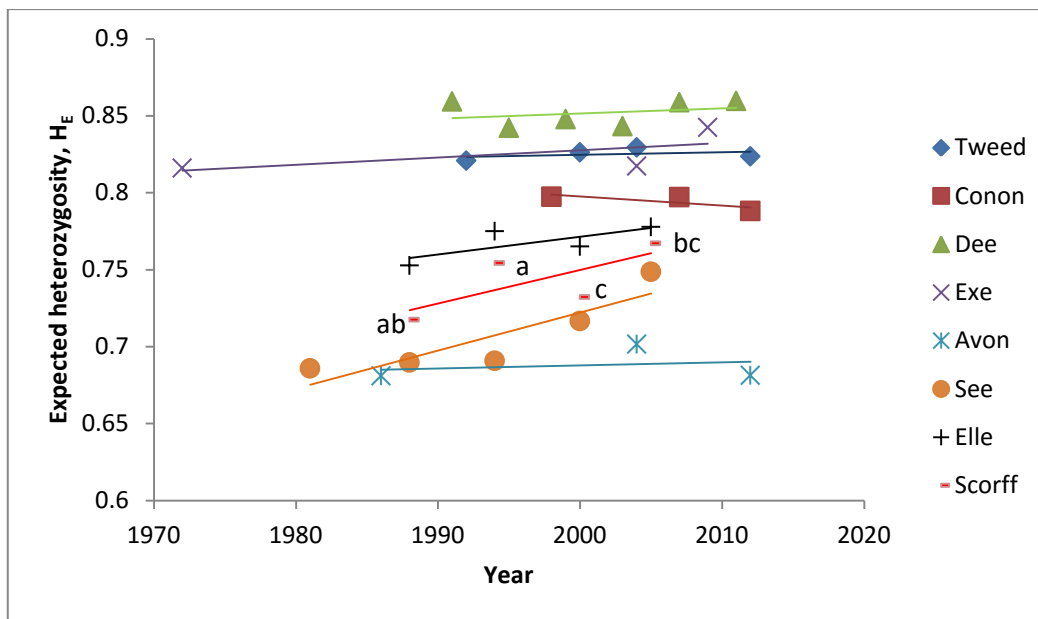
3c)
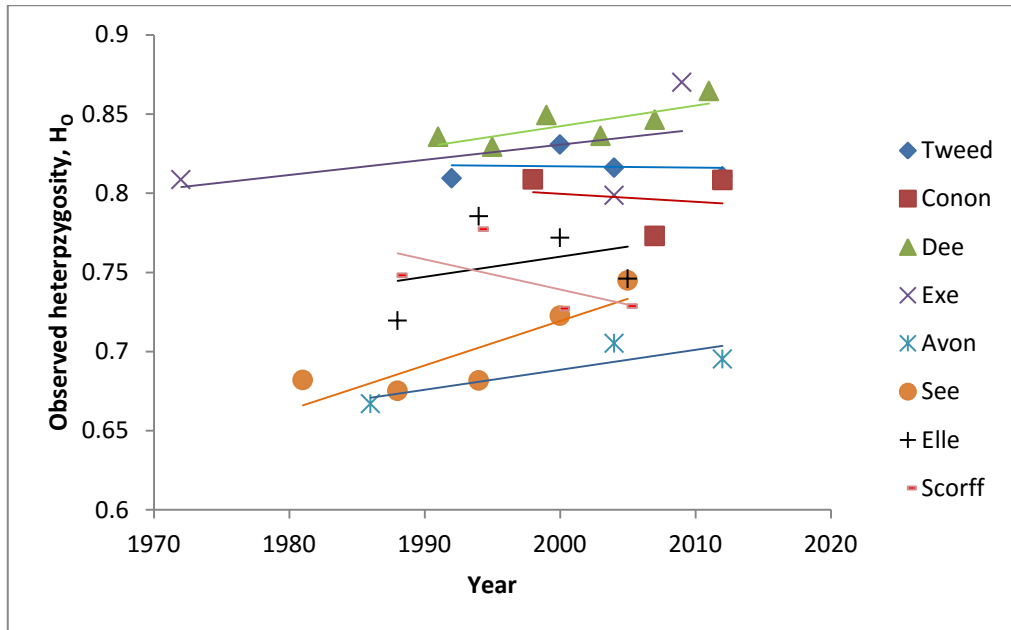
3di)
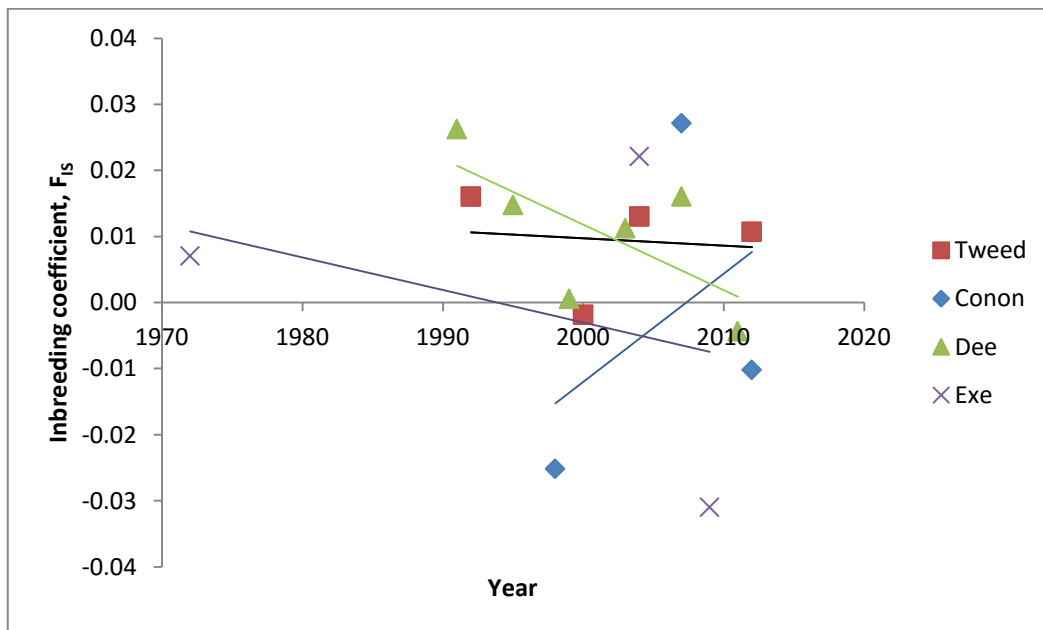
3dii)
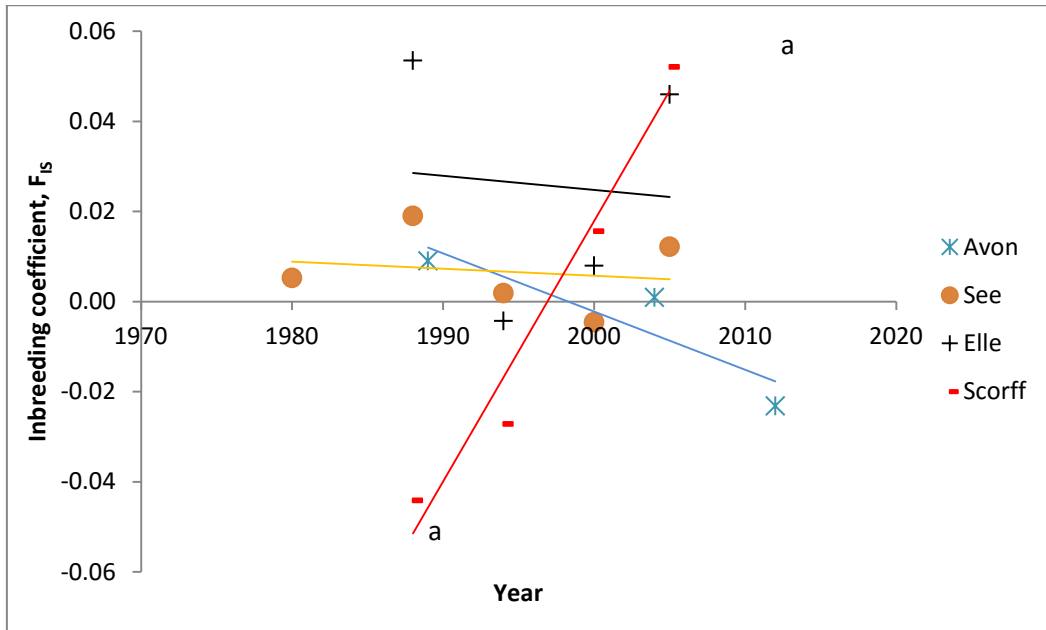
Figure 5.3- Graphs of temporal change in genetic diversity in populations using up to 16 microsatellites. Graphs depict the a) allelic richness A_R, b) expected heterozygosity H_E, c) observed heterozygosity H_O and d) inbreeding coefficient divided into i) and ii) for clarity, of salmon within each river at each time point. Lines of best fit are included to indicate the temporal trend within each river, and are coloured to match the data points they represent. Letters within 3a, 3b and 3dii indicate between which samples significant differences were found (Tukey test, $p < 0.05$)

5.4a)
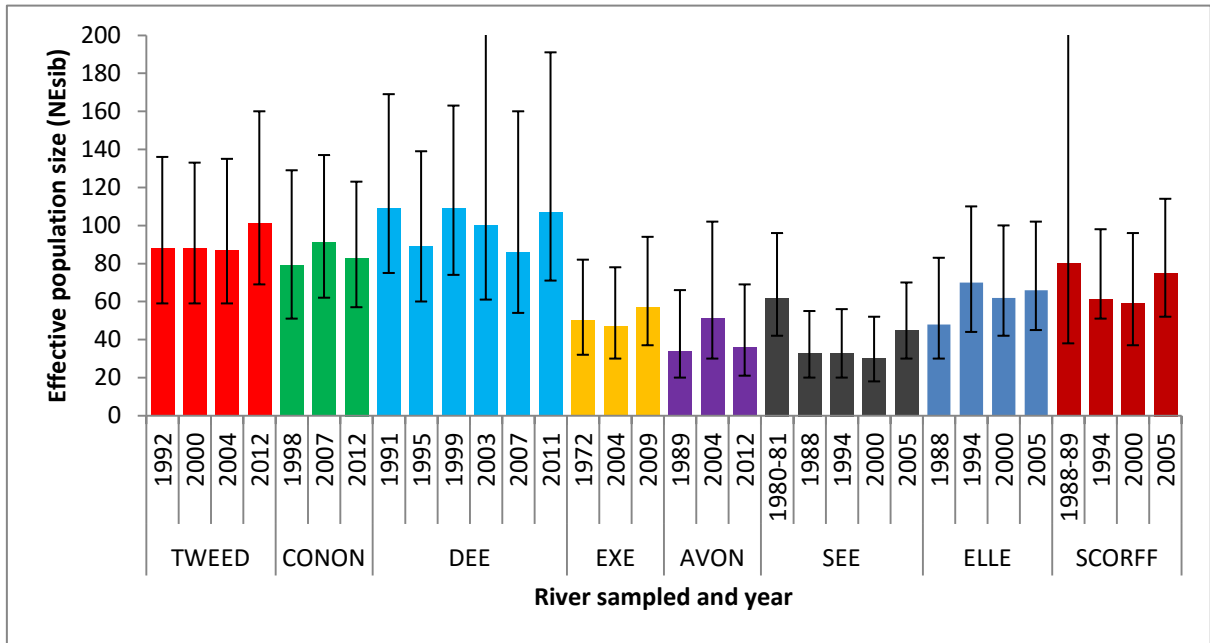
5.4b)
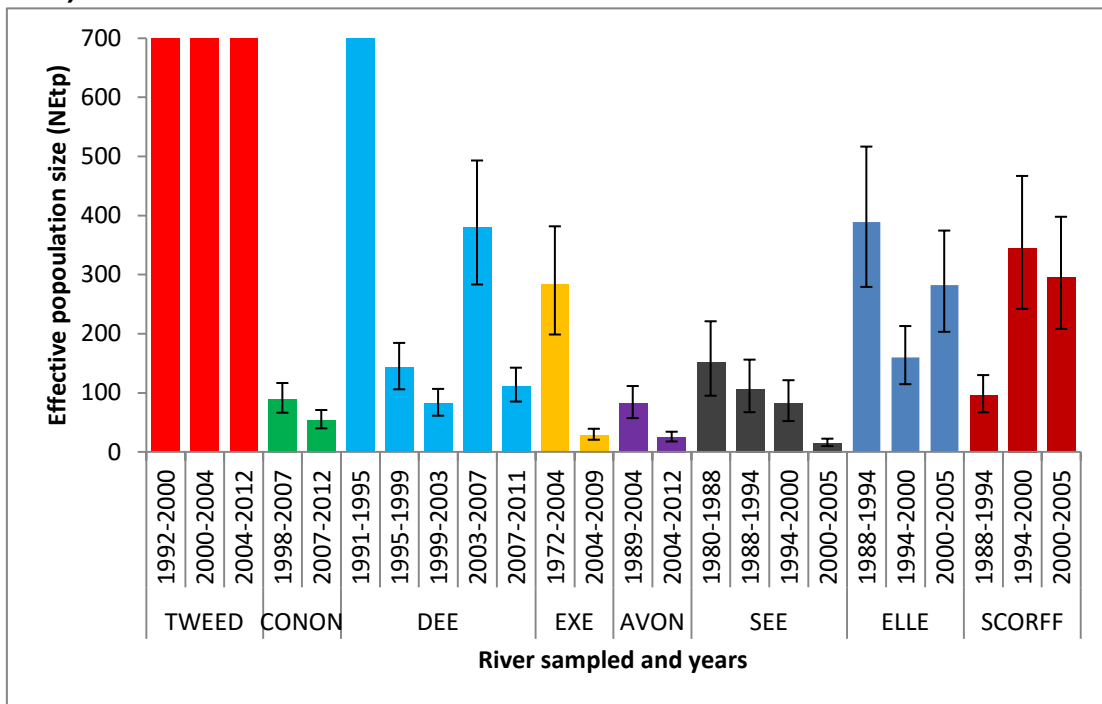
Figure 5.4 – Effective population size (N_E) estimates of temporal salmon samples. Graphs show N_E within each river at a) each time point calculated by COLONY and b) between adjacent time points calculated by the Jorde & Ryman (1995) temporal method.
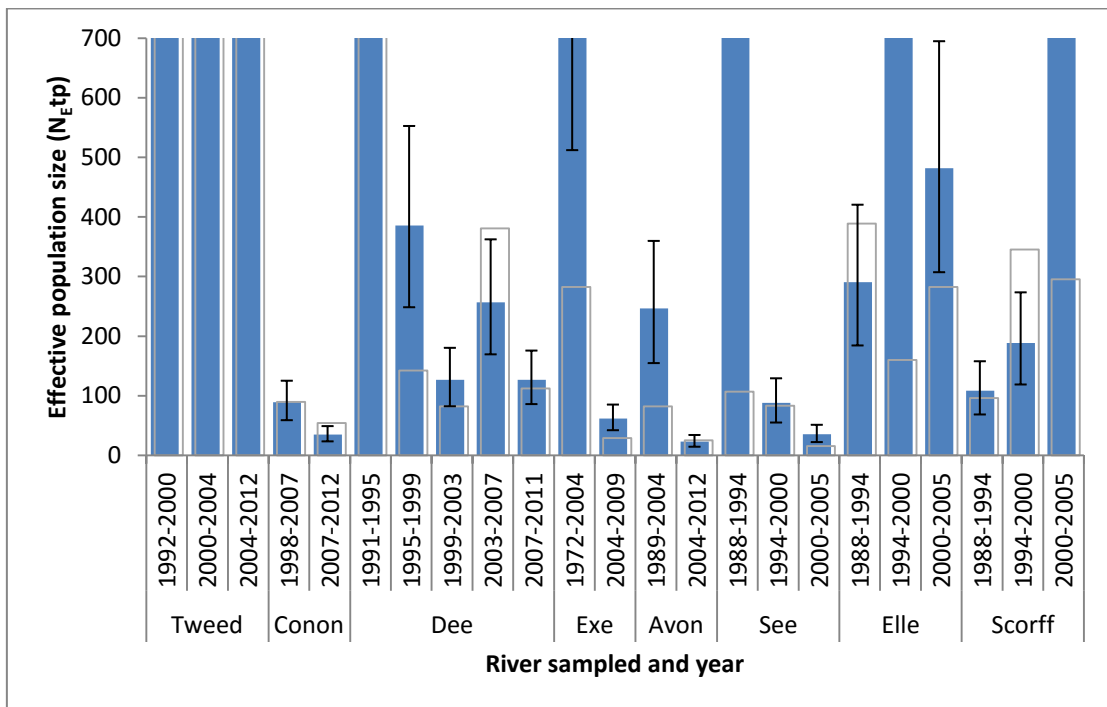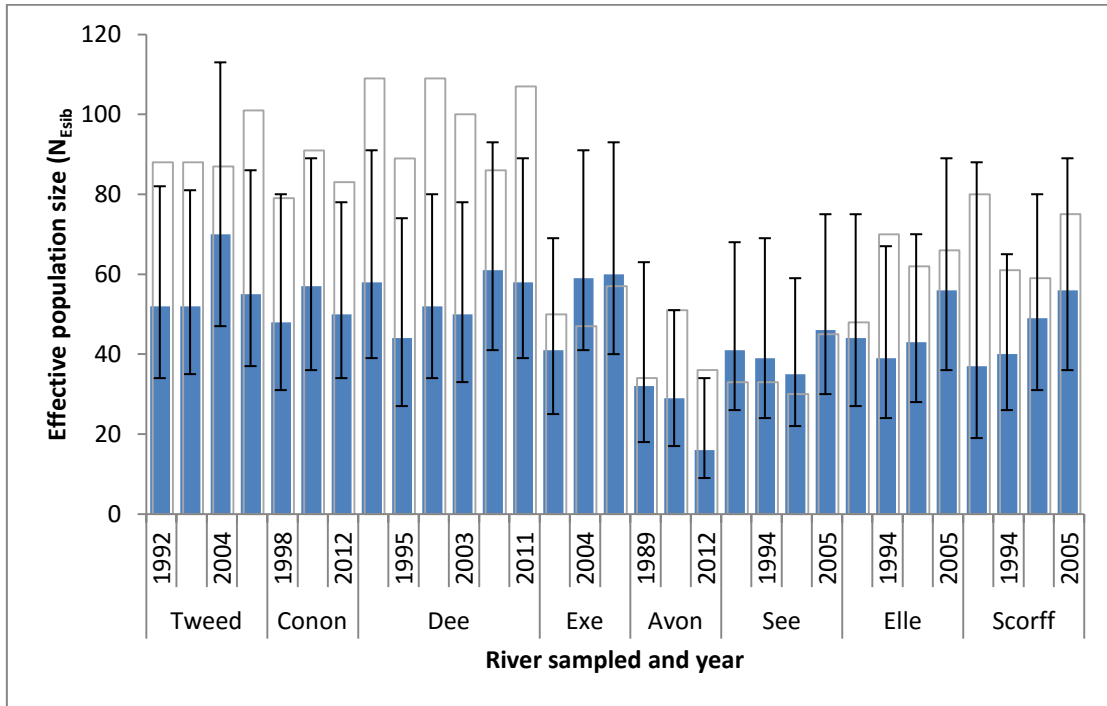
Figure 5.5- Effective population size (N_E) of salmon using nine consistent microsatellite loci. Graphs indicate the effect of using nine consistent loci (blue) rather than as many of the 16 as possible (grey outline) on calculations of at a) N_{Esib} each time point calculated by COLONY and b) N_{Etp} between adjacent time points calculated by the Jorde & Ryman (1995) temporal method.

Table 5.1- Status of salmon in Scotland, England & Wales and France as determined by the World Wildlife Federation (WWF 2001).

Country	Total number of historic salmon bearing rivers	Unknown status	Healthy	Vulnerable	Endangered	Critical	Extinct
Scotland	350	0%	63%	0%	37%	0%	0%
England & Wales	76	5%	33%	14%	25%	14%	9%
France	47	11%	0%	6%	21%	32%	30%

Table 5.2- Number of temporal samples that passed and failed to amplify. The table summarises which samples were obtained and how many were included within the final dataset due to the successful amplification within most individuals for nine or more loci. For adults, initial samples indicates the number of individuals for which genotyping was attempted.   For juveniles, the same column indicates the number of genotypes obtained from previous studies. Also included are the number of loci within each temporal dataset, the name of omitted loci and the average number of alleles(Na) for the remaining loci within each population. Samples from the Tyne, Axe and Frome failed to amplify consistently within crucial samples, and therefore were not used.

River	Sample year	Life stage	Initial sample size	Final sample size	No. loci	Omitted loci	Na
Conon	1998	Adult	42	36	15	Ssosl85	12.0
	2007	Adult	60	52	15		13.7
	2012	Adult	60	52	15		13.0
Tweed	1992	Adult	48	44	16	None	14.0
	2000	Adult	48	45	16		14.1
	2004	Adult	48	47	16		14.4
	2012	Adult	48	47	16		14.1
Dee	1991	Adult	50	46	14	417, Ssa202	14.5
	1995	Adult	50	41	14		14.1
	1999	Adult	50	46	14		14.6
	2003	Adult	50	49	14		15.1
	2007	Adult	50	50	14		15.4
	2011	Adult	50	47	14		15.1
Exe	1966	Adult	48	0	11	SSspG7, Ssa202, Ssa14, SSsp2201, Ssa157	NA
	1972	Adult	46	33	11		14.1
	2004	Juvenile	83	83	11		14.1
	2009	Juvenile	648	49	11		13.2
Avon	1951	Adult	47	0	NA	Ssa202	NA
	1986	Adult	26	23	15		7.1
	2004	Juvenile	43	43	15		7.9
	2012	Juvenile	113	50	15		7.9
Sée	1977-78	Adult	66	0	NA	SSspG7, Ssosl85, Ssa157, Ssa171, Ssa289, SsaD144, SSsp2201	NA
	1980-81	Adult	62	51	9		8.5
	1988	Adult	40	37	9		8.2
	1994	Adult	40	36	9		8.8
	2000	Adult	30	37	9		8.2
	2005	Adult	Unknown	46	9		10.3
Ellé	1968	Adult	42	0	NA	Ssosl85	NA
	1988	Adult	40	33	15		10.0
	1994	Adult	40	38	15		11.3
	2000	Adult	40	40	15		12.2
	2005	Adult	Unknown	47	15		12.6
Scorff	1972	Adult	72	0	NA	Ssosl85, Ssa157	NA
	1988-89	Adult	19	16	14		10.0
	1994	Adult	40	40	14		11.0

	2000	Adult	42	37	14		10.6
	2005	Adult	Unknown	45	14		12.6
Tyne	1991	Adult	50	0	NA	NA	NA
	1996	Adult	50	0	NA		NA
	2004	Adult	26	0	NA		NA
	2008	Adult	48	0	NA		NA
	2012	Adult	50	0	NA		NA
Axe	1963	Adult	50	0	NA	NA	NA
	1966	Adult	58	0	NA		NA
	1969	Adult	50	0	NA		NA
	1972	Adult	50	0	NA		NA
	1975	Adult	53	0	NA		NA
Frome	1954	Adult	60	0	NA	NA	NA
	1970	Adult	48	0	NA		NA
	1975	Adult	51	0	NA		NA
	2009	Juvenile	>200	0	NA		NA
	2011	Juvenile	>200	0	NA		NA

Table 5.3- Table depicting the adjustment required to convert microsatellite data from the SALSEA Merge database to match the Exeter format.

Microsatellite marker	Exeter to SALSEA baseline	SALSEA to Exeter baseline
SSspG7	2	-2
Ssosl417	-2	2
Ssa202	-2	2
Ssa197	5	-5
SSsp2210	2	-2
Ssa289	5	-5
SSsp2201	2	-2
Ssol85	-4	4
Ssa14	-1	1
SsaF43	4	-4
SSsa2216	-1	1

Table 5.4- Concentration of DNA within historic samples. Table shows the average fragment size in base pairs (bp), concentration and molarity within the DNA extractions taken from historic Axe samples and one contemporary sample.

	Size (bp)	Concentration (ng/µl)	Molarity (nmol/l)
Mer.adu02.02	3902	1.09	0.4
Axe.adu63.01	-	-	-
Axe.adu63.02	174	5.7	49.6
Axe.adu63.03	185	6.21	50.8
Axe.adu63.05	189	8.36	67.2
Axe.adu63.06	91	1.77	29.5
Axe.adu63.07	161	12.27	115.7
Axe.adu75.01	199	4.14	31.5
Axe.adu75.02	194	3.73	29.2
Axe.adu75.05	182	43.73	364.2
Axe.adu75.06	139	4.1	44.5
Axe.adu75.07	183	27.55	227.6

Table 5.5– The rate of change (slope) of allelic richness in temporal samples. Table shows allelic richness (A_R) using up to 16 loci (left columns) and nine loci common between all samples (right columns). * Sée samples with nine loci do not include samples from 1980-1 because these samples were missing too many microsatellite loci. RSQ indicates the goodness of fit between allelic richness and the calculated slope, where 0 = bad fit and 1 good fit.

River	9 - 16 loci			9loci		
	Slope	Correlation	RSQ	Slope	Correlation	RSQ
Tweed	-0.00304	-0.32102	0.103057	-0.01511	-0.86887	0.75494
Conon	-0.10094	-0.97122	0.943275	-0.01384	-0.26982	0.072802
Dee	0.029687	0.870105	0.757082	0.02824	0.593758	0.352549
Exe	0.004598	0.479839	0.230245	0.009129	0.999853	0.999705
Avon	-0.03657	-0.99559	0.991191	-0.00503	-0.39857	0.158855
Sée*	0.041075	0.706743	0.499486	0.014483	0.245172	0.06011
Ellé	0.082181	0.921077	0.848383	0.046091	0.768615	0.590769
Scorff	0.022706	0.56361	0.317656	0.02518	0.70775	0.50091

Table 5.6- Table to show difference in allele (left) and genotype (right) frequencies. Chi^2, degrees of freedom (df) and p value calculated by Genepop for populations using between 9 and 16 loci depending on the population. Significant pair-wise comparisons (Chi^2 P > 0.05) are in bold. HS indicates highly significant.

Populations		Allele Chi2	df	P value	Populations		Genotype Chi2	df	P value
Con98	**Con07**	**60.22**	**32**	**0.002**	**Con98**	**Con07**	**53222.00**	**32**	**0.010**
Con98	**Con12**	**58.72**	**32**	**0.003**	**Con98**	**Con12**	**52.50**	**32**	**0.013**
Con07	**Con12**	**Infinity**	**32**	**HS**	**Con07**	**Con12**	**60.50**	**32**	**0.002**
Twe92	Twe00	28.96	32	0.621	Twe92	Twe00	26.25	32	0.752
Twe92	Twe04	37.79	32	0.222	Twe92	Twe04	33.88	32	0.377
Twe92	Twe12	29.88	32	0.574	Twe92	Twe12	27.97	32	0.671
Twe00	Twe04	36.23	32	0.278	Twe00	Twe04	33.56	32	0.391
Twe00	Twe12	31.55	32	0.489	Twe00	Twe12	29.41	32	0.598
Twe04	Twe12	22.11	32	0.904	Twe04	Twe12	21.19	32	0.928
Dee91	Dee95	21.84	32	0.911	Dee91	Dee95	20.20	32	0.948
Dee91	Dee99	38.40	32	0.202	Dee91	Dee99	37.99	32	0.215
Dee91	Dee03	34.34	32	0.356	Dee91	Dee03	31.63	32	0.485
Dee91	Dee07	33.19	32	0.409	Dee91	Dee07	29.87	32	0.574
Dee91	Dee11	45.55	32	0.057	Dee91	Dee11	42.75	32	0.097
Dee95	Dee99	28.23	32	0.658	Dee95	Dee99	26.86	32	0.724
Dee95	Dee03	36.82	32	0.256	Dee95	Dee03	36.01	32	0.286
Dee95	Dee07	40.75	32	0.138	Dee95	Dee07	39.67	32	0.165
Dee95	Dee11	39.32	32	0.175	Dee95	Dee11	37.59	32	0.228
Dee99	Dee03	29.10	32	0.614	Dee99	Dee03	26.88	32	0.723
Dee99	Dee07	21.58	32	0.918	Dee99	Dee07	20.72	32	0.938
Dee99	Dee11	40.78	32	0.137	Dee99	Dee11	37.99	32	0.215
Dee03	Dee07	26.78	32	0.728	Dee03	Dee07	24.79	32	0.815
Dee03	Dee11	31.34	32	0.500	Dee03	Dee11	29.17	32	0.610
Dee07	Dee11	33.96	32	0.373	Dee07	Dee11	32.83	32	0.426
Exe72	**Exe04**	**Infinity**	**22**	**HS**	**Exe72**	**Exe04**	**Infinity**	**22**	**HS**
Exe72	**Exe07**	**37.32**	**22**	**0.022**	**Exe72**	**Exe07**	**36.88**	**22**	**0.024**
Exe04	**Exe07**	**Infinity**	**22**	**HS**	**Exe04**	**Exe07**	**Infinity**	**22**	**HS**
Avon86	**Avon04**	**Infinity**	**28**	**HS**	**Avon86**	**Avon04**	**Infinity**	**28**	**HS**
Avon86	**Avon12**	**Infinity**	**30**	**HS**	**Avon86**	**Avon12**	**Infinity**	**30**	**HS**
Avon04	**Avon12**	**Infinity**	**30**	**HS**	**Avon04**	**Avon12**	**Infinity**	**30**	**HS**
See8081	See88	27.78	20	0.115	See8081	See88	24.90	20	0.205
See8081	See94	30.96	20	0.056	See8081	See94	31.15	20	0.053
See8081	**See00**	**40.80**	**20**	**0.004**	**See8081**	**See00**	**37.07**	**20**	**0.011**
See8081	**See05**	**Infinity**	**20**	**HS**	**See8081**	**See05**	**Infinity**	**20**	**HS**
See88	See94	28.63	20	0.095	See88	See94	27.64	20	0.118
See88	See00	29.18	20	0.084	See88	See00	26.56	20	0.148
See88	**See05**	**Infinity**	**20**	**HS**	**See88**	**See05**	**Infinity**	**20**	**HS**
See94	See00	26.21	20	0.159	See94	See00	25.05	20	0.200
See94	**See05**	**Infinity**	**20**	**HS**	**See94**	**See05**	**Infinity**	**20**	**HS**
See00	**See05**	**Infinity**	**20**	**HS**	**See00**	**See05**	**Infinity**	**20**	**HS**
Elle88	Elle94	42.71	30	0.062	Elle88	Elle94	41.94	30	0.072

Elle88	Elle00	57.18	30	0.002	**Elle88**	**Elle00**	**55.02**	**30**	**0.004**
Elle88	**Elle05**	**Infinity**	**30**	**HS**	**Elle88**	**Elle05**	**Infinity**	**30**	**HS**
Elle94	Elle00	43.44	30	0.054	Elle94	Elle00	41.57	30	0.078
Elle94	**Elle05**	**Infinity**	**30**	**HS**	**Elle94**	**Elle05**	**Infinity**	**30**	**HS**
Elle00	**Elle05**	**Infinity**	**30**	**HS**	**Elle00**	**Elle05**	**Infinity**	**30**	**HS**
Sco88	Sco94	27.56	30	0.594	Sco88	Sco94	27.41	30	0.602
Sco88	Sco00	33.16	30	0.315	Sco88	Sco00	33.01	30	0.322
Sco88	**Sco05**	**Infinity**	**30**	**HS**	**Sco88**	**Sco05**	**Infinity**	**30**	**HS**
Sco94	Sco00	51.38	30	0.009	Sco94	Sco00	50.96	30	0.010
Sco94	**Sco05**	**Infinity**	**30**	**HS**	**Sco94**	**Sco05**	**Infinity**	**30**	**HS**
Sco00	**Sco05**	**Infinity**	**30**	**HS**	**Sco00**	**Sco05**	**Infinity**	**30**	**HS**

Table 5.7– Calculated trends in salmon rod-catch over time. The table shows the calculated slope (i.e. the rate of change) of salmon rod-catch over the years genotypes were successfully obtained. RSQ indicates the goodness of fit of the data to the slope. Pearson's indicates the Pearson's Correlation (r), and p indicates the probability of Pearson's Correlation with significance identified by $p < 0.05$.

River	Years	Slope	RSQ	Pearson	p
Tweed*	1992-2012	406.413	0.50422	0.710085	0.000252
Conon	1998-2012	7.821429	0.009528	0.09761	0.12199
Dee	1992-2011	7.061039	0.036948	0.192217	0.379581
Exe	1972-2009	4.07375	0.025858	0.160803	0.334485
Avon*	1989-2012	-8.0336	0.28537	-0.5342	0.00815
Sée	1980-2005	0.223077	0.000121	0.011015	0.957376
Ellé	1988-2005	-3.16718	0.023966	-0.15481	0.538665
Scorff	1988-2005	-0.01858	1.18E-05	-0.00343	0.986704

Table 5.8- Pair-wise F_{ST} values between temporal French samples. Table shows the pair-wise F_{ST} value between salmon in the Sée, Ellé and Scorff during four different time points

Year	Sée vs Ellé	Sée vs Scorff	Ellé vs Scorff
1988	0.051	0.049	0.011
1994	0.058	0.066	0.009
2000	0.047	0.040	0.009
2005	0.032	0.039	0.006

Table 5.9- Pair-wise FST values between salmon from all included rivers and time points. Values calculated using nine loci. Included are the pair-wise FST values (bottom diagonal) and the significance (upper diagonal) determined from 999 permutations. Bold values indicate non-significant comparisons (p >0.05) and underlined values indicate non-significant comparisons between samples from different rivers.

	Twe92	Twe00	Twe04	Twe12	Con98	Con07	Con12	Dee91	Dee95	Dee99	Dee03	Dee07	Dee11	Exe72	Exe04	Exe09	Avo89	Avo04	Avo12	Sée88	Sée94	Sée00	Sée05	Ell88	Ell94	Ell00	Ell05	Sco88	Sco94	Sco00	Sco05
Tweed92		**0.981**	**0.389**	**0.358**	0.001	0.001	0.001	0.010	0.001	0.002	0.001	**0.088**	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Tweed00	**0.004**		**0.727**	**0.641**	0.001	0.001	0.001	**0.150**	0.002	0.004	0.011	**0.362**	0.016	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Tweed04	**0.006**	**0.005**		**0.835**	0.001	0.001	0.001	0.005	0.001	0.001	0.002	**0.096**	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Tweed12	**0.006**	**0.006**	**0.005**		0.001	0.001	0.001	0.007	0.001	0.001	0.002	**0.114**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Conon98	0.018	0.017	0.020	0.021		**0.056**	**0.299**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Conon07	0.012	0.013	0.014	0.016	**0.009**		0.014	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Conon12	0.015	0.016	0.017	0.018	**0.007**	0.008		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee91	0.009	0.007	0.009	0.009	0.022	0.017	0.019		**0.661**	**0.139**	**0.060**	**0.423**	**0.500**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee95	0.012	0.012	0.013	0.013	0.023	0.019	0.022	**0.006**		**0.221**	**0.355**	**0.280**	**0.323**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee99	0.010	0.009	0.012	0.011	0.019	0.013	0.015	**0.007**	**0.007**		**0.222**	**0.316**	**0.056**	0.005	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee03	0.011	0.009	0.010	0.011	0.024	0.018	0.022	**0.007**	**0.006**	**0.006**		**0.576**	**0.341**	0.005	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee07	**0.007**	**0.006**	**0.007**	**0.007**	0.020	0.014	0.018	**0.006**	**0.007**	**0.006**	**0.005**		**0.169**	0.002	0.001	0.004	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Dee11	0.010	0.009	0.010	0.012	0.021	0.015	0.018	**0.006**	**0.007**	**0.007**	**0.006**	**0.006**		0.001	0.001	0.004	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Exe72	0.017	0.017	0.017	0.018	0.023	0.018	0.020	0.016	0.016	0.011	0.013	0.013	0.014		**0.249**	**0.139**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Exe04	0.014	0.016	0.014	0.014	0.027	0.017	0.021	0.016	0.019	0.015	0.015	0.013	0.016	**0.008**		0.030	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Exe09	0.011	0.011	0.010	0.011	0.022	0.017	0.019	0.010	0.014	0.011	0.011	0.009	0.009	**0.008**	0.007		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Avon89	0.051	0.053	0.058	0.051	0.071	0.068	0.072	0.050	0.046	0.054	0.051	0.044	0.051	0.071	0.068	0.061		**0.080**	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Avon04	0.048	0.049	0.051	0.046	0.067	0.067	0.066	0.045	0.043	0.051	0.048	0.042	0.047	0.066	0.065	0.054	**0.012**		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Avon12	0.044	0.048	0.049	0.047	0.059	0.057	0.060	0.044	0.042	0.046	0.042	0.038	0.045	0.056	0.055	0.050	0.017	0.016		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sée88	0.033	0.035	0.038	0.034	0.036	0.038	0.037	0.034	0.029	0.028	0.034	0.029	0.033	0.036	0.039	0.035	0.054	0.053	0.050		**0.851**	**0.164**	0.030	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sée94	0.037	0.039	0.043	0.036	0.040	0.042	0.041	0.036	0.032	0.030	0.036	0.032	0.035	0.040	0.045	0.038	0.049	0.050	0.047	**0.006**		**0.141**	0.053	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sée00	0.031	0.034	0.039	0.033	0.039	0.037	0.040	0.033	0.030	0.029	0.035	0.029	0.034	0.039	0.038	0.036	0.050	0.052	0.048	**0.008**	**0.009**		0.052	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Sée05	0.025	0.027	0.031	0.027	0.033	0.033	0.033	0.025	0.022	0.023	0.030	0.023	0.027	0.035	0.034	0.030	0.043	0.040	0.041	**0.009**	**0.009**	**0.008**		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Ellé88	0.030	0.031	0.035	0.032	0.043	0.038	0.042	0.028	0.022	0.027	0.031	0.025	0.031	0.038	0.036	0.035	0.048	0.053	0.055	0.030	0.031	0.024	0.022		**0.386**	**0.138**	**0.891**	**0.518**	0.037	**0.631**	**0.441**
Ellé94	0.025	0.025	0.029	0.028	0.036	0.033	0.034	0.022	0.016	0.021	0.023	0.021	0.022	0.031	0.033	0.028	0.048	0.050	0.050	0.031	0.032	0.026	0.022	**0.008**		**0.346**	**0.232**	**0.256**	0.024	**0.716**	**0.542**
Ellé00	0.029	0.031	0.032	0.030	0.046	0.041	0.041	0.026	0.020	0.027	0.027	0.022	0.026	0.035	0.032	0.030	0.042	0.045	0.043	0.029	0.031	0.025	0.021	**0.009**	**0.007**		**0.533**	**0.300**	0.004	**0.248**	**0.523**
Ellé05	0.028	0.029	0.032	0.028	0.044	0.038	0.041	0.025	0.020	0.025	0.027	0.022	0.026	0.034	0.031	0.030	0.044	0.049	0.050	0.028	0.030	0.023	0.021	**0.005**	**0.007**	**0.006**		**0.393**	0.007	**0.212**	**0.501**
Scorff88	0.038	0.040	0.044	0.040	0.051	0.047	0.047	0.034	0.029	0.036	0.042	0.033	0.036	0.047	0.042	0.042	0.059	0.063	0.064	0.035	0.038	0.032	0.028	**0.012**	**0.013**	**0.012**	**0.011**		**0.223**	**0.541**	**0.572**
Scorff94	0.030	0.030	0.034	0.031	0.038	0.035	0.037	0.028	0.023	0.028	0.030	0.024	0.029	0.034	0.032	0.031	0.056	0.061	0.059	0.031	0.033	0.027	0.024	**0.010**	**0.010**	**0.011**	**0.010**	**0.012**		**0.222**	**0.138**
Scorff00	0.032	0.033	0.037	0.036	0.046	0.041	0.044	0.032	0.024	0.028	0.032	0.027	0.032	0.039	0.037	0.036	0.052	0.057	0.055	0.034	0.036	0.028	0.024	**0.007**	**0.006**	**0.008**	**0.007**	0.011	0.008		**0.460**
Scorff05	0.026	0.027	0.031	0.027	0.038	0.034	0.035	0.024	0.021	0.023	0.027	0.021	0.025	0.035	0.031	0.029	0.046	0.049	0.048	0.025	0.028	0.020	0.018	**0.007**	**0.006**	**0.006**	**0.006**	0.010	0.008	0.007	

Chapter 6: General discussion

The aim of this thesis was to elucidate population structure of Atlantic salmon in the chalk streams of southern England. In order to achieve this, four studies were completed, each with separate aims. By considering the result in conjunction with published literature, we obtain a better understanding what salmon population structure and genetic diversity is like in these chalk streams, the forces that are responsible for their distinction and how these populations are likely to change.

What drives differentiation

This thesis identified conclusively that salmon from all the chalk streams in southern England are distinct and less diverse compared to nearby populations in non-chalk streams. From the results in Chapter 3 it appears that the differentiation has been enabled by the reduced migration into the chalk streams of salmon from the other regions. A recent study has investigated the relative effects of straying against adaptation and genetic drift (Bradbury *et al.* 2014) on Atlantic salmon population structure. The study (Bradbury *et al.* 2014) concluded that their population structure was influenced heavily by large amounts of genetic drift and low amounts of effective straying. The present findings support this conclusion.

The factors responsible for the reduced migration into the chalk streams cannot be determined from the results of this thesis, but they can be speculated. We propose that one of the most likely ecological factors is geology. Geology has been identified as a significant driving force in recent studies that have set out to determine drivers of genetic structure in Atlantic salmon (Perrier *et al.* 2013; Vincent *et al.* 2013). These studies have also identified other significant factors, such as temperature, coastal distance between river mouths, and river length for salmon populations in France (Perrier *et al.* 2013) and climate (including temperature and precipitation) for populations in North America (Vincent *et al.* 2013) The temperature of the water in the chalk streams is stabilized by the chalk sediment (Berrie 1998), so in this case at

least geology is still likely to be a larger driver. However, it is also likely to be some other aspect of the water chemistry that is affected by the geology, that prevents successful immigration into the chalk streams. A recent study by Bradbury *et al.* (2014) attempted to identify significant factors responsible for population structure in Atlantic salmon. Among the most important were watershed size, winter severity, and pH. Chalk streams share a number of characteristics that differentiate them from non-chalk streams, including watershed size and pH. Chalk streams are typically smaller than non-chalk rivers and they are also thought to be more alkaline (Berrie 1992), with estimates of pH 8. More investigation is necessary, but there are many reasons to expect pH to be a significant factor. This is because pH affects salmon survival, for example, acidification of rivers caused by acid rain, has notoriously caused a long term decline in salmon productivity in Norway (Gibson 1993; Sandøy & Langåker 2001) and Nova Scotia, Canada (Gibson 1993). There are naturally fewer studies of the effect of alkalinity on salmon, but within chinook salmon (*Oncorhynchus tshawytscha*), increased pH has been shown to reduce thermal tolerance (McCullough 1999).

Moving further into speculation, it is possible that pH is important because of gill ATPase activity. This membrane-bound ion pump is an important component of fish coping mechanisms in response to environmental change (Dalziel *et al.* 2014) and is known to play a crucial role during salmon acclimatisation to seawater and freshwater (Prunet *et al.* 1989; Bystriansky & Schulte 2011). Fish ATPase activity has been shown to be affected by pH (Nieminen *et al.* 1982), and crucially it is also thought to be an evolutionary hotspot. This enzyme has benefitted from the whole genome duplication event of Salmoniformes (Dalziel *et al.* 2014), and Atlantic salmon appear to have several ATPase paralogs (Dalziel *et al.* 2014). Thus it might be beneficial to investigate if there are any differences in the ATPase genes, paralogs or expression of ATPase genes between chalk and non-chalk salmon. A combination of tissue analysis with an RNA-seq approach, similar to that used in a recent study of metal tolerance in trout (Webster *et al.* 2013), would be one way to investigate this.

However there may be a much simpler explanation. The differentiation between these chalk and non-chalk salmon may not be physiological, it may

simply be that the "smell" of chalk streams is so different from non-chalk streams that salmon from non-chalk regions are much less inclined to stray towards it. This aversion to the chalk streams may have allowed genetic drift to become the dominant force, and the local adaptation described previously may not be a factor. Although an increasing number of studies are determining the ecological variables that are associated with population drift, it is still debatable whether subsequent local adaptation to these variables has driven population structure for both salmonids (Taylor 1991; Garcia de Leaniz *et al.* 2007; Vincent *et al.* 2013; Bradbury *et al.* 2014), and most other species (Allendorf *et al.* 2010; Manel *et al.* 2012). Alternatively, the widespread population structure may be the result of genetic drift.

It is important to know whether population differentiation is caused by local adaption or genetic drift for a number of reasons. For example, it would likely enhance the success rate of hatcheries if they knew which factors salmon needed to be adapted to. Also we could better predict how impending changes in climate will affect populations in the wild. Although the development of microsatellites and SNPs has allowed greater number of studies to detect correlations, reciprocal transplant experiments, where salmon from different habitats are swapped and their survival monitored, are still needed in order to determine local adaptation (Garcia de Leaniz *et al.* 2007). Despite calls for such studies since 1991 (Taylor 1991), there appear to be very few thus far. This should be the future focus of studies. In one recent study, Atlantic salmon, from one river (non-natives) were placed into a nearby river and also bred with the native population to form hybrids placed in the natives river (O'Toole *et al.* 2015). The overall lifetime success was found to be best in the natives, intermediate in the hybrids and lowest in the non-natives, which follows what might be the expected signature of local adaptation (O'Toole *et al.* 2015). However, in a previous reciprocal translocation study between three Atlantic salmon populations, the results indicated that the success of translocated individuals and their crossed could be highly variable, with only one sampling site showing signs of local adaptation.

The results of this thesis show that the chalk streams of southern England could be an ideal habitat to investigate local adaptation in the future.

Translocation experiments between one chalk stream and a nearby non-chalk stream, like the river Exe in Devon might be the one of the strongest indicators of local adaptation. Alternatively, in order to eliminate the risk of contamination, laboratory based translocations could be conducted with tanks containing water matching the chemistry of the Exe stream, and others matching the chemistry of the chalk streams. Monitoring the survival of fish at different life stages could be used as a measure of success. If possible, identifying the survival rate of hybrids in each water type would further help to support or refute the theory of local adaptation, following the guidelines given by Taylor (1991).

Another avenue that should be explored is whether this differentiation is unique to Atlantic salmon, or is it an ecological barrier for many species. A study by King and colleagues (unpublished) indicates that brown trout (*Salmo trutta*) in these chalk streams are also strongly differentiated from their non-chalk neighbours. It is possible that this extends to other species, outside of the Salmonidae family also. The relationship of salmon in the chalk streams in southern England with those in the chalk streams of Northern France, which includes the river Bresle, should also be investifated. From a scientific perspective, this could provide further support for the effect of geology on the differentiation of salmonids, as if it is a key factor, salmon from these two regions should be closer together than salmon from the other non-chalk regions. The French chalk salmon may also be crucial to identifying the phylogeographical history of these salmon, as it is possible that the English chalk salmon were colonised by the French chalk salmon either before or after the last glacial maximum. Investigation of the French salmon could also provide a conservation benefit, which are discussed in the following section.

Effects of low migration

While more studies would be needed to determine the environmental cause of low migration, multiple effects are evident on these populations. As well as the genetic differentiation described above, salmon from the chalk streams in southern England all have a reduced genetic diversity compared to their non-chalk counterparts. This is particularly interesting when compared to results from the literature.

Atlantic salmon populations from Spain are a particularly noteworthy comparison. These populations are at the lower limit of the species' range (MacCrimmon & Gots 1979), and evidence from several sources indicates that population sizes here are very low (Consuegra *et al.* 2005; Ribeiro *et al.* 2008; ICES 2014). Consensus indicates that these populations rapidly declined during the latter half of the 20th Century, there are no coastal or estuary fisheries remaining, and rod catches is amongst the lowest in Europe (ICES 2014). However, genetic diversity does not appear to be lowered across the board (King *et al.* 2001b; Consuegra *et al.* 2005; Ribeiro *et al.* 2008) contrary to their predictions. Populations on the north facing side of Spain, in populations such as the Asón, Pas, and Eo show no sign of a reduced diversity compared to other populations in Europe, or relative to their own population samples from before the apparent crash (King *et al.* 2001b; Consuegra *et al.* 2005; Ribeiro *et al.* 2008), Although one study appears to have identified decreases in genetic diversity in the rivers Esva, Narcea, Sella and Cares between the years 1993 and 1999 (Borrell *et al.* 2007), a greater number have identified stability, and one study identifies increases in allelic richness in all four of the rivers of which they have measurements spanning 1988 to 2007 (Horreo *et al.* 2011b).

The higher than expected genetic diversity for these Spanish populations contrasts against the low diversity identified in the chalk streams in the present study. We propose that this is an effect of the low migration into the chalk streams. From the results of the present study, only populations in the Ulla and Miño have a lowered genetic diversity. This possibly supports the theory that low migration has allowed the lower genetic diversity, as these two rivers are at the extreme southern end of the range and presumably would be the lease likely to receive strays based on geographic distance, however the amount of immigration into the region does require further study. It is however worth noting that diversity in the Ulla could also be low because the current population stems from supplementation with stock derived from a small number of breeders (Saura *et al.* 2008). At the same time it is worth noting that stocking into all the Spanish populations with exogenous fish was significant and considered to play a significant role in an evident break up of population structure and the recent increases in allelic richness (Horreo *et al.* 2011a).

Another effect of low migration is that the chalk populations are more susceptible to catastrophic events. As discussed in Chapter 3, being on an equal latitude to the populations in the southwest of England, it is likely that the chalk and southwest populations shared the same history of global sea level rises and falls of glacial maxima events (Clark *et al.* 2009), however there is only a signature of an historic bottleneck event in the chalk streams. This may be a significant finding for all population genetic studies. At the time of their proposal, simulations showed that the bottleneck test had enough power to detect population declines given a reasonable number of individuals and loci (Cornuet & Luikart 1996; Garza & Williamson 2001). However a significant number of published studies using real data, show this not to be the case (see review by Peery *et al.* 2012); bottleneck events are often not detected using genetic techniques when they are known to have happened (Peery *et al.* 2012). Outside the salmonidae family, this includes the California sea otter (*Enhydra utris nereis;* Aguilar *et al.* 2008), the Amur tiger (*Panthera tigris altaica;* Henry *et al.* 2009) and the Scandinavian lynx (*Lynx lynx;* Spong & Hellborg 2002). The results, which show evidence of a genetic bottlenecks in the chalk populations and not the southwest England could be an example of a widespread phenomena i.e. the signature of a population decline is masked when there is migration between the studied population and othersAt the contemporary level there was also no evidence to support recent decreases in population size using the heterozygote excess method in any of the populations. As populations in all four of the regions studied (southwest England, southern England, France and Spain) have all seen significant declines in recent decades, it does call in to question whether these methods are at all effective for the species if not salmonids as a whole and other species. It has been argued that more identifying changes in heterozygosity may need more loci, and for a sample size of 30, approximately 15 loci are needed to give an 80% probability of detecting a change in heterozygosity caused by a bottleneck (Luikart & Cornuet 1998; Säisä *et al.* 2003), so it is possible that more loci were needed; conversely the number of individuals should have been adequate as it always exceeded 30, with between 34 to 49 individuals for the French and Spanish populations and between 70 and 158 for the southwest England and chalk populations.

Single sample summary statistic methods such as these (M ratio and heterozygote excess) are predicted to be replaced (Peery *et al.* 2012) by Bayesian methods that detect posterior distribution of past effective sizes, like DIYABC (Cornuet *et al.* 2008) or VarEff (Nikolic & Chevalet 2014) used in this thesis. Indeed, VarEff detected signs of several bottlenecks, including a contemporary one in most of the studied populations. This method of analysis does hold promise and should be used more routinely following more testing.

Using the historic scales, we were able to see slight changes in genetic diversity across a wide range of salmon populations. Although very few changes were statistically significant and individually the results might be dismissed, together they should be considered noteworthy. The significant increases in allelic richness in two of the French populations, despite the worsening state of population sizes, can only support the effect of migration. Similar results have been identified in French populations previously (Perrier *et al.* 2013), and also in Spain (Horreo *et al.* 2011a; b). The fact, that the one population showing the steepest decline in allelic richness is a chalk stream, albeit with caveats (i.e. results not significant and stem from adults and juveniles), should not be ignored. This supports the predictions of the effect of a reduced population in the absence of immigration, and suggests that populations on the Avon if not other chalk streams require urgent attention.

One controversial solution to the reduction in population size and genetic diversity on the Avon might be to use brood stock from exogenous fish from the chalk streams of northern France. This technique has mostly been eradicated, in favour of stocking with the broodstock of local fish (e.g. Horreo *et al.* 2011a) due to the findings of poor success (e.g. Finnegan & Stevens 2008; Griffiths *et al.* 2011), a growing appreciation for local population structure and a determination to maintain the integrity of local stocks (Frankel 1974; ICES 2014). As suggested by the occasional study, which show lasting effects of exogenous stocking in Atlantic salmon (e.g. Horreo *et al.* 2011a) success might be improved if populations from source and sink are screen for compatibility. This clearly requires the hypothesis of local adaption in salmonids to be investigated, but the tools are finally available for that testing. The risk of eroding the integrity of a local populations genetic signature may also be reduced by this screening process to insure that as close as possible a genetic

match is used as brood stock. Conversely the risk of erosion may also be outweighed by the risk of continued decline in genetic diversity. Recently a small population (less than 70 individuals) of Florida panthers (*Puma concolor coryi*) in southern Florida were supplemented with eight Texas puma (*P. c. stanleyana*) in order to recover the population size and the reduced genetic diversity (Land & Lacy 2000). The pedigree of the population was studied and showed improvement in phenotypes thought to be associated with the low genetic diversity, for example a kinked tail (Land & Lacy 2000; Hedrick 2001). A number of management recommendations have been proposed that could also be useful if this idea was attempted (Storfer 1999).

Fine scale population structure of salmon

Thus far we have discussed how the results from this thesis and the literature indicate clearly that on a large scale the geological substrate and some unknown effect on water chemistry is likely the greater determinant of differentiation in Atlantic salmon. However the results also suggest that within a largely uniform environment, namely the five chalk streams, there are still patterns of population structure. Between the rivers there is a significant pattern of isolation by distance (IBD).

IBD between rivers has been identified previously, for example in Spain (Campos *et al.* 2007), Norway (Glover *et al.* 2012) and France (Perrier *et al.* 2011), but there have also been conflicting cases where IBD has not been identified between rivers. For example no significant IBD was detected in another study of Spanish rivers (Ayllon *et al.* 2006), or in a number of studies of rivers in Canada (Palstra *et al.* 2007; Bradbury *et al.* 2014).

From the results of this thesis and the literature (Ayllon *et al.* 2006; Campos *et al.* 2007; Perrier *et al.* 2011; Glover *et al.* 2012) we hypothesize that where IBD has not been identified, it may be due to sampling salmon from different genetic groups. Previous studies, aiming to elucidate population structure between rivers, have chosen rivers ad hoc, and have been unaware of the broad genetic groups. If salmon were sampled from different genetic groups (likely differentiated by different geological substrates), then ecological factors are likely to have impeded the migration necessary for IBD. This is demonstrated most clearly in one particular published study. Tonteri *et al.*

(2009) studied salmon populations across a large longitudinal range, including populations from Norway and Scotland. The study identified a pattern of IBD on small scales, but IBD disappeared at large distances (>100 km). Following our hypothesis, this is likely to be because salmon have been obtained from different groups, and it is not necessarily the case that two geographically proximate groups will be genetically more similar than two groups further apart. Future studies attempting to look at population structure between rivers would benefit from identifying the broad genetic groups first, before choosing either to investigate population structure between rivers or between genetic groups. Alternatively, researchers could choose to sample from rivers within a given area whilst noting the geological substrate or other indicators, as recent studies have done (Perrier *et al.* 2011). Either path would be beneficial and could contribute towards a more consistent picture of population structure in the species.

Within the river Frome, there was evidence of population sub-division. This is, to our knowledge, the smallest river in which subpopulations have been identified. In the majority of previous studies sub-divisions have been identified in only large rivers, for example the river Teno in Norway (Vähä *et al.* 2007) and the river Foyle in Ireland (Ensing *et al.* 2011). As the studies themselves identify (Vähä *et al.* 2007; Ensing *et al.* 2011), these rivers are large enough to have tributaries with different environmental parameters to drive local adaptation. The study by Vaha *et al.* (2007) also indicates that salmon along the main stem of the river Teno are less divergent, which suggests that if the river did not have tributaries, there would likely be no evidence of sub-division. Thus, finding this sub-division on the Frome, which lacks tributaries, is exceptional. We propose that this may be another effect of the chalk streams not receiving migrants from other regions.

On the other hand, it is important to assess whether the finding represents a biologically significant divide or if the markers offer so much variability as to infer slight differences which are not biologically significant (Hedrick 2001). As Hedrick (2001) describes, in order to identify biologically meaningful differences, "we need to define some measure or effect related to the likelihood of the accumulation of significant biological differences." Unfortunately, we (the scientific community) have not yet done this. However

temporal stability of that populations structure would be probably be one of those effects of biological significance. As this has not been identified in the two years studied, we believe this to not be a biologically significant difference. With the development of SNPs, which provide hundreds of loci and much greater statistical power, it is prudent that the effects of biologically significant differences are well defined.

The effect of geology on straying and assignment

Although not crucial to the aim of this thesis, the assignment of Mersey fish served its purpose as a training exercise. By clustering the baseline rivers into broad "reporting regions", following the examples of previous studies (Beacham *et al.* 2001, 2006), assignment confidence was increased to a point where the results were considered reliable. The river Mersey was found to be on the border between two reporting regions. The majority of salmon found on the Mersey clearly originated from rivers north of this border and in particular, the Solway & Northwest England reporting region. Surprisingly, very few salmon originated from the Southwest England & Wales region. Initially we hypothesised that the reason for this bias was the prevailing clockwise gyre in the eastern Irish Sea and an associated current (Heaps & Jones 1977; Ikediashi *et al.* 2012). We postulated that this current carried homing adult salmon past their natal rivers and southwards towards the Mersey, and at the same time acted to move fish from the rivers of north Wales away from the Mersey. However, following the findings from the rest of the thesis, it is perhaps more likely that geology (or water chemistry) was a more important factor in these results. The river Mersey flows west through Triassic and Carboniferous rock (Figure 6.1). Many rivers in northwest England also flow through Triassic and Carboniferous rock, whilst conversely, most rivers in Wales, like the river Dee, flow through Silurian and Ordovician rock (Figure 6.1). Therefore, it is possible that salmon from the Solway & Northwest England region are more likely to stray into the Mersey than salmon from the Southwest England & Wales region, due to their more similar water chemistry. This alternate hypothesis, we argue, is more congruent with evidence from the literature. It has been identified that in many species of salmonids, males are more likely to stray than females (Hard & Heard 1999; Hamann & Kennedy 2012). A positive male bias has also been identified in these Atlantic salmon recolonising the river Mersey (Miller

unpublished). This positive male bias suggests that straying is an active process, with males seeking an increased opportunity to successfully mate provided by the possibility that a different river might have reduced competition and greater access to females (Hamann & Kennedy 2012). The North Sea gyre hypothesis is slightly incongruent with the male bias, because together they also require a reason for males to be passively swept in the current more than females. The Geology Hypothesis, fits better with the male bias, as males are actively seeking novel habitats, but naturally choosing one more similar to their own.

It is worth noting that in the majority of salmon assignment studies to date, no consideration has been given to water chemistry or geology (Jonsson *et al.* 2003; Palstra *et al.* 2007; Griffiths *et al.* 2010; Ikediashi *et al.* 2012). Instead, the distance salmon stray is often the focus; for example a study of straying in Norway concludes that 96% of strays from the river Imsa migrated within 420 km of the river Imsa and 80% entered rivers within 60 km (Jonsson *et al.* 2003). It may be useful, if not imperative, to add ecological variables such as geology to assignment methods. As discovered during our own analysis (Chapter 2), there is an inverse relationship between the number of rivers in the baseline and confidence in assignment results. With technological advances, greater numbers of salmon in greater numbers of rivers are being genotyped. But without a novel approach, having more rivers sampled might paradoxically make genetic assignment more difficult. Identifying ecological factors may be that novel approach.

Thesis Conclusion

The overall aims of this thesis were to elucidate population structure in the chalk streams of southern England. This has been achieved on several levels by the completion of four separate studies. First and foremost, these salmon are distinct from their non-chalk stream neighbours, and they are also less genetically diverse. Low immigration from the surrounding regions detected in this study is likely to be the prime cause for both of these findings. This helps to answer a question posed previously by Bradbury *et al.* (2014); this results supports ecological factors are the prime driving force of population structure in Atlantic salmon.

Within the chalk region, where ecological factors are likely to be more constant, population structure was identified between the five chalk streams in the form of both a pattern of isolation by distance and a differentiation into three groups. Thus, we suspect that past inconsistent results regarding the presence and absence of population structure between rivers may be explained by a lack of consideration to the individual habitats each river was in. Sampling across wide areas inevitably lead to sampling salmon from rivers that were ecologically different. We predict that if the broad groups of genetically similar salmon are identified before further study of between-river population structure, then a more consistent pattern of Atlantic salmon population structure will begin to emerge.

Population subdivision was also identified within the river Frome, which appears to be the smallest river in which subdivision has been identified. We speculate that this may be an effect of having low immigration from neighbouring regions. However, we question the biological significance of the detected split, not least because the split was not consistent between the two years sampled.

Salmon in the Avon, the one chalk stream where relevant temporal samples were available, indicate a current negative trajectory in allelic richness. We postulate that this may be a signal of recent population decline in the Avon, and possibly all of the southern English chalk streams. We also postulate that this signal may be evident because of the lack of immigration from non-chalk stream regions. Thus, salmon in the Avon, and possibly all chalk streams should be monitored closely.

In the future, studies of Atlantic salmon population structure would benefit from incorporating key ecological variable(s), such as those affected by geology. This will help to build a more accurate picture of Atlantic salmon populations across their entire range, and address some of the current contradictions present. The question of local adaptation in the species still need to be answered, and depends on the use of traditional translocation studies and SNPs studies. For the purpose of understanding the history of salmon in the southern English chalk streams, and possibly for their own future survival, we suggest that their relationship with salmon in the chalk streams of France be investigated promptly.
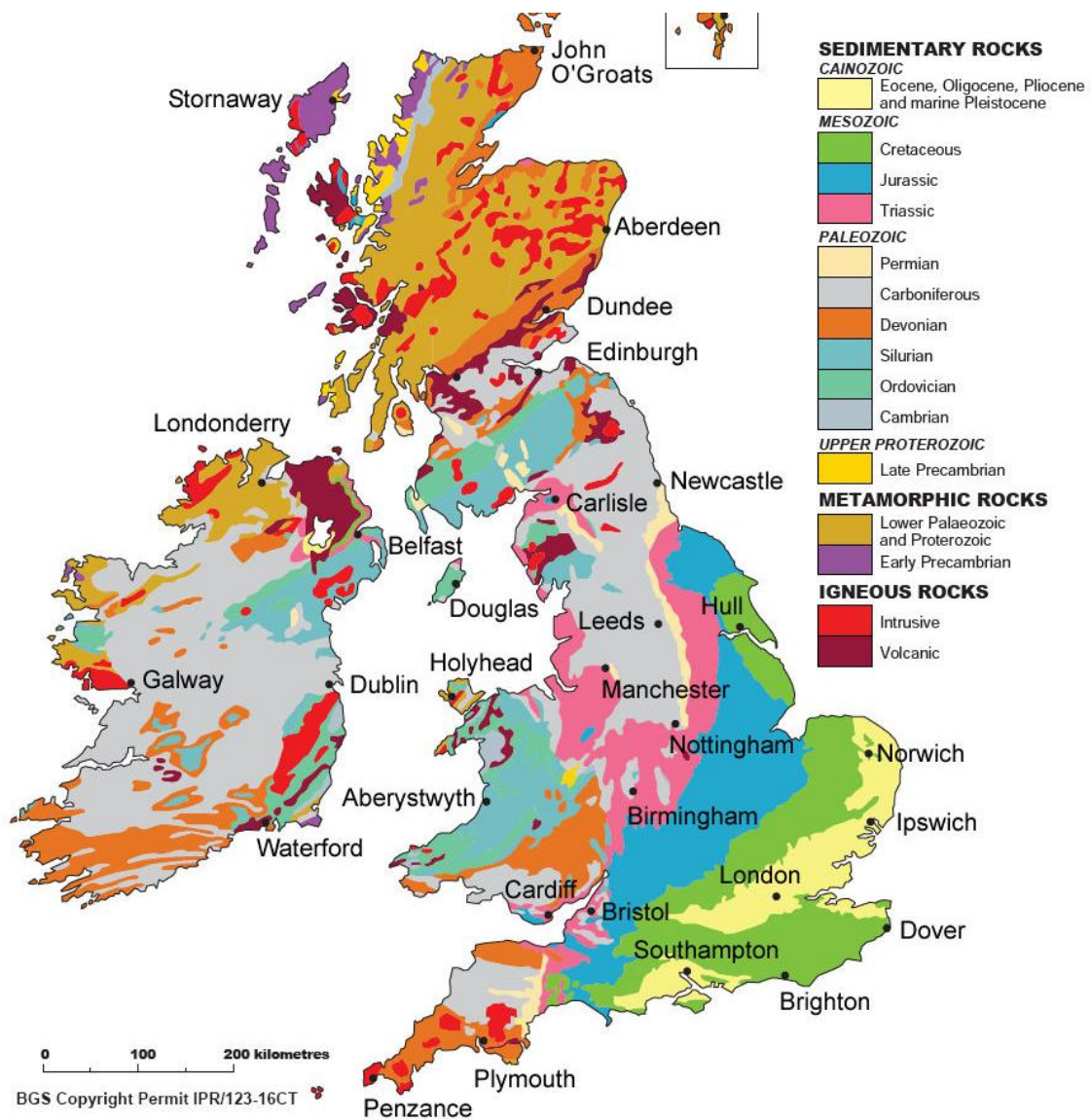
Figure 6.1- Geological map of the United Kingdom. Image created by the British Geological Survey.

Appendices

Appendix I- Relevant details of salmon caught on the Mersey and assignment results. Details of the Mersey samples included in this study are in the left wide column, the code assigned and the date of sampling. Details of the assignment results are found in the wide middle column (GeneClass 2) and the wide right column (ONCOR). Within each wide column, the result of assignment to reporting regions is found on the left, and the result of assignment to rivers is found on the right. Only the two most probable sources are shown and the relative probability that the individual belongs to it. S&NWE - Solway and northwest England, N.Ireland- Northern Ireland, SWE&W - Southwest England and Wales.

Sample ID	Date collected	Geneclass Assignment Reporting Region Probability 1st	2nd	River level 1st	2nd	ONCOR Assignment Reporting Region Probability 1st	2nd	River level 1st	2nd
Mer.adu 01.01	05/11/2001	Scotland 83.0	SWE&W 12.5	Conwy 77.2	Teifi 7.1	Scotland 0.8	S&NWE 0.1	Conwy 0.5	Nith 0.2
Mer.adu 01.02	05/11/2001	S&NWE 88.8	Scotland 9.7	Esk 47.5	Nith 14.8	S&NWE 1.0	Scotland 0.0	Nith 0.5	Esk 0.3
Mer.adu 01.03	05/11/2001	S&NWE 72.3	SWE&W 27.3	Annan 66.3	Nith 32.6	S&NWE 0.9	SWE&W 0.1	Nith 0.6	Annan 0.4
Mer.adu 01.04	01/10/2001	S&NWE 47.1	Scotland 31.8	Ehen 49.4	Luce 27.2	S&NWE 0.7	Scotland 0.2	Nith 0.4	Luce 0.3
Mer.adu 02.01	25/10/2002	Scotland 75.8	S&NWE 23.4	Luce 43.3	Nith 22.9	Scotland 0.5	S&NWE 0.5	Nith 0.6	Luce 0.3
Mer.adu 02.02	29/10/2002	Scotland 35.1	S&NWE 33.3	Ayr 21.2	Girvan 18.3	S&NWE 0.7	Scotland 0.3	Nith 0.7	Annan 0.2
Mer.adu 02.03	29/10/2002	S&NWE 51.0	Scotland 37.9	Girvan 49.7	Nith 12.9	S&NWE 0.8	Scotland 0.2	Nith 0.6	Annan 0.1
Mer.adu 02.04	30/10/2002	Scotland 60.0	SWE&W 28.6	Tamar 50.3	Nevern 31.6	Scotland 0.6	S&NWE 0.2	Tamar 0.6	Luce 0.2
Mer.adu 02.05	30/10/2002								
Mer.adu 02.06	31/10/2002	S&NWE 51.7	Scotland 47.2	Esk 58.1	Stinchar 29.0	S&NWE 0.8	Scotland 0.2	Esk 0.5	Stinchar 0.4
Mer.adu 02.07	31/10/2002	S&NWE 97.2	Scotland 2.2	Wye 99.7	Annan 0.1	S&NWE 1.0		Wye 1.0	Nith 0.0
Mer.adu 02.08	31/10/2002	S&NWE 87.9	Scotland 9.8	Annan 68.5	Nith 16.4	S&NWE 1.0	Scotland 0.0	Annan 0.6	Nith 0.4
Mer.adu 02.09	01/11/2002	Scotland 69.3	S&NWE 26.5	Clyde 28.0	Conwy 20.9	S&NWE 0.6	Scotland 0.4	Clyde 0.3	Luce 0.2
Mer.adu 02.10	03/11/2002	S&NWE 54.1	SWE&W 39.7	Dee 65.9	Nith 10.8	S&NWE 0.8	SWE&W 0.2	Nith 0.5	Dee 0.4
Mer.adu 02.11	03/11/2002	S&NWE 60.6	SWE&W 38.6	Dee 35.0	Esk 33.9	S&NWE 0.8	SWE&W 0.2	Esk 0.4	Dee 0.3
Mer.adu 02.12	03/11/2002	Scotland 48.1	SWE&W 35.1	Clyde 41.2	Exe 25.6	Scotland 0.4	S&NWE 0.4	Clyde 0.5	Luce 0.1
Mer.adu 02.13	04/11/2002	Scotland 52.7	SWE&W 40.0	Luce 86.2	Stinchar 1.6	Scotland 0.5	SWE&W 0.3	Luce 0.9	Nith 0.1
Mer.adu 02.14	05/11/2002	S&NWE 46.0	SWE&W 30.3	Dee 63.3	Nith 21.0	S&NWE 0.7	Scotland 0.2	Nith 0.6	Dee 0.3
Mer.adu 02.15	05/11/2002	S&NWE 88.9	Norway 7.0	Esk 76.9	Ribble 9.3	S&NWE 1.0	Scotland 0.0	Esk 0.7	Nith 0.2
Mer.adu 02.16	05/11/2002	S&NWE 82.8	Scotland 11.3	Tamar 43.1	Annan 42.4	S&NWE 0.9	Scotland 0.0	Annan 0.6	Tamar 0.2
Mer.adu 02.17	05/11/2002	SWE&W 76.9	S&NWE 10.9	Doon 36.5	Teign 29.5	SWE&W 0.7	S&NWE 0.3	Dee 0.6	Teifi 0.1
Mer.adu 02.18	06/11/2002	S&NWE 85.1	Scotland 14.5	Nith 44.3	Annan 26.9	S&NWE 0.9	Scotland 0.1	Nith 0.8	Annan 0.2
Mer.adu 02.19	06/11/2002	Scotland 44.2	N.Ireland 43.4	Clogh 84.7	Stinchar 5.9	Scotland 0.7	S&NWE 0.2	Stinchar 0.6	Clyde 0.1
Mer.adu 02.20	06/11/2002	S&NWE 64.2	Scotland 30.2	Nith 22.8	Stinchar 21.7	S&NWE 0.9	Scotland 0.1	Nith 0.6	Stinchar 0.2

ID	Date								
Mer.adu 02.21	31/11/2002	SWE&W 66.6	S&N WE 32.6	Dee 92.1	Duddon 5.9	S&N WE 0.6	SWE&W 0.4	Dee 0.9	Nith 0.1
Mer.adu 04.01	21/01/2004	SWE&W 40.3	S&N WE 32.8	Nith 62.5	Stinchar 10.3	S&N WE 0.6	SWE&W 0.2	Nith 0.9	Stinchar 0.1
Mer.adu 06.01	31/10/2006	N.Ireland 93.4	S&N WE 3.2	Leguer 24.6	Six_Mile 18.0	N.Ireland 0.5	S&N WE 0.4	Stinchar 0.3	Nith 0.3
Mer.adu 06.02	08/11/2006	SWE&W 79.1	Scotland 14.9	Exe 89.7	Taw 5.7	SWE&W 0.7	Scotland 0.1	Exe 0.9	Nith 0.0
Mer.adu 06.03	09/11/2006								
Mer.adu 06.04	09/11/2006	S&N WE 78.5	Scotland 17.9	Esk 94.1	Luce 2.9	S&N WE 0.9	Scotland 0.1	Esk 0.9	Luce 0.0
Mer.adu 06.05	14/11/2006	SWE&W 74.8	Scotland 9.4	Torridge 30.5	Tawe 28.2	SWE&W 0.7	S&N WE 0.2	Luce 0.5	Tamar 0.3
Mer.adu 06.06	14/11/2006	Scotland 74.7	S&N WE 24.0	Esk 52.3	Luce 23.6	Scotland 0.5	S&N WE 0.4	Esk 0.4	Luce 0.2
Mer.adu 06.07	21/11/2006	Scotland 47.1	SWE&W 41.5	Wye 48.8	Dee 11.4	Scotland 0.5	SWE&W 0.3	Nith 0.2	Clyde 0.2
Mer.adu 06.08	22/11/2006	S&N WE 92.9	Scotland 4.3	Annan 52.4	Nith 37.6	S&N WE 1.0	Scotland 0.0	Nith 0.7	Annan 0.3
Mer.adu 07.01	13/02/2007	SWE&W 96.9	S&N WE 3.0	Exe 81.0	Tawe 15.4	SWE&W 0.9	S&N WE 0.1	Exe 0.8	Nith 0.2
Mer.adu 07.02	12/10/2007	France 99.0	SWE&W 1.0	Sélune 96.7	Sée 3.3	France 0.9	SWE&W 0.1	Sélune 1.0	
Mer.adu 07.03	16/10/2007	Southern England 55.6	S&N WE 16.7	Test 50.7	Ehen 36.8	S&N WE 0.8	Scotland 0.2	Ehen 0.6	Nith 0.2
Mer.adu 07.04	16/10/2007	S&N WE 99.9	N.Ireland 0.1	Annan 72.9	Ribble 11.8	S&N WE 1.0		Annan 0.7	Nith 0.3
Mer.adu 07.05	17/10/2007	France 88.8	SWE&W 7.5	Camel 36.8	Laukhel 29.9	SWE&W 0.4	France 0.4	Camel 0.6	S?lune 0.2
Mer.adu 07.06	17/10/2007	France 94.5	SWE&W 5.5	Fowey 80.6	Elle 7.2	France 0.6	SWE&W 0.4	Fowey 0.9	Exe 0.1
Mer.adu 07.07	27/10/2007	SWE&W 75.2	Scotland 20.9	Dee 77.7	Luce 13.3	SWE&W 0.6	Scotland 0.4	Dee 0.7	Luce 0.2
Mer.adu 07.08	30/10/2007								
Mer.adu 07.09	30/10/2007	Norway 65.8	SWE&W 25.3	Vesterelva 41.3	Taw 18.5	SWE&W 0.6	S&N WE 0.3	Nith 0.2	Stinchar 0.2
Mer.adu 07.10	01/11/2007	S&N WE 92.9	Scotland 4.8	Esk 56.0	Nith 20.5	S&N WE 1.0	Scotland 0.0	Nith 0.6	Esk 0.3
Mer.adu 07.11	01/11/2007	S&N WE 60.8	N.Ireland 27.8	Annan 55.4	Black 22.7	S&N WE 0.9	SWE&W 0.0	Annan 0.6	Nith 0.4
Mer.adu 07.12	13/11/2007	Scotland 94.4	S&N WE 2.5	Clyde 86.2	Esk 9.9	Scotland 0.9	S&N WE 0.1	Clyde 0.9	Esk 0.1
Mer.adu 07.13	14/11/2007	N.Ireland 99.9	S&N WE 0.1	Grillagh 50.6	Agivey 45.6	N.Ireland 1.0	S&N WE 0.0	Grillagh 0.9	Nith 0.1
Mer.adu 07.14	14/11/2007	Scotland 54.6	N.Ireland 27.8	Grillagh 39.2	Luce 17.6	Scotland 0.6	S&N WE 0.4	Nith 0.4	Luce 0.2
Mer.adu 07.15	20/11/2007	S&N WE 98.1	SWE&W 1.5	Nith 51.9	Annan 20.2	S&N WE 1.0		Nith 0.9	Annan 0.1
Mer.adu 07.16	21/11/2007	S&N WE 73.5	SWE&W 25.2	Lune 74.7	Teign 19.7	S&N WE 0.9	SWE&W 0.1	Lune 0.7	Nith 0.2
Mer.adu 07.17	21/11/2007	Scotland 43.9	S&N WE 22.2	Kent 39.1	Dee 17.0	S&N WE 0.5	Scotland 0.4	Dee 0.3	Clyde 0.2
Mer.adu 07.18	21/11/2007	S&N WE 75.2	SWE&W 17.3	Nevern 39.0	Nith 34.2	S&N WE 0.9	SWE&W 0.1	Nith 0.9	Esk 0.0
Mer.adu 07.19	23/11/2007	SWE&W 61.6	S&N WE 36.9	Girvan 38.1	Garnock 20.5	S&N WE 0.7	SWE&W 0.3	Garnock 0.3	Urr 0.2
Mer.adu 07.20	27/11/2007	S&N WE 91.7	Scotland 4.3	Nith 79.4	Dee 5.8	S&N WE 1.0	Scotland 0.0	Nith 1.0	Annan 0.0
Mer.adu 07.21	28/11/2007	Scotland 83.6	S&N WE 8.0	Ehen 54.6	Clyde 18.3	Scotland 0.8	S&N WE 0.2	Luce 0.3	Ehen 0.3
Mer.adu 07.22	28/11/2007	S&N WE 96.3	Scotland 3.0	Lune 60.5	Esk 29.2	S&N WE 1.0	Scotland 0.0	Esk 0.5	Nith 0.3
Mer.adu 07.23	28/11/2007	SWE&W 73.4	France 13.0	Conwy 19.4	Taw 17.3	SWE&W 0.6	S&N WE 0.3	Nith 0.2	Conwy 0.2
Mer.adu 07.24	28/11/2007	Scotland 43.8	S&N WE 41.8	Bladnoch 57.6	Cree 13.0	S&N WE 0.7	Scotland 0.2	Esk 0.6	Luce 0.1

ID	Date	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8
Mer.adu 07.25	30/11/2007	S&NWE 37.8	SWE&W 34.8	Leguer 90.1	Clyde 2.7	S&NWE 0.7	SWE&W 0.2	Nith 0.7	Clyde 0.1
Mer.adu 07.26	30/11/2007	S&NWE 67.9	Scotland 29.6	Eden 95.5	Esk 1.4	S&NWE 0.8	Scotland 0.1	Eden 0.6	Nith 0.2
Mer.adu 07.27	30/11/2007	S&NWE 68.5	SWE&W 14.8	Annan 38.4	Cree 28.3	S&NWE 0.9	SWE&W 0.1	Nith 0.5	Annan 0.4
Mer.adu 07.28	30/11/2007	N.Ireland 100.0	Scotland 0.0	Blac 40.5	Grillagh 18.5	N.Ireland 1.0		Grillagh 0.9	Garnock 0.1
Mer.adu 07.29	06/12/2007	S&NWE 70.1	Scotland 29.8	Ribble 43.9	Annan 24.2	S&NWE 0.9	Scotland 0.1	Annan 0.4	Nith 0.3
Mer.adu 07.30	07/12/2007	SWE&W 76.0	Scotland 13.6	Conwy 90.4	Luce 3.6	SWE&W 0.6	S&NWE 0.3	Conwy 0.8	Luce 0.1
Mer.adu 07.31	11/12/2007	S&NWE 65.1	Scotland 28.7	Ehen 34.7	Eden 29.1	S&NWE 0.8	Scotland 0.2	Nith 0.4	Ehen 0.2
Mer.adu 07.32	11/12/2007								
Mer.adu 07.33	21/12/2007	SWE&W 89.5	Scotland 7.9	Luce 46.6	Dart 20.1	SWE&W 0.8	Scotland 0.1	Luce 0.8	Stinchar 0.1
Mer.adu 08.01	03/10/2008								
Mer.adu 08.02	03/10/2008	Scotland 53.5	SWE&W 36.4	Stinchar 70.3	Teifi 8.0	Scotland 0.7	SWE&W 0.2	Stinchar 0.8	Nith 0.1
Mer.adu 08.03	07/10/2008	S&NWE 69.2	Scotland 24.6	Kent 20.4	Nith 19.5	S&NWE 0.9	Scotland 0.1	Nith 0.7	Annan 0.2
Mer.adu 08.04	14/10/2008	Scotland 86.1	Norway 10.7	Luce 58.8	Cree 27.7	Scotland 0.9	S&NWE 0.1	Luce 0.8	Nith 0.1
Mer.adu 08.05	15/10/2008	S&NWE 82.2	Scotland 12.6	Bladnoch 50.2	Esk 3	S&NWE 0.9	Scotland 0.1	Annan 0.6	Esk 0.4
Mer.adu 08.06	15/10/2008	Scotland 80.8	S&NWE 10.3	Ehen 62.0	Luce 17.7	Scotland 0.8	S&NWE 0.2	Ehen 0.4	Luce 0.4
Mer.adu 08.07	15/10/2008	S&NWE 94.9	Scotland 3.8	Bladnoch 79.7	Annan 11.0	S&NWE 1.0	Scotland 0.0	Nith 0.7	Annan 0.3
Mer.adu 08.08	16/10/2008	S&NWE 98.7	SWE&W 1.0	Nith 55.1	Conwy 20.0	S&NWE 1.0		Nith 0.9	Annan 0.1
Mer.adu 08.09	17/10/2008	N.Ireland 79.6	S&NWE 19.0	Nith 62.3	Kent 18.3	S&NWE 0.8	N.Ireland 0.1	Nith 1.0	Annan 0.0
Mer.adu 08.10	17/10/2008	S&NWE 92.4	Scotland 6.0	Annan 57.6	Clyde 11.4	S&NWE 1.0	Scotland 0.0	Annan 0.6	Nith 0.2
Mer.adu 08.11	17/10/2008	S&NWE 99.8	Scotland 0.1	Esk 49.5	Nith 33.3	S&NWE 1.0		Nith 0.7	Esk 0.2
Mer.adu 08.12	17/10/2008	Norway 85.1	S&NWE 14.8	Daleelva 99.3	Namsen 18.3	S&NWE 1.0		Daleelva 1.0	Nith 0.0
Mer.adu 08.13	17/10/2008	S&NWE 83.1	Scotland 12.4	Ribble 87.0	Eden 10.1	S&NWE 0.9	Scotland 0.0	Ribble 0.9	Eden 0.0
Mer.adu 08.14	20/10/2008	SWE&W 77.8	S&NWE 14.1	Tamar 63.8	Conwy 15.2	SWE&W 0.6	S&NWE 0.4	Nith 0.4	Tamar 0.3
Mer.adu 08.15	20/10/2008	SWE&W 94.3	Scotland 3.3	Girvan 27.3	Daleelva 26.9	SWE&W 0.9	Scotland 0.0	Stinchar 0.3	Daleelva 0.2
Mer.adu 08.16	20/10/2008	S&NWE 83.7	Scotland 10.4	Nith 28.5	Daleelva 18.5	S&NWE 1.0	Scotland 0.0	Nith 0.8	Annan 0.1
Mer.adu 08.17	20/10/2008	Scotland 87.4	S&NWE 7.9	Clyde 59.0	Garnock 15.1	Scotland 0.7	S&NWE 0.2	Clyde 0.6	Nith 0.2
Mer.adu 08.18	22/10/2008	S&NWE 88.3	SWE&W 11.7	Nith 54.5	Conwy 23.2	S&NWE 1.0	SWE&W 0.0	Nith 0.9	Annan 0.1
Mer.adu 08.19	22/10/2008	S&NWE 51.6	Norway 46.3	Ribble 100.0	Urr 0.0	S&NWE 1.0	Scotland 0.0	Ribble 1.0	
Mer.adu 08.20	23/10/2008	Scotland 81.1	SWE&W 17.5	Urr 61.2	Bladnoch 20.9	Scotland 0.8	SWE&W 0.2	Urr 0.6	Dee 0.1
Mer.adu 08.21	23/10/2008	Scotland 76.8	SWE&W 13.9	Luce 30.2	Stinchar 27.8	Scotland 0.8	SWE&W 0.1	Stinchar 0.4	Luce 0.3
Mer.adu 08.22	23/10/2008	Scotland 32.0	SWE&W 31.7	Stinchar 27.0	Annan 23.1	S&NWE 0.5	Scotland 0.3	Stinchar 0.4	Annan 0.3
Mer.adu 08.23	23/10/2008	S&NWE 68.5	N.Ireland 29.9	Ribble 34.3	Grillagh 22.5	S&NWE 1.0	N.Ireland 0.0	Nith 0.6	Ribble 0.2
Mer.adu 08.24	23/10/2008	S&NWE 99.1	SWE&W 0.8	Nith 73.6	Urr 14.6	S&NWE 1.0		Nith 0.9	Annan 0.0
Mer.adu 08.25	05/11/2008	S&NWE 46.9	N.Ireland 32.3	Conwy 66.4	Nith 9.5	S&NWE 0.9	Scotland 0.1	Nith 0.5	Conwy 0.3
Mer.adu 08.26	07/11/2008	N.Ireland 56.2	Norway 31.4	Moyola 35.1	Grillagh 31.9	S&NWE 0.7	N.Ireland 0.2	Nith 0.5	Grillagh 0.4

ID	Date																
Mer.adu 08.27	12/11/2008	SWE&W	62.5	Norway	36.0	Teifi	39.3	Exe	23.8	SWE&W	0.9	S&N WE	0.1	Teifi	0.4	Exe	0.2
Mer.adu 08.28	12/11/2008	Scotland	47.5	S&N WE	33.1	Urr	45.7	Cree	31.4	S&N WE	0.6	Scotland	0.3	Urr	0.3	Luce	0.3
Mer.adu 08.29	12/11/2008	Scotland	42.4	S&N WE	26.7	Luce	42.7	Esk	18.4	S&N WE	0.5	Scotland	0.3	Luce	0.5	Esk	0.2
Mer.adu 08.30	12/11/2008	SWE&W	97.8	Scotland	1.4	Exe	29.7	Teifi	20.8	SWE&W	1.0	Scotland	0.0	Exe	0.2	Teifi	0.2
Mer.adu 08.31	12/11/2008																
Mer.adu 08.32	12/11/2008	S&N WE	63.4	Scotland	34.4	Eden	40.2	Esk	21.8	S&N WE	0.9	Scotland	0.1	Nith	0.5	Esk	0.2
Mer.adu 08.33	12/11/2008	Scotland	60.2	S&N WE	32.8	Conwy	69.4	Stinchar	13.3	S&N WE	0.5	Scotland	0.5	Conwy	0.4	Stinchar	0.3
Mer.adu 08.34	13/11/2008	Scotland	86.8	S&N WE	11.7	Garnock	63.9	Ayr	16.0	Scotland	0.7	S&N WE	0.3	Garnock	0.5	Esk	0.2
Mer.adu 08.35	13/11/2008	Norway	87.5	SWE&W	4.5	Namsen	71.6	Dee	13.5	S&N WE	0.6	Scotland	0.2	Dee	0.5	Nith	0.2
Mer.adu 08.36	14/11/2008	S&N WE	90.8	N.Ireland	4.1	Cree	73.9	Nith	13.6	S&N WE	1.0	Scotland	0.0	Nith	0.9	Annan	0.1
Mer.adu 08.37	14/11/2008	Norway	79.4	Scotland	9.3	Urr	81.4	Leguer	5.9	S&N WE	0.6	Scotland	0.3	Urr	0.9	Tamar	0.0
Mer.adu 08.38	14/11/2008	SWE&W	97.0	S&N WE	1.9	Tamar	97.8	Leguer	1.6	SWE&W	0.9	S&N WE	0.1	Tamar	1.0	Nith	0.0
Mer.adu 08.39	14/11/2008	S&N WE	98.7	SWE&W	1.1	Annan	54.7	Nith	34.0	S&N WE	1.0			Nith	0.6	Annan	0.3
Mer.adu 08.40	18/11/2008	S&N WE	90.2	Scotland	7.9	Nith	19.3	Luce	19.1	S&N WE	1.0	Scotland	0.0	Nith	0.6	Luce	0.2
Mer.adu 08.41	18/11/2008	S&N WE	78.9	Scotland	15.3	Stinchar	73.4	Nith	9.7	S&N WE	0.9	Scotland	0.1	Stinchar	0.7	Nith	0.3
Mer.adu 08.42	30/11/2008	SWE&W	65.5	S&N WE	27.9	Urr	37.0	Dee	14.9	S&N WE	0.6	SWE&W	0.4	Nith	0.5	Urr	0.1
Mer.adu 08.43	30/11/2008	SWE&W	54.7	Norway	26.2	Camel	33.2	Aulne	18.5	SWE&W	0.7	Scotland	0.2	Camel	0.3	Clyde	0.3
Mer.adu 09.01	14/01/2009																
Mer.adu 09.02	05/10/2009	SWE&W	39.4	S&N WE	34.0	Stinchar	52.4	Ehen	19.5	S&N WE	0.6	SWE&W	0.2	Stinchar	0.7	Nith	0.1
Mer.adu 09.03	02/11/2009	S&N WE	82.6	SWE&W	16.9	Teifi	78.1	Neven	8.5	S&N WE	0.9	SWE&W	0.1	Teifi	0.8	Annan	0.1
Mer.adu 09.04	03/11/2009	SWE&W	73.0	S&N WE	18.1	Kent	29.9	Esk	21.2	SWE&W	0.5	S&N WE	0.4	Esk	0.5	Nith	0.2
Mer.adu 09.05	05/11/2009	S&N WE	98.7	Scotland	1.2	Eden	57.4	Annan	20.5	S&N WE	1.0			Annan	0.5	Nith	0.4
Mer.adu 10.01	01/09/2010	SWE&W	97.4	Scotland	1.7	Teifi	98.7	Taw	0.3	SWE&W	1.0	Scotland	0.0	Teifi	1.0	Stinchar	0.0
Mer.adu 10.02	07/09/2010	S&N WE	88.6	Scotland	6.4	Ehen	33.7	Teifi	31.2	S&N WE	1.0	Scotland	0.0	Nith	0.7	Teifi	0.1
Mer.adu 10.03	07/09/2010																
Mer.adu 10.04	08/09/2010	Scotland	70.9	S&N WE	24.9	Stinchar	51.9	Bladnoch	11.2	Scotland	0.5	S&N WE	0.5	Stinchar	0.6	Nith	0.3
Mer.adu 10.05	08/09/2010	S&N WE	69.4	Scotland	17.4	Ehen	85.4	Ayr	3.6	S&N WE	0.9	Scotland	0.1	Ehen	0.7	Stinchar	0.1
Mer.adu 10.06	09/09/2010	SWE&W	67.9	Scotland	27.7	Garnock	77.3	Taw	12.5	SWE&W	0.6	Scotland	0.3	Garnock	0.6	Stinchar	0.2
Mer.adu 10.07	09/09/2010	Scotland	87.4	S&N WE	6.0	Girvan	34.8	Luce	16.8	Scotland	0.8	S&N WE	0.2	Luce	0.3	Clyde	0.2
Mer.adu 10.08	09/09/2010	S&N WE	69.2	Scotland	27.9	Annan	39.6	Luce	26.7	S&N WE	0.9	Scotland	0.1	Annan	0.4	Luce	0.2
Mer.adu 10.09	09/09/2010																
Mer.adu 10.10	09/09/2010																
Mer.adu 10.11	09/09/2010	Norway	57.2	Scotland	21.7	Stinchar	30.8	Nith	20.0	S&N WE	0.6	Scotland	0.4	Nith	0.6	Stinchar	0.3
Mer.adu 10.12	11/09/2010	S&N WE	39.8	Scotland	33.2	Doon	47.9	Stinchar	25.8	S&N WE	0.7	Scotland	0.2	Stinchar	0.8	Conwy	0.1
Mer.adu 10.13	11/09/2010	Scotland	87.8	S&N WE	8.3	Stinchar	61.3	Ayr	18.4	Scotland	0.8	S&N WE	0.2	Stinchar	0.9	Nith	0.1

ID	Date																
Mer.adu 10.14	12/09/2010	Scotland	98.4	SWE&W	1.4	Stinchar	71.9	Luce	22.1	Scotland	1.0	SWE&W	0.0	Stinchar	0.8	Luce	0.2
Mer.adu 10.15	14/09/2010	S&NWE	79.5	SWE&W	10.4	Nith	48.9	Annan	19.7	S&NWE	0.9	SWE&W	0.0	Nith	0.9	Annan	0.1
Mer.adu 10.16	16/09/2010	S&NWE	45.1	Scotland	39.4	Clyde	59.1	Annan	33.1	S&NWE	0.7	Scotland	0.2	Clyde	0.5	Annan	0.5
Mer.adu 10.17	17/09/2010	S&NWE	90.9	Scotland	8.9	Annan	44.7	Nith	33.2	S&NWE	1.0	Scotland	0.0	Nith	0.6	Annan	0.3
Mer.adu 10.18	21/09/2010	SWE&W	90.7	S&NWE	8.3	Taw	79.5	Teifi	14.5	SWE&W	0.8	S&NWE	0.2	Teifi	0.4	Taw	0.4
Mer.adu 10.19	21/09/2010	Scotland	52.1	N.Ireland	24.4	Clyde	65.2	Ehen	12.7	Scotland	0.7	SWE&W	0.2	Clyde	0.9	Ehen	0.1
Mer.adu 10.20	21/09/2010	Scotland	62.3	N.Ireland	31.2	Girvan	53.2	Garnock	18.8	Scotland	0.8	S&NWE	0.1	Garnock	0.2	Girvan	0.2
Mer.adu 10.21	22/09/2010	SWE&W	87.8	Scotland	11.9	Stinchar	60.4	Conwy	33.2	SWE&W	0.8	Scotland	0.2	Stinchar	0.9	Conwy	0.1
Mer.adu 10.22	23/09/2010	SWE&W	50.9	S&NWE	48.0	Usk	28.0	Kent	27.9	S&NWE	0.8	SWE&W	0.2	Nith	0.4	Annan	0.3
Mer.adu 10.23	23/09/2010	S&NWE	75.6	SWE&W	14.8	Kent	26.7	Moyola	20.3	S&NWE	0.9	SWE&W	0.1	Nith	0.7	Annan	0.2
Mer.adu 10.24	23/09/2010	S&NWE	53.3	SWE&W	32.4	Kent	69.2	Ribble	17.0	S&NWE	0.8	SWE&W	0.1	Ribble	0.4	Kent	0.2
Mer.adu 10.25	23/09/2010	S&NWE	55.4	SWE&W	25.7	Urr	27.1	Ribble	23.8	S&NWE	0.8	SWE&W	0.1	Nith	0.6	Ribble	0.1
Mer.adu 10.26	23/09/2010	Scotland	68.5	SWE&W	16.2	Stinchar	46.1	Luce	14.4	Scotland	0.7	S&NWE	0.3	Stinchar	0.6	Nith	0.2
Mer.adu 10.27	24/09/2010	Scotland	81.1	S&NWE	10.4	Fowey	69.4	Ehen	9.8	Scotland	0.7	S&NWE	0.3	Fowey	0.4	Esk	0.2
Mer.adu 10.28	29/09/2010	Scotland	34.7	SWE&W	30.8	Clyde	39.1	Stinchar	19.5	Scotland	0.4	SWE&W	0.3	Clyde	0.4	Stinchar	0.4
Mer.adu 10.29	05/10/2010																
Mer.adu 10.30	06/10/2010	Scotland	73.8	S&NWE	22.4	Ayr	43.3	Nith	20.8	Scotland	0.5	S&NWE	0.5	Nith	0.8	Clyde	0.1
Mer.adu 11.01	06/01/2011																
Juveniles																	
Mer.juv03.01	14/10/2003																
Mer.juv06.01	15/08/2006																
Mer.juv06.02	15/08/2006	S&NWE	64.3	SWE&W	29.3	Eden	43.1	Nith	23.8	S&NWE	0.9	SWE&W	0.1	Nith	0.9	Annan	0.1

Appendix II- Details of sample sites and rivers used for the salmon baseline. Sampling sites are grouped into the rivers from which they were collected, and then into the reporting regions used in this study. Geographic coordinates of sampling sites, the date of collection and the number of individuals sampled are on the right.

| Report Region | River | River mouth | | Sampling year | Sampling site | Sampling site | | Life stage | Sample size (n) |
		Longitude	Latitude			Longitutde	Latitutude		
Scotland	Ayr	-4.68	55.5	2003	Guelt Water	-4.13	55.44	Juv	20
				2003	Glenmuir Water	-4.15	55.45	Juv	20
				2002	Greenock water	-4.08	55.56	Juv	31
				2009	Lower River Ayr mainstem	-4.47	55.48	Juv	49
				2009	Upper River Ayr mainstem	-4.19	55.52	Juv	50
	Bladnoch	-4.4	54.87	2008	Mainstem	-4.55	54.87	Juv	34
	Clyde	-5	55.67	2008	Allander Water	-4.28	55.93	Juv	60
				2008	Allander Water	-4.37	55.98	Juv	50
				2009	River Calder	-4.64	55.80	Juv	50
				2008	River Clyde Mainstem	-3.91	55.73	Juv	50
				2008	Glazert Water	-4.22	55.98	Juv	30
				2009	River Kelvin Mainstem	-4.21	55.94	Juv	29
				2008	River Gryffe	-4.64	55.88	Juv	68
	Cree	-4.4	54.85	2008	Penkiln Burn	-4.48	54.97	Juv	46
	Doon	-4.65	55.44	2005	Muck Water	-4.41	55.32	Juv	27
				2005	Doon Mainstem	-4.55	55.39	Juv	2
				2008	Garpel Burn	-4.40	55.25	Juv	31
	Garnock	-4.69	55.61	2003	Dusk Water	-4.70	55.69	Juv	20
				2003	R.Garnock Main	-4.71	55.71	Juv	20
	Girvan	-4.85	55.25	2005	Water of Girvan Mainstem	-4.60	55.34	Juv	30
				2005	Upper Water of Girvan	-4.54	55.28	Juv	31
				2004	Water of Girvan Mainstem	-4.78	55.26	Juv	31
	Luce	-4.83	54.85	2008	Cross Water of Luce	-4.84	54.97	Juv	50
				2008	Main Water of Luce	-4.92	55.00	Juv	50
				2008	Main Water of Luce	-4.85	54.94	Juv	100
				2008	Main Water of Luce	-4.84	54.91	Juv	50
				2008	Main Water of Luce	-4.82	54.88	Juv	50
	Stinchar	-5	55.1	2003	R.Stinchar Main	-4.91	55.13	Juv	42
				2004	Lower Water of Assel	-4.82	55.19	Juv	24
				2005	Duisk River	-4.73	55.09	Juv	37
				2004	Stinchar Mainstem	-4.64	55.23	Juv	31
				2003	Muck water	-4.75	55.15	Juv	30
England	Annan	-3.27	54.97	2009	Birnock Water	-3.44	55.33	Juv	49

	River			Year	Site				
				2009	Dryfe Water	-3.30	55.26	Juv	35
				2005	Evan Water	-3.54	55.40	Juv	35
				2009	Mein Water	-3.28	55.04	Juv	31
				2005	Kinnel Water	-3.47	55.28	Juv	29
				2009	Mainstem near Millhouse Bridge	-3.41	55.15	Juv	50
				2009	Moffat Water	-3.32	55.40	Juv	31
				2009	Water of Ae	-3.60	55.19	Juv	35
				2005	Water of Milk	-3.20	55.17	Juv	32
				2005	Wamphray Water	-3.40	55.25	Juv	31
	Duddon	-3.12	54.42	2008	R. Lickle	-3.21	54.30	Juv	31
	Eden	-3.06	54.95	2008	Scandal Beck (2 sites)	-1.01	54.47197/ 54.49311	Juv	49
	Ehen	-3.52	54.43	2008	Ehen site A	-3.50	54.52	Juv	11
				2008	Ehen site B	-3.53	54.44	Juv	9
				2008	Ehen site C	-3.53	54.45	Juv	10
				2008	Ehen site D	-3.43	54.53	Juv	16
	Esk	-3.43	54.34	2008	Hermitage Water	-2.86	55.26	Juv	43
				2008	Wauchope Water	-3.02	55.14	Juv	46
	Kent	-2.78	54.25	2004	Sprint	-2.80	54.44	Juv	23
				2004	Sprint	-2.79	54.44	Juv	18
	Lune	-2.88	53.99	2004	Greta, A65 road bridge	-2.48	54.15	Juv	30
	Nith	-3.6	55	2009	Cample Water	-3.75	55.23	Juv	33
				2005	Carin Water	-3.80	55.15	Juv	29
				2009	Carron Water	-3.78	55.28	Juv	48
				2005	mainstem	-3.63	55.11	Juv	21
				2005	mainstem	-4.01	55.00	Juv	29
				2009	River Nith	-3.63	55.12	Juv	50
				2008	Scar Water	-3.96	55.31	Juv	50
				2008	Upper Nith	-4.30	55.38	Juv	48
				2009	Crawick Water	-3.89	55.43	Juv	50
				2009	Spango Water	-3.96	55.44	Juv	50
	Ribble	-2.77	53.75	2004	Hammerton Hall	-2.43	53.98	Juv	31
	Urr	-3.83	54.83	2008	Mainstem	-3.89	54.98	Juv	48
Southwest England & Wales	Camel	-4.89	50.55	2005	De Lank	-4.70	50.53	Juv	30
				2005	Gam	-4.70	50.57	Juv	28
				2005	Kenning-Stock	-4.69	50.60	Juv	30
	Conwy	-3.83	53.3	2007	Trefriw	-3.82	53.15	Juv	50
	Dart	-3.56	50.34	2006	East Dart (Postbridge)	-3.91	50.60	Juv	39
				2005	East Dart (Postbridge)	-3.91	50.60	Juv	40
	Dee	-3.23	53.35	2008	Main River 5.2	-2.98	52.97	Juv	30
				2008	Main River 5.4	-3.05	52.97	Juv	30
				2005	Ceiriog	-3.20	52.91	Juv	38
	Exe	-3.43	50.61	2004	Danes Brook	-3.63	51.06	Juv	43

Region	River	Lon	Lat	Year	Site	Lon	Lat	Stage	N
				2004	Sherdon water	-3.71	51.11	Juv	31
				2004	Barle	-3.75	51.14	Juv	30
				2005	Barle	-3.75	51.14	Juv	38
	Fowey	-4.63	50.35	2005	Margate Ford	-4.68	50.47	Juv	20
				2004	Treverbyn	-4.53	50.48	Juv	35
	Nevern	-4.84	52.02	2009		-4.73	52.00	Juv	44
	Tamar	-4.17	50.36	2003	Lyd	-4.29	50.63	Juv	33
				2003	Inny	-4.32	50.57	Juv	33
				2003	Ottery	-4.56	50.71	Juv	29
	Taw	-4.22	51.08	2004	Bray	-3.89	51.13	Juv	25
				2004	Twitchen stream	-3.72	51.03	Juv	32
	Tawe	-3.93	51.62	2009		-3.71	51.80	Juv	48
	Teifi	-4.67	52.08	2005	Clettwr	-4.27	52.04	Juv	8
				2005	Nant Egnant	-3.78	52.28	Juv	15
				2005	Lampeter	-4.07	52.11	Juv	11
				2009		-4.09	52.11	Juv	42
	Teign	-3.5	50.54	2004	Leigh Bridge			Juv	23
				2004	Leigh House	-3.87	50.68	Juv	21
	Torridge	-4.2	51.05	2004	East Oakement	-3.98	50.74	Juv	21
				2005	West Oakement	-4.01	50.73	Juv	29
	Usk	-2.98	51.58	2004	Ysgir	-3.45	51.97	Juv	30
	Wye	-2.66	51.61	2004	Edw	-3.30	52.13	Juv	27
Southern England	Avon	-1.63	50.79	2004	Upper Avon	-1.82	51.10	Juv	23
				2004	Bugmoor Hatches	-1.79	51.01	Juv	20
	Itchen	-1.39	50.91	2005	Bishopstoke Barge	-1.34	50.97	Juv	27
				2006	Bishopstoke Barge	-1.34	50.97	Juv	24
	Test	-1.48	50.93	2004	Across catchment			Juv	49
Northern Ireland	Upper Bann	-6.76	55.16	2008		-6.08	54.11	Juv	85
	Agivey	-6.76	55.16	2008		-6.37	55.01	Juv	91
	Blac	-6.76	55.16	2008		-7.09	54.25	Juv	96
	Clogh	-6.76	55.16	2008		-6.14	54.57	Juv	73
	Grillagh	-6.76	55.16	2008		-6.37	54.52	Juv	82
	Kells water	-6.76	55.16	2008		-6.03	54.49	Juv	78
	Moyola	-6.76	55.16	2008		-6.42	54.49	Juv	77
	Six Mile	-6.76	55.16	2008		-6.02	54.43	Juv	88
Ungroupable	Barrow	-6.97	52.23	2006	Greese	-6.92	52.91	Juv	47
				2006	Mountain	-6.90	52.60	Juv	43
	Boyne	-6.25	53.72	2006	Trimblestown	-6.91	53.62	Juv	48
				2006	Balckwater	-6.86	53.73	Juv	46
	Suir	-7	52.27	2006	Anner	-7.65	52.43	Juv	48
				2005	Drish	-7.78	52.67	Juv	48
				2006	Nire	-7.73	52.27	Juv	37

France	Aulne	-4.33	48.31	2005	Across catchment			Adult	38
	Blavet	-3.37	47.71	2005	Across catchment			Adult	47
	Elle	-3.55	47.87	2005	Across catchment			Adult	47
	Elorn	-4.41	48.39	2005	Across catchment			Adult	47
	Leguer	-3.41	48.57	2005	Across catchment			Adult	47
	Scorff	-3.35	47.74	2005	Across catchment			Adult	45
	Sée	-1.41	48.67	2005	Across catchment			Adult	47
	Sélune	-1.4	48.65	2005	Across catchment			Adult	48
Norway	Daleelva	61.01	4.09	2002		6.07	61.20	Adut	105
	Laukhell evassdra get	69.23	17.85	2006		17.59	69.25	Juv	87
	Namsen	64.46	11.5	2007		12.33	64.46	Juv	92
	Vesterelv a	28.56	70.10	2009		28.56	70.10	Juv	93

Appendix III- Pair-wise F_{ST} values between each salmon population of the Mersey assignment baseline (calculated in Fstat). Number are displayed to 2 decimal points. Colours are used to identify pair-wise F_{ST} values between rivers within each reporting region. For the purpose of readability, the table has been divided into sections which are displayed individually as a, b or c.

Section a

	Ayr	Bladn	Clyde	Cree	Doon	Garno	Girvan	Luce	Stinch	Annan	Duddo	Eden	Ehen	Esk	Kent	Lune	Nith	Ribble	Urr	Camel	Conwy	Dart	Dee	Exe	Fowey	Nevern	Tamar	Taw	Tawe	Teifi	Teign	Torrid	Usk	Wye
Bladnoch	0.01																																	
Clyde	0.01	0.01																																
Cree	0.02	0.00	0.01																															
Doon	0.02	0.03	0.02	0.03																														
Garnock	0.01	0.01	0.01	0.01	0.02																													
Girvan	0.01	0.01	0.01	0.02	0.02	0.01																												
Luce	0.01	0.01	0.01	0.01	0.02	0.01	0.01																											
Stinchar	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01																										
Annan	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01																									
Duddon	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.02	0.02	0.02																								
Eden	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.01	0.03																							
Ehen	0.02	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01	0.03	0.01																						
Esk	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.00	0.02	0.01																					
Kent	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01																				
Lune	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.02	0.01	0.03	0.02	0.02	0.01	0.01																			
Nith	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.00	0.01	0.01																		
Ribble	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.01																	
Urr	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.01	0.02	0.02	0.01	0.03																
Camel	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02															
Conwy	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01														
Dart	0.04	0.04	0.03	0.02	0.03	0.03	0.03	0.03	0.02	0.01	0.03	0.03	0.03	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.02													
Dee	0.02	0.01	0.01	0.02	0.03	0.01	0.01	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02												
Exe	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.02											
Fowey	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.01	0.02										
Nevern	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.03	0.02											
Tamar	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.01									
Taw	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01									
Tawe	0.02	0.02	0.02	0.02	0.04	0.02	0.03	0.02	0.02	0.02	0.03	0.03	0.02	0.01	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.02	0.03	0.03						
Teifi	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.01	0.00	0.00	0.01	0.02				
Teign	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.01	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01				
Torridge	0.03	0.02	0.02	0.04	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.04	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.02	0.02	
Usk	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.01	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.02	0.02		
Wye	0.03	0.03	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.04	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.03	0.01	0.02	0.04	0.00							

a

Section b / c

	Ayr	Bladn	Clyde	Cree	Doon	Garno	Girvan	Luce	Stinch	Annan	Duddo	Eden	Ehen	Esk	Kent	Lune	Nith	Ribble	Urr	Camel	Conwy	Dart	Dee	Exe	Fowey	Nevern	Tamar	Taw	Tawe	Teifi	Teign	Torrid	Usk	Wye	Avon	Itchen	Test	Upper Bann	Agivey	Blackw	Clogh	Grillag	Kells w	Moyol	Six Mil	Barrow	Boyne	Suir	Aulne	Blavet	Elle	Elorn	Legue	Scorff	Sée	Sélune	Daleel	Laukh	Namse	Vester
Avon	0.07	0.08	0.07	0.08	0.08	0.07	0.07	0.06	0.07	0.07	0.08	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.07	0.07	0.07	0.09	0.07	0.09	0.07	0.07	0.07	0.08	0.08	0.07	0.09	0.09	0.06	0.08																										
Itchen	0.07	0.08	0.07	0.08	0.08	0.07	0.07	0.06	0.07	0.07	0.09	0.08	0.06	0.07	0.08	0.07	0.06	0.08	0.07	0.07	0.07	0.09	0.07	0.09	0.07	0.07	0.07	0.07	0.07	0.09	0.08	0.06	0.08	0.03	0.02																									
Test	0.07	0.08	0.07	0.09	0.08	0.08	0.07	0.06	0.07	0.07	0.09	0.08	0.06	0.07	0.09	0.08	0.06	0.08	0.06	0.07	0.07	0.09	0.07	0.09	0.07	0.07	0.07	0.08	0.09	0.06	0.07	0.08	0.03	0.02																							
Upper Bann	0.06	0.05	0.04	0.04	0.06	0.04	0.06	0.04	0.05	0.05	0.07	0.06	0.06	0.05	0.05	0.05	0.05	0.08	0.05	0.06	0.07	0.06	0.07	0.13	0.12	0.13																									
Agivey	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.02	0.02	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.03	0.04	0.03	0.04	0.10	0.09	0.10	0.04																				
Blackwater	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.04	0.03	0.03	0.03	0.03	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.10	0.09	0.10	0.03	0.01																	
Clogh	0.04	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.03	0.03	0.04	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.05	0.04	0.04	0.05	0.10	0.11	0.04	0.02	0.02															
Grillagh	0.03	0.03	0.02	0.02	0.03	0.01	0.02	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.02	0.03	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.02	0.04	0.09	0.08	0.09	0.04	0.01	0.01	0.02													
Kells Water	0.04	0.04	0.03	0.04	0.03	0.04	0.03	0.03	0.03	0.04	0.04	0.03	0.03	0.04	0.03	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.05	0.10	0.10	0.10	0.04	0.03	0.02	0.02	0.02												
Moyola	0.03	0.02	0.02	0.02	0.03	0.02	0.04	0.02	0.01	0.01	0.03	0.03	0.02	0.01	0.02	0.02	0.02	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.03	0.08	0.08	0.08	0.04	0.02	0.02	0.02	0.02	0.02											
Six Mile	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.03	0.02	0.02	0.03	0.03	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.04	0.09	0.09	0.10	0.02	0.02	0.01	0.01	0.01	0.02	0.02										
Barrow	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.07	0.07	0.07	0.04	0.02	0.02	0.02	0.01	0.02	0.02	0.02									
Boyne	0.02	0.02	0.02	0.02	0.03	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.03	0.02	0.02	0.02	0.07	0.07	0.07	0.05	0.03	0.03	0.04	0.02	0.04	0.03	0.03	0.01							
Suir	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.02	0.07	0.07	0.07	0.05	0.03	0.03	0.04	0.02	0.03	0.02	0.03	0.01	0.02						
Aulne	0.04	0.03	0.03	0.03	0.05	0.03	0.03	0.02	0.03	0.03	0.04	0.03	0.05	0.02	0.04	0.03	0.04	0.05	0.03	0.04	0.04	0.02	0.04	0.05	0.03	0.04	0.08	0.08	0.08	0.04	0.05	0.03	0.06	0.03										
Blavet	0.05	0.04	0.04	0.04	0.06	0.04	0.04	0.03	0.04	0.04	0.06	0.05	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.03	0.02	0.04	0.04	0.03	0.05	0.04	0.04	0.04	0.04	0.08	0.09	0.09	0.09	0.06	0.06	0.07	0.05	0.07	0.05	0.07	0.04	0.05	0.04	0.01				
Elle	0.04	0.04	0.03	0.04	0.05	0.04	0.04	0.03	0.04	0.04	0.05	0.04	0.06	0.04	0.05	0.04	0.04	0.06	0.03	0.03	0.04	0.05	0.03	0.03	0.04	0.05	0.03	0.03	0.03	0.08	0.08	0.08	0.09	0.07	0.06	0.07	0.06	0.06	0.04	0.07	0.04	0.05	0.04	0.01	0.02	0.00		
Elorn	0.05	0.05	0.04	0.05	0.06	0.05	0.05	0.04	0.05	0.07	0.06	0.04	0.04	0.05	0.05	0.04	0.06	0.03	0.03	0.04	0.04	0.03	0.06	0.03	0.04	0.02	0.02	0.03	0.08	0.08	0.08	0.09	0.07	0.06	0.07	0.06	0.06	0.04	0.07	0.04	0.05	0.04	0.01	0.02	0.01	0.00		
Leguer	0.04	0.03	0.03	0.04	0.05	0.03	0.03	0.03	0.03	0.03	0.05	0.03	0.03	0.03	0.03	0.04	0.03	0.02	0.02	0.04	0.02	0.04	0.02	0.02	0.03	0.03	0.03	0.03	0.07	0.08	0.07	0.08	0.06	0.05	0.06	0.06	0.03	0.06	0.03	0.04	0.03	0.01	0.01	0.00	0.01	0.00		
Scorff	0.04	0.03	0.03	0.03	0.05	0.03	0.04	0.03	0.03	0.04	0.06	0.05	0.03	0.04	0.04	0.04	0.04	0.04	0.05	0.03	0.03	0.05	0.03	0.03	0.05	0.04	0.04	0.03	0.05	0.08	0.08	0.07	0.08	0.06	0.05	0.06	0.05	0.06	0.03	0.06	0.03	0.04	0.03	0.01	0.01	0.00	0.01	0.00		
Sée	0.06	0.06	0.05	0.05	0.07	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.07	0.05	0.03	0.05	0.04	0.06	0.04	0.06	0.04	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.09	0.10	0.09	0.11	0.07	0.07	0.08	0.07	0.08	0.04	0.08	0.04	0.04	0.04	0.02	0.03	0.04	0.05	0.04	0.04	0.00				
Sélune	0.05	0.05	0.04	0.05	0.07	0.04	0.04	0.04	0.04	0.04	0.07	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.06	0.04	0.04	0.05	0.03	0.04	0.05	0.04	0.04	0.04	0.04	0.09	0.09	0.09	0.09	0.06	0.06	0.07	0.06	0.07	0.04	0.06	0.04	0.04	0.02	0.03	0.03	0.03	0.03	0.03	0.00			
Daleelva	0.04	0.03	0.03	0.02	0.05	0.03	0.04	0.03	0.03	0.03	0.05	0.04	0.04	0.04	0.04	0.04	0.03	0.05	0.04	0.04	0.04	0.03	0.04	0.03	0.04	0.05	0.04	0.04	0.04	0.11	0.10	0.11	0.06	0.04	0.05	0.06	0.04	0.05	0.05	0.05	0.04	0.05	0.05	0.06	0.07	0.06	0.06	0.06	0.06			
Laukhelleva	0.05	0.05	0.04	0.03	0.06	0.04	0.05	0.04	0.04	0.04	0.06	0.05	0.06	0.05	0.05	0.05	0.04	0.06	0.05	0.05	0.04	0.06	0.05	0.04	0.06	0.05	0.05	0.04	0.06	0.10	0.11	0.11	0.08	0.05	0.06	0.06	0.05	0.06	0.06	0.06	0.04	0.05	0.05	0.05	0.07	0.05	0.05	0.05	0.05	0.01		
Namsen	0.04	0.04	0.04	0.03	0.06	0.03	0.05	0.03	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.06	0.04	0.06	0.05	0.05	0.04	0.05	0.04	0.06	0.05	0.05	0.04	0.04	0.06	0.11	0.11	0.11	0.07	0.05	0.05	0.06	0.05	0.06	0.06	0.04	0.04	0.06	0.06	0.06	0.08	0.06	0.06	0.06	0.06	0.01	0.01	
Vesterelva	0.06	0.07	0.05	0.05	0.08	0.05	0.06	0.05	0.05	0.06	0.08	0.07	0.06	0.06	0.06	0.08	0.06	0.08	0.06	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.06	0.07	0.11	0.11	0.10	0.10	0.06	0.07	0.08	0.06	0.08	0.07	0.08	0.05	0.06	0.06	0.04	0.05	0.06	0.07	0.06	0.06	0.05	0.05	0.05	0.03	0.04

b   *c*

217

a	Ayr	Bladnoch	Clyde	Cree	Doon	Garnock	Girvan	Luce	Stinchar	Annan	Duddon	Eden	Ehen	Esk	Kent	Lune	Nith	Ribble	Urr	Camel	Conwy	Dart	Dee	Exe	Fowey	Nevern	Tamar	Taw	Tawe	Teifi	Teign	Torridge	Usk
Bladnoch	0.01																																
Clyde	0.01	0.01																															
Cree	0.02	0.00	0.01																														
Doon	0.02	0.03	0.02	0.03																													
Garnock	0.01	0.01	0.01	0.01	0.02																												
Girvan	0.01	0.01	0.01	0.02	0.02	0.01																											
Luce	0.01	0.01	0.01	0.02	0.02	0.01	0.01																										
Stinchar	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01																									
Annan	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01																								
Duddon	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.02	0.02	0.02																							
Eden	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.01	0.03																						
Ehen	0.02	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01	0.03	0.01																					
Esk	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.02	0.01	0.01																				
Kent	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01																			
Lune	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.02	0.01	0.03	0.02	0.02	0.01	0.01																		
Nith	0.02	0.01	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.00	0.01	0.01																	
Ribble	0.03	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.01																
Urr	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.01	0.02	0.02	0.01	0.03															
Camel	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02														
Conwy	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01													
Dart	0.04	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.02	0.02	0.03	0.03	0.03	0.02	0.02	0.03	0.02	0.03	0.03	0.02	0.02												
Dee	0.02	0.01	0.01	0.02	0.03	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02											
Exe	0.03	0.02	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.02										
Fowey	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.01	0.02									
Nevern	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.03	0.02								
Tamar	0.03	0.02	0.01	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.01							
Taw	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01						
Tawe	0.02	0.02	0.02	0.02	0.04	0.02	0.03	0.02	0.02	0.02	0.03	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.02	0.03	0.03					
Teifi	0.02	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.02	0.01	0.00	0.00	0.01	0.02				
Teign	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01			
Torridge	0.03	0.02	0.02	0.03	0.04	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.02	0.03	0.03	0.03	0.02	0.04	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02		
Usk	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02		
Wye	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.03	0.01	0.02	0.04	0.00

b	Ayr	Bladnoch	Clyde	Cree	Doon	Garnock	Girvan	Luce	Stinchar	Annan	Duddon	Eden	Ehen	Esk	Kent	Lune	Nith	Ribble	Urr	Camel	Conwy	Dart	Dee	Exe	Fowey	Nevern	Tamar	Taw	Tawe	Teifi	Teign	Torridge	Usk
Avon	0.07	0.08	0.07	0.08	0.08	0.07	0.07	0.06	0.07	0.07	0.08	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.06	0.07	0.07	0.09	0.07	0.09	0.07	0.07	0.07	0.08	0.08	0.07	0.09	0.09	0.06
Itchen	0.07	0.08	0.07	0.08	0.08	0.07	0.07	0.06	0.06	0.06	0.07	0.08	0.07	0.06	0.06	0.07	0.06	0.07	0.06	0.07	0.07	0.08	0.06	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.09	0.08	0.06
Test	0.07	0.08	0.07	0.09	0.08	0.08	0.07	0.06	0.07	0.07	0.09	0.08	0.06	0.07	0.09	0.08	0.06	0.08	0.06	0.07	0.07	0.09	0.07	0.09	0.07	0.07	0.07	0.08	0.08	0.07	0.09	0.09	0.06
Upper Bann	0.06	0.05	0.04	0.04	0.06	0.04	0.06	0.04	0.05	0.05	0.07	0.06	0.06	0.05	0.05	0.05	0.05	0.08	0.05	0.07	0.05	0.07	0.06	0.06	0.07	0.05	0.05	0.06	0.06	0.05	0.06	0.07	0.06
Agivey	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.04	0.04	0.03	0.03	0.02	0.03	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.04	0.03	0.03	0.04	0.03	0.03	0.04	0.04
Blackwater	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.03	0.04	0.03
Clogh	0.04	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.03	0.03	0.04	0.03	0.04	0.04	0.04	0.03	0.05	0.03	0.04	0.05	0.04	0.03	0.04	0.04	0.03	0.04	0.05	0.03
Grillagh	0.03	0.03	0.02	0.02	0.03	0.01	0.02	0.01	0.02	0.02	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.04	0.03	0.03	0.02	0.03	0.02	0.03	0.03	0.03	0.03	0.02	0.03	0.02	0.02	0.03	0.03
Kells Water	0.04	0.04	0.03	0.03	0.04	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.03	0.05	0.04	0.04	0.03	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.05	0.03
Moyola	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.03	0.02	0.01	0.02	0.02	0.02	0.03	0.02	0.03	0.02	0.03	0.01	0.03	0.03	0.02	0.02	0.02	0.03	0.01	0.02	0.03	0.02
Six Mile	0.03	0.03	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.03	0.03	0.02	0.02	0.03	0.03	0.04	0.03	0.04	0.03	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.02	0.03	0.04	0.03
Barrow	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.02
Boyne	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.02	0.03	0.02	0.01	0.01	0.01	0.03	0.02	0.03	0.02	0.02	0.01	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02
Suir	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.01	0.01	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.03	0.01	0.01	0.02	0.02
Aulne	0.04	0.03	0.03	0.03	0.05	0.03	0.03	0.02	0.03	0.03	0.06	0.04	0.03	0.03	0.03	0.04	0.03	0.05	0.02	0.02	0.03	0.04	0.02	0.04	0.03	0.03	0.03	0.04	0.04	0.03	0.02	0.03	0.03
Blavet	0.05	0.04	0.04	0.04	0.06	0.04	0.04	0.03	0.04	0.04	0.06	0.05	0.04	0.04	0.04	0.05	0.04	0.05	0.03	0.02	0.04	0.04	0.03	0.05	0.03	0.05	0.04	0.03	0.05	0.04	0.04	0.04	0.04
Elle	0.04	0.04	0.03	0.04	0.05	0.03	0.04	0.03	0.04	0.04	0.06	0.06	0.03	0.04	0.05	0.04	0.04	0.05	0.03	0.03	0.04	0.05	0.03	0.05	0.03	0.04	0.03	0.04	0.05	0.03	0.03	0.03	0.03
Elorn	0.05	0.05	0.04	0.05	0.06	0.05	0.04	0.03	0.04	0.05	0.07	0.06	0.04	0.04	0.05	0.05	0.04	0.06	0.03	0.03	0.04	0.05	0.03	0.06	0.03	0.04	0.03	0.05	0.06	0.03	0.04	0.04	0.04
Leguer	0.04	0.03	0.03	0.04	0.05	0.03	0.03	0.03	0.03	0.03	0.05	0.04	0.03	0.03	0.04	0.03	0.03	0.05	0.02	0.02	0.02	0.04	0.02	0.04	0.02	0.03	0.02	0.03	0.04	0.02	0.02	0.03	0.02
Scorff	0.04	0.03	0.03	0.03	0.05	0.03	0.04	0.03	0.03	0.04	0.06	0.05	0.03	0.04	0.04	0.04	0.04	0.05	0.03	0.03	0.03	0.04	0.03	0.05	0.03	0.03	0.02	0.04	0.04	0.03	0.03	0.03	0.03
Sée	0.06	0.06	0.05	0.05	0.07	0.04	0.05	0.05	0.05	0.05	0.08	0.05	0.05	0.05	0.05	0.06	0.05	0.07	0.05	0.03	0.05	0.05	0.04	0.06	0.04	0.06	0.05	0.05	0.05	0.05	0.04	0.05	0.04
Sélune	0.05	0.05	0.04	0.05	0.07	0.04	0.04	0.04	0.04	0.04	0.07	0.04	0.03	0.04	0.04	0.05	0.04	0.06	0.04	0.03	0.04	0.04	0.03	0.05	0.03	0.05	0.04	0.04	0.04	0.04	0.03	0.04	0.03
Daleelva	0.04	0.03	0.03	0.02	0.05	0.03	0.04	0.03	0.03	0.03	0.05	0.04	0.04	0.04	0.03	0.04	0.03	0.04	0.04	0.04	0.03	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.03	0.03	0.03	0.05	0.04
Laukhellevas	0.05	0.05	0.04	0.03	0.06	0.04	0.05	0.04	0.04	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.04	0.06	0.05	0.05	0.04	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.05	0.04
Namsen	0.04	0.04	0.04	0.03	0.06	0.03	0.05	0.03	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.06	0.04	0.06	0.05	0.04	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.04	0.04	0.06	0.06	0.06	0.05
Vesterelva	0.06	0.07	0.05	0.05	0.08	0.05	0.06	0.05	0.05	0.06	0.08	0.07	0.06	0.06	0.06	0.08	0.06	0.08	0.06	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.06	0.06

	Wye	Avon	Itchen	Test	Upper Bann	Agivey	Blackwater	Clogh	Grillagh	Kells water	Moyola	Six Mile	Barrow	Boyne	Suir	Aulne	Blavet	Elle	Elorn	Leguer	Scorff	Sée	Sélune	Daleelva	Laukhellevassdraget	Namsen
Avon	0.08																									
Itchen	0.08	0.03																								
Test	0.08	0.03	0.02																							
Upper Bann	0.07	0.13	0.12	0.13																						
Agivey	0.04	0.10	0.09	0.10	0.04																					
Blackwater	0.04	0.10	0.09	0.10	0.03	0.01																				
Clogh	0.04	0.10	0.10	0.11	0.04	0.02	0.02																			
Grillagh	0.03	0.09	0.08	0.09	0.04	0.01	0.01	0.02																		
Kells Water	0.04	0.10	0.10	0.10	0.04	0.03	0.02	0.02	0.02																	
Moyola	0.03	0.08	0.08	0.08	0.04	0.02	0.02	0.02	0.02	0.02																
Six Mile	0.04	0.09	0.09	0.10	0.02	0.02	0.01	0.01	0.01	0.02	0.02															
Barrow	0.02	0.07	0.07	0.07	0.04	0.02	0.02	0.02	0.01	0.02	0.02	0.02														
Boyne	0.02	0.08	0.07	0.07	0.05	0.03	0.03	0.04	0.02	0.04	0.03	0.03	0.01													
Suir	0.02	0.07	0.07	0.07	0.05	0.03	0.03	0.04	0.02	0.03	0.02	0.03	0.01	0.02												
Aulne	0.04	0.08	0.08	0.08	0.08	0.05	0.05	0.06	0.04	0.05	0.03	0.06	0.03	0.03	0.03											
Blavet	0.04	0.08	0.09	0.09	0.09	0.06	0.06	0.07	0.05	0.07	0.05	0.07	0.04	0.05	0.04	0.01										
Elle	0.05	0.08	0.08	0.08	0.09	0.06	0.06	0.06	0.05	0.06	0.04	0.06	0.03	0.05	0.04	0.01	0.01									
Elorn	0.05	0.08	0.08	0.08	0.09	0.07	0.06	0.07	0.06	0.06	0.04	0.07	0.04	0.05	0.04	0.01	0.02	0.00								
Leguer	0.04	0.07	0.08	0.07	0.08	0.06	0.05	0.06	0.05	0.06	0.03	0.06	0.03	0.04	0.02	0.01	0.02	0.01	0.01							
Scorff	0.05	0.08	0.08	0.07	0.08	0.06	0.05	0.06	0.05	0.06	0.04	0.06	0.03	0.04	0.03	0.01	0.01	0.00	0.01	0.00						
Sée	0.05	0.09	0.10	0.09	0.11	0.07	0.07	0.08	0.05	0.07	0.06	0.08	0.04	0.04	0.04	0.03	0.03	0.04	0.05	0.04	0.04					
Sélune	0.05	0.09	0.09	0.09	0.09	0.06	0.06	0.07	0.05	0.06	0.05	0.07	0.04	0.03	0.04	0.02	0.03	0.03	0.03	0.03	0.03	0.00				
Daleelva	0.04	0.11	0.10	0.11	0.06	0.04	0.05	0.05	0.04	0.05	0.04	0.05	0.03	0.04	0.03	0.05	0.06	0.06	0.07	0.06	0.06	0.06	0.06			
Laukhellevas	0.05	0.10	0.11	0.11	0.08	0.05	0.06	0.06	0.05	0.06	0.05	0.06	0.04	0.05	0.04	0.05	0.05	0.05	0.07	0.05	0.05	0.05	0.05	0.01		
Namsen	0.05	0.11	0.11	0.11	0.07	0.05	0.05	0.06	0.05	0.06	0.04	0.06	0.04	0.04	0.04	0.06	0.06	0.06	0.08	0.06	0.06	0.06	0.06	0.01	0.01	
Vesterelva	0.07	0.11	0.11	0.10	0.10	0.06	0.07	0.08	0.06	0.08	0.07	0.08	0.05	0.06	0.06	0.04	0.05	0.06	0.07	0.06	0.06	0.05	0.05	0.05	0.03	0.04

Appendix IV- Assessment of Hardy–Weinberg equilibrium at each locus for each baseline river (calculated in Arlequin). Probability of conformance to Hardy Weinberg as calculated by Arlequin. Number in bold show signficant deviation away from Hardy-Weinberg after Bonferroni correction. Table has been divided into sections for the purpose of readability, and is order to the convention described in the chapter.

	Ayr	Bladnoch	Clyde	Cree	Doon	Garnock	Girvan	Luce	Stinchar	Annan	Duddon	Eden	Ehen	Esk	Kent	Lune	Nith	Ribble	Urr	Camel	Conwy	Dart	Dee	Exe	Fowey	Nevern	Tamar	Taw	Tawe	Teifi
SSspG7	0.88	0.27	0.09	0.17	0.26	0.95	0.50	0.03	0.05	0.08	0.90	0.06	0.07	0.10	0.02	0.42	0.01	0.00	0.90	0.02	0.87	0.02	0.01	**0.00**	0.01	0.27	0.09	0.27	0.05	0.09
Ssa14	0.03	0.45	0.53	0.72	0.92	0.42	0.21	0.34	0.52	0.60	0.14	0.04	1.00	0.08	0.72	0.06	0.83	0.51	0.06	0.71	1.00	0.56	0.01	0.32	0.58	0.26	**0.00**	0.47	0.66	0.91
Ssa202	0.98	0.13	0.13	0.14	0.00	0.07	0.07	0.18	0.06	0.17	0.95	0.44	0.13	0.36	0.73	0.36	0.95	0.03	0.67	**0.00**	0.67	0.23	0.89	0.00	0.25	0.37	0.27	0.89	0.48	0.46
SSsp3016	0.30	0.32	0.38	0.11	0.08	0.65	0.06	0.17	0.06	0.01	0.44	0.11	0.51	0.16	0.01	0.38	0.01	0.34	0.94	0.07	0.59	0.01	0.06	0.15	0.07	0.63	0.84	0.73	0.12	0.11
Ssa197	0.02	0.79	0.37	0.97	0.08	0.78	0.44	0.11	**0.00**	0.06	0.08	0.20	0.41	**0.00**	0.01	0.59	0.02	0.85	0.24	0.00	0.13	0.14	0.28	0.00	0.16	0.00	**0.00**	0.05	0.06	0.50
SsaF43	0.77	0.37	0.70	0.95	0.49	0.15	0.57	0.02	0.88	0.99	0.76	0.01	0.21	0.11	0.07	0.04	0.40	0.71	0.13	0.10	0.84	0.00	0.27	0.07	0.09	0.00	0.56	0.43	0.19	0.59
SSsp1605	0.18	0.81	0.53	0.58	0.07	0.65	0.30	0.98	0.29	0.13	0.84	0.38	0.78	0.19	0.30	0.32	0.07	0.44	0.13	0.10	0.43	0.06	0.40	0.01	0.32	0.00	0.04	0.27	0.04	0.10
SSsp2210	0.04	0.40	0.35	0.94	0.84	0.35	0.22	0.52	0.34	0.61	0.84	0.49	0.84	0.33	0.75	0.23	0.15	0.41	0.32	0.06	0.28	0.20	0.07	0.09	0.08	0.05	0.05	0.34	0.07	0.03
SSsp2216	0.21	0.90	0.41	0.93	0.17	0.62	0.76	0.39	0.04	0.22	0.22	0.11	0.80	0.29	0.08	0.84	0.62	0.74	0.84	0.15	0.96	0.13	0.12	0.10	0.75	0.58	0.65	0.64	0.52	0.88
SsaD157	0.24	0.09	0.11	0.86	**0.00**	0.42	0.84	0.69	**0.00**	0.32	0.30	0.23	**0.00**	0.07	0.76	0.37	0.93	0.75	0.83	0.20	0.60	0.76	0.23	0.02	0.22	0.40	0.04	0.19	0.00	0.02
Ssa171	0.06	0.49	0.01	0.27	0.01	0.29	0.04	0.04	0.17	0.15	0.57	0.07	0.51	0.69	0.93	0.27	**0.00**	0.87	0.28	0.19	0.71	0.33	0.16	0.40	0.44	0.39	0.09	0.40	0.14	0.98
Ssa289	0.41	0.52	1.00	1.00	0.16	0.23	0.02	0.22	0.02	0.04	0.11	0.02	0.13	0.44	0.34	0.61	0.50	0.61	0.15	0.88	0.52	0.74	0.66	0.03	0.72	0.22	0.39	0.84	0.35	0.80
SsaD144	**0.00**	0.75	**0.00**	0.37	**0.00**	0.53	0.16	0.58	**0.00**	**0.00**	0.75	0.55	0.24	**0.00**	0.15	0.35	0.04	0.82	0.56	0.03	0.68	0.45	0.18	0.20	**0.00**	0.29	0.76	0.17	0.01	0.37
SSsp2201	0.67	0.43	0.31	0.24	**0.00**	0.45	0.55	0.39	**0.00**	0.48	0.02	0.08	0.23	0.01	0.21	0.30	0.26	0.89	0.26	0.13	0.59	0.02	0.59	0.05	**0.00**	0.00	0.30	0.27	0.14	0.23

	Teign	Torridge	Usk	Wye	Avon	Itchen	Test	Upper Bann	Agivey	Black Water	Clogh	Grillagh	Kells Water	Moyola	Six Mile	Barrow	Boyne	Suir	Aulne	Blavet	Elle	Elorn	Leguer	Scorff	See	Selune	Daleelva	Laukhelle vassdrage	Namsen	Vesterelva
SSspG7	0.83	0.03	0.56	0.14	0.31	0.13	0.01	0.84	0.55	0.05	0.61	0.31	0.30	0.12	0.09	0.14	0.02	0.03	0.03	0.81	0.00	0.18	**0.00**	0.10	**0.00**	**0.00**	0.24	0.78	0.16	0.59
Ssa14	0.34	0.24	0.19	0.16	NA	NA	1.00	1.00	0.04	0.41	0.44	0.13	0.18	1.00	0.83	0.09	0.83	0.56	0.02	0.73	0.71	0.86	0.29	0.06	0.32	0.56	0.43	0.48	1.00	1.00
Ssa202	0.06	0.33	0.48	0.55	0.07	0.28	0.27	0.04	0.69	0.01	0.08	0.15	0.77	0.61	0.27	0.31	0.11	0.03	0.07	0.05	0.60	0.24	0.95	0.31	0.47	0.29	0.02	0.01	0.44	0.47
SSsp3016	0.31	0.16	0.29	0.91	0.71	0.92	0.07	0.69	0.84	0.05	0.88	0.38	0.41	0.72	0.19	0.05	0.02	0.38	0.74	0.89	0.13	0.37	0.33	0.06	0.75	0.56	0.24	0.27	0.89	0.16
Ssa197	0.65	0.22	0.89	0.35	0.85	0.26	0.05	0.53	0.01	0.76	0.35	0.11	0.18	0.80	0.52	0.11	0.20	0.63	0.13	0.21	0.66	**0.00**	0.00	0.00	0.03	0.01	0.45	0.96	1.00	0.04
SsaF43	0.03	0.53	0.18	0.10	0.58	0.12	0.26	0.51	0.26	0.07	0.73	0.90	0.93	0.42	0.23	0.37	0.82	0.70	0.14	0.77	0.80	0.10	0.56	0.48	0.30	0.02	0.47	0.70	0.88	0.33
SSsp1605	0.08	0.44	0.34	0.38	0.09	0.18	0.17	0.03	0.64	0.27	0.66	0.88	0.04	0.54	0.27	0.96	0.15	0.37	0.07	0.04	0.04	0.48	0.43	0.47	0.22	0.16	0.94	0.03	0.99	0.32
SSsp2210	0.92	0.95	0.19	0.88	0.04	0.09	0.01	0.76	0.72	0.02	0.63	0.52	0.81	0.43	0.23	0.06	0.42	0.38	0.58	0.05	**0.00**	0.42	0.11	0.01	0.06	0.21	0.88	0.06	0.12	0.02
SSsp2216	0.34	0.10	0.25	0.34	0.17	0.22	0.68	0.33	0.37	0.34	0.32	0.90	0.22	0.37	0.31	0.00	0.42	0.50	0.50	0.45	0.69	0.58	0.34	1.00	0.66	0.13	0.22	0.95	0.26	
SsaD157	0.56	0.25	0.06	0.30	0.21	0.39	0.16	0.96	0.74	0.89	0.23	0.98	0.04	0.01	0.00	0.19	**0.00**	0.61	0.34	0.84	0.83	0.05	0.24	0.24	0.34	0.31	0.03	0.61	0.64	0.95
Ssa171	0.71	0.00	0.89	0.28	0.55	0.71	0.94	0.30	0.48	0.16	0.21	0.97	0.66	0.01	0.33	0.79	0.04	0.03	0.21	0.17	0.06	0.01	0.10	0.55	1.00	0.33	0.02	0.66	0.76	0.48
Ssa289	0.24	0.19	0.31	0.12	0.20	0.02	0.87	0.10	0.36	0.90	0.49	0.23	0.10	0.20	0.10	0.53	0.80	0.17	0.48	0.69	0.13	0.04	0.40	0.40	0.76	0.33	0.17	0.10	0.41	0.19
SsaD144	0.51	0.43	0.05	0.60	0.54	0.15	0.42	0.16	0.43	0.25	0.43	0.82	0.10	0.03	0.03	0.12	0.13	0.03	0.23	0.87	0.37	0.02	0.19	0.61	0.72	0.18	0.01	0.20	0.07	0.03
SSsp2201	0.42	0.16	0.06	0.98	0.59	0.44	0.03	0.42	**0.00**	0.00	0.24	0.04	0.42	0.69	0.04	0.15	0.01	0.60	0.44	0.18	0.24	0.05	0.96	0.43	0.56	0.03	0.03	0.95	0.37	0.63

Appendix V- Exclusion assignment of Mersey salmon. Analysis performed using the Cornuet *et al.* (1999) algorithm, which simulated 10000 individuals. A reference population is excluded only when the probability of assigning to the population is below 0.05, according to Vasemagi *et al.* (2001). Highlighted in red are individuals who fail to assign to any of the reporting regions

Sample ID	Scotland	Solway & Northwest England	Southwest England & Wales	Southern England	Northern Ireland	France	Norway
Mer.adu01.01	0.104	0.0322	0.0608	0	0	0.0017	0.0001
Mer.adu01.02	0.7827	0.9445	0.4793	0	0.2766	0.2383	0.3313
Mer.adu01.03	0.0067	0.0866	0.0681	0	0.001	0	0
Mer.adu01.04	0.1817	0.2328	0.1842	0	0.0387	0	0.0069
Mer.adu02.01	0.1046	0.0701	0.0045	0	0.0004	0	0.0034
Mer.adu02.02	0.6971	0.7625	0.6259	0	0.5104	0.0117	0.1846
Mer.adu02.03	0.7111	0.8034	0.7004	0	0.044	0.0237	0.004
Mer.adu02.04	0.6057	0.4946	0.6352	0	0.1161	0.0325	0.0847
Mer.adu02.06	0.142	0.1639	0.0356	0	0.0002	0.0004	0.0039
Mer.adu02.07	0.0038	0.0268	0.0022	0	0	0	0
Mer.adu02.08	0.1651	0.3474	0.0658	0	0.0338	0	0.0186
Mer.adu02.09	0.7554	0.7525	0.595	0	0.1766	0.116	0.1041
Mer.adu02.10	0.0428	0.1072	0.1084	0	0	0	0.0031
Mer.adu02.11	0.0136	0.1501	0.1412	0	0.0002	0.0037	0.003
Mer.adu02.12	0.467	0.4118	0.5152	0	0.0143	0.0031	0.0419
Mer.adu02.13	0.5063	0.3674	0.5423	0	0.0045	0.0089	0.0425
Mer.adu02.14	0.2132	0.2899	0.2801	0	0.0005	0	0.003
Mer.adu02.15	0.0698	0.2637	0.0793	0	0.0005	0.0034	0.0373
Mer.adu02.16	0.0129	0.0316	0.0104	0	0	0	0
Mer.adu02.17	0	0.0001	0.0005	0	0	0	0.0001
Mer.adu02.18	0.171	0.325	0.0374	0	0.0127	0	0
Mer.adu02.19	0.5474	0.4288	0.4276	0	0.3628	0.0044	0.0044
Mer.adu02.20	0.7813	0.8933	0.7484	0	0.115	0.0879	0.0377
Mer.adu02.21	0.0106	0.1381	0.1968	0	0.0002	0.0064	0.003
Mer.adu04.01	0.2513	0.337	0.3807	0	0.1025	0.0014	0.0205
Mer.adu06.01	0.0158	0.0427	0.0398	0	0.0789	0.0079	0.0004
Mer.adu06.02	0.849	0.8506	0.9641	0	0.4566	0.396	0.2801
Mer.adu06.04	0.3415	0.5073	0.2502	0	0.0116	0.0196	0.0011
Mer.adu06.05	0.3885	0.4013	0.6587	0	0.0049	0.1526	0.0105
Mer.adu06.06	0.2035	0.1604	0.0514	0	0.0071	0	0.0016
Mer.adu06.07	0.0667	0.0377	0.0763	0	0.0002	0	0.0013
Mer.adu06.08	0.2617	0.5733	0.2617	0	0.0632	0.0026	0.0004
Mer.adu07.01	0.0069	0.1149	0.3943	0	0.0018	0.0006	0.0079
Mer.adu07.02	0.0206	0.0561	0.232	0	0	0.266	0.0015
Mer.adu07.03	0.4422	0.5114	0.2146	0.0021	0.0241	0.1828	0.1296
Mer.adu07.04	0.0026	0.182	0.0026	0	0.0016	0	0
Mer.adu07.05	0.124	0.1772	0.3152	0	0.0659	0.201	0.0376
Mer.adu07.06	0.0143	0.0086	0.2467	0	0.0002	0.1707	0
Mer.adu07.07	0.0542	0.0248	0.1047	0	0	0.0008	0.0008
Mer.adu07.09	0.1764	0.1999	0.3659	0	0.0136	0.0302	0.2045
Mer.adu07.10	0.1911	0.4633	0.0902	0	0.0359	0.007	0.0534
Mer.adu07.11	0.1069	0.3412	0.1927	0	0.1242	0.0152	0.0474
Mer.adu07.12	0.3917	0.1612	0.138	0	0.0421	0.0082	0.0352
Mer.adu07.13	0.1107	0.353	0.0853	0	0.8142	0	0.0001
Mer.adu07.14	0.7976	0.7907	0.4721	0	0.6302	0.061	0.1644
Mer.adu07.15	0.1767	0.7139	0.3452	0	0.0451	0.0269	0.0289
Mer.adu07.16	0.0001	0.0029	0.0024	0	0	0	0
Mer.adu07.17	0.5125	0.4978	0.519	0	0.2407	0.0014	0.0555
Mer.adu07.18	0.1075	0.2754	0.1928	0	0.03	0.0023	0.0003
Mer.adu07.19	0.0017	0.0143	0.0206	0	0.0002	0	0.0003
Mer.adu07.20	0.0994	0.2925	0.0968	0	0.0255	0.0003	0.0038
Mer.adu07.21	0.4378	0.2746	0.196	0	0.0023	0.0726	0.0415
Mer.adu07.22	0.0909	0.3055	0.0633	0	0.0004	0	0.0006
Mer.adu07.23	0.1087	0.2501	0.4319	0	0.0032	0.0759	0.0147

Mer.adu07.24	0.1105	0.1124	0.0429	0	0.017	0.0004	0.015
Mer.adu07.25	0.0107	0.014	0.0156	0	0	0	0.0011
Mer.adu07.26	0.1222	0.171	0.0568	0	0.0002	0	0.0013
Mer.adu07.27	0.0403	0.0949	0.0608	0	0.0125	0.0006	0
Mer.adu07.28	0.0802	0.0634	0.0903	0	0.6918	0	0.0015
Mer.adu07.29	0.0109	0.0158	0.0001	0	0	0	0
Mer.adu07.30	0.002	0.0008	0.0053	0	0	0	0
Mer.adu07.31	0.4494	0.5741	0.0922	0	0.1705	0	0.0273
Mer.adu07.33	0.0426	0.0254	0.1313	0	0.0004	0	0
Mer.adu08.02	0.6597	0.4789	0.7149	0	0.0573	0.017	0.2705
Mer.adu08.03	0.7156	0.8567	0.6206	0	0.1521	0.2094	0.2911
Mer.adu08.04	0.7368	0.4634	0.2922	0	0.2045	0.0463	0.3525
Mer.adu08.05	0.0106	0.0244	0.0077	0	0	0	0
Mer.adu08.06	0.5041	0.3618	0.3034	0	0.0045	0.061	0.109
Mer.adu08.07	0.067	0.2189	0.0511	0	0.0042	0.0016	0.0007
Mer.adu08.08	0.1425	0.6787	0.2265	0	0.0001	0.0042	0.0164
Mer.adu08.09	0.0036	0.022	0.0054	0	0.0183	0	0.0004
Mer.adu08.10	0.0745	0.2076	0.0323	0	0	0	0.0101
Mer.adu08.11	0.0081	0.1431	0.002	0	0.0005	0	0.0004
Mer.adu08.12	0.002	0.0832	0.0045	0	0.0002	0	0.0592
Mer.adu08.13	0.1601	0.3176	0.1375	0	0.0048	0.0058	0
Mer.adu08.14	0.101	0.1306	0.2577	0	0.0002	0	0
Mer.adu08.15	0.0004	0	0.0014	0	0	0	0
Mer.adu08.16	0.0154	0.04	0.0056	0	0.0023	0	0.0008
Mer.adu08.17	0.3195	0.181	0.1502	0	0.0315	0.0025	0.0021
Mer.adu08.18	0.0038	0.248	0.1411	0	0	0	0.0012
Mer.adu08.19	0.0021	0.0163	0.0026	0	0.0002	0	0.0039
Mer.adu08.20	0.01	0.0003	0.0055	0	0	0	0
Mer.adu08.21	0.4534	0.1988	0.3655	0	0.0197	0.0289	0.1141
Mer.adu08.22	0.2243	0.2229	0.2751	0	0.0823	0.0041	0.0056
Mer.adu08.23	0.1124	0.4905	0.1841	0	0.209	0.0006	0.0034
Mer.adu08.24	0.0015	0.046	0.0055	0	0.0004	0	0
Mer.adu08.25	0.1749	0.3202	0.0785	0	0.1318	0	0.0817
Mer.adu08.26	0.6393	0.8369	0.6351	0	0.8119	0.2887	0.6681
Mer.adu08.27	0.0045	0.0192	0.1106	0	0	0	0.0218
Mer.adu08.28	0.1442	0.1552	0.0968	0	0.0009	0	0.0321
Mer.adu08.29	0.2582	0.2535	0.2702	0	0.0602	0.0008	0.0088
Mer.adu08.30	0.4908	0.4284	0.8907	0	0.0502	0.0083	0.1891
Mer.adu08.32	0.7074	0.822	0.5017	0	0.1874	0.0104	0.1841
Mer.adu08.33	0.156	0.1413	0.0898	0	0.0005	0.0029	0.0008
Mer.adu08.34	0.302	0.1908	0.0808	0	0.0024	0	0.0194
Mer.adu08.35	0.1516	0.1837	0.2028	0	0.0008	0	0.1988
Mer.adu08.36	0.0447	0.1666	0.0352	0	0.0162	0.0014	0.0035
Mer.adu08.37	0.0032	0.003	0.0015	0	0	0.0001	0.0019
Mer.adu08.38	0.0017	0.0017	0.0194	0	0	0	0
Mer.adu08.39	0.0055	0.0881	0.0131	0	0	0	0
Mer.adu08.40	0.5767	0.8374	0.5279	0	0.1047	0.011	0.0044
Mer.adu08.41	0.2168	0.3759	0.198	0	0.0046	0.0006	0.001
Mer.adu08.42	0.0796	0.1681	0.2407	0	0.0212	0.0008	0.0007
Mer.adu08.43	0.5028	0.4005	0.7018	0	0.1439	0.0516	0.338
Mer.adu09.02	0.0774	0.0878	0.1056	0	0	0.0004	0.0003
Mer.adu09.03	0.0137	0.1569	0.0981	0	0.003	0	0.0004
Mer.adu09.04	0.0687	0.1649	0.2769	0	0.043	0	0.0036
Mer.adu09.05	0.161	0.551	0.0914	0	0.0005	0	0.0003
Mer.adu10.01	0.1555	0.0859	0.5378	0	0.0006	0.0234	0.0326
Mer.adu10.02	0.0489	0.1393	0.05	0	0.0002	0.0014	0
Mer.adu10.04	0.2268	0.1871	0.1012	0	0.0014	0.0048	0.0098
Mer.adu10.05	0.0947	0.1687	0.0719	0	0.0005	0	0.0225
Mer.adu10.06	0.0311	0.0112	0.053	0	0.0008	0	0
Mer.adu10.07	0.5502	0.3548	0.1256	0	0.1554	0.0054	0.1093
Mer.adu10.08	0.1395	0.2073	0.0743	0	0.0028	0.0016	0.0002
Mer.adu10.11	0.3716	0.3712	0.2368	0	0.1411	0.0008	0.2603
Mer.adu10.12	0.1786	0.2139	0.1586	0	0.0589	0.0018	0.0004

Mer.adu10.13	0.7476	0.6138	0.4117	0	0.2898	0	0.0807
Mer.adu10.14	0.0463	0.0013	0.0075	0	0	0	0
Mer.adu10.15	0.2541	0.4966	0.3384	0	0.077	0.0227	0.0456
Mer.adu10.16	0.0558	0.0621	0.0443	0	0.0008	0	0.0007
Mer.adu10.17	0.1786	0.3843	0.0544	0	0.0014	0	0.0059
Mer.adu10.18	0	0	0.0007	0	0	0	0
Mer.adu10.19	0.8884	0.744	0.8964	0	0.7977	0.3919	0.1517
Mer.adu10.20	0.1792	0.0711	0.0417	0	0.0623	0.0003	0.0167
Mer.adu10.21	0.0067	0.0003	0.0199	0	0	0	0
Mer.adu10.22	0.1234	0.4045	0.4357	0	0.003	0	0.0004
Mer.adu10.23	0.8408	0.97	0.9386	0	0.6019	0.3748	0.4269
Mer.adu10.24	0.2428	0.3974	0.3778	0	0.0006	0.0005	0.0538
Mer.adu10.25	0.0124	0.0338	0.0295	0	0.0044	0	0
Mer.adu10.26	0.1028	0.064	0.0754	0	0	0.0012	0.0012
Mer.adu10.27	0.4946	0.3496	0.2194	0	0.0006	0.0992	0
Mer.adu10.28	0.5811	0.4686	0.6571	0	0.0287	0	0.35
Mer.adu10.30	0.7003	0.6646	0.2612	0	0.0341	0.0908	0.2431
Mer.juv06.02	0.0199	0.0977	0.084	0	0.0015	0.0024	0.0108

Appendix VI- Divergence dates between paired salmon groups. Divergence dates (mean, median and mode) calculated under each of three scenarios, as calculated by DIYABC. CA = common ancestor. q050 indicates the lower 95% confidence limit and q950 indicates the upper 95% confidence limit. Direct and logistic probabilities indicate the relative probability of each scenario.

| Scenario	Mean	Median	Mode	q050	q950	Direct probability	Logistic probability
chalk vs France MA	20600	9080	3300	1688	77600	0.2520	0.0888
chalk colonise France	**27120**	**17360**	**7840**	**3272**	**83600**	**0.5240**	**0.7927**
France colonise chalk	18240	7080	1708	1172	67600	0.2240	0.1185
chalk vs Spain MA	180800	94000	42800	18360	664000	0.1760	0.0265
chalk colonise spain	46400	27640	10120	5040	157600	0.3900	0.4756
Spain colonise chalk	**36480**	**21600**	**11160**	**3768**	**125200**	**0.4340**	**0.4979**
chalk vs southwest England MA	20920	8880	4440	1552	76800	0.0980	0.0926
chalk colonise southwest England	**34880**	**25600**	**15640**	**5960**	**92800**	**0.7720**	**0.7286**
southwest England colonise chalk	21760	10040	4240	1568	87600	0.1300	0.1788
Spain vs France MA	5840	2164	1472	344	19720	0.0640	0.0355
France colonise Spain	8840	4240	2328	724	28120	0.1880	0.1800
Spain colonise France	**17000**	**11400**	**4680**	**2412**	**47200**	**0.7480**	**0.7845**
southwest England vs France MA	10800	3520	892	564	39640	0.0880	0.0163
southwest England colonises France	**6440**	**3644**	**1128**	**644**	**17400**	**0.3260**	**0.6213**
France colonise southwest England	**8320**	**4520**	**1700**	**828**	**23720**	**0.5860**	**0.3624**
southwest England vs Spain MA	7080	2424	648	364	21800	0.0760	0.0477
Spain colonise southwest England	**30400**	**24240**	**14880**	**4880**	**76400**	**0.7880**	**0.7486**
southwest England colonises Spain	13760	7240	2680	1272	38480	0.1360	0.2036
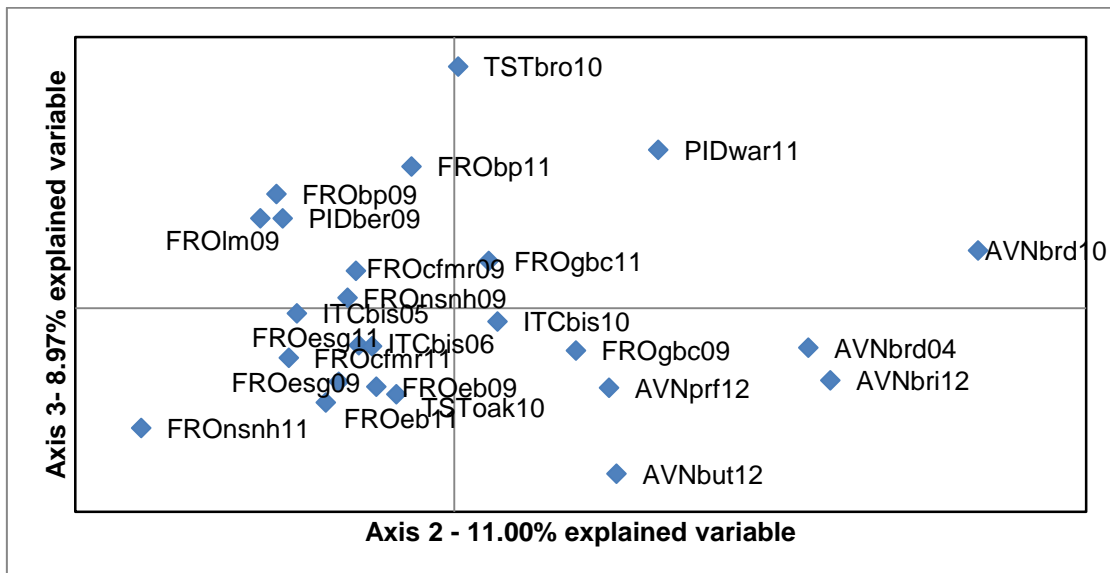
Appendix VII – Principal coordinate analysis axis 2 and 3 of contemporary chalk sample sites . Only sample sites with at least 20 individuals were included.

Appendix VIII- Pair-wise F_{ST} values between all chalk salmon sample sites. F_{ST} values are below the diagonal. Probability, P(rand >= data) based on 999 permutations are above the diagonal. Sample site IDs on top row and furthest right. Numbers in bold indicate non-significant pair-wise F_{ST} values.

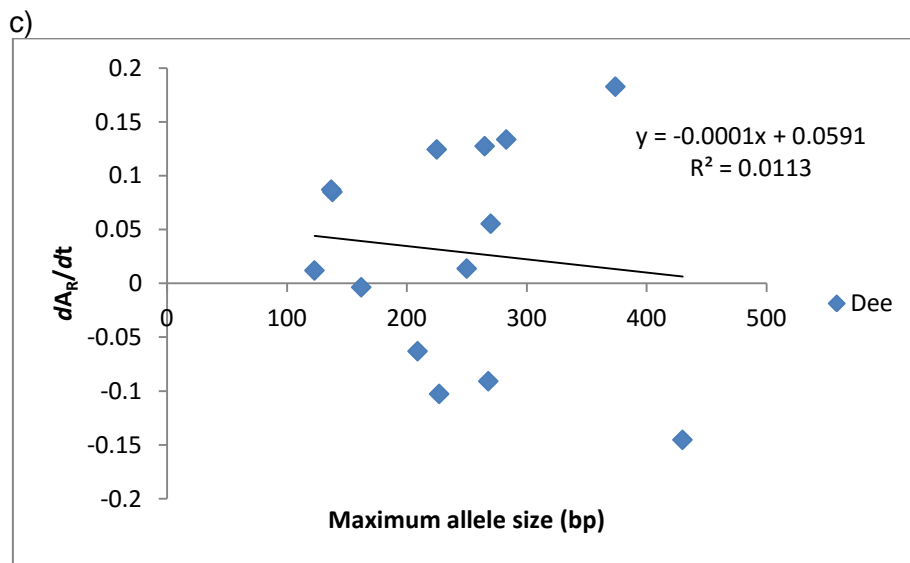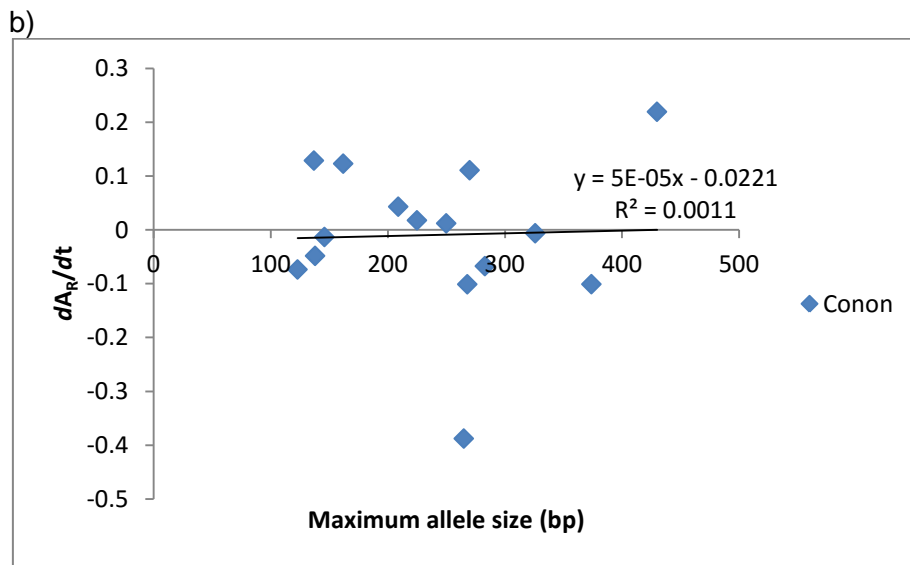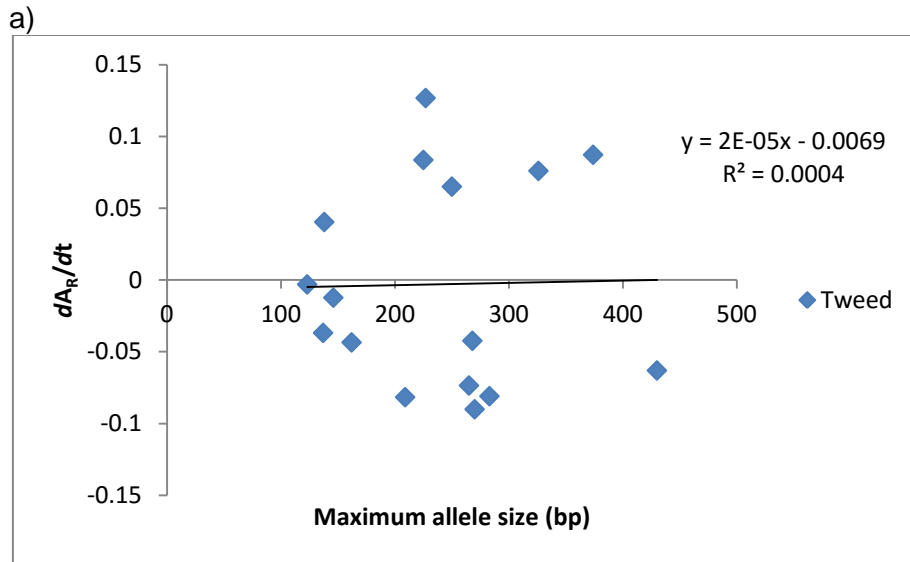| FRObp09	FROgbc09	FROnsnh0	FROlm09	FROcfmr0	FROeb09	FROesg09	FRObp11	FROgbc11	FROnsnh1	FROcfmr1	FROeb11	FROesg11	PIDber09	PIDwar11	AVNbrd04	AVNbrd10	AVNbri12	AVNbut12	AVNprf12	TSTbro10	TSToak10	ITCbis05	ITCbis06	ITCbis10	
	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FRObp09
0.020		0.012	0.001	0.001	0.007	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.003	0.006	0.001	0.004	0.002	0.001	0.001	0.001	0.001	0.001	0.001	FROgbc09
0.023	0.015		0.004	0.001	0.001	0.001	0.001	0.005	0.001	0.002	0.002	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FROnsnh0
0.018	0.019	0.018		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FROlm09
0.021	0.018	0.018	0.022		0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FROcfmr09
0.019	0.013	0.019	0.020	0.015		0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001	FROeb09
0.023	0.020	0.016	0.019	0.020	0.013		0.001	0.001	0.001	0.011	0.001	0.001	0.002	0.001	0.003	0.001	0.001	0.003	0.002	0.001	0.015	0.001	0.030	0.001	FROesg09
0.027	0.027	0.027	0.029	0.021	0.026	0.028		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FRObp11
0.016	0.013	0.014	0.017	0.014	0.015	0.017	0.016		0.001	0.004	0.001	0.001	0.001	0.001	0.006	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	FROgbc11
0.027	0.024	0.022	0.023	0.021	0.020	0.022	0.030	0.019		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FROnsnh1
0.017	0.014	0.014	0.015	0.011	0.011	0.010	0.021	0.010	0.016		0.012	0.002	0.001	0.001	0.001	0.001	0.001	0.004	0.001	0.001	0.001	0.001	0.001	0.001	FROcfmr1
0.027	0.016	0.017	0.020	0.016	0.016	0.014	0.024	0.015	0.019	0.010		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	FROeb11
0.022	0.015	0.014	0.019	0.017	0.012	0.014	0.029	0.013	0.021	0.011	0.015		0.013	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.036	0.001	0.024	0.001	FROesg11
0.024	0.024	0.023	0.019	0.021	0.021	0.019	0.027	0.018	0.023	0.017	0.021	0.014		0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.003	0.001	PIDber09
0.029	0.021	0.022	0.023	0.020	0.021	0.023	0.031	0.018	0.036	0.022	0.031	0.022	0.021		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	PIDwar11
0.027	0.017	0.024	0.025	0.028	0.020	0.019	0.033	0.017	0.032	0.019	0.021	0.018	0.028	0.024		0.002	**0.054**	0.003	0.025	0.001	0.007	0.001	0.004	0.001	AVNbrd04
0.037	0.025	0.031	0.034	0.030	0.033	0.030	0.034	0.022	0.037	0.029	0.030	0.031	0.034	0.031	0.024		0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	AVNbrd10
0.031	0.018	0.024	0.028	0.025	0.021	0.024	0.030	0.017	0.028	0.020	0.022	0.018	0.026	0.025	**0.016**	0.024		0.010	0.021	0.001	0.003	0.001	0.001	0.001	AVNbri12
0.026	0.018	0.025	0.027	0.020	0.016	0.018	0.027	0.018	0.025	0.014	0.021	0.023	0.026	0.027	0.023	0.030	0.018		0.005	0.001	0.001	0.001	0.001	0.001	AVNbut12
0.026	0.017	0.023	0.022	0.025	0.021	0.018	0.032	0.018	0.024	0.016	0.022	0.018	0.021	0.025	0.016	0.031	0.016	0.018		0.001	0.001	0.001	0.001	0.001	AVNprf12
0.030	0.031	0.030	0.027	0.032	0.031	0.026	0.035	0.025	0.039	0.025	0.028	0.023	0.025	0.030	0.026	0.037	0.029	0.041	0.028		0.001	0.001	0.002	0.001	TSTbro10
0.029	0.022	0.022	0.025	0.027	0.019	0.015	0.033	0.019	0.027	0.016	0.019	0.013	0.021	0.031	0.019	0.033	0.022	0.026	0.024	0.027		0.015	0.040	0.006	TSToak10
0.031	0.027	0.029	0.025	0.035	0.028	0.022	0.036	0.021	0.027	0.021	0.026	0.023	0.024	0.036	0.030	0.033	0.030	0.034	0.028	0.027	0.017		**0.109**	0.005	ITCbis05
0.030	0.023	0.020	0.024	0.027	0.022	0.014	0.035	0.018	0.025	0.016	0.019	0.013	0.021	0.027	0.021	0.032	0.024	0.028	0.020	0.022	0.016	**0.014**		0.016	ITCbis06
0.024	0.020	0.025	0.026	0.023	0.019	0.019	0.027	0.013	0.028	0.016	0.018	0.015	0.022	0.027	0.020	0.026	0.021	0.023	0.024	0.026	0.016	0.014	0.014		ITCbis10
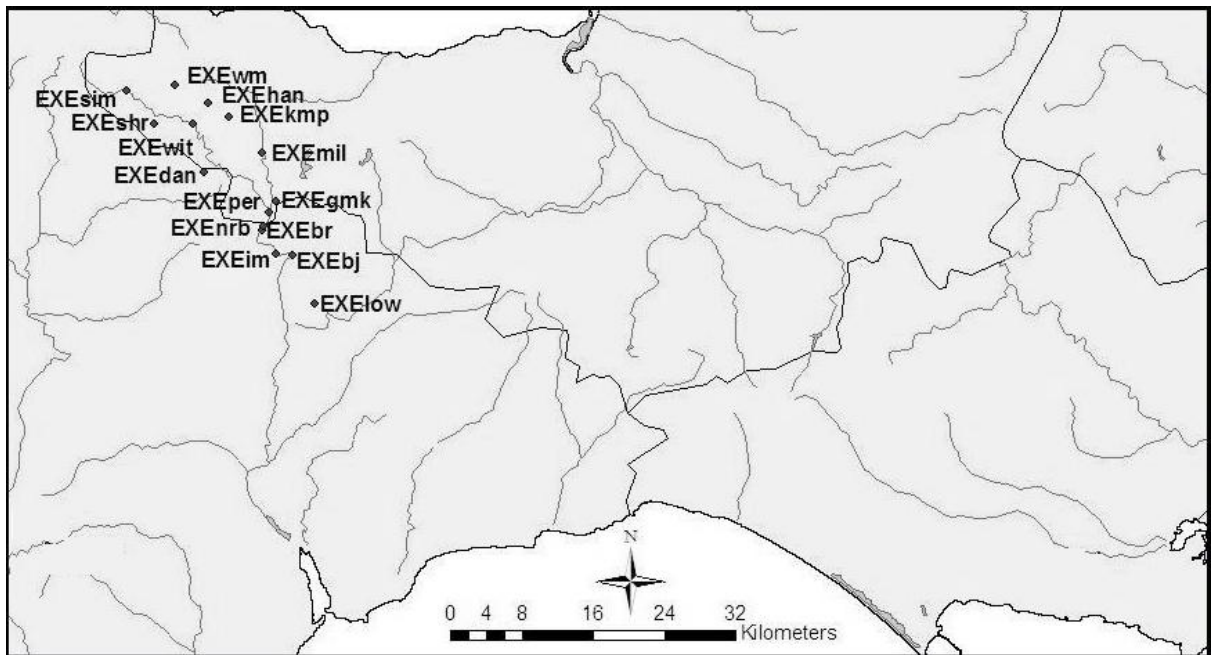
Appendix IX- Map of juvenile salmon sampling sites on the river Exe. In 2004 only the EXEdan, EXEsim and EXEshr were sampled while in 2009 samples were collected from all of the sampling sites. Edited from Counter thesis 2012.
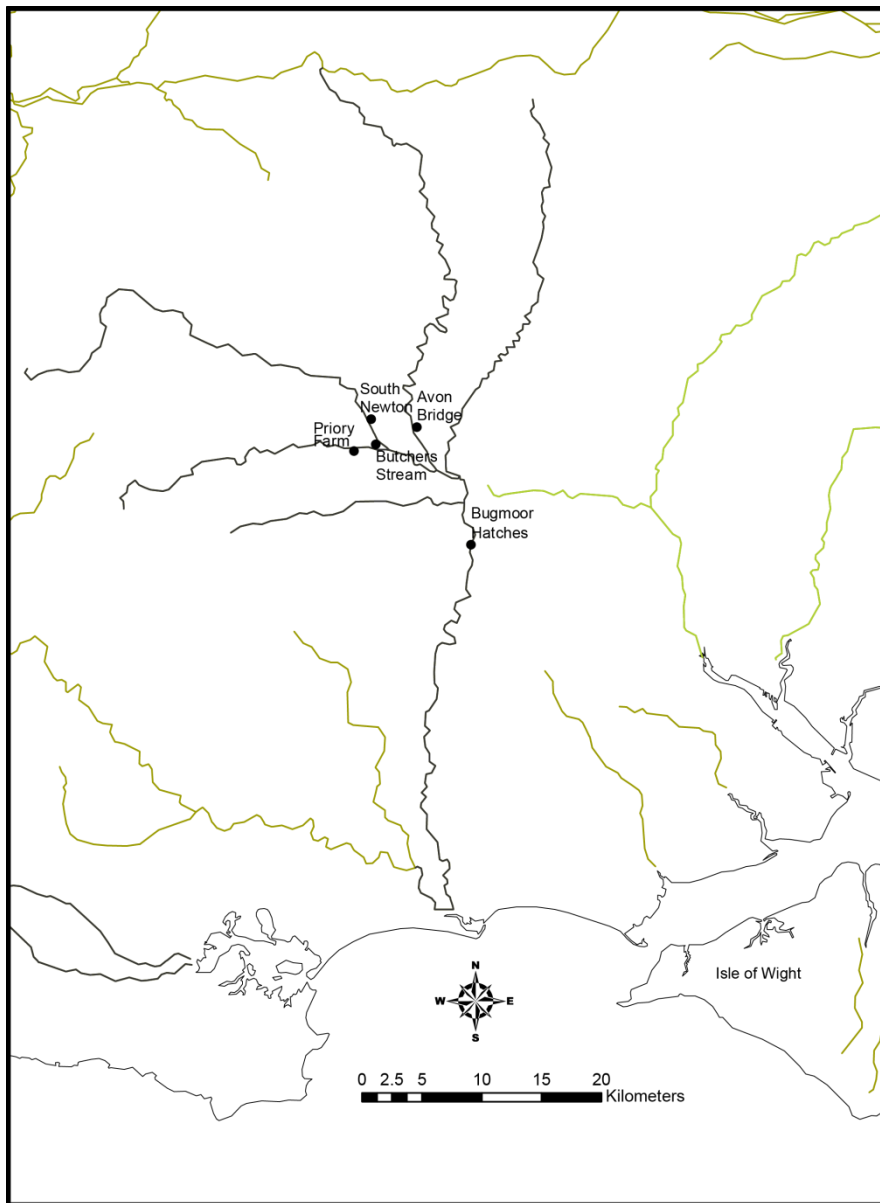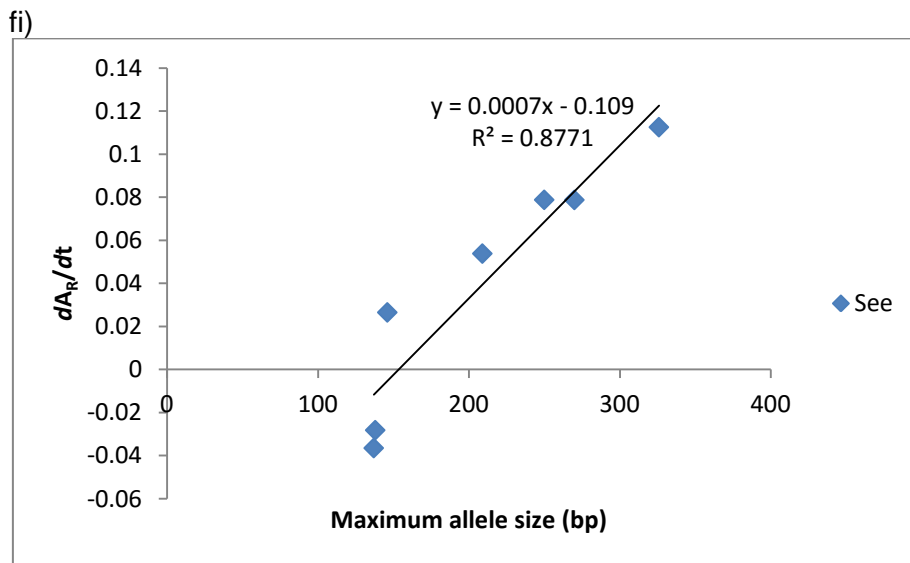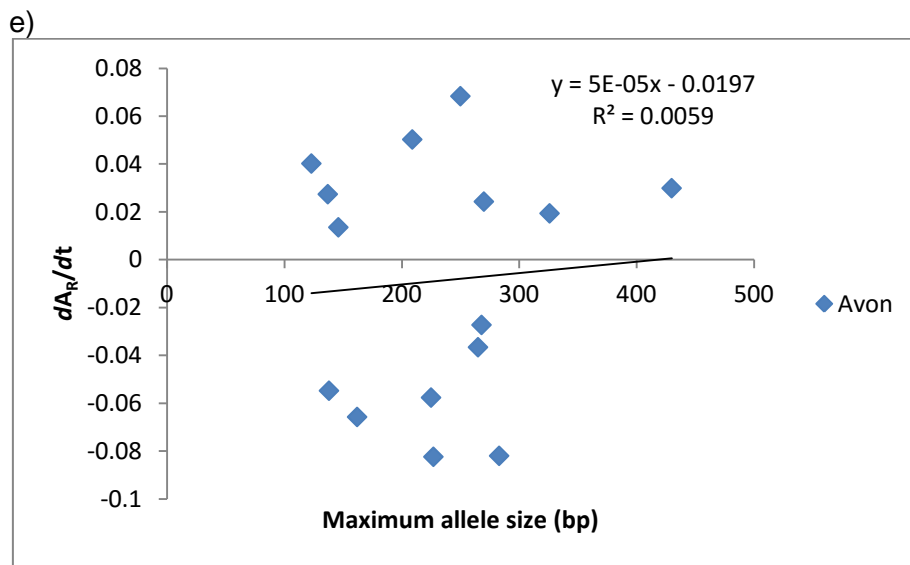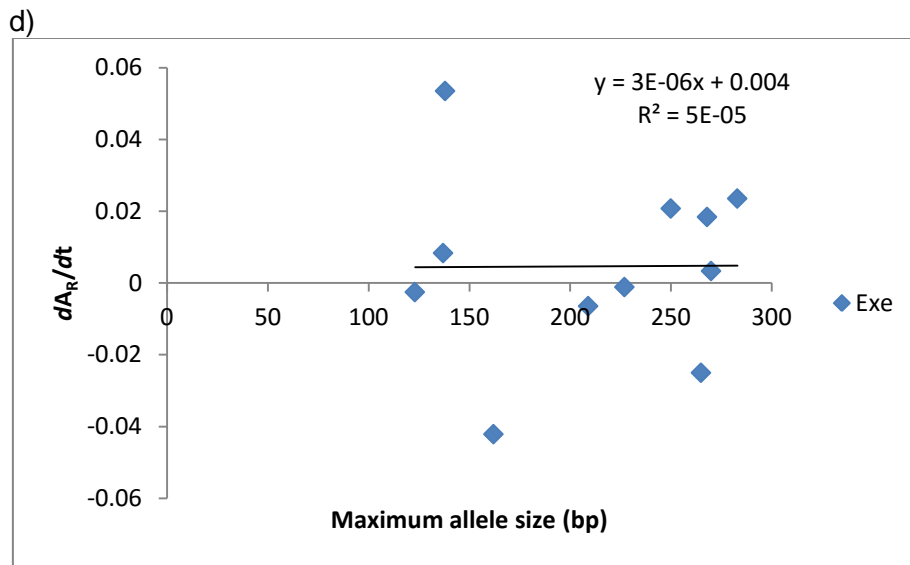
Appendix X- Map of sampling sites on the river Avon. Bugmoor Hatches was sampled in 2004 only. Avon Bridge was sampled in 2004 and 2010. Priory Farm, South Newton, and Butchers Stream were sampled in 2010 only.
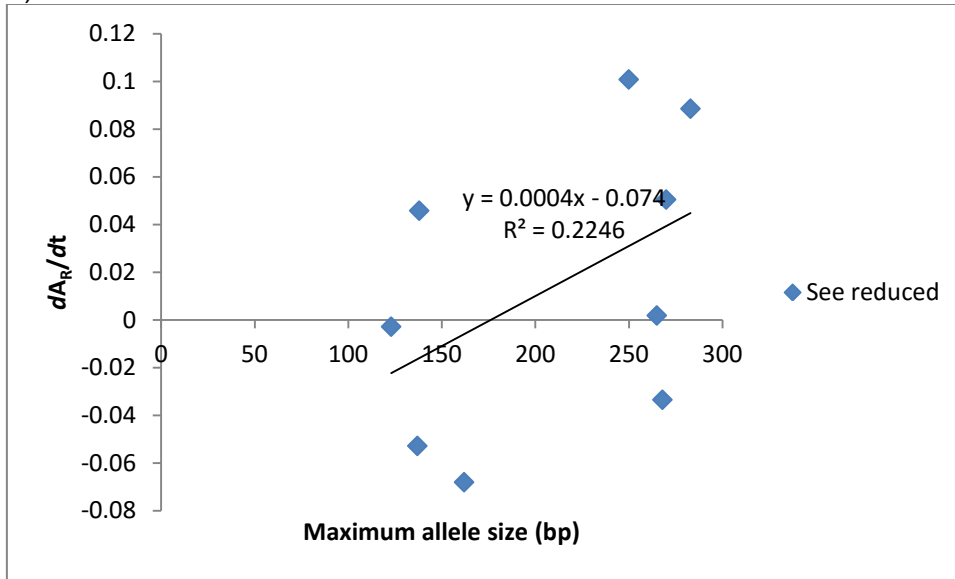
Appendix XI- Reported salmon catch by rod-and-line anglers on rivers sampled within this study.

Year	River							
	Tweed	Conon	Dee	Exe	Avon	See	Elle	Scorff
1966	-	-	-	2055	-	-	-	-
1967	-	-	-	1405	-	-	-	-
1968	-	-	-	978	-	-	-	-
1969	-	-	-	475	-	-	-	-
1970	-	-	-	312	-	-	-	-
1971	-	-	-	294	-	-	-	-
1972	-	-	-	383	-	-	-	-
1973	-	-	-	228	-	-	-	-
1974	-	-	-	234	-	-	-	-
1975	-	-	-	109	-	-	-	-
1976	-	-	-	48	-	-	-	-
1977	-	-	-	233	-	-	-	-
1978	-	-	-	227	-	-	-	-
1979	-	-	-	399	-	-	-	-
1980	-	-	-	423	-	250	140	145
1981	-	-	-	437	-	290	600	300
1982	-	-	-	252	-	310	145	70
1983	-	-	-	341	-	235	210	85
1984	-	-	-	292	-	300	320	150
1985	-	-	-	707	-	135	200	140
1986	-	-	-	672	-	430	390	150
1987	-	-	-	824	-	160	250	110
1988	-	-	-	1030	-	170	460	90
1989	-	-	-	331	441	242	163	35
1990	-	1304	-	321	295	280	123	40
1991	-	873	-	260	171	140	166	35
1992	9549	2822	721	762	321	270	50	35
1993	9245	1602	1134	642	137	250	120	25
1994	9919	2698	860	1466	89	400	199	56
1995	9607	3119	548	453	75	200	110	86
1996	9347	1919	456	396	128	212	172	96
1997	9518	1134	472	702	31	119	86	43
1998	9190	2105	639	676	52	80	101	100
1999	9262	757	429	477	40	220	121	33
2000	9061	652	395	649	14	600	202	50
2001	8958	1618	678	187	43	500	125	39
2002	10320	1121	542	210	104	600	63	27
2003	13886	1169	444	146	109	66	46	7
2004	15257	1475	1188	594	125	140	400	100
2005	13488	1455	893	269	132	88	152	57
2006	14034	1579	683	187	147	-	-	-
2007	16042	1289	751	340	85	-	-	-
2008	13502	1162	974	540	67	-	-	-
2009	10324	1180	566	355	74	-	-	-
2010	22824	1620	880	404	46	-	-	-
2011	16410	1275	856	601	138	-	-	-
2012	12876	1466	744	438	-	-	-	-

a)

y = 2E-05x - 0.0069
R² = 0.0004

dA_R/dt

Maximum allele size (bp)

♦ Tweed

b)

y = 5E-05x - 0.0221
R² = 0.0011

dA_R/dt

Maximum allele size (bp)

♦ Conon

c)

y = -0.0001x + 0.0591
R² = 0.0113

dA_R/dt

Maximum allele size (bp)

♦ Dee

d)

$y = 3E\text{-}06x + 0.004$
$R^2 = 5E\text{-}05$

Maximum allele size (bp)

dA_R/dt

◆ Exe

e)

$y = 5E\text{-}05x - 0.0197$
$R^2 = 0.0059$

Maximum allele size (bp)

dA_R/dt

◆ Avon

fi)

$y = 0.0007x - 0.109$
$R^2 = 0.8771$

Maximum allele size (bp)

dA_R/dt

◆ See

fii)

g)
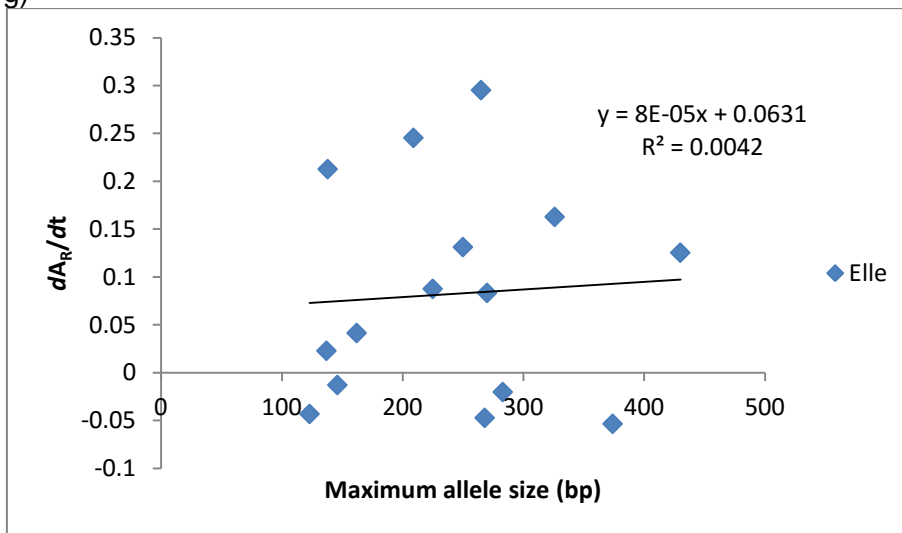
h)
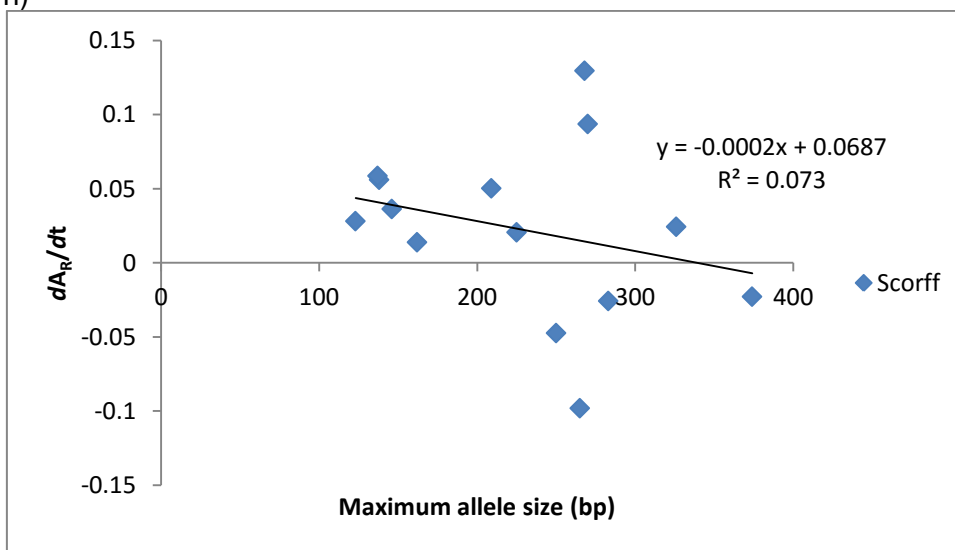
Appendix XII- Graphs of rate of change in allelic richness (dA_R/dt) against maximum allele size.

Appendix XIII- Pair-wise F_{ST} values between temporally sampled salmon within each temporal dataset. Numbers below the clear diagonal indicate the pair-wise F_{ST} values, while numbers above the diagonal indicate the significance determined from 999 permutations. Values in bold indicate a significant difference ($p < 0.05$) between temporal samples

5.1a)

Tweed92	Tweed00	Tweed04	Tweed12	
	0.941	0.320	0.396	Tweed92
0.005		0.777	0.519	Tweed00
0.006	0.005		0.982	Tweed04
0.006	0.006	0.004		Tweed12

5.1b)

Conon98	Conon07	Conon12	
	0.006	0.309	Conon98
0.009		**0.006**	Conon07
0.006	**0.007**		Conon12

5.1c)

Dee91	Dee95	Dee99	Dee03	Dee07	Dee11	
	0.564	0.283	0.141	0.510	0.486	Dee91
0.006		0.299	0.194	0.154	0.050	Dee95
0.006	0.007		0.264	0.669	0.008	Dee99
0.006	0.007	0.006		0.640	0.086	Dee03
0.006	0.007	0.005	0.005		0.047	Dee07
0.006	0.007	0.007	0.006	0.007		Dee11

5.1d)

Exe72	Exe04	Exe09	
	0.004	0.034	Exe72
0.009		**0.001**	Exe04
0.009	**0.009**		Exe09

5.1e)

Avon89	Avon04	Avn12	
	0.006	**0.004**	Avon89
0.013		**0.001**	Avon04
0.014	**0.017**		Avon12

5.1f)

Sée80-81	Sée88	Sée94	Sée00	Sée05	
	0.209	0.146	**0.003**	**0.001**	Sée80-81
0.007		0.174	0.150	**0.004**	Sée88
0.008	0.009		0.132	0.027	Sée94
0.011	0.009	0.009		0.074	Sée00
0.012	**0.011**	0.010	0.008		Sée05

5.1g)

Ellé88	Ellé94	Ellé00	Ellé05	
	0.438	0.100	0.817	Ellé88
0.008		0.113	0.323	Ellé94
0.009	0.008		0.508	Ellé00
0.006	0.007	0.006		Ellé05

5.1h)

Scorff88	Scorff94	Scorff00	Scorff05	
	0.119	0.433	0.573	Scorff88
0.013		0.269	0.127	Scorff94
0.012	0.007		0.195	Scorff00
0.011	0.007	0.007		Scorff05

Appendix XIV- Effective population size estimates for all samples, as calculated by the heterozygote excess (He excess), linkage disequilibrium (LD), molecular co-ancestry (MC), sib-ship (Sib), and temporal (Temp) methods. Lower and upper 95% confidence values are also given (L95 and U95 respectively)

River	Year	He excess N_E	LD N_E	LD L95	LD U95	MC N_E	MC L95	MC U95	Sib N_E	Sib L95	Sib U95	Temp dates (years)	Temp N_E	Temp L95	Temp U95
Tweed	1992	∞	289.2	144.4	4219	22.5	12.1	26	88	59	136	1992-2000	∞	134.4	∞
Tweed	2000	∞	761	216.7	∞	47.5	1.2	175.2	88	59	133	2000-2004	∞	∞	∞
Tweed	2004	∞	∞	403.7	∞	76.9	12.8	197.3	87	59	135	2004-2012	∞	∞	∞
Tweed	2012	∞	∞	398.3	∞	∞	∞	∞	101	69	160				
Conon	1998	∞	∞	516	∞	32.1	5.3	82.5	79	51	129	1998-2007	89.6	66.2	116.5
Conon	2007	∞	872.7	244.4	∞	38.5	9.3	88	91	62	137	2007-2012	54.2	39.8	70.8
Conon	2012	∞	243	134.7	939.3	∞	∞	∞	83	57	123				
Dee	1991	∞	∞	382.7	∞	86.1	0.1	432.1	109	75	169	1991-1995	∞	∞	∞
Dee	1995	∞	2697.7	257.8	∞	131.5	3.3	485	89	60	139	1995-1999	142.4	105.9	184.3
Dee	1999	∞	1929	272.8	∞	919.3	0.9	4615.1	109	74	163	1999-2003	82.3	61.2	106.5
Dee	2003	∞	1923.2	278.9	∞	∞	∞	∞	100	69	150	2003-2007	380.8	283.1	492.8
Dee	2007	∞	1195	293.7	∞	275.9	0.3	1384.9	86	59	133	2007-2011	112	85.1	142.4
Dee	2011	∞	∞	477.7	∞	∞	∞	∞	107	72	169				
Exe	1972	∞	129.7	61	32446	∞	∞	∞	50	32	82	1972-2004	282.7	198.5	381.5
Exe	2004	∞	355.6	189.22	1769.6	57.8	6.9	161.1	47	30	78	2004-2009	29.1	20.5	39.1
Exe	2009	∞	∞	396.1	∞	∞	∞	∞	57	37	94				
Avon	1989	∞	37.4	24.3	70.7	12.2	3.7	25.8	34	20	66	1989-2004	117.6	82.8	158.3
Avon	2004	∞	164.8	88.1	746.6	∞	∞	∞	51	30	102	2004-2012	22.1	15.3	74
Avon	2012	∞	35.9	29.1	45.3	21.7	5.9	47.5	36	21	69				
See	1980-81	∞	213	79	∞	20.4	0.5	75.3	62	42	96	1980-1988	151.5	95	220.9
See	1988	∞	∞	156.3	∞	∞	∞	∞	33	20	55	1988-1994	106.9	67.1	156
See	1994	∞	4841	112.7	∞	5.9	2.5	10.8	33	20	56	1994-2000	83.1	52.1	121.2
See	2000	∞	∞	163.4	∞	∞	∞	∞	30	18	52	2000-2005	15.4	9.8	22.3
See	2005	∞	∞	576	∞	18.5	4.5	42.3	45	30	70				
Elle	1988	∞	82.2	53.5	187	19.3	3.2	49.4	48	30	83	1988-1994	388.7	279	516.3
Elle	1994	∞	141.4	79.4	480.7	∞	∞	∞	70	44	110	1994-2000	159.9	114.5	212.8
Elle	2000	∞	45961.7	219.2	∞	∞	∞	∞	62	42	100	2000-2005	282.2	203	374.3
Elle	2005	∞	676.6	203	∞	∞	∞	∞	66	45	102				
Scorff	1988-89	∞	∞	88.7	∞	∞	∞	∞	80	38	555	1988-1994	95.9	67	129.9
Scorff	1994	∞	199.9	103.5	1329.8	25.9	9	51.5	61	51	98	1994-2000	345.2	241.8	466.7
Scorff	2000	∞	68.8	48.9	109.2	25.3	3	70.4	59	37	96	2000-2005	295.3	208	397.6
Scorff	2005	∞	208.2	112.1	953.5	∞	∞	∞	75	52	114				

Appendix XV- Effective population size estimates from repeated resampling of salmon from the river Exe and Avon using i) The temporal method and ii) the sib-ship method

i)

Repeat		Original			1			2			3			4		
River	Years	NE	L95	U95	NE	L95	U95	NE	L95	U95	NE	L95	U95	NE	L95	U95
Exe	1972-2004	282.7	198.5	381.5	394	276.7	531.8	241	167.9	327.2	336.6	233.8	457.8	198.4	138.2	269.4
	2004-2009	29.1	20.5	39.1	52	35.9	70.2	25.2	17.8	33.8	40.3	28.4	54.3	27.1	19.2	36.3
Avon	1989-2004	117.6	82.8	158.3	117.6	82.8	158.3	117.6	82.8	158.3	117.6	82.8	158.3	117.6	82.8	158.3
	2004-2012	22.1	15.3	74	2.7	1.9	3.5	2.7	1.9	3.5	2.6	1.9	3.5	2.7	1.9	3.5

ii)

Repeat		Original			1			2			3			4		
River	Years	NE	L95	U95	NE	L95	U95	NE	L95	U95	NE	L95	U95	NE	L95	U95
Exe	2004	47	30	78	61	41	95	64	43	97	60	41	92	69	48	105
	2009	57	37	94	60	41	91	65	46	100	62	41	95	67	46	103
Avon	2012	36	21	69	29	18	50	40	26	64	30	18	51	28	17	49

References

Aarestrup K, Koed A (2003) Survival of migrating sea trout (*Salmo trutta*) and Atlantic salmon (*Salmo salar*) smolts negotiating weirs in small Danish rivers. *Ecology of Freshwater Fish*, **12**, 169–176.

Aguilar A, Jessup DA, Estes J, Garza JC (2008) The distribution of nuclear genetic variation and historical demography of sea otters. *Animal Conservation*, **11**, 35–45.

Alerstam T, Gudmundsson GA, Green M, Hedenstro A (2001) Migration Along Orthodromic Sun Compass Routes by Arctic Birds. *Science*, **291**, 300–304.

Allendorf FW, Hohenlohe P a, Luikart G (2010) Genomics and the future of conservation genetics. *Nature reviews. Genetics*, **11**, 697–709.

Allendorf F, Phelps S (1981) Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences*, **38**, 1507–1514.

Allendorf F, Seeb L (2000) Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mitochondrial DNA markers. *Evolution*, **54**, 640–651.

Anderson JH, Quinn TP (2007) Movements of adult coho salmon (*Oncorhynchus kisutch*) during colonization of newly accessible habitat. *Canadian Journal of Fisheries and Aquatic Sciences*, **64**, 1143–1154.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC bioinformatics*, **9**, 323.

Ardren W, Kapuscinski A (2003) Demographic and genetic estimates of effective population size (Ne) reveals genetic compensation in steelhead trout. *Molecular Ecology*, **12**, 35–49.

Ayllon F, Martinez J, Garciavazquez E (2006) Loss of regional population structure in Atlantic salmon, *Salmo salar* L., following stocking. *ICES Journal of Marine Science*, **63**, 1269–1273.

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**, 485–97.

Beacham TD, Candy JR, Jonsen KL *et al.* (2006) Estimation of Stock Composition and Individual Identification of Chinook Salmon across the Pacific Rim by Use of Microsatellite Variation. *Transactions of the American Fisheries Society*, **135**, 861–888.

Beacham TD, Candy JR, Supernault KJ *et al.* (2001) Evaluation and Application of Microsatellite and Major Histocompatibility Complex Variation for Stock Identification of Coho Salmon in British Columbia. *Transactions of the American Fisheries Society*, **130**, 1116–1149.

Beacham TD, McIntosh B, Wallace CG (2011) A comparison of polymorphism of genetic markers and population sample sizes required for mixed-stock analysis of

sockeye salmon (Oncorhynchus nerka) in British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, **68**, 550–562.

Beaumont MA (2008) Joint Determination of Topology, Divergence Time and Immigration in Population Trees. In: *Simulation, Genetics and Human Prehistory*, pp. 135–154. Cambridge.

Beebee TJC (2009) A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: genetic compensation and population size-genetic diversity correlations. *Molecular ecology*, **18**, 4790–7.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4563–4568.

Bellinger MR, Johnson J a., Toepfer J, Dunn P (2003) Loss of Genetic Variation in Greater Prairie Chickens Following a Population Bottleneck in Wisconsin, U.S.A. *Conservation Biology*, **17**, 717–724.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*, **57**, 289–300.

Berrie A. (1992) The chalk-stream environment. *Hydrobiologia*, **248**, 3–9.

Binkley D, Brown T (1993) Forest practices as nonpoint sources of pollution in north america. *Water Resources Bulletin*, **29**, 729–740.

Boeuf G, Marc A, Prunet P, Bail P Le, Smal J (1994) Stimulation of parr-smolt transformation by hormonal treatment in Atlantic salmon (*Salmo salar* L.). *Aquaculture*, **121**, 195–208.

Borrell YJ, Bernardo D, Blanco G, Vázquez E, Sánchez JA (2007) Spatial and temporal variation of genetic diversity and estimation of effective population sizes in Atlantic salmon (*Salmo salar*, L.) populations from Asturias (Northern Spain) using microsatellites. *Conservation Genetics*, **9**, 807–819.

Bracegirdle C (1973) *The dark river* (John Sherratt and Sons Ltd, Ed,). Institute of Technology for the Citizen, Altrincham.

Bradbury IR, Hamilton LC, Robertson MJ *et al.* (2014) Landscape structure and climatic variation determine Atlantic salmon genetic connectivity in the Northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, **258**, 246–258.

Brook BW, Sodhi NS, Bradshaw CJ a (2008) Synergies among extinction drivers under global change. *Trends in ecology & evolution*, **23**, 453–60.

Broquet T, Angelone S, Jaquiery J *et al.* (2010) Genetic bottlenecks driven by population disconnection. *Conservation biology : the journal of the Society for Conservation Biology*, **24**, 1596–605.

Brumfield RT, Beerli P, Nickerson D a., Edwards S V. (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution*, **18**, 249–256.

Buetow KH, Edmonson MN, Cassidy AB (1999) Reliable identification of large numbers of candidate SNPs from public EST data. *Nature genetics*, **21**, 323–325.

Burton L (2003) The Mersey Basin: an historical assessment of water quality from an anecdotal perspective. *The Science of The Total Environment*, **314-316**, 53–66.

Bystriansky JS, Schulte PM (2011) Changes in gill H+-ATPase and Na+/K+-ATPase expression and activity during freshwater acclimation of Atlantic salmon (*Salmo salar*). *The Journal of experimental biology*, **214**, 2435–42.

Campagna L, St Clair JJH, Lougheed SC *et al.* (2012) Divergence between passerine populations from the Malvinas - Falkland Islands and their continental counterparts: a comparative phylogeographical study. *Biological Journal of the Linnean Society*, **106**, 865–879.

Campos JL, Posada D, Morán P (2007) Introgression and genetic structure in northern Spanish Atlantic salmon (*Salmo salar* L.) populations according to mtDNA data. *Conservation Genetics*, **9**, 157–169.

Carlsson J (2008) Effects of microsatellite null alleles on assignment testing. *The Journal of heredity*, **99**, 616–23.

Cavalli-Sforza LL, Edwards a W (1967) Phylogenetic analysis. Models and estimation procedures. *American journal of human genetics*, **19**, 233–57.

Cefas, Environment Agency (2013) *Annual assessment of salmon stocks and fisheries in England and Wales 2013*.

Chadwick JW, Canton SP (1986) Recovery of benthic invertebrate communities in silver bow creek, Montana, following improved metal mine wastewater treatment. *Water, Air, & Soil Pollution*, **28**, 427–438.

Chaput G, Legault C, Reddin D, Caron F, Amiro P (2005) Provision of catch advice taking account of non-stationarity in productivity of Atlantic salmon (*Salmo salar* L.) in the Northwest Atlantic. *ICES Journal of Marine Science*, **62**, 131–143.

Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA (2010) The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*, **186**, 983–95.

Child A, Burnell A, Wilkins N (1976) The existence of two races of Atlantic salmon (*Salmo salar* L.) in the British Isles. *Journal of Fish Biology*, **8**, 35–43.

Clark P, Dyke A, Shakun J *et al.* (2009) The last glacial maximum. *Science*, **325**, 710–715.

Collins JP (2010) Amphibian decline and extinction: what we know and what we need to learn. *Diseases of aquatic organisms*, **92**, 93–9.

Consuegra S, Garcia De Leaniz C, Serdio a. *et al.* (2002) Mitochondrial DNA variation in Pleistocene and modern Atlantic salmon from the Iberian glacial refugium. *Molecular Ecology*, **11**, 2037–2048.

Consuegra S, Verspoor E, Knox D, García de Leániz C (2005) Asymmetric gene flow and the evolutionary maintenance of genetic diversity in small, peripheral Atlantic salmon populations. *Conservation Genetics*, **6**, 823–842.

Corander J, Waldmann P, Sillanpa MJ (2003) Bayesian Analysis of Genetic Differentiation Between Populations. *Differentiation*, **374**, 367–374.

Cornuet J, Luikart G (1996) Description and power analysis of two tests for detecting

recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cornuet JM, Piry S, Luikart G, Estoup a, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Cornuet J-M, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics (Oxford, England)*, **24**, 2713–9.

Crawford A (2010) *Fifth otter survey of England 2009 – 2010. Technical Report*. Bristol.

Crozier WW (1993) Evidence of genetic interaction between escaped farmed salmon and wild Atlantic salmon (*Salmo salar* L.) in a Northern Irish river. *Aquaculture*, **113**, 19–29.

Crozier W (2001) Relationship between freshwater angling catch of Atlantic salmon and stock size in the River Bush, Northern Ireland. *Journal of Fish Biology*, **58**, 240–247.

Dalziel AC, Bittman J, Mandic M, Ou M, Schulte PM (2014) Origins and functional diversification of salinity-responsive Na(+) , K(+) ATPase α1 paralogs in salmonids. *Molecular ecology*, **23**, 3483–503.

Dillane E, Cross MC, Mcginnity P *et al.* (2007) Spatial and temporal patterns in microsatellite DNA variation of wild Atlantic salmon, *Salmo salar*, in Irish rivers. *Fisheries Management and Ecology*, **14**, 209–219.

Dionne M, Caron F, Dodson JJ, Bernatchez L (2008a) Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Molecular ecology*, **17**, 2382–96.

Dionne M, Caron F, Dodson JJ, Bernatchez L (2008b) Comparative survey of within-river genetic structure in Atlantic salmon; relevance for management and conservation. *Conservation Genetics*, **10**, 869–879.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular ecology resources*, **14**, 209–14.

Doiron S, Bernatchez L, Blier PU (2002) A Comparative Mitogenomic Analysis of the Potential Adaptive Value of Arctic Charr mtDNA Introgression in Brook Charr Populations (*Salvelinus fontinalis* Mitchill). *Molecular biology and evolution*, **19**, 1902–1909.

Drenner SM, Clark TD, Whitney CK *et al.* (2012) A synthesis of tagging studies examining the behaviour and survival of anadromous salmonids in marine environments. *PloS one*, **7**, e31311.

Dunkley D, Shearer W (1982) An assessment of the performance of a resistivity fish counter. *Journal of fish biology*, **20**, 717–737.

Eatherley D, Thorley J, Stephen A *et al.* (2005) *Trends in Atlantic salmon: the role of automatic fish counter data*.

Ellegren H (2000) Microsatellite mutations in the germline:: implications for evolutionary inference. *Trends in genetics*, **16**, 551–558.

Ellis JS, Gilbey J, Armstrong a et al. (2011) Microsatellite standardization and evaluation of genotyping error in a large multi-partner research programme for conservation of Atlantic salmon (Salmo salar L.). *Genetica*, **139**, 353–67.

Ellis J, Sumner K, Griffiths A, Bright D, Stevens J (2010) Population genetic structure of Atlantic salmon, *Salmo salar* L., in the River Tamar, southwest England. *Fisheries Management and Ecology*, **18**, 233–245.

Ensing D, Crozier WW, Boylan P, O'Maoiléidigh N, McGinnity P (2013) An analysis of genetic stock identification on a small geographical scale using microsatellite markers, and its application in the management of a mixed-stock fishery for Atlantic salmon Salmo salar in Ireland. *Journal of fish biology*, **82**, 2080–94.

Ensing D, Prodöhl P a, McGinnity P et al. (2011) Complex pattern of genetic structuring in the Atlantic salmon (*Salmo salar* L.) of the River Foyle system in northwest Ireland: disentangling the evolutionary signal from population stochasticity. *Ecology and evolution*, **1**, 359–72.

Environment Agency (2004) *The State of England's Chalk Rivers*.

Environment Agency (2012) *Salmon behaviour in the Mersey catchment. The Mersey Life Project.* Bristol.

Estoup A, Jarne P, Cornuet J (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular ecology*, **11**, 1591–1604.

Estoup A, Largiader CR, Perrot E, Chourrout D (1996) Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology*, **5**, 295–298.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14**, 2611–20.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, **10**, 564–7.

FAO (2014) *The state of world fisheries and aquaculture*. Rome.

Faubet P, Gaggiotti OE (2008) A new Bayesian method to identify the environmental factors that influence recent migration. *Genetics*, **178**, 1491–504.

Finnegan AK, Griffiths AM, King RA et al. (2013) Use of multiple markers demonstrates a cryptic western refugium and postglacial colonisation routes of Atlantic salmon (*Salmo salar* L.) in Northwest Europe. *Heredity*, **111**, 34–43.

Finnegan AK, Stevens JR (2008) Assessing the long-term genetic impact of historical stocking events on contemporary populations of Atlantic salmon, *Salmo salar*. *Fisheries Management and Ecology*, **15**, 315–326.

Fleming I (1998) Pattern and variability in the breeding system of Atlantic salmon (*Salmo salar*), with comparisons to other salmonids. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 59–76.

Fleming I, Jonsson B, Gross M, Lamberg A (1996) An experimental study of the reproductive behaviour and success of farmed and wild Atlantic salmon (*Salmo

salar). *Journal of Applied Ecology*, **33**, 893–905.

Fleming I, Lamberg A, Jonsson B (1997) Effects of early experience on the reproductive performance of Atlantic salmon. *Behavioral Ecology*, **8**, 470–480.

Frankel O (1974) Genetic conservation: our evolutionary responsibility. *Genetics*, **78**, 53–65.

Frankham R, Lees K, Montgomery ME *et al.* (1999) Do population size bottlenecks reduce evolutionary potential? *Animal Conservation*, **2**, 255–260.

Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications*, **1**, 535–586.

Fraser DJ, Hansen MM, Ostergaard S *et al.* (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular ecology*, **16**, 3866–89.

Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity*, **106**, 404–20.

Frazer KK, Russello M a (2013) Lack of parallel genetic patterns underlying the repeated ecological divergence of beach and stream-spawning kokanee salmon. *Journal of evolutionary biology*, **26**, 2606–21.

Friedland KD (1998) Ocean climate influences on critical Atlantic salmon (<I>Salmo salar</I>) life history events. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 119–130.

Friedland K, Hansen L, Dunkley D, MacLean J (2000) Linkage between ocean climate, post-smolt growth, and survival of Atlantic salmon (*Salmo salar* L.) in the North Sea area. *ICES Journal of Marine Science*, **57**, 419–429.

Friedland KD, Shank B V., Todd CD, McGinnity P, Nye J a. (2014) Differential response of continental stock complexes of Atlantic salmon (*Salmo salar*) to the Atlantic Multidecadal Oscillation. *Journal of Marine Systems*, **133**, 77–87.

Garant D, Dodson J, Bernatchez L (2000) Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar* L.)*. *Molecular Ecology*, **9**, 615–628.

Garcia de Leaniz C, Fleming IA, Einum S *et al.* (2007) A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological reviews of the Cambridge Philosophical Society*, **82**, 173–211.

Garstang W (1900) The Impoverishment of the Sea. A Critical Summary of the Experimental and Statistical Evidence bearing upon the Alleged Depletion of the Trawling Grounds. *Journal of the Marine Biological Association of the United Kingdom*, **6**, 1–69.

Garza J, Williamson E (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular ecology*, **10**, 305–318.

Gibson RJ (1993) The Atlantic salmon in fresh water: spawning, rearing and production. *Reviews in Fish Biology and Fisheries*, **73**, 39–73.

Glover K a, Quintela M, Wennevik V *et al.* (2012) Three decades of farmed escapees

in the wild: a spatio-temporal analysis of Atlantic salmon population genetic structure throughout Norway. *PloS one*, **7**, e43129.

Goldberg CS, Waits LP (2010) Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. *Molecular ecology resources*, **10**, 304–13.

Goldstein D, Ruiz Linares A, Cavalli-Sforza LL, Feldman M (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 6723–6727.

Goudet J (1995) FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of heredity*, **86**, 1994–1995.

Grandjean F, Verne S, Cherbonnel C, Richard A (2009) Fine-scale genetic structure of Atlantic salmon (*Salmo salar*) using microsatellite markers: effects of restocking and natural recolonization. *Freshwater Biology*, **54**, 417–433.

Gregory S, Lawton R, Learmonth ATA (1953) The middle Mersey and the chemical area. *A scientific survey of Merseyside. Univ. of Liverpool Press, Liverpool*, 299.

Griffiths a. M, Bright D, Stevens JR (2009) Comparison of patterns of genetic variability in wild and supportively bred stocks of brown trout, Salmo trutta. *Fisheries Management and Ecology*, **16**, 514–519.

Griffiths AM, Ellis JS, Clifton-Dey D *et al.* (2011) Restoration versus recolonisation: The origin of Atlantic salmon (*Salmo salar* L.) currently in the River Thames. *Biological Conservation*, **144**, 2733–2738.

Griffiths AM, Machado-Schiaffino G, Dillane E *et al.* (2010) Genetic stock identification of Atlantic salmon (*Salmo salar*) populations in the southern part of the European range. *BMC genetics*, **11**, 31.

Guerin A, Jackson A, Bowyer P, Youngson A (2014) *Hydrodynamic models to understand salmon migration in Scotland.*

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

Hamann E, Kennedy B (2012) Juvenile dispersal affects straying behaviors of adults in a migratory population. *Ecology*, **93**, 733–740.

Hamilton PB, Nicol E, De-Bastos ESR *et al.* (2014) Populations of a cyprinid fish are self-sustaining despite widespread feminization of males. *BMC biology*, **12**, 1.

Hampshire Biodiversity Partnership (2000) *Chalk Streams.*

Handley JF, Wood RWS (1999) The consequences of landscape change: principles and practice, problems and opportunities. *Ecology and landscape development: a history of the Mersey Basin. Liverpool University Press, Liverpool*, 243–261.

Hansen L, Jonsson B (1994) Homing of Atlantic salmon: effects of juvenile learning on transplated post-spawners. *Animal behaviour*, **47**, 220–222.

Hansen L, Jonsson N, Jonsson B (1993) Oceanic migration in homing Atlantic salmon. *Animal Behaviour*, **45**, 927–941.

Hard JJ, Heard WR (1999) Analysis of straying variation in Alaskan hatchery chinook salmon (*Oncorhynchus tshawytscha*) following transplantation. *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 578–589.

Hardy O, Charbonnel N, Fréville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **1482**, 1467–1482.

Heaps NS, Jones JE (1977) Density currents in the Irish Sea. *Geophysical Journal International*, **51**, 393–429.

Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology and Evolution*, **16**, 629–636.

Heggenes J, Beere M (2011) Estimation of genetic diversity within and among populations of *Oncorhynchus mykiss* in a coastal river experiencing spatially variable hatchery augmentation. *Transactions of the American Fisheries Society*, **140**, 37–41.

Henry P, Miquelle D, Sugimoto T *et al.* (2009) In situ population structure and ex situ representation of the endangered Amur tiger. *Molecular ecology*, **18**, 3173–84.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular biology and evolution*, **27**, 905–20.

Hey J (2011) Documentation for IMa2. , 1–60.

Hoar WS (1988) The physiology of smolting. *Fish Physiology*, **XI**, 275–342.

Hoban S, Arntzen J a., Bruford MW *et al.* (2014) Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. *Evolutionary Applications*, **7**, 984–998.

Holland DG, Harding JPC (1984) Mersey. In: *Ecology of European Rivers*, pp. 113–144. Blackwell Publishing Ltd, Oxford.

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature reviews. Genetics*, **10**, 639–50.

Horreo JL, Machado-Schiaffino G, Ayllon F *et al.* (2011a) Impact of climate change and human-mediated introgression on southern European Atlantic salmon populations. *Global Change Biology*, **17**, 1778–1787.

Horreo JL, Machado-Schiaffino G, Griffiths a. M *et al.* (2011b) Atlantic Salmon at Risk: Apparent Rapid Declines in Effective Population Size in Southern European Populations. *Transactions of the American Fisheries Society*, **140**, 605–610.

Hovey SJ, F KDP, Thomson D, Scott A (1989) Mitochondrial DNA and allozyme analysis of Atlantic salmon, *Salmo salar* L., in England and Wales. *Journal of Fish Biology*, **35**, 253–260.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular ecology resources*, **9**, 1322–32.

Hutchings J a, Fraser DJ (2008) The nature of fisheries- and farming-induced evolution. *Molecular ecology*, **17**, 294–313.

ICES (2013) *Report of the ICES advisory committee 2013- Book 10 North Atlantic Salmon Stocks*. Denmark.

ICES (2014) *Report of the Working Group on North Atlantic Salmon (WGNAS) International Council for the Exploration of the Sea*.

Ikediashi C, Billington S, Stevens JR (2012) The origins of Atlantic salmon (*Salmo salar* L.) recolonizing the River Mersey in northwest England. *Ecology and evolution*, **2**, 2537–48.

Jacobsen MW, Hansen MM, Orlando L *et al.* (2012) Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus spp.*). *Molecular ecology*, **21**, 2727–42.

Jarne P, Lagoda P (1996) Microsatellites, from molecules to populations and back. *Trends in ecology & evolution*, **11**, 424–429.

Jepsen N, Aarestrup K, Rasmussen G (1998) Survival of radio-tagged Atlantic salmon (*Salmo salar* L.) and trout (*Salmo trutta* L.) smolts passing a reservoir during seaward migration. *Hydrobiologia*, **371/372**, 347–353.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)*, **24**, 1403–5.

Jones D (2000) The Mersey Estuary - Back from the Dead? Solving a 150 - Year Old Problem. *Water and Environment Journal*, **14**, 124–130.

Jones PD (2006) Water quality and fisheries in the Mersey estuary, England: A historical perspective. *Marine pollution bulletin*, **53**, 144–54.

Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular ecology resources*, **10**, 551–5.

Jonsson I, Antonsson T, Gudjonsson S (2008) Relation between stock size and catch data of Atlantic salmon (*Salmo salar*) and Arctic charr (*Salvelinus alpinus*). *Icelandic Agricultural Science*, **21**, 61–68.

Jonsson B, Jonsson N, Hansen L (2003) Atlantic salmon straying from the River Imsa. *Journal of Fish Biology*, **62**, 641–657.

Jordan W, Cross T, Crozier WW *et al.* (2005) Allozyme variation in Atlantic salmon from the British Isles: associations with geography and the environment. *Journal of Fish Biology*, **67**, 146–168.

Jorde PE, Ryman N (2007) Unbiased estimator for genetic drift and effective population size. *Genetics*, **177**, 927–35.

Jr SR, McCleave J (1973) Sensitivity of American eels (*Anguilla rostrata*) and Atlantic salmon (*Salmo salar*) to weak electric and magnetic fields. *Journal of the Fisheries Board of Canada*, **30**, 657–663.

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

Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, **106**, 625–32.

Kalinowski S, Manlove K, Taper M (2007) *ONCOR: a computer program for genetic stock identification*.

Kiffney PM, Pess GR, Anderson JH, Faulds P (2009) Changes in fish communities following recolonization of the Cedar River, WA, USA by Pacific salmon after 103 years of local extirpation. *River Research and Applications*, **452**, 438–452.

Kimura M, Crow J (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725–738.

King T (2000) Mitochondrial DNA diversity in North American and European Atlantic salmon with emphasis on the Downeast rivers of Maine. *Journal of Fish Biology*, **57**, 614–630.

King TL, Eackles MS, Letcher BH (2005) Microsatellite DNA markers for the study of Atlantic salmon (*Salmo salar*) kinship, population structure, and mixed-fishery analyses. *Molecular Ecology Notes*, **5**, 130–132.

King TL, Kalinowski ST, Schill WB, Spidle a P, Lubinski B a (2001a) Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular ecology*, **10**, 807–21.

King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski B (2001b) Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular ecology*, **10**, 807–21.

Klemetsen A, Amundsen P-A, Dempson J *et al.* (2003) Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): a review of aspects of their life histories. *Ecology of freshwater fish*, **12**, 1–59.

Knox D, Verspoor E (1991) A mitochondrial DNA restriction fragment length polymorphism of potential use for discrimination of farmed Norwegian and wild Atlantic salmon populations in Scotland. *Aquaculture*, **98**, 249–257.

Knutsen H, Knutsen J, Jorde P (2001) Genetic evidence for mixed origin of recolonized sea trout populations. *Heredity*, **87**, 207–214.

Koljonen M, Pella J, Masuda M (2005) Classical individual assignments versus mixture modeling to estimate stock proportions in Atlantic salmon (*Salmo salar*) catches from DNA microsatellite data. *Canadian Journal of Fisheries and Aquatic Sciences*, **62**, 2143–2158.

Krkosek M, Lewis M a, Volpe JP (2005) Transmission dynamics of parasitic sea lice from farm to wild salmon. *Proceedings. Biological sciences / The Royal Society*, **272**, 689–96.

Kuparinen A, Tufto J, Consuegra S *et al.* (2009) Effective size of an Atlantic salmon (*Salmo salar* L.) metapopulation in Northern Spain. *Conservation Genetics*, **11**, 1559–1565.

Lage C, Kornfield I (2006) Reduced Genetic Diversity and Effective Population Size in an Endangered Atlantic Salmon (*Salmo Salar*) Population from Maine, USA. *Conservation Genetics*, **7**, 91–104.

Laikre L, Allendorf F, Aroner L *et al.* (2009) Neglect of genetic diversity in implementation of the convention on biological diversity. *Conservation Biology*, **24**, 86–88.

Land ED, Lacy RC (2000) Introgression level achieved through Florida panther genetic restoration. *Endangered Species Update*, **17**, 100–105.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, **7**, 295–302.

Leo SST, Cheng L, Sperling FAH (2012) Genetically separate populations of the ocean-skater Halobates sericeus (Heteroptera: Gerridae) have been maintained since the late Pleistocene. *Biological Journal of the Linnean Society*, **105**, 797–805.

Lin Q (2011) Influence of Dams on River Ecosystem and Its Countermeasures. *Journal of Water Resource and Protection*, **03**, 60–66.

Luikart G, Cornuet J-M (1998) Empirical Evaluation of a Test for Identifying Recently Bottlenecked Populations from Allele Frequency Data. *Conservation Biology*, **12**, 228–237.

Lundqvist H, Clarke W, Johansson H (1988) The influence of precocious sexual maturation on survival to adulthood of river stocked Baltic salmon, Salmo salar, smolts. *Holarctic Ecology*, **11**, 60–69.

Lynch M, Koskella B, Schaack S (2006) Mutation pressure and the evolution of organelle genomic architecture. *Science*, **311**, 1727–1731.

Lynch PD, Shertzer KW, Latour RJ (2012) Performance of methods used to estimate indices of abundance for highly migratory species. *Fisheries Research*, **125-126**, 27–39.

MacCrimmon H, Gots B (1979) World distribution of Atlantic salmon, *Salmo salar*. *Journal of the Fisheries Board of Canada*, **36**, 422–457.

MacIntyre F, Estep K, Noji T (1995) Is it deforestation or desertification when we do it to the ocean? *Naga, the ICLARM Quarterly*.

MacKenzie KM, Palmer MR, Moore A *et al.* (2011) Locations of marine animals revealed by carbon isotopes. *Scientific reports*, **1**, 21.

Mackenzie KM, Trueman CN, Palmer MR *et al.* (2012) Stable isotopes reveal age-dependent trophic level and spatial segration during adult marine feeding in populations of salmon. *ICES Journal of Marine Science*, **69**, 1637–1645.

Manel S, Gugerli F, Thuiller W *et al.* (2012) Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. *Molecular Ecology*, **21**, 3729–3738.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer research*, **27**, 209–220.

Martinez J, Dumas J, Bealle E, García-Vázquez E (2001) Assessing introgression of foreign strains in wild Atlantic salmon populations: variation in microsatellites assessed in historic scale collections. *Freshwater Biology*, **46**, 835–844.

Massiot-Granier F, Prevost E, Chaput G *et al.* (2014) Embedding stock assessment within an integrated hierarchical Bayesian life cycle modelling framework: an application to Atlantic salmon in the Northeast Atlantic. *ICES Journal of Marine Science.*

Mawle G, Peirson G (2001) *Economic evaluation of inland fisheries: Managers report from science project SC050026/SR2*. Bristol.

Mayr E (1954) Changes of genetic environment and evolution. In:, pp. 157–180.

McConnell S, Hamilton L, Morris D *et al.* (1995) Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture*, **137**, 19–30.

McConnell S, Ruzzante D, O'Reilly P, Hamilton L, Wright J (1997) Microsatellite loci reveal highly significant genetic differentiation among Atlantic salmon (*Salmo salar* L.) stocks from the east coast of Canada. *Molecular Ecology*, **6**, 1075–1089.

McCormick SD, Hansen LP, Quinn TP, Saunders RL (1998) Movement, migration, and smolting of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 77–92.

McCullough D (1999) *A review and synthesis of effects of alterations to the water temperature regime on freshwater life stages of salmonids, with special reference to Chinook salmon*. United States EPA, Seattle.

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

McGinnity P, Jennings E, DeEyto E *et al.* (2009) Impact of naturally spawning captive-bred Atlantic salmon on wild populations: depressed recruitment and increased risk of climate-mediated extinction. *Proceedings. Biological sciences / The Royal Society*, **276**, 3601–10.

McGinnity P, Prodöhl P, Ferguson A *et al.* (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings. Biological sciences / The Royal Society*, **270**, 2443–50.

McGinnity P, Stone C, Taggart JB *et al.* (1997) Genetic impact of escaped farmed Atlantic salmon (*Salmo salar* L .) on native populations : use of DNA profiling to assess freshwater performance of wild , farmed , and hybrid progeny in a natural river environment. *Journal of Marine Science*, **54**, 998–1008.

Menotti-Raymond M, O'Brien S (1993) Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 3172–3176.

Mills D, Atlantic Salmon Federation, Trust AS (2003) *Salmon at the Edge.*

Mills KE, Pershing AJ, Sheehan TF, Mountain D (2013) Climate and ecosystem linkages explain widespread declines in North American Atlantic salmon populations. *Global Change Biology*, **19**, 3046–3061.

Milner NJ, Russell IC, Aprahamian M, Inverarity R, Shelley J (2004) Environment Agency The role of stocking in recovery of the River Tyne salmon. *Fisheries (Bethesda).*

Moore A, Freake S, Thomas I (1990) Magnetic particles in the lateral line of the Atlantic salmon (*Salmo salar* L.). *Philosophical Transactions: Biological Sciences*, **329**, 11–15.

Moran P, Perez J, Dumas J, Bealle E, García-Vázquez E (2005) Stocking-related patterns of genetic variation at enzymatic loci in south European Atlantic salmon populations. *Journal of Fish Biology*, **67**, 185–199.

Morin P a., Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.

Morin PA, Mccarthy M (2007) Highly accurate SNP genotyping from historical and low-quality samples. *Molecular Ecology Notes*, **7**, 937–946.

Murton JB, Lautridou J-P (2003) Recent advances in the understanding of Quaternary periglacial features of the English Channel coastlands. *Journal of Quaternary Science*, **18**, 301–307.

Myers R, Hutchings J (1987) Mating of anadromous Atlantic salmon, *Salmo salar* L., with mature male parr. *Journal of Fish Biology*, **31**, 143–146.

NASCO (2008) *Council CNL(08)17 Interim Report of the Socio-Economics Working Group*.

NASCO (2009) *NASCO Guidelines for the Management of Salmon Fisheries*.

NASCO (2013) *Report of the ICES Advisory Committee*.

Naylor RL, Eagle J, Smith WL (2003) Salmon Aquaculture in the Pacific Northwest A Global Industry with Local Impacts. *Environment: Science and Policy for Sustainable Development*, **45**, 18–39.

Nei M, Tajima F (1981) Genetic drift and estimation of effective population size. *Genetics*, **98**, 625–640.

Nielsen EE, Hansen MM (2008) Waking the dead: The value of population genetic analyses of historical samples. *Fish and Fisheries*, **9**, 450–461.

Nielsen E, Hansen M, Loeschcke V (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology*, **6**, 487–492.

Nielsen EE, Hansen MM, Loeschcke V (1999) Analysis of DNA from old scale samples: Technical aspects, applications and perspectives for conservation. *Hereditas*, **130**, 265–276.

Nieminen M, Korhonen I, Laitinen M (1982) Effects of pH on the gill ATPase activity and blood composition of whitefish (*Coregonus peled*) and trout (*Salmo trutta* fario). *Comparative Biochemistry and Physiology*, **72**, 637–642.

Nikolic N, Butler JR a, Baglinière J-L *et al.* (2009) An examination of genetic diversity and effective population size in Atlantic salmon populations. *Genetics research*, **91**, 395–412.

Nikolic N, Chevalet C (2014) Detecting past changes of effective population size. *Evolutionary applications*, **7**, 663–81.

Nilsson J, Gross R, Asplund T *et al.* (2001) Matrilinear phylogeography of Atlantic

salmon (*Salmo salar* L.) in Europe and postglacial colonization of the Baltic Sea area. *Molecular Ecology*, **10**, 89–102.

Nomura T (2008) Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, **1**, 462–474.

O'Connell M (2003) An examination of the use of angling data to estimate total returns of Atlantic salmon, *Salmo salar*, to two rivers in Newfoundland, Canada. *Fisheries Management and ecology*, **1**, 201–208.

O'Reilly P, Hamilton L, McConnell S, Wright J (1996) Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Canadian Journal of Fisheries and Aquatic Sciences*, **2298**, 2292–2298.

O'Toole CL, Reed TE, Bailie D *et al.* (2015) The signature of fine scale local adaptation in Atlantic salmon revealed from common garden experiments in nature. *Evolutionary applications*, **8**, 881–900.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical research*, **22**, 201–204.

Olafsson K, Pampoulie C, Hjorleifsdottir S, Gudjonsson S, Hreggvidsson GO (2014) Present-day genetic structure of Atlantic salmon (*Salmo salar*) in Icelandic rivers and ice-cap retreat models. *PloS one*, **9**, e86809.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes*, **4**, 535–538.

Østergaard S, Hansen MM, Loeschcke V, Nielsen EE (2008) Long-term temporal changes of genetic composition in brown trout (*Salmo trutta* L.) populations inhabiting an unstable environment. *Molecular Ecology*, **12**, 3123–3135.

Palstra FP, O'Connell MF, Ruzzante DE (2007) Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. *Molecular ecology*, **16**, 4504–22.

Palstra FP, O'Connell MF, Ruzzante DE (2009) Age structure, changing demography and effective population size in Atlantic salmon (*Salmo salar*). *Genetics*, **182**, 1233–49.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**, 3428–3447.

Parrish DL, Behnke RJ, Gephard SR, McCormick SD, Reeves GH (1998) Why aren't there more Atlantic salmon (*Salmo salar*)? *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 281–287.

Paterson S, Piertney SB, Knox D, Gilbey J, Verspoor E (2004) Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Molecular Ecology Notes*, **4**, 160–162.

Peakall R, Smouse P (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, **1**, 6–8.

Peery MZ, Kirby R, Reid BN *et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular ecology*, **21**, 3403–18.

Perrier C, Evanno G, Belliard J, Guyomard R, Baglinière J-L (2010) Natural recolonization of the Seine River by Atlantic salmon (*Salmo salar*) of multiple origins. *Canadian Journal of Fisheries and Aquatic Sciences*, **67**, 1–4.

Perrier C, Guyomard R, Bagliniere J-L, Evanno G (2011) Determinants of hierarchical genetic structure in Atlantic salmon populations: environmental factors vs. anthropogenic influences. *Molecular ecology*, **20**, 4231–45.

Perrier C, Guyomard R, Bagliniere J-L, Nikolic N, Evanno G (2013) Changes in the genetic structure of Atlantic salmon populations over four decades reveal substantial impacts of stocking and potential resiliency. *Ecology and evolution*, **3**, 2334–49.

Picoult-Newberg L, Ideker T, Pohl MG *et al.* (1999) Mining SNPs From EST Databases. *Genome Research*, **9**, 167–174.

Pinfold G (2011) Economic value of wild Atlantic salmon. *Atlantic Salmon Federation, Canada*.

Piry S, Alapetite a, Cornuet J-M *et al.* (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *The Journal of heredity*, **95**, 536–9.

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pollak E (1983) A new method for estimating the effective population size from allele frequency changes. *Genetics*, **104**, 531–548.

Potter EC., Maoileidigh N., Chaput G (2003) *Marine mortality of Atlantic salmon, Salmo salar L: methods and measures*.

Pountney A, Stevens J, Sykes T, Tyler C (2009) *Population genetics and PBDE analysis of English and Welsh otters.* Bristol.

Primmer CR, Veselov AJ, Zubchenko A *et al.* (2006) Isolation by distance within a river system: genetic population structuring of Atlantic salmon, Salmo salar, in tributaries of the Varzuga River in northwest Russia. *Molecular ecology*, **15**, 653–66.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–59.

Prunet P, Boeuf G, Bolton J, Young G (1989) Smoltification and seawater adaptation in Atlantic salmon (*Salmo salar*): plasma prolactin, growth hormone, and thyroid hormones. *General and comparative Endocrinology*, **74**, 355–364.

Pudovkin A, Zaykin D, Hedgecock D (1996) On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics*, **144**, 383–387.

Putman NF, Lohmann KJ, Putman EM *et al.* (2013) Evidence for geomagnetic imprinting as a homing mechanism in Pacific salmon. *Current Biology*, **23**, 312–316.

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

Reed DH, Frankham R (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*, **17**, 230–237.

Ribeiro Â, Morán P, Caballero A (2008) Genetic diversity and effective size of the Atlantic salmon *Salmo salar* L. inhabiting the River Eo (Spain) following a stock collapse. *Journal of Fish Biology*, **72**, 1933–1944.

Riccioni G, Landi M, Ferrara G *et al.* (2010) Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic bluefin tuna of the Mediterranean Sea. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 2102–7.

Rienzo A Di, Peterson A (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 3166–3170.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Russell IC, Ives MJ, Potter ECE, A BA, Duckett L (1995) *Salmon and migratory trout statistics for England and Wales, 1951-1990.*

Säisä M, Koljonen M-L, Gross R *et al.* (2005) Population genetic structure and postglacial colonization of Atlantic salmon (*Salmo salar*) in the Baltic Sea area based on microsatellite DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, **62**, 1887–1904.

Säisä M, Koljonen M, Tähtinen J (2003) Genetic changes in Atlantic salmon stocks since historical times and the effective population size of a long-term captive breeding programme. *Conservation Genetics*, **4**, 613–627.

Saltveit S (2006) The effects of stocking Atlantic salmon, *Salmo salar*, in a Norwegian regulated river. *Fisheries Management and Ecology*, **13**, 197–205.

Sánchez J a, Clabby C, Ramos D *et al.* (1996) Protein and microsatellite single locus variability in Salmo salar L. (*Atlantic salmon*). *Heredity*, **77**, 423–32.

Sandlund OT, Karlsson S, Thorstad EB *et al.* (2014) Spatial and temporal genetic structure of a river-resident Atlantic salmon (*Salmo salar*) after millennia of isolation. *Ecology and evolution*, **4**, 1538–54.

Sandøy S, Langåker R (2001) Atlantic salmon and acidification in southern Norway: a disaster in the 20th Century, but a hope for the Future? *Water, Air, and Soil Pollution*, **130**, 1343–1348.

Saunders R, Schom C (1985) Importance of the variation in life history parameters of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **42**, 615–618.

Saura M, Caballero P, Caballero A, Moran P (2006) Genetic variation in restored Atlantic salmon (*Salmo salar* L.) populations in the Ulla and Lérez rivers, Galicia, Spain. *ICES Journal of Marine Science*, **63**, 1290–1296.

Saura M, Caballero P, Morán P (2008) Are there Atlantic salmon in the River Tambre? *Journal of Fish Biology*, **72**, 1223–1229.

Schreiber A, Diefenbach G (2005) Population genetics of the European trout (Salmo trutta L.) migration system in the river Rhine: Recolonisation by sea trout. *Ecology of Freshwater Fish*, **14**, 1–13.

Selly S-LC, Hickey J, Stevens JR (2014) A tale of two hatcheries: Assessing bias in the hatchery process for Atlantic salmon (*Salmo salar* L.). *Aquaculture*, **434**, 254–263.

Serbezov D, Jorde PE, Bernatchez L, Olsen EM, Vøllestad LA (2012) Short-term genetic changes: evaluating effective population size estimates in a comprehensively described brown trout (*Salmo trutta*) population. *Genetics*, **191**, 579–92.

Skaala O, Wennevik V, Glover K (2006) Evidence of temporal genetic change in wild Atlantic salmon, *Salmo salar* L., populations affected by farm escapees. *ICES Journal of Marine Science*, **63**, 1224–1233.

Slettan A, Olsaker I, Lie Ø (1995) Atlantic salmon, *Salmo salar*, microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci. *Animal genetics*, **24**, 281–282.

Smith S, Dobson M, Fenn C *et al.* (2014) *The State of England' s Chalk Streams*.

Smith MJ, Pascal CE, Grauvogel Z *et al.* (2011) Multiplex preamplification PCR and microsatellite validation enables accurate single nucleotide polymorphism genotyping of historical fish scales. *Molecular Ecology Resources*, **11**, 268–277.

Solomon DJ (1978) Some observations on salmon smolt migration in a chalkstream. *Journal of Fish Biology*, **12**, 571–574.

Spidle A, Kalinowski S, Lubinski B *et al.* (2003) Population structure of Atlantic salmon in Maine with reference to populations from Atlantic Canada. *Transactions of the American Fisheries Society*, **132**, 196–209.

Spong G, Hellborg L (2002) A near-extinction event in lynx: do microsatellite data tell the tale? *Conservation Ecology*, **6**.

Stabell O (1984) Homing and olfaction in salmonids: a critical review with special reference to Atlantic salmon. *Biological Reviews*, **59**, 333–388.

Stahl G (1987) Genetic population structure of Atlantic salmon. In: *Population genetics and fishery management*, pp. 121–140.

Storfer A (1999) Gene flow and endangered species translocations: A topic revisited. *Biological Conservation*, **87**, 173–180.

Taberlet P, Griffin S, Goossens B *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic acids research*, **24**, 3189–3194.

Takahashi M, Kato Y, Mukoyama H (1997) Evaluation of five polymorphic microsatellite markers for typing DNA from decomposed human tissues:–Correlation between the size of the alleles and that of the template DNA. *Forensic Science International*, **90**, 1–9.

Taylor E (1991) A review of local adaptation in Salmonidac, with particular reference to Pacific and Atlantic salmon. *Aquaculture*, **98**, 185–207.

Thompson D, C RI (1991) Allele frequency variation at the AAT-4 locus as a potential measure of the relative performance of native and introduced Atlantic salmon in the River Test. *Aquaculture*, **98**, 243–247.

Thorley J, Eatherley D, Stephen A *et al.* (2005) Congruence between automatic fish counter data and rod catches of Atlantic salmon (*Salmo salar*) in Scottish rivers. *ICES Journal of Marine Science*, **62**, 809–817.

Thorstad EB, Økland F, Aarestrup K, Heggberget TG (2007) Factors affecting the within-river spawning migration of Atlantic salmon, with emphasis on human impacts. *Reviews in Fish Biology and Fisheries*, **18**, 345–371.

Tonteri A, Veselov AJ, Zubchenko A V., Lumme J, Primmer CR (2009) Microsatellites reveal clear genetic boundaries among Atlantic salmon (*Salmo salar*) populations from the Barents and White seas, northwest Russia. *Canadian Journal of Fisheries and Aquatic Sciences*, **66**, 717–735.

Truett G, Heeger P, Mynatt R (2000) Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques*, **29**, 52–54.

Ugedal O, Næsje TF, Thorstad EB *et al.* (2008) Twenty years of hydropower regulation in the River Alta: long-term changes in abundance of juvenile and adult Atlantic salmon. *Hydrobiologia*, **609**, 9–23.

Vähä J-P, Erkinaro J, Niemelä E, Primmer CR (2007) Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular ecology*, **16**, 2638–54.

Vähä J-P, Erkinaro J, Niemelä E, Primmer CR (2008) Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. *Evolutionary Applications*, **1**, 137–154.

Valiente AG, Beall E, Garcia-Vazquez E (2010) Population genetics of south European Atlantic salmon under global change. *Global Change Biology*, **16**, 36–47.

Vasemägi a, Gross R, Paaver T *et al.* (2001) Identification of the origin of an Atlantic salmon (*Salmo salar* L.) population in a recently recolonized river in the Baltic Sea. *Molecular ecology*, **10**, 2877–82.

Vasemägi A, Gross R, Paaver T, Koljonen M-L, Nilsson J (2005) Extensive immigration from compensatory hatchery releases into wild Atlantic salmon population in the Baltic sea: spatio-temporal analysis over 18 years. *Heredity*, **95**, 76–83.

Venables W, Smith D (2005) An Introduction to R. *R Development Core Team*.

Verspoor E, Beardmore J, Consuegra S *et al.* (2005) Population structure in the Atlantic salmon: insights from 40 years of research into genetic protein variation. *Journal of Fish Biology*, **67**, 3–54.

Verspoor E, Mccarthy EM, Knox D, Bourke EA, Cross TF (1999) The phylogeography of European Atlantic salmon. *Biological Journal of the Linnean Society*, **68**, 129–146.

Vincent B, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene-environment interactions driving local adaptation. *Evolution; international journal of organic evolution*, **67**, 3469–87.

Wang J (2009) A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular ecology*, **18**, 2148–64.

Wang IJ (2011) Choosing appropriate genetic markers and analytical methods for testing landscape genetic hypotheses. *Molecular Ecology*, **20**, 2480–2482.

Waples RS, Yokota M (2007) Temporal estimates of effective population size in species with overlapping generations. *Genetics*, **175**, 219–33.

Webster TU, Bury N, van Aerle R, Santos E (2013) Global transcriptome profiling reveals molecular mechanisms of metal tolerance in a chronically exposed wild population of brown trout. *Environmental science & technology*, **47**, 8869–8877.

Willing E-M, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PloS one*, **7**, e42649.

Wilson KW, D'Arcy BJ, Taylor S (1988) The return of fish to the Mersey estuary. *Journal of Fish Biology*, **33**, 235–238.

Wilson AB, Eigenmann Veraguth I (2010) The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Molecular Ecology*, **19**, 4535–4553.

Wilson G, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.

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

WWF (2001) *The Status of Wild Atlantic Salmon: A River by River Assessment*.