The genetics of a managed Atlantic salmon stock and implications for conservation.

by
Karen Frake

Submitted for the degree of Doctor of Philosophy
April 2007
University of Stirling
Declaration

I declare that this thesis has been compiled by myself and is the result of my own investigations. It has not been submitted for any other degree and all sources of information have been duly acknowledged.

Signed

Karen Frake

Date

Printed on 100% recycled paper.
Acknowledgements

Thanks to my supervisors John Taggart and Alan Teale.

Many, many thanks to John Taggart for assistance, guidance and criticism throughout the project as well as a running commentary on the trials and tribulations of bringing up twins.

Thanks to the following people at Fisheries Research Services (FRS) Freshwater Laboratory, Pitlochry: Alan Youngson, Joe (detox) Thorley and Alastair Thorne. Special thanks to Iain McLaren for continually replying to my pestering emails and supplying me with information, and to Phil Bacon for invaluable assistance with statistics. Thanks also to John Webb, Colin Bull and David Hopkins for their input in the initial stages of the project.

Thanks to Peter Nonacs and Mark Beaumont for answering my pleas for help in utilising and interpreting results from their respective computer packages. Thanks to James Bron for assistance in carrying out a statistical test.

Thanks to people in and around the laboratory: Helen Whitaker, Matteo Minghetti, Amer Diab, Ann Gilmour, Anu Frank-Lawale, Almas Gheyas, Gareth Butterfield, Marine Herlin and Fazel.

Many thanks to people of the ex-ecology department: Diana Bowler, Mairead Maclean, Jenny Bright and Chris Pendlebury.

Thanks to friends and family for much-needed support and to my sister Liz for proofreading sections.

Many thanks to my partner, Peter Watson, for everything… especially for putting up with papers strewn around the house and for seeing the brighter side of things on the many occasions when I didn’t think I would ever reach this stage.

This work was funded by NERC and The University of Stirling.
Abstract
Numerous populations of wild Atlantic salmon have declined in recent years. The Atlantic salmon in Girnock Burn, an upland spate tributary of the River Dee, Scotland have been monitored intensely by government scientists since 1966. The burn is equipped with upstream and downstream traps, which have enabled monitoring of juveniles leaving the burn and adults returning to it since 1966. Recently, due to a decline in numbers of female returns, a supportive breeding program was instigated. Using microsatellite-based DNA profiling, this study exploited existing and novel tissue samples to investigate aspects of Atlantic salmon biology and conservation. A panel of up to 12, mainly highly polymorphic, microsatellite loci were employed to derive allele frequency data and to resolve parentage in egg, parr, smolt and anadromous adult samples taken between 1991 and 2004. Genotyping error was investigated and rectified where possible. Overall, the detected error was low (c.0.5%), providing confidence in subsequent population and parentage analyses. The error rate involved in estimating the age of salmon in Girnock Burn from scale readings was also estimated (c.2-8%).
A study of the dynamics of natural spawning, based on the parentage of parr, confirmed that multiple matings by anadromous returns of both sexes were prevalent. Not all anadromous returns were apparently successful spawners; data from parr and existing redd samples failed to detect a contribution from 35% of males and 29% of females.
An important aspect of the work was to determine the success of the supportive breeding program. Results showed that, in comparison to natural spawning, the program gave a more complete and even representation of adult spawners in offspring. In addition, there was no detectable difference in the output (number of
smolts) of the two schemes when the number of eggs used in each was taken into account.

The distribution of juvenile kin (parr aged 1+) within the burn was determined, which revealed clustering of full and half sib groups. This was found to impact on standard population genetic analyses. Adjacent samples (n = 50), each sampled over a c.1.5 km stretch of river were shown to exhibit significant allelic differentiation, while samples from individuals selected at random over a 7.5km stretch did not. Parentage analysis of adult returns showed that the number of returns likely to be philopatric was higher than would be predicted solely from physical tagging data. This was attributed to ‘leakage’ of the downstream parr/smolt trap. An initial investigation into the role of mature parr in adaptation of populations to the environment was made, although sire type (i.e. anadromous male or mature parr) was not found to affect survival in the freshwater environment in this case. More research into this aspect is warranted, particularly with the possible impact of predicted climate change on male parr maturity. A comparison of genetic diversity through time (measured by allelic richness) revealed no detectable change between 1991 and 2004. Estimates of the effective population size using different genetic (temporal) methods were associated with a large degree of uncertainty, and were surprisingly high (ranging from 595 to 1992) c.f. demographic based estimates (ranging from 95 to 144), which was likely to be due in part to violation of assumptions made in the calculations.

These findings have highlighted a range of avenues for future lines of research, should aid in the management of Atlantic salmon within Girnock Burn and assist in the design of sampling regimes.
# Table of Contents

Declaration .......................................................................................................................... ii
Acknowledgements ............................................................................................................. iii
Abstract ................................................................................................................................ iv
Table of Contents ............................................................................................................... vi

## 1 Introduction .................................................................................................................. 1

1.1 The status of wild Atlantic salmon .............................................................................. 1
1.2 Conservation genetics - a rapidly evolving field ....................................................... 5
1.3 Introducing Girnock Burn ......................................................................................... 11
1.4 Aims of the study ...................................................................................................... 13
  1.4.1 Accuracy of genotyping and other techniques .................................................. 13
  1.4.2 Sampling juveniles- the spatial distribution of kin ........................................ 13
  1.4.3 Natural and quasi-natural spawning ................................................................. 14
  1.4.4 Straying and trap leakage ................................................................................. 15
  1.4.5 Implications of reproductively successful mature parr .................................. 16
  1.4.6 Genetic diversity over time and effective population size ............................. 16
  1.5 Main objectives ..................................................................................................... 18

## 2 Materials and methods .............................................................................................. 19

2.1 Study site .................................................................................................................. 19
  2.1.1 Salmon stock in Girnock Burn .......................................................................... 20
  2.1.2 Traps and routine sampling .............................................................................. 22
  2.1.3 Management of the population ...................................................................... 24
2.2 Sampling ................................................................................................................... 25
  2.2.1 Anadromous fish .............................................................................................. 25
  2.2.2 Eggs sampled in 1995 ..................................................................................... 27
  2.2.3 Parr sampled in 1997 ..................................................................................... 27
  2.2.4 Smolts sampled in 2004 .................................................................................. 28
2.3 DNA extraction ........................................................................................................ 28
2.4 Genotyping ................................................................................................................ 29
  2.4.1 Use of different fluorescent dyes and fragment analyzers .............................. 29
  2.4.2 Controls and estimation of error rate ............................................................... 30
  2.4.3 Estimation of the number of loci required ....................................................... 31
  2.4.4 Choice of Loci ................................................................................................. 31
  2.4.5 Optimization of multiplexes .......................................................................... 35
  2.4.6 Empirical confirmation of the number of loci required for robust parental assignment .............................................................................................................. 37
2.4.7 PIG tailing ........................................................................................................ 38
2.4.8 Final PCR conditions ....................................................................................... 39
2.4.9 Detection and Sizing of PCR products ............................................................. 41
2.4.10 ABI fragment analyzers ................................................................................. 41
2.4.11 Beckman Coulter fragment analyzers ............................................................. 44
2.4.12 Allele callings ................................................................................................ 45
2.5 Parentage Assignment ............................................................................................. 45
  2.5.1 Exclusion versus maximum likelihood methods ........................................... 45
  2.5.2 Software employed in parentage analysis ....................................................... 46

## 3 Data quality and accuracy ............................................................................................ 48

3.1 Genotyping Accuracy ............................................................................................... 48
  3.1.1 Error rates and their sources in genotyping studies ........................................ 51
  3.1.2 Monitoring of genotyping errors in this study ................................................ 52
  3.1.3 Re-typing samples ............................................................................................ 52
  3.1.4 Controls ........................................................................................................... 55
  3.1.5 Parentage analysis ............................................................................................ 55
  3.1.6 Null alleles/ allele dropout ............................................................................. 59
  3.1.7 Mutation .......................................................................................................... 61
  3.1.8 Conclusion- levels of accuracy in genotyping ................................................ 62
3.2 Assignment of parents -power of exclusion ............................................................. 65
  3.2.1 False inclusion in an ideal population ............................................................. 66
7 Mature parr: their reproductive success and potential impact on adaptation 190
  7.1 Introduction ........................................................................................................ 190
  7.1.1 Mature male parr ............................................................................................ 190
  7.1.2 Spawning success of mature parr .................................................................... 192
  7.1.3 Genetic diversity and mature parr ..................................................................... 196
  7.1.4 Consequences of the mature parr ART for adaptation to the environment .... 198
  7.1.5 Issues investigated ......................................................................................... 203
  7.1.6 Mature parr at Girnock Burn ........................................................................... 203
  7.1.7 Objectives ........................................................................................................ 205
  7.2 Methods .............................................................................................................. 208
  7.2.1 Sampling ......................................................................................................... 208
  7.2.2 Spawning success of mature parr; effects of differing sex ratios of anadromous returns ................................................................................................................. 210
  7.2.3 Survival of offspring from alternative sire types in freshwater .................... 211
  7.2.4 Survival of offspring from alternative sire types in the marine environment ... 212
  7.2.5 Overall survival............................................................................................... 213
  7.3 Results .................................................................................................................. 214
  7.3.1 Spawning success of mature parr; effects of differing sex ratios of anadromous returns ................................................................................................................. 214
  7.3.2 Survival of offspring from alternative sire types in freshwater .................... 216
  7.3.3 Survival of offspring from alternative sire types in the marine environment ... 218
  7.3.4 Overall survival of offspring from alternative sire types in the marine environment 221
  7.4 Discussion ............................................................................................................ 224
  7.4.1 Reproductive success of mature parr with differing sex ratios of anadromous spawners ................................................................................................................. 224
  7.4.2 Survival of offspring from alternative sire types in freshwater .................... 228
  7.4.3 Survival of offspring from alternative sire types in the marine environment ... 229
  7.4.4 Overall survival of offspring from alternative sire types in the marine environment 231
  7.4.5 Future directions ............................................................................................ 232

8 Genetic diversity over time and estimation of the effective population size 234
  8.1 Introduction .......................................................................................................... 234
  8.1.1 Genetic diversity over time .............................................................................. 234
  8.1.2 Estimation of effective population size .......................................................... 239
  8.1.3 Objectives ........................................................................................................ 243
  8.2 Methods .............................................................................................................. 244
  8.2.1 Samples .......................................................................................................... 244
  8.2.2 Analysis of temporal stability of genetic diversity between 1991 and 2004 .... 244
  8.2.3 Testing for a recent bottleneck ........................................................................ 245
  8.2.4 Estimating the effective population size ....................................................... 247
  8.3 Results .................................................................................................................. 252
  8.3.1 Genetic diversity between 1991 and 2004 ....................................................... 252
  8.3.2 Testing for a recent bottleneck ........................................................................ 264
  8.3.3 Estimate of effective population size ............................................................. 265
  8.4 Discussion .......................................................................................................... 269
  8.4.1 Genetic diversity through time ....................................................................... 269
  8.4.2 Estimating the effective population size ....................................................... 271

9 Concluding Remarks .............................................................................................. 279

A1 Appendix .............................................................................................................. 289
  1.1 DNA extraction- simplified chelex protocol ..................................................... 289
  1.2 DNA extraction using Wizard® SV Genomic DNA Purification System (Promega, Cat. No. A2360) ........................................................................................................... 290
  1.3 PCR components and thermocycling conditions ............................................. 292
  1.4 Example traces of multiplexed loci ................................................................. 298

Bibliography ............................................................................................................... 305
1 Introduction

An environmentally conscious human being in the year 2150 might well look back to the current time and ponder what our generation could possibly have been thinking when, through lack of self-restraint on population growth and environmental degradation, we thoroughly fouled our planet and precipitated one of the greatest mass extinctions the Earth had ever known: “Those societies could not have been completely ignorant— they did develop DNA technologies.”

Avise (1996)

1.1 The status of wild Atlantic salmon

Atlantic salmon (Salmo salar L.) are renowned for their ability to home to particular river systems. This reputation has stemmed from their documented high fidelity of homing, despite extensive journeys taken between freshwater sites to feeding grounds in the ocean, and back again in order to spawn. An additional aspect of the species that is perhaps less well known is the high level of heterogeneity in life history characteristics that they may adopt.

Atlantic salmon spawn in gravel beds in freshwater during autumn or early winter. Collections of eggs are deposited in successive depressions in the gravel, forming ‘redds’. Eggs hatch the following spring, giving rise to the life stage known as alevins. For the first four to six weeks the yolk sac provides the newly hatched fish with nutrients (Youngson and Hay 1996). After approximately a month, when the yolk sac is completely absorbed and they begin external feeding, the young fish are classified as fry, and after further growth, they are termed parr. Following a period in freshwater migrations to the ocean occur, where growth is more rapid than in freshwater. Across their range, Atlantic salmon may spend between one and eight
years in freshwater, and between one and five years at sea (Klemetsen et al. 2003). Variation in the time spent in freshwater and at sea is also seen within single populations. Following the marine phase maturing salmon return to freshwater to spawn. In some populations, males may become sexually mature in freshwater without undergoing a seaward migration. These fish are termed mature male parr, or precocious parr. Although often less than a hundredth of the weight of their anadromous male counterparts (Fleming 1996), they can collectively account for the fertilisations of large proportions of eggs. A fuller description of the dynamics and complexity of spawning is given in Chapter 4.

Atlantic salmon belong to the Salmonidae family, which is composed of three subfamilies, the Coregoninae (whitefishes), Thymallinae (graylings) and Salmoninae (lenok, mekous, belvica, huchen, taimen, chars, trout and salmons) (Stearley and Smith 1993). The native range of Atlantic salmon extends in North America from approximately the Connecticut River (USA) in the south to Ungava Bay (Canada) in the north, and on the eastern side of the Atlantic, from northern Portugal to regions near the Barents and White seas in Russia (Klemetsen et al. 2003). Unfortunately, wild Atlantic salmon have suffered a dramatic decline in numbers since 1970, and have now disappeared from Germany, Switzerland, the Netherlands, Belgium, the Czech Republic, and Slovakia, and are on the verge of extinction in the USA and parts of southern Canada (World Wildlife Fund for Nature- WWF 2001). Ironically, farmed Atlantic salmon are plentiful, Parrish et al. (1998) stated that they made up approximately 98% of the biomass of Atlantic salmon. Reasons for the decline are manifold. Threats to wild Atlantic salmon include overfishing, construction of dams, river engineering schemes, pollution and salmon aquaculture (WWF 2001). There have also been declines in other salmonids,
with 20 species being categorized as threatened in the 2006 IUCN (International Union for the Conservation of Nature and Natural Resources or the World Conservation Union) Red List of Threatened Species, and five species have been listed as extinct (IUCN 2006).

As a result of declines in numbers of Atlantic salmon, a range of management strategies to restore populations has been developed, such as improvement of habitat, construction of fish ladders to bypass dams, liming rivers to combat acidification, regulation or cessation of fishing in the high seas (WWF 2001) and stocking various life stages that have been hatchery-reared (Youngson et al. 2003). Other measures that have been developed to reduce the loss of genetic diversity of wild stocks of Atlantic salmon involve ‘living’ and ‘frozen’ gene banks (i.e. maintaining subsets of stocks in tanks and cryopreservation of milt, respectively). Such schemes are being implemented in Norway and Finland (Walso 1998; Piironen and Heinimaa 1998). A ‘frozen’ gene bank of Atlantic salmon has also been set up in Iceland, in addition to the introduction of measures to reduce interactions between wild and farmed salmon (Ísaksson 1998).

Stocking may be carried out in order to, for example, mitigate the effects of lost habitat, or to re-establish populations (for instance following a rise in quality of the freshwater environment) (Aprahamian et al. 2003). Hatchery breeding programs are not without their drawbacks, however. It is currently recognized that stocking should be carried out with native fish in order to preserve local adaptations and population sub-structure, although historically, this has not been widely practised. In the past large numbers of fish were transferred among locations and much stocking was carried out with non-native fish (Youngson et al. 2003; Utter 2004). Even when native broodstock are utilised, there still remains the possibility of reducing the
overall effective population size (Waples and Do 1994; Ryman et al. 1995) and inadvertently promoting domestication selection (Heath et al. 2003). Monitoring is therefore important to determine the impact of such programs (Cuenco et al. 1993; Laikre 1999). Previously, monitoring of introduced fish using allozymes (which generally have a low level of polymorphism) required substantial differences in allele frequencies to be detectable between introduced and recipient populations, or the creation of genetically ‘tagged’ fish for stocking that were identifiable by otherwise rare alleles (e.g. Taggart and Ferguson 1984 in brown trout *Salmo trutta*; Chilcote et al. 1986 in steelhead trout *Salmo gairdneri*). Monitoring is now facilitated by genetic screening for highly polymorphic (informative) DNA markers, which allows for non-destructive sampling and greater ease of storing sampled tissue.

A number of challenges are faced when researching wild Atlantic salmon. They are aquatic and as such ‘are much more difficult to observe, study and enumerate than their terrestrial counterparts’ (Ryman et al. 1995). This is particularly relevant to the study of mating systems, as direct observation is generally not feasible. They are also migratory; work on Atlantic salmon during the marine phase is expensive and logistically difficult (Hawkins 2000). Physical markers have been developed, but in some cases identification of individuals using genetic methods has proven especially useful. Genetic ‘markers’ are permanent, inherited and present in all life stages. They are particularly powerful in the study of mating systems, providing a means to determine familial relationships. However, they are also technically difficult to screen, require laboratory work and are relatively expensive compared to physical markers.
1.2 Conservation genetics- a rapidly evolving field

Conservation genetics is a multifaceted subject, and is a subdiscipline of conservation biology (Avise 2004). The first text to include the role of genetics in conservation biology was published in 1981 by Frankel and Soulé (Frankham et al. 2002), and more recently, in 2000, a journal devoted to the topic (Conservation Genetics) was initiated (Avise 2004).

The IUCN recognises three levels of diversity that should be conserved- ecosystem diversity, species diversity, and genetic diversity (Frankham et al. 2002). Genetic diversity is ‘arranged’ by a number of factors, including selection, drift, inbreeding, migration and speciation (Avise 2004). In small populations drift and inbreeding occur at higher levels. Drift (the shift in allele frequencies due to chance sampling of gametes) results in the loss of genetic diversity, as by chance some alleles are not passed on to the next generation. Inbreeding (the breeding of closely related individuals) can result in a reduction in fitness known as inbreeding depression. An important concept in conservation genetics is the effective population size ($N_e$). It is this, rather than the census size, that determines the rate of inbreeding and drift. The effective population size is usually lower than the census population size due to factors such as unequal sex ratio and variance in reproductive success. The effective population size is the size of an ideal population (one which has for example random mating, constant size and random variation in reproductive success, Stearns and Hendry 2004) where the value ‘of some specified genetic property (Wang 2005)’ is equal to that in the actual population. One of the main aims of conservation programs involved with threatened species is to halt declines in, or where possible, to increase effective population size (Frankham et al. 2002). Another factor that affects genetic diversity is migration, or specifically the
migration of individuals between populations that successfully reproduce—also known as gene flow. An isolated population (one without gene flow) is unable to receive new genetic variation through migrants. Therefore maintaining connectivity between nearby populations in fragmented habitats can be important (Segelbacher et al. 2003). Genetic methods can be used to estimate both the effective population size and the level of gene flow among populations (for example, Fraser et al. 2004; Consuegra et al. 2005; Poulsen et al. 2006). This is however, just the tip of the iceberg.

Genetic methods can be applied to a diverse range of conservation issues. They can be utilized directly by monitoring the degree of genetic diversity in past and present populations. Genetic methods can be employed to determine taxonomic status and genetic differences among populations, which can be important in determining which populations or units to conserve. They can be used to gain a greater understanding of organisms, for example to provide insights into mating systems in the wild. Additional applications include monitoring of artificial breeding schemes and the introgression of non-native organisms, determining the true origin of illegally traded goods or animals and in determining the most beneficial procedure for captive breeding programs. Specific examples from salmonids include:

- Identification of the (mistaken) use of non-winter-run spawners in a supplementation program of winter-run chinook salmon. This resulted in measures being taken to prevent its re-occurrence (Hedrick et al. 2000).
- Identification of indigenous Atlantic salmon individuals in populations subject to stocking from exogenous populations. This would enable indigenous individuals to be chosen for restoration programs (Nielsen et al. 2001).
• Determination of the scale of population structure among tributaries of the Varzuga River in Russia. Isolation by distance of populations inhabiting different tributaries was found and genetic diversity was lower further from the river mouth. This led to: a) the recommendation that in order for long term preservation, multiple spawning and nursery sites should be maintained; and b) the suggestion that populations further from the river mouth might be more susceptible to negative effects of population reductions (Primmer et al. 2006).

Central to many of the aforementioned techniques have been relatively recent advances in screening highly variable genetic markers and in computational analyses of the resultant data. The development of suitable molecular markers has been crucial in advancing our understanding of biological systems. Allozymes were the first widely applied molecular markers. Allozymes are protein variants distinguishable by their electrophoretic mobility, that are coded for by single loci (Beebee and Rowe 2004). They have low numbers of variants, are the products of expressed genes so are not necessarily neutral and require fresh or frozen tissue samples for screening (Schlötterer 2004). The next step in the development of molecular markers involved surveying DNA-based variation directly. In the first instance restriction fragment length polymorphisms (RFLPs) were applied. RFLPs are detected using restriction enzymes to cut DNA at specific sites (restriction sites) and separating the resulting fragments using electrophoresis. Base variability at the restriction site, generally 4-6 base pairs (bp) in length, can cause the addition or loss of a restriction cleavage point, which in turn alters the resultant fragment banding pattern. RFLPs have been widely applied to mitochondrial DNA in phylogeographic studies and widely used in salmonid studies (e.g. Palva et al. 1989; Hynes et al.)
Introduction

1996; Nielsen et al. 1996; Prodöhl et al 1997; Nilsson et al. 2001). Nuclear DNA may also be assayed with RFLPs, either by using Southern blotting and hybridisation of a probe specific to a particular sequence, or more recently by amplifying a portion of nuclear DNA using PCR and subsequent RFLP analysis of the resulting product. However, despite the utility of allozymes and RFLPs in studying differences between populations, they are seldom adequately polymorphic to detect differences between individuals (Beebee and Rowe 2004). Minisatellite loci, on the other hand, are ‘hypervariable,’ generally exhibiting a high degree of polymorphism. Minisatellites are tandemly repeated stretches of DNA, with the repeat unit ranging from ~10-70bp (Avise 2004) and the total length being up to ~30kb or more (Chambers and MacAvoy 2000). Minisatellites are generally too large to be amplified with PCR, so require large amounts of isolated DNA for analysis. In addition, the banding patterns (‘fingerprints’) from multilocus methods are difficult to interpret, and although single locus minisatellites can be developed this proved ‘technically complex and difficult’ (Beebee and Rowe 2004). Both multilocus and single locus minisatellite screening methodologies have been developed and applied to salmonid studies (e.g. Fields et al. 1989; Taggart and Ferguson 1990 a,b; Prodöhl et al. 1994 a,b; Taggart et al. 1995; O’Reilly and Wright 1995; Galvin et al. 1996; Prodöhl et al. 1997). More recently developed markers rely on PCR amplification. These include microsatellites, SNPs (single nucleotide polymorphisms), AFLPs (amplified fragment length polymorphisms) and RAPDs (randomly amplified polymorphic DNAs).

SNPs can be highly informative when large numbers are typed and have the potential for automated high-throughput typing, but are expensive to develop for a particular organism and so are generally more applicable to model organisms.
(Schlötterer 2004). In the future SNPs may become more widely applied to well-studied organisms such as Atlantic salmon, where large amounts of sequence data are available; a project known as GRASP (Genomics Research on Atlantic Salmon Project) has identified \(>51000\) sequences from Atlantic salmon (Rise et al. 2004). Such data may also be applied in order to develop SNPs in other salmonids (Smith et al. 2005).

AFLPs are analysed by restriction of the DNA with restriction endonucleases, addition of linkers to the ends of the fragments and subsequent selective amplification using PCR. RAPDs are essentially randomly amplified sections of DNA - primers of arbitrary sequence are used to amplify regions of DNA. A limitation of AFLPs and RAPDs is that they are generally dominant - heterozygotes cannot be distinguished from homozygotes (Avise 2004). In salmonids the use of AFLPs has mainly been confined to linkage mapping (e.g. Nichols et al. 2003; Moen et al. 2004).

The marker of choice for many studies is the microsatellite. Microsatellites are short tandemly repeated stretches of DNA, with each repeat unit being between 2 and 6bp long (Chambers and MacAvoy 2000), and usually being twenty to a few hundred base pairs in total (Chistiakov et al. 2006). Their popularity is due to a number of features, which include being widespread throughout genomes, highly polymorphic, relatively easy to screen through PCR amplification and relatively easy to isolate (Schlötterer 2004; Chistiakov et al. 2006; Selkoe and Toonen 2006). They are also generally considered neutral markers, due to their distribution mainly in non-coding regions of DNA, although there are also numerous examples of functional microsatellites (Chistiakov et al. 2006). For example microsatellites have been implicated in chromatin organization, recombination, and in the regulation of gene
expression (Selkoe and Toonen 2006). In humans at least eleven human diseases are associated with the expansion of microsatellites in genes, including Huntington’s disease and Fragile-X syndrome (Mitas 1997). In addition, microsatellites that are not directly under selection may appear to be so due to hitchhiking i.e. being in the vicinity of a region of DNA that is under selection. Since microsatellites can be amplified using PCR, they can often still be amplified from small amounts of DNA, even when this is partially degraded (Selkoe and Toonen 2006). They can therefore be used to obtain information from DNA extracted from a wide range of samples, including historical (e.g. Pertoldi et al. 2001; Nyström et al. 2006), scat (e.g. Piggott et al. 2006; Smith et al. 2006), and hair samples (e.g. Marshall and Ritland 2002). For studies on populations where there are low numbers of individuals, microsatellite markers can be particularly pertinent because they provide a non-destructive, non-invasive way of obtaining information. Indeed, the study organism may not even have to be handled- Marshall and Ritland (2002) used ‘hair snares’, which consist of barbed wire placed around a scent lure, to obtain hair samples from Kermode bears Ursus americanus kermodei.

In addition to the development of molecular markers, there have been advances in the methods available to analyse the datasets generated. Such datasets are now often large and complex. Increases in computational power have allowed the application of computer intensive statistical methods, namely in the use of simulation techniques in model-based analyses (Marjoram and Tavare 2006). As such, there is now an array of software available for the analysis of population genetics data, with applications ranging from the estimation of effective population size to assigning individuals to populations to estimating the time of divergence of populations (reviewed in Excoffier and Heckel 2006).
1.3 **Introducing Girnock Burn**

‘The wise procedure is for law to delay, science to evaluate, and familiarity to preserve. There is an implicit principle of human behaviour important to conservation: the better an ecosystem is known, the less likely it will be destroyed.’

E.O Wilson (1994)

Wilson eloquently describes how a more detailed understanding of organisms can be intrinsically important, aside from being applicable to current and future management and conservation schemes.

Girnock Burn provides a natural setting in which to study numerous aspects of Atlantic salmon biology. Permanent traps have allowed extensive monitoring of fish travelling both upstream and downstream, since the traps were constructed in 1966. Girnock Burn is an upland spate tributary of the Aberdeenshire Dee, in Scotland (a map is given in Chapter 2). The stock of Atlantic salmon in the Dee is of particular importance because it has an early running component. It is the higher reaches of the Dee, such as Girnock Burn, that are likely to be spawning grounds for these fish, as shown by radio-tracking of adult sea-returns in a study carried out by Laughton and Smith (1992).

A series of dedicated teams have monitored the traps since 1966, despite the often unfavourable weather conditions. Unfortunately the traps are not maintenance free, particularly the trap catching fish travelling downstream; its configuration must be tailored to the flow, and leaves and debris cleared from the trap in an attempt to prevent ‘leakage’. If the trap does become clogged with leaves and debris, then water and possibly fish may ‘leak’ over the top of the trap. In addition to maintaining the traps in a functional state, there is the task of processing the fish.
Introduction

Since 1966 the length and sex of anadromous returns entering the upstream trap have been recorded and scale samples taken (the upstream trap is so-called because of the direction of travel of the fish it intercepts, rather than its position in relation to the downstream trap). Since 1991 adipose fin clips of anadromous returns have also been retained for potential DNA profiling studies. At the downstream trap, parr and smolts trapped since 1966 have also had their length recorded and a scale sample has been routinely taken from one in every five individuals.

Such long-term monitoring has allowed changes to be identified over the years. Generally, there has been a trend towards earlier out-migration of smolts from the burn (Langan et al. 2001), and a downward trend in the numbers of 2SW females returning to the burn (Gurney 2004, unpublished report). In addition, a vast amount of information has been collated on various aspects of Atlantic salmon biology, including; consideration of the homing of sea returns to Girnock Burn (Youngson et al. 1994), the spawning success of mature parr (Taggart et al. 2001), the growth rates of parr (Jones et al. 2002), a detailed appraisal of the numbers of returning adults and outgoing parr and smolts (Gurney 2004, unpublished report), the distribution of spawning in relation to geomorphic characteristics (Moir et al. 2004) and the effects of groundwater on embryo development (Youngson et al. 2004).

Such studies have been complemented by the evaluation of physical characteristics: temperature trends over 30 years (Langan et al. 2001), the effect of woodland on temporal and spatial variation in temperature (Malcolm et al. 2004), and the interaction between groundwater and surface water (Soulsby et al. 2005). However, despite the plethora of work, many questions remain unanswered, and in some cases more have arisen as a result of new insights.
1.4  **Aims of the study**

An existing collection of archived tissue samples, obtained from anadromous returns to Girnock Burn and parr sampled in 1997 by electrofishing, combined with the high discriminatory power of microsatellites, provided a valuable resource that was exploited in the current study. Additional samples of smolts were also examined. The intention was to extract data that might assist in the future management and conservation of Atlantic salmon in Girnock Burn, in addition to providing further insights into the characteristics of the stock. The five main aims are described below.

1.4.1  **Accuracy of genotyping and other techniques**

A substantial portion of the analyses conducted for this thesis was based on parentage assignment. Previously Taggart *et al.* (2001) utilized single locus minisatellites to determine the parentage of egg samples from Girnock Burn. However, this was a lengthy process and there was ambiguity and redundancy at some loci. With the development of techniques to type microsatellite loci with a high throughput (using multiplex PCR and automated size determination of alleles), along with the high polymorphism of these markers, microsatellites offered a more favourable approach. Despite these advantages, it was still imperative that parentage assignment was accurate. One aim was therefore to evaluate the accuracy of genotyping and subsequent parentage analysis. An investigation was also made into the accuracy of ageing fish using scale readings.

1.4.2  **Sampling juveniles- the spatial distribution of kin**

Sampling juveniles for population analyses can be problematic as it can result in the comparison of families rather than populations. This issue was investigated in
Girnock Burn by examining the distribution of juvenile kin within the burn, and how sampling strategy affected the results of standard population genetic analyses.

1.4.3 Natural and quasi-natural spawning

Taggart et al. (2001) studied the spawning success of anadromous returns in 1991, 1992 and 1995 by determining the parentage of eggs collected from redds, using minisatellite-based analyses. The authors revealed that, in three years of spawning that were examined, a proportion of anadromous males and females were not represented in the eggs sampled (for example in two of the years, 62% and 64% of potential female spawners were not detected). Possible reasons for non-detection included that they had not spawned, that redds had not been sampled, or that sampling of redds was not sufficiently extensive to have included them. One objective of the current study was to investigate further the spawning success of anadromous returns in one of the years studied, but using samples obtained from outwith the redds. The intention was to determine which of the possible reasons was (or were) most likely to account for non-representation of anadromous spawners, in order to gain further information on spawning dynamics.

An additional aspect of this project was to compare the output of a supplementary spawning program with the output from natural spawning in the same region. At Girnock Burn a process known as ‘quasi-natural’ spawning was introduced in the year 2000 following low numbers of anadromous adults returning to the burn (the number of females had ranged between 4 and 11 between 1997-1999). Quasi-natural spawning involves stripping fish of milt and eggs, constructing artificial crosses, incubating fertilised eggs over winter and planting out as eggs the following spring. In the years 2000 and 2001 this was conducted alongside natural spawning.
In order to monitor the program, the reproductive success of spawners (up to smolt stage) was determined and compared between the two schemes.

### 1.4.4 Straying and trap leakage

The reproduction of animals at the same ‘site’ as their parents is known as philopatry and is a common feature of salmonids. This can result in the evolution of populations adapted to the local environment. In anadromous salmonids this necessitates the ability to home to the natal site prior to spawning. However, homing is not perfect, with straying occurring to some degree in the vast majority of studies on anadromous salmonids to date (for example Appendix 1 in Hendry and Stearns 2004). At Girnock Burn the ability to identify philopatric anadromous returns is intrinsically related to the degree of trap leakage. All parr and smolts that are caught in the trap are marked as they leave the burn by clipping (removal) of the adipose fin. Youngson et al. (1994) found approximately 45% of anadromous returns to the burn bore this mark. However, the origin of unclipped fish is unknown. They could originate from just below the traps within Girnock Burn, or from further afield. Another possibility is that they originate above the traps, but due to trap leakage have passed over the downstream trap and escaped clipping. Part of the current study was concerned with determining the origin of unclipped returns and therefore inferring the degree of trap leakage. Such information is necessary in order to interpret the trap data reliably. Parentage analysis was used to determine whether any unclipped anadromous adults could actually be assigned to putative parents from previous years that had been placed above the traps. If they could, it suggested that trap leakage had occurred. In addition, an assignment technique (that assigned individuals to populations) was used to determine whether it was likely that any anadromous returns had not originated from the stock in Girnock Burn.
1.4.5 **Implications of reproductively successful mature parr**

In Atlantic salmon, mature male parr have the potential to fertilise eggs without having experienced the marine environment, smoltification, or the return migration. Despite considerable interest in the influence of mature male parr on effective population size, and factors affecting the sexual maturation of males as parr, the potential role of sexually mature parr in the adaptive response of populations to the environment has not been addressed in detail. It was hypothesised that since mature parr are able to reproduce without having survived the marine environment, they may pass on alleles that confer adaptation to the freshwater environment, but which are not necessarily favourable for anadromy. Samples from Girnock Burn provided the opportunity for an initial exploration into the evolutionary implications of mature male parr in the adaptation of populations to the environment.

An additional aspect of male parr maturity investigated was the relationship between the sex ratio of anadromous spawners and the collective reproductive success of mature parr.

1.4.6 **Genetic diversity over time and effective population size**

Comparisons of the genetic composition of temporally spaced samples have been conducted in the past in order to monitor conservation efforts and to gain further insights into the evolutionary dynamics of populations. Genetic analysis of samples obtained at a single point in time can also be used to test for a recent bottleneck. A further aspect of this project was a temporal evaluation of genetic diversity. This was of interest given the decline in numbers of returning 2SW females and the low numbers of female returns between 1997 and 2003 (not more than 27). In addition it could provide a useful baseline for future studies. Two methods were also employed in order to test for evidence of a recent bottleneck.
Different methods of estimating $N_e$ have recently been developed that are based on genetic data obtained from two or more temporally spaced samples. A range of methods was applied to the data obtained from Girnock Burn. The methods were compared and the utility of such estimates was evaluated.
1.5 Main objectives

The main objectives of the study were to:

- Develop a robust multilocus assay
- Investigate the level of accuracy of genotyping
- Evaluate the accuracy of ageing adults through scale reading
- Examine how sampling juveniles might affect the results of analyses of population structure
- Provide more information on spawning success in the wild, to complement previous data
- Compare the output of quasi-natural and natural spawning
- Examine in greater detail the origin of anadromous adults returning to the burn that were unclipped (i.e. that were supposedly non-philopatric)
- Consider the possible evolutionary implications of mature parr in adaptation to the environment and the possible effects of climate change (with respect to mature parr)
- Examine genetic diversity through time
- Compare different methods of estimating the effective population size and consider the utility of estimating this.
2 Materials and methods

2.1 Study site

Sampling took place in Girnock Burn, a tributary of the Dee on the East coast of Scotland, approximately 80km from Aberdeen (Figure 2.1). It is an upland spate tributary, with a mean width of approximately 6.6m (Webb and Hawkins 1989) and a length of c. 9km. The lower reaches are bordered by cultivated fields and rough pasture while the upper reaches are largely encompassed by heather (Buck and Hay 1984). Trees are only present in the lower 3km of the burn, and result in decaying leaves settling in this region (Buck and Hay 1984). In addition to salmon, brook lampreys (Petromyzon planeri), eels (Anguilla anguilla), minnows (Phoxinus phoxinus) and a small number of freshwater resident and anadromous trout (Salmo trutta) are present in the burn (Buck and Hay 1984). The mean temperature of the burn in summer is 12.5°C and 2.12°C in winter (Langan et al. 2001). The mean temperature of water in spring increased significantly between 1968 and 1997, which may have been due to a decrease in snowfall leading to a reduction in cool meltwater entering the burn in spring (Langan et al. 2001).
Figure 2.1 Map of Girnock Burn and its position in relation to the Dee and Aberdeen (inset). The position of the fish traps is marked by an open box. Taken from Taggart et al. (2001).

2.1.1 Salmon stock in Girnock Burn

Girnock Burn supports both grilse (~36%) and multi-seawinter fish (~64%). The majority of fish at Girnock Burn spend two to three years in freshwater and between one and three years at sea (according to scale reading this was 100% of fish that returned to the burn between 1991 and 2004). Grilse tend to be males while the females are mainly multi-sea-winter fish. The stock in the Dee is of particular importance because it has an early running component. It is the higher reaches of the Dee, such as Girnock Burn, that are likely to be spawning grounds for these fish, as shown by radio-tracking of adult sea-returns in a study carried out by Laughton and Smith (1992). They
found that spring running multi-sea winter fish and early running grilse migrated to higher reaches of the river. Anadromous fish ascending the burn tend to enter the upstream trap in October and November. This is partly dependent on water flow, with increased flow usually coinciding with salmon moving up into the burn. Spawning takes place in November and eggs generally hatch in April (Bacon et al. 2005).

The majority of smolts leave the burn in spring (March-May), whereas the remainder, approximately one third, leave in autumn and early winter (September-January), although small numbers of fish leave the burn throughout the winter (Youngson et al. 1983). Those that leave in autumn still have a parr like appearance (Youngson et al. 1983) and include mature parr (Youngson and Hay 1996). Precise numbers of male parr that mature in freshwater have not been quantified. Tagging studies have shown that parr leaving in the autumn reach the mouth of the Dee around the same time as smolts that leave Girnock Burn the following spring (Youngson and Hay 1996).

Numbers of females caught in the Girnock Burn traps during the annual spawning runs, since 1966 are shown in Figure 2.2. Data on the numbers of smolts leaving the traps and adults returning to them in different years have been analyzed by Gurney (2004, unpublished report). This review showed that numbers of smolts leaving the burn were variable, but did not show a downward trend. In comparison the numbers of female 2SW sea returns were also variable but showed a significant downward trend between 1970 and 2003. The burn reaches full capacity with eggs from 30-40 fish (FRS 2005), but fewer than 30 females returned each year between 1997 and 2003.
2.1.2 Traps and routine sampling

An upstream trap is in operation that allows continuous monitoring and sampling of sea returns. A downstream trap permits parr and smolts traveling downstream to be captured. The traps are sited approximately 0.9km upstream from where the Girnock Burn joins the Dee mainstem, leaving approximately 8km above the traps available for navigation/ spawning. The upstream trap (Figure 2.3) consists of a barrier across the burn through which water may flow but adult fish should not be able to pass. These are funneled into a holding cage in which they should be trapped. The barrier is not perfect, as anadromous fish are occasionally found above the upstream trap (but prevented from further movement upstream by the downstream trap).
Figure 2.3 The upstream trap at Girnock Burn. The holding cage is visible on the far left hand side of the figure. The downstream trap is just visible in the top right hand corner.

Figure 2.4 The downstream trap at Girnock Burn. Only the ‘sieve’ at the end furthest from the viewer is in action.
The downstream trap (Figure 2.4) is based on a Wolf trap (Wolf 1951). It is a converted weir, the water flowing down through a meshed screen. This acts as a sieve, preventing small fish from flowing through. Enough water remains above the sieve to allow the fish to be transferred to a channel of water and subsequently a holding box. The area of ‘sieve’ in action can be varied to match water flow. However, the trap can overflow in times of very high water flow, particularly when clogged with fallen leaves and other debris.

The traps have been in operation since 1966 and there are data on the numbers of adult returns since this time. Routine sampling has involved recording the length of sea returns, their weight and taking scale samples (for aging). Adipose fin clips have also been taken and archived since 1991, for DNA profiling analyses. Routine sampling of fish caught in the downstream trap involves recording the length and weight and taking scale samples from one in five fish. In addition, all fish are microtagged and have the adipose fin removed (but this is not routinely kept) as a way of tagging fish leaving the burn. Since the adipose fin does not completely grow back, it allows the identification of anadromous adults returning to the burn that are philopatric (i.e. originated above the traps at Girnock Burn).

2.1.3 Management of the population

Until 2000, adult returns to the upstream traps were placed upstream (above the downstream trap) to spawn naturally or occasionally released downstream (e.g. if spent or diseased). Due to declines in numbers of 2SW females returning to the burn the decision was made to assist natural spawning by ‘quasi-natural’ spawning. This is described in more detail in Chapter 4. Briefly, sea returns are stripped of eggs and milt,
eggs fertilised and ‘incubated’ in the burn in baskets over winter and then planted throughout the length of the burn in the following spring.

### 2.2 Sampling

Sampling was carried out by Fisheries Research Services (FRS) staff, based at Pitlochry. When this project was initiated in April 2003 the majority of samples required for the study were already archived. The samples utilized, how they are related and the aspects of the project they are associated with, are shown in Figure 2.5.

#### 2.2.1 Anadromous fish

Since 1991 tissue samples from sea returns entering the upstream trap have been sampled. These biopsy samples comprise a punch 5mm in diameter removed from the adipose fin. In cases where the adipose fin had been removed as a juvenile a tissue sample of the remaining ‘stub’ was removed with a scalpel or single edge blade.

Samples were preserved in 99% ethanol (at 4°C) until extraction of DNA.

Samples from sea-returns utilized in this study were those that returned to Girnock Burn in 1991, 1995, 1999, 2000, 2001 and 2004. Only ten fish aged 2.1 (or that possibly belonged to this age group), as determined by scale reading, were utilized from 1999. These were selected because they could have been offspring of anadromous adults that returned in 1995 and thus were used in the estimation of the extent of trap leakage. Table 2.1 shows the number of samples from each year that were genotyped and the numbers that were either released above the traps to spawn naturally or were used in quasi-natural spawning.
Materials and methods

Chapter 2

Figure 2.5 Project samples: numbers of fish genotyped, their relationships and how they were utilised.

Key

The age is the number of years spent in freshwater followed by the number of years in seawater. For example 2.1 relates to two years in freshwater followed by one at sea. The type of box on the left shows the age of fish linked to different samples by the arrows. For example the top arrow shows fish spawned by returns to Girnock Burn in 1991, aged 2.1, may have returned to Girnock Burn in 1995.

The type of box on the left represents samples of anadromous adults. It shows the year they returned to Girnock Burn and the number of samples and how they were utilised (code A-E) in parentheses.

The type of box on the far right represents samples of fish taken before they had left freshwater. It shows the year they were sampled, the life stage of the fish and the number of samples and how they were utilised (code A-E) in parentheses.

A Comparison of success of quasi-natural and natural spawning
B Estimation of the level of downstream trap leakage.
C Comparison of the survival of offspring of mature parr versus offspring of anadromous adults in freshwater.
D Assessment of the level of genetic diversity over time and estimation of $N_e$.
E Assessment of the distribution of kin throughout the Burn, at the parr life stage.

Arrows show which samples may contain individuals that are related, the start of the arrow shows putative parents while the arrowhead points to putative offspring.
Table 2.1 Numbers of potential anadromous salmon spawners used in the study.

<table>
<thead>
<tr>
<th>Year</th>
<th>Numbers of samples genotyped</th>
<th>Numbers of potential contributors placed above traps&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number females</td>
<td>Number males</td>
</tr>
<tr>
<td>1991</td>
<td>56</td>
<td>39</td>
</tr>
<tr>
<td>1995</td>
<td>71</td>
<td>45</td>
</tr>
<tr>
<td>1999&lt;sup&gt;3&lt;/sup&gt;</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2000</td>
<td>27</td>
<td>62</td>
</tr>
<tr>
<td>2001</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>2004</td>
<td>61</td>
<td>56</td>
</tr>
</tbody>
</table>

<sup>1</sup> Includes natural and/or quasi-natural spawning

<sup>2</sup> Tissue sample was lost from one female.

<sup>3</sup> Only those aged 2.1, or ?.1 (according to scale readings) were included.

In the majority of years (1995, 2000, 2001 and 2004) more fish were genotyped than were released above the traps or used in quasi-natural spawning. This provided a full dataset for population analyses and insurance against sample mis-labelling.

2.2.2 Eggs sampled in 1995

Archived DNA from a subset of eggs sampled in 1995 by Taggart <i>et al.</i> (2001) was used. The subset included DNA from a total of 60 eggs (i.e. 10 eggs from each of six redds). Taggart <i>et al.</i> (2001) described how the eggs were sampled and the method of DNA extraction. Briefly, eggs preserved in 99% ethanol were dissected and the embryo teased from the yolk sac. DNA was extracted from the embryo using a phenol-chloroform based method (Taggart <i>et al.</i> 1992).

2.2.3 Parr sampled in 1997

A total of 274 parr were sampled by electrofishing along the length of the burn above the traps, in September 1997. Samples were picked by size to be presumed 1+ parr (i.e. offspring of adults that returned to the burn in 1995). One fish was sampled
approximately every 30m, the adipose fin removed, scale samples taken, and the fish returned to the burn. Samples were preserved in 99% ethanol until extraction of DNA.

### 2.2.4 Smolts sampled in 2004

Smolts were sampled between 3\textsuperscript{rd} February and 6\textsuperscript{th} May in 2004. The adipose fin clip of one in every five smolts passing through the traps was removed and preserved in 99% ethanol. Their length and weight was recorded and scale samples taken. A total of 410 smolts had adipose fin clips preserved.

### 2.3 DNA extraction

Three different methods of DNA extraction were employed on samples in this study. DNA had already been extracted using a simplified phenol-chloroform based protocol (described in Taggart \textit{et al.} 1992) from adults that returned in 1991 and 1995 and from eggs sampled in 1995. DNA was extracted from samples of parr collected in 1997 using a simplified chelex protocol (detailed in Appendix, section A1.1). The chelex protocol was favoured because of the absence of harmful organic solvents and being substantially less labour intensive. The majority of extractions were performed in a 96-well plate format. However, PCR amplification was less consistent when using chelex extracted DNA template compared to phenol-chloroform extracted DNA. A more reliable method that would be more resilient to long term storage, but did not involve phenol and chloroform, was sought. Finally a commercially available kit was used (the Wizard® SV genomic purification system produced by Promega; method detailed in the Appendix, section A1.2). For samples that had not already been extracted, a biopsy punch (diameter 3mm, Krusse, Catalogue number 273690) was used to remove a standard sized piece of tissue for extraction, from each archived sample. This was done
in an effort to standardize the amount of DNA extracted. To check that DNA from
different extractions was of high molecular weight, 5-7µL were run on an agarose gel
(2% agarose, 200µg/ml ethidium bromide in 1× TBE) and viewed under UV.

2.4 Genotyping

2.4.1 Use of different fluorescent dyes and fragment analyzers.
During the course of the project the equipment used to determine microsatellite
genotypes was upgraded. Initially two ABI 377 slab gel systems were provided.
Subsequently, in November 2004, these were replaced by two Beckman Coulter
CEQ8800 capillary gel fragment analyzers. The transition between platforms not only
necessitated changes in working practice, but also the calibration of allele sizes (due to
differing gel matrix, software, detection dyes and size standard).

Two main parental assignments were carried out during the project. In order to reduce
potential problems, DNA profiling required for each investigation was carried out using
the same genotyping platform. Thus, the assignment of parr sampled in 1997 to adult
returns of 1995 was carried out with genotypes scored using the ABI 377 hardware and
associated software. Subsequently the assignment of smolts sampled in 2004 to adult
returns in 2000 and 2001 was carried out with genotypes obtained using the Beckman-
Coulter CEQ8800 system.

Other parts of the analysis required the alleles to be comparable between samples typed
on different platforms. This included the comparison of genetic diversity over years (for
all adult returns) and the assignment of sea returns in 1999, 2000 and 2001 to returns in
1995. In order be able to compare genotypes gained from different platforms, samples
from adults that returned in 1995 were re-typed using the Beckman Coulter CEQ8800
fragment analyzers and the ‘new’ set of multiplexes. Bins were then set up and the alleles given names that corresponded to those for the original typing. Bins did not overlap and were centered on the size of the most commonly observed fragment for a particular allele. Eight random samples from adults that returned in 1991 were also re-typed using the new set of multiplexes and Beckman Coulter CEQ8800 fragment analyzers. This acted as an additional test of the similarity of typing using the two systems.

2.4.2 Controls and estimation of error rate
In order to gauge the accuracy of genotyping, control samples were employed and a subset of samples was re-typed. For the genotyping carried out on ABI fragment analyzers, control samples were primarily used to verify correct loading and tracking (on slab gels) i.e. to ensure that each lane was identified correctly. Four samples of known genotype were run on each gel.

Three control samples were routinely analysed on every genotyping run carried out on the Beckman-Coulter fragment analysers. This provided a test for the consistency of size calling for alleles (albeit for a limited number of alleles). There was less concern with regards to sample tracking since the capillary system did not require software-based identification of lanes. Nevertheless the controls did provide a check that samples were correctly identified within the 96 well plate. Since the Beckman Coulter fragment analyzers were new, controls were especially useful in monitoring their accuracy and the way in which they were used.

An estimate of the error rate of genotyping was obtained by re-typing samples. Randomly selected samples from smolts sampled in 2004 and adult returns in 1991
were re-amplified and re-typed. This was carried out with approximately 5-6% of samples. (Re-typing a subset from the returns of 1991 also acted as a test of the similarity of genotypes obtained by typing with different multiplexes on different machines).

2.4.3 Estimation of the number of loci required
Prior to genotyping, simulations were carried out in PAPA (Duchesne et al. 2002) and FAP (Taggart 2007) to estimate the number of loci required for confident parentage assignment. In the absence of microsatellite information, genotypes from single locus minisatellites already typed for this stock (6-22 alleles segregating) in a previous study (Taggart et al 2001) were utilised. These suggested that 9-12 loci would be needed for unambiguous assignment. These were regarded as approximations, as neither the distribution nor number of alleles of the loci actually being employed, nor the total number of candidate parents, were known precisely.

2.4.4 Choice of Loci
No single large panel of microsatellite loci was available for Atlantic salmon. In selecting a panel of microsatellite markers suitable for parentage analysis in this study the following criteria were deemed important. The loci should;

- Amplify clearly and produce reliable results.
- Eventually be amplified in multiplexes (whereby more than one locus was amplified in a single PCR, also termed ‘true multiplexing’). Multiplexing requires fewer freeze-thaw cycles of samples (in comparison to amplifying loci singly) and there is reduced opportunity for error in typing (compared to combining PCR product from different amplifications when running on a gel). It is also less labour
intensive and for screening large numbers of individuals it is also more cost effective.

The following loci (Table 2.2 and Table 2.3) were initially amplified to determine their suitability.
Table 2.2 Initial panel of microsatellites (not known to be combined in true multiplexes) that were screened for use in this project.

<table>
<thead>
<tr>
<th>Loci tested</th>
<th>Primer sequence (F-Forward/R-Reverse)</th>
<th>Source</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ssa401UOS</td>
<td>ACTGGTTGTTGCAAGAGTTTGATGC (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa401UOS</td>
<td>AAACATACCTGATTCCCCGAACCAG (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa404UOS</td>
<td>ATGCAGTGAAGAGGGGTAAAC (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa404UOS</td>
<td>CTCTGCTCTCCTCTGACTCTC (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa407UOS</td>
<td>TGTGTAGGCAGGTGGAC (F)</td>
<td>Cairney <em>et al.</em> (2000)</td>
<td>Clear tetranucleotide repeats, wide allele range</td>
</tr>
<tr>
<td>Ssa407UOS</td>
<td>CACTGCTGTTACTTGGTGATTCC (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa408UOS</td>
<td>AATGGATTACGGGTACGTTAGACA (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa408UOS</td>
<td>CTCTTGTGCAAGGTTCTTCATCTGGT (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa410UOS</td>
<td>GGAAATAATCAATGCTGCTGGTT (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa410UOS</td>
<td>CTACAATCTGGACTATTTCTCTCA (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JMS1</td>
<td>GGATGGGAATCTCTTTGTTG (F)</td>
<td>Estoup <em>et al.</em> (2000)</td>
<td>Di-nucleotide repeats, unable to distinguish all alleles clearly</td>
</tr>
<tr>
<td>JMS1</td>
<td>GGACAGAGACGGATGTG (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SsaG7</td>
<td>CTTGGTCCCGTTACTGACAACC (F)</td>
<td>Verspoor (unpublished)</td>
<td>Di-nucleotide repeats, unable to distinguish all alleles clearly</td>
</tr>
<tr>
<td>SsaG7</td>
<td>TGCACGCTGCTTGGTCCTTG (R)</td>
<td>Paterson <em>et al.</em> (2004)</td>
<td></td>
</tr>
<tr>
<td>One9ASC</td>
<td>GCATGTCTCTGACAGCCTACAGCT (R)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.3 Initial panel of microsatellites (known to be combined in truemultiplexes) that were screened for use in this project. Three multiplexes were tested, which each included two or three loci.

<table>
<thead>
<tr>
<th>Loci tested</th>
<th>Primer sequence (F-Forward/R-Reverse)</th>
<th>Source</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ssa171</td>
<td>GAGGTCGCTGGGTTTACTAT (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa171</td>
<td>TTATTATCCAAAGGGGTCAAAA (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa197</td>
<td>TGGCAGGGATTTGACATAAC (R)</td>
<td>O’Reilly et al. (1996)</td>
<td>Tetranucleotide repeats, robust multiplex</td>
</tr>
<tr>
<td>Ssa197</td>
<td>GGGTTGAGTAGGGAGGCTTG (F)</td>
<td>Plant (2000)</td>
<td></td>
</tr>
<tr>
<td>Ssa202</td>
<td>TTCATGTGTAATGTTGCGTG (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa202</td>
<td>CTGGAATATCTAGAATAGGC (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa2215</td>
<td>ACTAGCCAGGTCCTGCCGGT (F)</td>
<td>Verspoor (unpublished)</td>
<td>Problematic spurious product in multiplex</td>
</tr>
<tr>
<td>Ssa2215</td>
<td>AGGGTCGTCAGTCACACCATGCAC (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa2216</td>
<td>GGCACACAGATAAACAAACACGC (F)</td>
<td>Paterson et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>Ssa2216</td>
<td>GCCAACAGCAGCATCTACACCCAG (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa2210</td>
<td>AAGTATTCATGCACACACATTCTACG (F)</td>
<td>Verspoor (unpublished)</td>
<td>Clear tetranucleotide repeats, robust multiplex</td>
</tr>
<tr>
<td>Ssa2210</td>
<td>CAAGACCCTTTTCCAAATGGGATTTC (R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa1605</td>
<td>CGCAATGGAAGTCAGTGACTGG (F)</td>
<td>Verspoor (unpublished)</td>
<td></td>
</tr>
<tr>
<td>Ssa1605</td>
<td>CTGATTAGCTTTTATAGTGCCCAATGC (R)</td>
<td>Paterson et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>Ssa2201</td>
<td>TTAGATGTTGGAGTACTGGGAGGC (F)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssa2201</td>
<td>CCGGAGCCCCATAACCCTACTAATAAC (R)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The multiplex of Ssa2215 and Ssa2216 was found to produce a spurious product, which was the size of an actual allele and was therefore not used. In general, the di-nucleotide microsatellites tested were not considered suitable for the current study. More pronounced stutter bands made routine calling of alleles difficult, particularly for heterozygotes with similar sized alleles. However, one exception was One9 (a di-
nucleotide microsatellite originally isolated from sockeye salmon, *Oncorhynchus nerka*). The allelic PCR products resolved very clearly (only 5 alleles in total) and this locus was included in the DNA profiling panel. The final loci were selected based on their compatibility in terms of annealing temperature and size range for true multiplexing, and only loci that could be amplified clearly were employed.

2.4.5 Optimization of multiplexes.

Octavian Henegariu’s online information ([http://info.med.yale.edu/genetics/ward/tavi/Guide.html](http://info.med.yale.edu/genetics/ward/tavi/Guide.html)) was used as a guideline in optimizing multiplex PCRs. Generally, the final buffer and dNTP concentrations were kept constant, while the concentration of MgCl$_2$, template DNA and primers were varied along with thermocycling conditions. If product from a particular locus was low, the concentration of the primer was increased. If it remained low the annealing temperature was altered or the number of cycles increased. If unspecific products were obtained annealing temperature was increased.

A number of combinations of loci were considered for co-amplification. Combinations were initially chosen based on their compatibility in terms of allele size range, annealing temperature and the MgCl$_2$ concentration found to be optimal for their amplification individually. Combinations of primers were checked for complimentary sequences, using the program QUICKPRI (DNASTAR) to inform about possible primer dimer production during amplification. The final multiplexes used and the loci they were composed of are presented in Table 2.4. The PCR components and conditions for these multiplex reactions are detailed in Appendix A1.3. The rationale for constructing the multiplexes in this manner are described later in section 2.4.6.
Table 2.4 Loci amplified in the various multiplex PCR combinations used in this study.

<table>
<thead>
<tr>
<th>Multiplex number</th>
<th>Loci included</th>
<th>Total number of loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2201 1605 2210</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>410 407 One9</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>197 202 171</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>404 408</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2201 1605 2210 410 407</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>404 408 171 One9</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>197 202 401</td>
<td>3</td>
</tr>
</tbody>
</table>

Multiplexes 5, 6 and 7 required re-optimization for use with DNA extracted from smolts (using a commercially available kit), as opposed to DNA extracted from adult fish (using phenol-chloroform based methods). Many conditions were altered in an attempt to optimize the reaction. Generally, extra cycles were included in the PCR and the concentration of Taq polymerase and or DNA was increased. Touchdown PCR was used in initial stages of the construction of multiplex 2, but this was finally optimized to a single annealing temperature.

The numbers of samples typed at each locus, from different years and at different life stages are shown in Table 2.5.
Materials and methods

Table 2.5 Number of samples from year groups that were typed with each locus. Unless otherwise stated, year groups refer to sea-returns

<table>
<thead>
<tr>
<th>Year &amp; sample taken</th>
<th>Locus 2210</th>
<th>1605</th>
<th>2201</th>
<th>407</th>
<th>410</th>
<th>One9</th>
<th>408</th>
<th>404</th>
<th>171</th>
<th>197</th>
<th>202</th>
<th>401</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991 (A)</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>18</td>
<td>18</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999 (A)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2000 (A)</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>2001 (A)</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>2004 (A)</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td>117</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 sample type: A - anadromous adult; P – parr; S - smolt

2.4.6 Empirical confirmation of the number of loci required for robust parental assignment

Initially the nine loci included in multiplexes 1-3 (Table 2.4) were tested empirically by attempting to assign 60 eggs to parents that returned in 1995. This gave 100% concurrence with a previous assignment using minisatellites, described in Taggart et al. (2001). Following this, 274 parr (also offspring of adults that returned in 1995) were typed with this set of nine loci and parentage analysis was carried out.

Assignment was ambiguous in nine cases. These samples and candidate dams were typed with two extra loci (included in multiplex 4, Table 2.4), which resolved all ambiguities. A third assignment sample set was then considered; the assignment of adults that returned in 1995 to potential parents that spawned in 1991 (i.e. treating adults that returned in 1995 as offspring and adults that returned in 1991 as parents). In this analysis, not all putative dams were collected, so potential ambiguous assignments would not necessarily be apparent. Therefore dams assigned with no corresponding
anadromous male (the sire assumed to be a mature parr), using nine loci, were typed with two extra loci (included in multiplex 4, Table 2.4). The same two extra loci were also used when assigning returns in 1999, 2000 and 2001 to parents that spawned in 1995, when only a dam was assigned.

This extra genotyping was time consuming as the DNA samples had to be picked manually from 96 well plates for re-amplification, and inevitably increased the possibility of errors due to sample misidentification. The number of freeze-thaw cycles the samples underwent was also increased, which was likely to degrade the DNA. Therefore, for the remaining samples 12 loci were used from the onset. This entailed the development of multiplexes 5, 6 and 7 (Table 2.4), in order to accommodate the extra loci while retaining the compatibility of loci within multiplexes.

2.4.7 PIG tailing.

Taq polymerases (without proofreading abilities) can sometimes add an extra nucleotide (usually adenine) onto the 3’ end of the DNA strand being synthesized. This can result in two different sized PCR products (differing by one base pair) being generated for each allele. Consistent genotyping is facilitated by forcing the predominance of only one of these products; either the addition of a nucleotide or its omission. Having only one type of product leads to a larger overall number of products for that particular size and so a stronger signal. The degree to which the addition occurs can be manipulated by altering the final extension time or the sequence of the unlabelled primer, as described in Brownstein et al. (1996).

When constructing multiplexes 5, 6 and 7 attention was paid to the addition or omission of an extra nucleotide to the PCR product. The majority of loci in multiplex 5 resulted
in single products with a final extension time of 7 minutes. However, loci Ssa1605 and Ssa407 tended to have PCR products of two sizes for each allele, one base pair apart. Products were then forced to either have an extra nucleotide added (Ssa407) or to have this omitted (Ssa1605) by changing the primer sequences as detailed in Brownstein et al. (1996). An extra T was added to the 5’ end of the reverse primer for Ssa1605 and an extra G added to the 5’ end for the reverse primer of Ssa407. This modification proved successful for Ssa407 but not for Ssa1605. (Longer additional sequences may have resulted in a greater effect, but these were not used as they may have given rise to additional primer dimer due to the large number of primers in the reaction). The PCR products of all loci in multiplexes 6 and 7 tended to be of a single type (with or without an extra non-templated nucleotide), thus primer sequences were not altered.

2.4.8 Final PCR conditions.
Thermocycling was carried out on Biometra®-T gradient machines. Water was nuclease free. Non-fluorescent primers and ‘ABI’ fluorescent primers were ordered from MWG Biotech while Beckman fluorescent dyes were supplied by Proligo France SAS. Other PCR components were supplied by ABGENE i.e. Thermoprime Plus DNA Polymerase (5U / µL ) Cat No. AB-0301 (inc 10× Buffer II solution (100 mM Tris-HCl, pH8.3; 500 mM KCl) and separate 25 mM MgCl₂ solution); dNTP set (high concentration, 100 mM each) Cat No. AB-0315.

Fluorescent tagged primers for ABI gel fragment analyzers were re-suspended and diluted, when necessary, in nuclease free water. Fluorescent tagged primers for Beckman Coulter fragment analysers were initially re-suspended to 100 µM in 1×TE buffer (10 mM Tris, 1 mM EDTA; pH 8.0). Subsequent dilutions were made in
nuclease free water. Standard (non-fluorescent) primers were suspended and diluted in nuclease free water. All primers were stored at -25°C at 100µM stock concentration. At any one time one aliquot of a particular primer would be diluted to the ‘working’ concentration of 10µM.

For DNA extracted using phenol chloroform, c.100ng DNA template (as determined by optical density) was used in each reaction. For DNA template obtained using chelex extractions, 1µl of undiluted extraction product was used in each reaction (except for multiplex 3, where 1µl of a 1:10 dilution gave optimal resolution). DNA template extracted using the commercially available kit (the Wizard® SV genomic purification system) was used undiluted: 2 or 3µl were used in each reaction, depending on the particular multiplex (as detailed in the PCR conditions in the Appendix, Table A5 to Table A10).

All constituents of the PCR assay bar DNA were combined on ice. Water and buffer were initially combined, then dNTP, MgCl₂ and primers followed by Taq polymerase. This was mixed by hand, briefly spun, and then added to the DNA, already aliquotted in 96 well plates. (During initial trials a few samples were amplified at a time, in which case thin walled flat cap 0.2ml tubes were used instead of 96 well plates). After addition of all components, plates or tubes were spun at 1000g for one minute prior to placement in the thermocycler. No mineral oil was required as the thermocyclers had heated lids. When 96 well plates were employed, an adhesive top seal (ABGENE) was used and an insulating pad was placed between this plate seal and the heated lid of the thermocycler. Thermocycler blocks were brought to 95°C before inserting the reaction tubes or plates. Theoretically this should minimize potential PCR artifacts caused by
‘non-specific’ amplification during initial temperature ramping. Lid temperature was maintained at 105°C throughout. Primers and fluorescent labels, where applicable, are shown in Table 2.6 (sequences were given in Table 2.2 and Table 2.3). Details of the constituents of PCRs and thermocycling conditions are given in Tables A1 to Table A10 in the Appendix.

2.4.9 Detection and Sizing of PCR products

PCR products from adult samples taken in 1991 and 1995 and parr samples from 1997 were detected and sized using two 64 lane ABI 377 DNA fragment analyzers. Here separation of products of different sizes is based on electrophoresis through a slab polyacrylamide gel poured by the user. PCR products from adult samples taken in 1999, 2000, 2001 and 2004 and smolt samples taken in 2004 were screened on two Beckman Coulter CEQ 8800 fragment analyzers. These are capillary electrophoresis based analysers, each with a capacity to run eight capillaries simultaneously. Size separation is conducted through a gel matrix in individual capillaries, each corresponding to a ‘lane’ on the slab gel. Both machines detected fluorescently labeled markers, but different sets. ABI 377 DNA fragment analyzers detected UV fluorescent TAMRA, TET, HEX and C6FAM dyes. Beckman Coulter CEQ 8800 fragment analyzers detected IR fluors; WellRED D1, D2, D3 and D4.

2.4.10 ABI fragment analyzers

Plates were washed in hot water, rinsed in distilled water, dried with a hairdryer and finally assembled into the proprietary casting frame. For each plate, 50ml of acrylamide gel was mixed. Initially, 14.4g urea, 4.8ml acrylamide stock (LongRanger acrylamide stock, Cambrex Bio Science Inc.), 23ml MilliQ water and 0.5g mixed bead resin were
Table 2.6 Primers used to amplify the 12 routinely screened microsatellite loci. Reverse primer for \textit{Ssa1605} ‘PLUS T’ was not used in genotyping. Reverse primer for \textit{Ssa407} ‘PLUS G’ was only used in multiplex 5.

<table>
<thead>
<tr>
<th>Locus Name</th>
<th>Forward/ Reverse (F/R) and any modifications</th>
<th>Fluor 1</th>
<th>Fluor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Ssa2201}</td>
<td>F</td>
<td>FAM</td>
<td>D4</td>
</tr>
<tr>
<td>\textit{Ssa2201}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa1605}</td>
<td>F</td>
<td>FAM</td>
<td>D4</td>
</tr>
<tr>
<td>\textit{Ssa1605}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa1605}</td>
<td>R (PLUS T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa2210}</td>
<td>F</td>
<td>FAM</td>
<td>D4</td>
</tr>
<tr>
<td>\textit{Ssa2210}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa410}</td>
<td>F</td>
<td>TET</td>
<td>D2</td>
</tr>
<tr>
<td>\textit{Ssa410}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa407}</td>
<td>F</td>
<td>HEX</td>
<td>D3</td>
</tr>
<tr>
<td>\textit{Ssa407}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa407}</td>
<td>R (PLUS G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa404}</td>
<td>F</td>
<td>HEX</td>
<td>D2</td>
</tr>
<tr>
<td>\textit{Ssa404}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa408}</td>
<td>F</td>
<td>TET</td>
<td>D3</td>
</tr>
<tr>
<td>\textit{Ssa408}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa171}</td>
<td>R</td>
<td>FAM</td>
<td>D4</td>
</tr>
<tr>
<td>\textit{Ssa171}</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{One9}</td>
<td>F</td>
<td>FAM</td>
<td>D4</td>
</tr>
<tr>
<td>\textit{One9}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa197}</td>
<td>R</td>
<td>HEX</td>
<td>D3</td>
</tr>
<tr>
<td>\textit{Ssa197}</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa202}</td>
<td>R</td>
<td>TET</td>
<td>D2</td>
</tr>
<tr>
<td>\textit{Ssa202}</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{Ssa401}</td>
<td>F</td>
<td></td>
<td>D4</td>
</tr>
<tr>
<td>\textit{Ssa401}</td>
<td>R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1\textsuperscript{1} Fluorescent dye used with ABI fragment analyzer.

2\textsuperscript{2} Fluorescent dye used with Beckman Coulter fragment analyser.
Materials and methods

placed in a beaker and mixed with a stirrer for c.30 minutes. This was filtered, 5ml filtered 10 × TBE added and then de-gassed prior to addition of 200µL 10% ammonium persulphate solution and 25µL Temed. The gel mix was gently swirled and a syringe used to inject the mixture into the assembled casting frame. Gels were left for 2.5 hours to polymerise. The electrophoresis buffer was 1× TBE. Prior to addition of samples, plates were pre-conditioned, i.e. heated for 10 minutes on the ABI fragment analyzer until the gel temperature was at least 40°C. Two methods were used to load samples onto the gel interface. The first method was more time consuming than the second, so when the second method was introduced the remaining samples were loaded in this manner. In both cases 1µL of diluted or undiluted PCR product (see Table 2.7) was mixed with 1.7µL of loading dye (4 parts deionised formamide to 1 part blue dextran dye) and 0.3µL TAMRA size marker (genSIZE T350, Genetix) prior to denaturing.

<table>
<thead>
<tr>
<th>Multiplex</th>
<th>Dilution of PCR product before loading onto fragment analyzer</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA from Phenol Chloroform extractions</td>
<td>DNA from Chelex extractions</td>
</tr>
<tr>
<td>1</td>
<td>1:10</td>
</tr>
<tr>
<td>2</td>
<td>no dilution</td>
</tr>
<tr>
<td>3</td>
<td>1:5</td>
</tr>
<tr>
<td>4</td>
<td>no dilution</td>
</tr>
</tbody>
</table>

In the first method a shark tooth comb was used to produce wells into which samples were pipetted, using ‘duck-billed’ tips. The PCR samples were denatured by heating at 95°C for 3 min. 1µL of each sample was then pipetted into alternate lanes. The gel was
run for 4min, after which the remaining samples were loaded. This was in order to reduce contamination between adjacent lanes and to aid later lane tracking. The gel was subsequently run for two hours at 3000V at a temperature of 50°C.

In the second method a membrane comb (Web Scientific) was used to load samples. A loading solution of 10% Ficoll was used to fill the comb area to a depth of approximately 10mm. The PCR samples were denatured by heating at 95°C for 3min. Samples were then placed on ice and 0.8µL loaded onto a chilled loading tray (Web Scientific, a plastic plate with 64 sample indentations matched to membrane comb teeth spacing). The membrane comb was placed in these wells for 10s, allowing transfer of sample to the comb. The comb was then transferred to the gel, with tips of the comb teeth just touching the gel interface. The upper buffer tray was installed, filled with buffer and the gel run for 60s. The membrane comb was removed and the Ficoll solution washed out of the comb area using a pipette. The gel was subsequently run for two hours at 3000V at a temperature of 50°C.

The gel matrix was often reused for a second separation run (with different samples). While the slab gel was still warm, the comb area was washed out with buffer and the loading procedure repeated.

2.4.11 Beckman Coulter fragment analyzers

Beckman proprietary gel, running buffer, size standard and loading buffer were used. PCR products (max 3ul) were added to 30µL sample loading solution (SLS, deionised formamide) containing 0.25µL size standard (SS400). The sample plates were then spun at 1000g for 1min and one drop of mineral oil was added on top of each sample. Samples were run under the standard conditions recommended for use with size
standard 400 (i.e. ‘FRAG-3’ method). To minimize waste and cost, both sample and buffer plates were re-used. To clean, they were soaked in 5% Decon overnight, rinsed well in hot water followed by distilled water and then dried.

### 2.4.12 Allele callings

Data generated from the ABI and Beckman Coulter fragment analyzers were interpreted using proprietary software (GENESCAN v3.1.2 and GENOTYPER v2.5 for the ABI fragment analyzer and FRAGMENT ANALYSER for the Beckman Coulter fragment analyzer). Automated allele calling was not implemented in either case. Each chromatogram was inspected visually and allele labels assigned manually (using predefined bins, generally of c.2bp width).

### 2.5 Parentage Assignment

#### 2.5.1 Exclusion versus maximum likelihood methods

Parentage analysis based on maximum likelihood and exclusion methods were considered for use in this study. Maximum likelihood methods require less genetic information than exclusion methods (Duchesne et al. 2002), but were not considered suitable. Assumptions used in maximum likelihood approaches were either unclear or known to be violated (namely Hardy-Weinberg equilibrium). The population structure was known to be complex, with input from mature parr (Taggart et al. 2001), overlapping generations and straying. The impact of these characteristics on maximum likelihood methods was not known. In addition Jones and Ardren (2003) stated that maximum likelihood methods could cause an upward or downward bias in the variance of reproductive success (depending on the method used) and can be sensitive to the total number of candidate parents. In this study the total number of sires was not known.
(due to the contribution of mature parr). It was imperative that if any reliable conclusions were to be made from the study, parentage assignment was accurate, due to the random variation involved in a natural system and the restricted numbers of samples in some aspects of the study. When a maximum likelihood method (CERVUS) was employed on samples in the current study, results differed to those gained from exclusion analysis and were also dependent on the parameters used. Varying the proportion of males that were ‘collected’ (i.e. sampled) from 0.2 to 0.5 gave different results, a problem as this value is unknown due to the input of mature parr in natural spawning. Exclusion methodology was therefore employed and since multiplexes were utilized the cost of typing extra loci was minimized.

2.5.2 Software employed in parentage analysis

Parentage was determined using two DOS based programs written in QUICKBASIC language, DAMFIND and SIREFIND (J. Taggart unpublished). These were based on simple exclusion principles and made maximum use of known properties of the samples.

Initially, an attempt was made to assign the dam using DAMFIND. This program analyzed potential offspring and dam genotypes and resulted in the assignment of a dam if, for each of the loci tested, either allele of the dam matched either allele of the offspring. In the vast majority of analyses undertaken all possible dams had been sampled and genotyped, thus exclusion of all potential dams at any locus suggested a genotyping error, mutation, or mis-classified offspring (i.e. of the wrong age-class). Subsequently, assignment of sires was carried out using SIREFIND. This program took into account the dam that had been attributed, and resulted in the assignment of a sire if
the remaining ‘sire’ allele, at all the loci tested, matched either of those of a potential anadromous sire.

In addition to perfect matches, the output of both programs permitted ‘relaxed assignments’ to be explored (i.e. allowing for one or more mismatched alleles to be accommodated). These relaxed assignments were flagged and the problematic loci identified. This ‘relaxed approach’ was particularly useful in identifying mismatches that were due to genotyping error since a mismatched allele would be identified and the corresponding chromatogram could be re-examined to look for potential errors in allele calling. Finally if, following all checks, an individual was assigned a female but no anadromous male, the male parent was assumed to be a mature parr.
3  Data quality and accuracy

Genotypic data quality was particularly important in this study to ensure that parentage analyses could be relied on to give meaningful, accurate results. The project also afforded the opportunity to assess, by independent means, the accuracy of scale readings, which has rarely been performed.

In the following sections:

• The accuracy of genotyping in this study is explored and discussed.
• The power of parentage analyses employed in this study is considered.
• The degree of concordance between ages of fish determined by scale readings and through genetic profiling is investigated.

3.1  Genotyping Accuracy

A survey of articles published in Molecular Ecology, carried out in 2003, revealed that very few microsatellite-based studies quantified genotyping errors. Bonin et al. (2004) found that only 6% of articles that involved microsatellites also reported error rates, allelic dropout rates or false allele amplifications. Although an awareness of genotyping error is widely recognised to be of importance in interpreting studies based on limited amounts of and / or degraded DNA, sources or rates of error are not so widely reported in other studies (Hoffman and Amos 2005a). However, consideration of error rates can be extremely important. For example, Vigilant et al. (2001) believed the findings of a previous study, on the paternity of wild chimpanzees, were incorrectly biased as a result of genotyping error. Hoffman and Amos (2005b) also concluded that the results of a study, which found that Antarctic fur seals preferentially suckled pups of female relatives, could largely be attributed to genotyping error. Examining the efficacy of genotyping is especially important when primers have not been extensively tested previously. For
studies on humans, sets of loci with primers specifically chosen for robustness and reliability are available thus minimising the number of errors made (Ewen et al. 2000). However, this is not the case for Atlantic salmon, or indeed for the majority of other species that have been the subject of microsatellite based studies.

Errors can occur at a variety of stages in the genotyping process, and for a number of different reasons. They can be caused by sample misidentification or contamination - either in the extraction process, PCR amplification or separation of samples through electrophoresis. Errors may also be generated during the typing process due to human error when scoring, or to technical artefacts such as allele dropout, null alleles and stutter. Another potential source of error is the occurrence of alleles that are too similar in size to be accurately distinguished and which may be assigned inconsistently to different adjacent bins as a result. Incorrectly configured genotyping software and misplaced confidence in the accuracy of automated allele scoring are also areas of concern. Amplification of multiple loci in a single PCR can be more problematical than single locus amplifications. Yields of specific amplicons may be less consistent, which can increase the likelihood of incorrect allele calling. For example lower than expected yield may increase the occurrence of large allele dropout. At the other extreme, higher than expected amplicon yields can lead to masking of weaker fluor signals or ‘allelic pull up’, where a strong signal from one fluor may also be mistakenly detected as signal from a second fluor.

Errors can be identified and quantified in different ways. A straightforward method is to repeat the genotyping of some (or all) individuals. Bonin et al. (2004) advocate this approach for future studies, suggesting that 5-10% of samples should be re-
typed and results compared. Although a worthwhile strategy, it should be noted that this gives a measure of consistency rather than accuracy.

Another approach is to analyse existing genotypes for properties that may suggest true genotypes have not been recorded. The program MICROCHECKER (Van Oosterhout et al. 2004) has been written specifically for this task. High rates of allele dropout or the presence of null alleles may be indicated by a lower degree of heterozygosity than would be expected. Consistent heterozygote deficiency at one or few loci would suggest the existence of null alleles or allele dropout, whereas consistency over all loci would suggest population factors are more likely to be the cause of the deficit, possibly due to violation of assumptions made when testing for Hardy Weinberg equilibrium. However, a deficit will only be evident if the number of errors is large (Paetkau 2003). Both CERVUS and MICROCHECKER can estimate the frequency of null alleles, but under the assumption that the population is in HW equilibrium.

Hoarau et al. (2005) used a different approach to test for null alleles. These authors lowered the annealing temperatures for a set of samples that were found to be homozygous, to test whether more alleles would be amplified. However, this did not result in the identification of additional heterozygote genotypes.

A more direct method of assessing genotyping errors is to compare pedigree genotypes, where transmission of specific alleles is expected. For example, in parentage analysis where assignment is based on exclusion, parent-offspring pairs that mis-match at few alleles are candidates for genotyping errors. Once these individuals are flagged they can be investigated further and the genotypes either confirmed or errors identified.
3.1.1 Error rates and their sources in genotyping studies

The main sources of error determined by Bonin et al. (2004) in a study on brown bears were allelic dropout and human error in, for example, sample misidentification. O’Reilly et al. (1998) found a higher error rate when typing di-nucleotide microsatellites in Atlantic salmon (compared to tetra-nucleotides), with stutter error and allele designation error being the causes. Castro et al. (2004) in a study of turbot (Scophthalmus maximus) also found that di-nucleotide microsatellites were prone to difficulties in typing when the individual was heterozygous for two adjacent alleles.

In previous studies employing microsatellites, error rates have been reported of 0.1-0.7% (Hoffman and Amos 2005a), 0.8-2% (Bonin et al. 2004), less than 4% (Shrimpton and Heath 2003) and 2.5% (Castro et al. 2004). Errors identified by Bonin et al. (2004) were compiled from different sample types, with a lower rate of error for DNA extracted from tissue compared to faecal samples. The initial three estimates cited above were of the percentage of alleles that were typed incorrectly i.e. the percentage of alleles typed that were incorrect, as opposed to the percentage of genotypes that included an incorrect allele. The distinction between percentage of incorrect alleles or genotypes was not made in the last estimate.

Hoffman and Amos (2005a) reported that different methods of detection gave different error rates when used on the same set of samples (from Antarctic fur seals). They estimated error rates of 0.0013 per reaction from assignment of dams, and 0.0074 per reaction for individuals that were re-typed unintentionally (this occurred when more than one sample had been inadvertently obtained from individuals). According to Hoffman and Amos (2005a), the best method of error detection was probably to re-type a random sample.
3.1.2 Monitoring of genotyping errors in this study
When errors were identified in the current study, practises were modified to reduce or negate the chances of their reoccurrence.

Three complimentary methods were used to assess potential error rates throughout this study:

- randomly selected samples were re-typed
- control samples were typed throughout the study
- parentage analysis was used to flag possible errors

These are detailed below. All error rates referred to are allelic error rates i.e. the percentage of alleles checked that was incorrect. The occurrence of mutations was also examined because although not a ‘true’ error, they can create problems with parentage analysis.

3.1.3 Re-typing samples
Randomly selected samples from smolts sampled in 2004 and adult returns in 1991 were re-typed. Approximately 5-6% of samples were re-amplified. This is within the boundaries proposed by Bonin et al. (2004). They suggested that 5-10% of samples should go through a blind test for accuracy, and that this should include all steps in the typing (from extraction to calling alleles). This was not feasible in this study because the amount of tissue sampled was limited. Therefore only processes subsequent to DNA extraction were repeated. Re-typing a subset from the 1991 returns acted as a test of the similarity of genotypes obtained by typing with different multiplexes on different machines (the samples were initially typed with an ABI fragment analyser, whereas the re-typing took place on the Beckman Coulter fragment analyser). Reasons for using the different machines and the samples typed on each machine are given in Chapter 2.
Table 3.1 and Table 3.2 show the number of detected errors in each set of samples. Overall rates of 0.8% for smolt samples taken in 2004 and 3.8% for adult returns in 1991 were found. Two errors at locus Ssa171 in the sample of smolts were due to the same two adjacent alleles being incorrectly typed. These alleles were too close to be reliably distinguished, and so bins for these were combined in all subsequent typing. The errors in the sample of returns in 1991 were due to a ‘shift’ of the size standard in one of the multiplexes (multiplex 6) in one sample, which was apparently a software problem. Thus the two loci being checked in this sample were incorrectly typed. Subsequently (and retrospectively), the size standard in all chromatograms was checked to ensure that the peaks had been correctly labelled and that they were in the correct position in relation to the scale used to identify the size of amplified peaks.
Table 3.1 Number of incorrect alleles identified by re-typing 32 random samples from smolts that were sampled in 2004. Overall numbers relate to the total number of genotypes assessed.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$Ssa$ 171</th>
<th>$Ssa$ 197</th>
<th>$Ssa$ 401</th>
<th>$Ssa$ 407</th>
<th>$Ssa$ 408</th>
<th>$Ssa$ 410</th>
<th>$Ssa$ 1605</th>
<th>$Ssa$ 2201</th>
<th>$Ssa$ 2210</th>
<th>One9</th>
<th>$Ssa$ 404</th>
<th>$Ssa$ 202</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td># samples re-typed successfully</td>
<td>30</td>
<td>12</td>
<td>22</td>
<td>14</td>
<td>30</td>
<td>10</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>30</td>
<td>28</td>
<td>28</td>
<td>258</td>
</tr>
<tr>
<td># incorrect alleles</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>% incorrect alleles</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Table 3.2 Number of incorrect alleles identified by re-typing 8 random samples from adults that returned in 1991. Overall numbers relate to the total number of genotypes assessed.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$Ssa$ 171</th>
<th>$Ssa$ 197</th>
<th>$Ssa$ 407</th>
<th>$Ssa$ 410</th>
<th>$Ssa$ 1605</th>
<th>$Ssa$ 2201</th>
<th>$Ssa$ 2210</th>
<th>One9</th>
<th>$Ssa$ 404</th>
<th>$Ssa$ 202</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td># samples re-typed successfully</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td># incorrect alleles</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>% incorrect alleles</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>3.8</td>
<td></td>
</tr>
</tbody>
</table>
3.1.4 Controls

For the genotyping carried out on ABI fragment analysers, control samples were primarily used to verify correct loading and tracking (on slab gels) i.e. to ensure that each lane was identified correctly. Four samples of known genotype were run on each gel. In every case the tracking was confirmed to be accurate and allele sizes were assigned to the correct bin sizes. This gave reassurance as to the provenance of samples and the consistency of the genotype scoring.

For the remaining genotyping Beckman Coulter fragment analysers were employed. Three control samples were routinely analysed. This provided a test for the consistency of sizes of alleles (albeit for a limited number of alleles). There was less concern with regard to sample tracking as all samples were analyzed in a 96 well plate format. The second set was run on new machines and so controls were especially useful in monitoring the accuracy of the machines and how they were used.

Table 3.3 shows the number of controls used in the second set of genotyping and the number of alleles that were incorrectly typed. The overall percentage of incorrect alleles, averaged over controls and loci, was 0.5%. The percentage of alleles typed that were incorrect for each control is shown in the right-hand column and the percentage of alleles that were incorrect for each individual locus is shown in the bottom row.

3.1.5 Parentage analysis

The third method of error detection involved parentage analysis. Assignments with one mis-matched allele, or with multiple mis-matches in a single multiplex (to screen for possible misidentified samples) were scrutinised for errors.
Table 3.3 Details of incorrect alleles identified by repeat typing of controls. Three individuals were used as controls. The number of incorrect alleles found at each locus are shown for each individual control.

| Control # | Locus                  | Ssa 171 | Ssa 197 | Ssa 401 | Ssa 407 | Ssa 408 | Ssa 410 | Ssa 1605 | Ssa 2201 | Ssa 2210 | One9 | Ssa 404 | Ssa 202 | Total Typed | Total Error | % incorrect alleles |
|-----------|------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|      |        |        |             |             |                           |
| 1         | # times typed          | 12      | 11      | 11      | 6       | 10      | 5       | 7       | 7       | 6       | 11    | 10     | 10     | 106          |             |                           |
|           | successfully           |         |         |         |         |         |         |         |         |         |       |        |        |                           |             |                           |
|           | # incorrect alleles    | 0       | 1       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0     | 0      | 0      | 1            | 0.4          |                           |
| 2         | # times typed          | 12      | 9       | 8       | 4       | 12      | 5       | 6       | 6       | 6       | 12    | 11     | 8      | 99           |             |                           |
|           | successfully           |         |         |         |         |         |         |         |         |         |       |        |        |                           |             |                           |
|           | # incorrect alleles    | 1       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 0     | 0      | 0      | 1            | 0.5          |                           |
| 3         | # times typed          | 12      | 8       | 7       | 5       | 10      | 5       | 5       | 5       | 5       | 10    | 9      | 7      | 88           |             |                           |
|           | successfully           |         |         |         |         |         |         |         |         |         |       |        |        |                           |             |                           |
|           | # incorrect alleles    | 0       | 0       | 0       | 0       | 0       | 0       | 1       | 0       | 0       | 0     | 0      | 0      | 1            | 0.6          |                           |
| total     | # samples typed        | 36      | 28      | 26      | 15      | 32      | 15      | 18      | 18      | 17      | 33    | 30     | 25     |                           |             |                           |
| total     | # incorrect alleles    | 1       | 1       | 0       | 0       | 0       | 0       | 0       | 0       | 0       | 1     | 0      | 0      |                           |             |                           |
| % incorrect alleles | 1 for each locus       | 1       | 2       | 0       | 0       | 0       | 0       | 3       | 0       | 0       | 3     | 0      | 0      |                           |             |                           |

1 for each locus
2 for each control
Errors detected, and their causes, were explored in depth for two datasets involving:

- The assignment of smolts sampled in 2004 to 2000 and 2001 adult spawners
- The assignment of adult returns in 1999, 2000 and 2001 to 1995 adult spawners

Parentage analyses in the earlier sets of genotyping (assigning parr sampled in 1997 to returns in 1995 and assigning returns in 1995 to returns in 1991) also identified errors. However, though corrected at the time, these were not rigorously recorded for error estimates and thus are not detailed here.

The error rate for both sets of assignments that were explored in detail (Table 3.4) was estimated by calculating the total number of alleles typed and tested (in both the parents and offspring) and determining the percentage of these that required correction. Only alleles that were involved in an assignment were included i.e. when an offspring was assigned only to a female adult return (presumed to be sired by a mature parr) just half of the offspring’s alleles were tested.

For the assignment of smolts sampled in 2004, a total of 130 parents were screened, 400 offspring were assigned to female returns and 239 were assigned to male returns. For 12 loci a total of 7596 alleles were tested in offspring and approximately 1560 alleles were tested in parents. In total c. 0.4% (37/9156) of alleles were initially identified as genotype scoring errors.

For the assignment of adult returns in 1999, 2000 and 2001 to 1995 adult spawners, a total of 43 different parents were assigned, 51 offspring were assigned to females and 32 were assigned to males. For 12 loci a total of approximately 1512 alleles were tested. In total c. 0.1% (2/1512) of alleles were initially incorrect. This is probably lower than the previous estimate of 0.4% because the adult returns of 1995 had already been purged of errors through the assignment of parr sampled in 1997.

Table 3.4 shows the types of errors that were found.
Table 3.4 Error types and frequencies identified through parentage analysis, assigning smolts sampled in 2004 to returns in 2000 and 2001, and assigning returns in 1999, 2000 and 2001 to returns in 1995. Possible errors were flagged by mismatches at only one locus (or at all loci in the same multiplex), and the correct genotype obtained by reanalysis of chromatograms or by re-typing.

<table>
<thead>
<tr>
<th>Error type</th>
<th>Locus</th>
<th>Ssa 171</th>
<th>Ssa 197</th>
<th>Ssa 401</th>
<th>Ssa 407</th>
<th>Ssa 408</th>
<th>Ssa 410</th>
<th>Ssa 1605</th>
<th>Ssa 2201</th>
<th>Ssa 2210</th>
<th>One9</th>
<th>Ssa 404</th>
<th>Ssa 202</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygote incorrectly typed as homozygote</td>
<td></td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Homozygote incorrectly typed as heterozygote</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Incorrect allele input manually(^1)</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Trisomy(^2)</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pull up(^3)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Size standard shifted(^4)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Incorrect sample(^5)</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Unknown/not recorded</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>% of total errors</td>
<td></td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>13</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>13</td>
<td>33</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^1\) In some cases the software would not recognise a peak. In these instances the allele name was input manually.

\(^2\) One or more loci were found to have 3 alleles.

\(^3\) Due to PCR products from different loci being run in single lanes, alleles of similar size sometimes overlapped. This was not normally a problem because different fluoros were used for different loci. However, a strong signal with particular fluoros could result in a low signal for other fluoros, which could have been interpreted incorrectly as an allele at a different locus.

\(^4\) The CEQ genotyping software occasionally did not match the scale from which allele sizes were calculated to the size standard peaks. This resulted in shifts in allele size.

\(^5\) An incorrect sample was used in one PCR amplification, leading to all the alleles in a multiplex being incorrectly scored.
3.1.6 Null alleles/ allele dropout

Exact tests for deviation from Hardy Weinberg equilibrium were carried out using GENEPOP version 3.4 (Raymond and Rousset 1995). No consistent deviation was found at any one locus (Table 3.5). This suggests that neither null alleles nor allele dropouts were a major problem. This observation was corroborated by parentage analysis results, where again there was no evidence of allele dropout or null alleles.
Table 3.5 Significance of tests for Hardy Weinberg equilibrium (HW p), F\textsubscript{IS}, observed (H\textsubscript{o}) and expected (H\textsubscript{e}) number of heterozygotes for all sample groups. Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th>Sample Group</th>
<th>Locus</th>
<th>2210</th>
<th>1605</th>
<th>2201</th>
<th>410</th>
<th>One9</th>
<th>407</th>
<th>202</th>
<th>171</th>
<th>197</th>
<th>408</th>
<th>404</th>
<th>401</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult 1991</td>
<td>HW p</td>
<td>0.121</td>
<td>0.765</td>
<td>0.402</td>
<td>0.388</td>
<td>0.579</td>
<td>0.837</td>
<td><strong>0.005</strong></td>
<td>0.458</td>
<td>0.419</td>
<td>0.156</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>-0.010</td>
<td>-0.020</td>
<td>0.015</td>
<td>0.015</td>
<td>0.021</td>
<td>-0.013</td>
<td>0.092</td>
<td>-0.013</td>
<td>0.042</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>75.3</td>
<td>78.4</td>
<td>89.3</td>
<td>89.4</td>
<td>59.2</td>
<td>86.9</td>
<td>84.7</td>
<td>83.0</td>
<td>87.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>76</td>
<td>80</td>
<td>88</td>
<td>88</td>
<td>58</td>
<td>88</td>
<td>77</td>
<td>84</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult 1995</td>
<td>HW p</td>
<td>0.074</td>
<td>0.955</td>
<td>0.832</td>
<td><strong>0.001</strong></td>
<td>0.485</td>
<td>0.139</td>
<td>0.290</td>
<td>0.057</td>
<td><strong>0.039</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.002</strong>***</td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>0.022</td>
<td>-0.019</td>
<td>-0.034</td>
<td>0.042</td>
<td>-0.007</td>
<td>-0.034</td>
<td>-0.038</td>
<td>-0.017</td>
<td>0.005</td>
<td>-0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>97.1</td>
<td>94.2</td>
<td>108.3</td>
<td>108.6</td>
<td>70.5</td>
<td>107.4</td>
<td>102.1</td>
<td>100.3</td>
<td>107.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>95</td>
<td>96</td>
<td>112</td>
<td>104</td>
<td>71</td>
<td>111</td>
<td>106</td>
<td>102</td>
<td>107</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parr 1997</td>
<td>HW p</td>
<td><strong>0.019</strong></td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.497</td>
<td>&lt;0.001</td>
<td>0.864</td>
<td>&lt;0.001</td>
<td>0.074</td>
<td></td>
<td></td>
<td><strong>&lt;0.001</strong>***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>-0.052</td>
<td>0.020</td>
<td>0.008</td>
<td>-0.014</td>
<td>-0.037</td>
<td>0.001</td>
<td>-0.010</td>
<td>0.005</td>
<td>0.002</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>227.3</td>
<td>223.6</td>
<td>256.1</td>
<td>256.4</td>
<td>165.9</td>
<td>251.2</td>
<td>239.7</td>
<td>235.2</td>
<td>251.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>239</td>
<td>219</td>
<td>254</td>
<td>260</td>
<td>172</td>
<td>251</td>
<td>242</td>
<td>234</td>
<td>251</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult 2000</td>
<td>HW p</td>
<td><strong>0.018</strong></td>
<td>0.080</td>
<td>0.307</td>
<td>0.575</td>
<td><strong>0.035</strong></td>
<td><strong>0.011</strong></td>
<td>0.569</td>
<td>0.617</td>
<td><strong>0.044</strong></td>
<td>0.189</td>
<td>0.322</td>
<td>0.251</td>
<td><strong>0.002</strong>***</td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>0.041</td>
<td>0.036</td>
<td>0.046</td>
<td>-0.037</td>
<td>0.068</td>
<td>-0.022</td>
<td>0.033</td>
<td>-0.024</td>
<td>0.013</td>
<td>0.031</td>
<td>-0.004</td>
<td>0.035</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>73.0</td>
<td>72.6</td>
<td>83.8</td>
<td>83.9</td>
<td>52.5</td>
<td>83.2</td>
<td>78.6</td>
<td>76.2</td>
<td>82.0</td>
<td>80.4</td>
<td>85.7</td>
<td>82.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>70</td>
<td>70</td>
<td>80</td>
<td>87</td>
<td>49</td>
<td>85</td>
<td>76</td>
<td>78</td>
<td>81</td>
<td>78</td>
<td>86</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Adult 2001</td>
<td>HW p</td>
<td>0.730</td>
<td><strong>0.046</strong></td>
<td>0.781</td>
<td><strong>0.028</strong></td>
<td>0.929</td>
<td>0.328</td>
<td>0.263</td>
<td>0.210</td>
<td><strong>0.005</strong></td>
<td>0.864</td>
<td>0.205</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>0.001</td>
<td>0.145</td>
<td>-0.011</td>
<td>-0.005</td>
<td>-0.022</td>
<td>-0.003</td>
<td>-0.012</td>
<td>-0.010</td>
<td>0.101</td>
<td>0.048</td>
<td>0.004</td>
<td>0.036</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>33.0</td>
<td>33.9</td>
<td>38.6</td>
<td>38.8</td>
<td>26.4</td>
<td>37.9</td>
<td>35.6</td>
<td>35.6</td>
<td>37.8</td>
<td>35.7</td>
<td>39.1</td>
<td>38.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>33</td>
<td>29</td>
<td>39</td>
<td>39</td>
<td>27</td>
<td>38</td>
<td>36</td>
<td>36</td>
<td>34</td>
<td>34</td>
<td>39</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Adult 2004</td>
<td>HW p</td>
<td>0.095</td>
<td>0.479</td>
<td><strong>0.009</strong></td>
<td>0.393</td>
<td>0.690</td>
<td>0.359</td>
<td>0.506</td>
<td>0.062</td>
<td>0.463</td>
<td>0.188</td>
<td><strong>0.006</strong></td>
<td>0.238</td>
<td><strong>0.006</strong>*</td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>0.055</td>
<td>0.104</td>
<td>0.041</td>
<td>0.001</td>
<td>0.061</td>
<td>-0.002</td>
<td>-0.020</td>
<td>0.089</td>
<td>-0.012</td>
<td>0.014</td>
<td>0.029</td>
<td>0.054</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>97.4</td>
<td>93.8</td>
<td>109.5</td>
<td>110.1</td>
<td>74.6</td>
<td>107.8</td>
<td>101.0</td>
<td>101.0</td>
<td>108.8</td>
<td>106.5</td>
<td>111.2</td>
<td>109.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>92</td>
<td>84</td>
<td>105</td>
<td>110</td>
<td>70</td>
<td>108</td>
<td>103</td>
<td>92</td>
<td>110</td>
<td>105</td>
<td>108</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>Smolt 2004</td>
<td>HW p</td>
<td>0.220</td>
<td><strong>0.032</strong></td>
<td><strong>0.003</strong></td>
<td>0.001</td>
<td>0.404</td>
<td>&lt;0.001</td>
<td><strong>0.018</strong></td>
<td>0.458</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>F\textsubscript{IS}</td>
<td>0.015</td>
<td>0.011</td>
<td>-0.019</td>
<td>-0.016</td>
<td>-0.006</td>
<td>0.013</td>
<td>-0.015</td>
<td>-0.015</td>
<td>0.024</td>
<td>0.025</td>
<td>-0.011</td>
<td>-0.002</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{o}</td>
<td>332.9</td>
<td>330.6</td>
<td>377.8</td>
<td>380.1</td>
<td>252.5</td>
<td>378.1</td>
<td>352.7</td>
<td>351.9</td>
<td>376.1</td>
<td>354.8</td>
<td>387.8</td>
<td>382.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H\textsubscript{e}</td>
<td>328</td>
<td>327</td>
<td>385</td>
<td>386</td>
<td>254</td>
<td>373</td>
<td>358</td>
<td>357</td>
<td>367</td>
<td>346</td>
<td>392</td>
<td>383</td>
<td></td>
</tr>
</tbody>
</table>

* significant at 0.05 level following sequential Bonferroni correction, for overall values only, initial \( \alpha = 0.05/7 \)
** significant at 0.01 level following sequential Bonferroni correction, for overall values only, initial \( \alpha = 0.01/7 \)
3.1.7 Mutation

In six cases in this study mismatches were attributed to mutations (Table 3.6). The offspring were assigned (at 11 or 12 loci) to an adult pair combination, with a mismatch at a single allele. Parental and offspring chromatograms were re-inspected and the genotypic scoring was confirmed. As further confirmation, the parental and offspring sets were genotyped again and the scores once more confirmed. In all cases the mismatching offspring allele was one or two repeats larger or smaller than expected. The mutation mode of microsatellites usually follows a stepwise pattern, so that mutations to alleles one repeat apart from the original are more common (Beebee and Rowe 2004). In cases one and two (Table 3.6), where the mutation is presumed to be paternal in origin (due to the size of alleles), a mature parr could have been the true sire. However, if this was the case the mature parr parent would have been from a different year class to the anadromous male and would have been unlikely to be so closely related. The ages of the offspring as identified by scale readings match with the time of spawning of the assigned parents in all cases except one (case three).

Table 3.6 Details of single mis-matches attributed to mutations.

<table>
<thead>
<tr>
<th>Case</th>
<th>Sample parents belong to</th>
<th>Sample offspring belongs to</th>
<th>Details of mismatching allele</th>
<th>Number of loci tested in mismatching pair</th>
<th>Mismatch in male/female parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1995 sea returns</td>
<td>2000 sea returns</td>
<td>Locus</td>
<td>Allele in parent</td>
<td>Allele in offspring</td>
</tr>
<tr>
<td>2</td>
<td>2001 sea returns</td>
<td>2004 smolts</td>
<td>11</td>
<td>2201</td>
<td>303</td>
</tr>
<tr>
<td>3</td>
<td>1995 sea returns</td>
<td>2000 sea returns</td>
<td>12</td>
<td>171</td>
<td>251</td>
</tr>
<tr>
<td>4</td>
<td>2000 sea returns</td>
<td>2004 smolts</td>
<td>12</td>
<td>410</td>
<td>249</td>
</tr>
<tr>
<td>5</td>
<td>2000 sea returns</td>
<td>2004 smolts</td>
<td>12</td>
<td>407</td>
<td>248</td>
</tr>
<tr>
<td>6</td>
<td>2000 sea returns</td>
<td>2004 smolts</td>
<td>12</td>
<td>401</td>
<td>308</td>
</tr>
</tbody>
</table>
In total, 12675 meioses were observed, (over 9-12 loci, and accounting for meioses that were not observed due to paternity attributed to mature parr). This gives an average mutation rate over all loci of $4.7 \times 10^{-4}$.

### 3.1.8 Conclusion- levels of accuracy in genotyping

Considering all the types of errors that were identified, the two most prevalent were heterozygotes that had been incorrectly scored as homozygotes and the lack of distinction between two similar sized alleles of $Ssa171$. The errors caused by the imperfect tetra nucleotide repeats at $Ssa171$ were eliminated in the early stages by combining the appropriate bins.

The three different methods of identification of error also gave varying error rates, summarised in Table 3.7.

**Table 3.7 Summary of error rate estimated using different methods**

<table>
<thead>
<tr>
<th>Type of analysis</th>
<th>% incorrect alleles (total alleles tested)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-typing smolts sampled in 2004</td>
<td>0.6 (516)</td>
</tr>
<tr>
<td>Re-typing adults that returned in 1991</td>
<td>3.8 (104)</td>
</tr>
<tr>
<td>Controls</td>
<td>0.5 (656)</td>
</tr>
<tr>
<td>Parentage analysis</td>
<td>0.4 (9156)</td>
</tr>
</tbody>
</table>

The largest discrepancies were found in re-typed samples that returned in 1991. However, this was due to the size standard being incorrectly called in one multiplex for one sample. This meant that 2 loci were incorrectly typed, which was noticeable in the final error rate due to the small numbers (8 individuals) that were re-typed in this set. As a result of this finding all chromatograms were checked for potential size standard recognition errors. Very few other cases were actually identified.

The error rate estimated from parentage analysis is likely to be lower than the other estimates because errors identified though re-typing smolts and in typing the
controls had been removed (bins were combined at Ssa171 and size standards were checked).

In a study on Antarctic fur seals, Hoffman and Amos (2005a) found that maternity analysis gave the lowest error rate for the different error detection methods they employed and concluded that maternity analysis was not the best method for error rate estimation. No error rate for paternity assignment was given, but it can be derived from the data they report. This computes to an error rate of 0.3% (19 incorrect alleles/9 loci × 388 pairs of sires and offspring × 2 alleles in each pair). This falls between the error rates estimated from other detection methods they used. Although it may not be the best way to estimate the error rate when only one parent is available for assignment, it offers a relatively sensitive test for the presence of null alleles or allele dropout. In addition to this it is a way of flagging possible errors in the whole data set that merit further attention. Randomly re-typing samples allows the eradication of errors found only in the particular sub-set that is re-typed. In cases where all female and male parents are sampled it offers a means of identifying the vast majority of errors rather than solely calculating an error rate. Controls are useful as a continual check of the consistency of PCRs and electrophoresis, but only test a limited number of alleles. They also offer a constant reminder of the importance of precision in the genotyping process and a check that sample labelling has been correctly carried out. Undoubtedly using a variety of error detection methods is to be recommended, as this is likely to give a better estimate of overall error levels.

Compared to other studies mentioned in the introduction, error rate estimates derived in this study were reassuringly low. They were comparable to that reported in Hoffman and Amos (2005a) (excluding the technical error found in the sample of
8 returns in 1991 that were re-typed). One reason for this is likely to be the reliance on tetra-nucleotide repeat loci. Only one di-nucleotide was scored, which had just three common alleles and these exhibited minimal stutter and were typed clearly. Another reason could be the merging of four bins into two at locus Ssa171. Imperfect repeats at Ssa171 were responsible for the majority of errors found at this locus by O’Reilly et al. (1998). Overall, O’Reilly et al. (1998) found that Ssa197 had the highest rate of error of the tetra-nucleotides they employed, and this error was mainly due to dropout of larger alleles. This was not found to be a problem in the current study, possibly due to the length of extension in PCR. Longer extension times were used, of 30s and 40s in duration (in different multiplexes) compared to 20s as employed by O’Reilly et al. (1998).

The study by Hoffman and Amos et al. (2005a) relied exclusively on data from di-nucleotide repeat microsatellite markers, but nevertheless maintained low error rates. This may be due to prior selection of particularly robust loci that gave clear banding patterns. Six of the nine loci had already been typed for a sub-set of samples, so sources of error may already have been identified and at least partly rectified.

In this study the majority of errors were due to heterozygotes being incorrectly assigned as homozygotes. The vast majority of locus-specific errors occurred in locus Ssa2210. Particularly error free loci were One9, Ssa404 and Ssa202, where no locus-specific errors were identified.

An average mutation rate over all loci of $4.7 \times 10^{-4}$ was obtained. This is slightly higher than that found by O’Reilly et al. (1998) of approximately $3.4 \times 10^{-4}$ (in Atlantic salmon), and slightly lower than that calculated by Norris et al. (2000) for Atlantic salmon, of $7.8 \times 10^{-4}$, and that determined by Castro et al. (2004) for turbot,
of $6.7 \times 10^{-4}$. O'Reilly et al. (1998) used \( \text{Ssa}171 \), \( \text{Ssa}197 \) and \( \text{Ssa}202 \) and identified mutations in \( \text{Ssa}171 \) and \( \text{Ssa}197 \). Norris et al. (2000) found a mutation at \( \text{Ssa}171 \) and \( \text{Ssa}202 \). The distribution of mutations across studies suggests that none of the three loci (\( \text{Ssa}171 \), \( \text{Ssa}202 \) and \( \text{Ssa}197 \)) are particularly more prone to mutation than the other two.

### 3.2 Assignment of parents - power of exclusion

Using exclusion methodologies there are four outcomes of parentage assignment:

1. Correct assignment of a true parent
2. Incorrect assignment of a non-parent (false inclusion)
3. Correct exclusion of a non-parent
4. Incorrect exclusion of a true parent (false exclusion)

It is important to be aware of, and to quantify where possible, potential assignment errors. The predominant factors that can lead to the false exclusion of a true parent are genotyping errors and mutations. These have been covered in detail in the preceding sections of this chapter. Detected error rates were low and, as such, false exclusion was considered unlikely to be a significant factor in the current study. Furthermore, the exclusion programs employed (SIRFIND & DAMFIND) allowed this to be monitored by flagging mismatched alleles. The occurrence and extent of possible false inclusion errors are more difficult to assess. Though a possible contributing factor, genotyping errors and mutation are less of a concern for driving false inclusions. False inclusion rates are primarily modulated by the power of the loci used in the parentage analysis – a factor that is difficult to quantify precisely under field conditions. Factors that determine the power of a set of loci include the
number of loci (correctly) typed, the number of putative parents that have not been sampled (‘collected’) and the relatedness of putative parents.

3.2.1 False inclusion in an ideal population

One method to estimate the power of loci to distinguish between individuals, and therefore their power to exclude incorrect parents, is to estimate the probability of exclusion, based on an idealised population. Assuming Hardy Weinberg equilibrium, it is possible to estimate the power of exclusion when assigning the first or second parent, for each locus used individually, and all loci combined, based on the allele frequencies. As described in Marshall et al. (1998), the exclusion probability, \( P_l \) for the first parent at locus \( l \) with \( k \) codominant alleles and where \( p_i \) is the frequency of allele \( i \), is given by:

\[
P_l = a_1 - 4a_2 + 4a_3 - 3a_4 + 2a_5^2
\]

where

\[
a_n = \sum_{i=1}^{k} p_i^n
\]

The exclusion probability \( P_l \) for the second parent is given by:

\[
P_l = a_1 - 2a_2 + a_3 + 3(a_2a_3 - a_5) - 2(a_5^2 - a_4)
\]

where

\[
a_n = \sum_{i=1}^{k} p_i^n
\]

The overall average probability of exclusion (\( P \)) over \( n \) independently inherited loci is calculated by:

\[
P = 1 - \prod_{i=1}^{n} [1 - P_i]
\]
The equations above, as employed in CERVUS (Marshall et al. 1998), were used to calculate the probability of exclusion of an unrelated individual for the first parent (Table 3.8 and Table 3.9) and the second parent (Table 3.9). The exclusion probabilities for each locus used individually are shown in Table 3.8, and the combined exclusion probabilities in Table 3.9.

Table 3.8 Exclusion probability of individual loci. Excl(1)- exclusion probability for first parent. Calculated in Cervus (Marshall 1998).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Sample sets * used in calculation of Excl(1)</th>
<th>Excl(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2210</td>
<td>1991 1995 2000 2001 2004</td>
<td>0.493</td>
</tr>
<tr>
<td>2201</td>
<td>1991 1995 2000 2001 2004</td>
<td>0.778</td>
</tr>
<tr>
<td>410</td>
<td>1991 1995 2000 2001 2004</td>
<td>0.782</td>
</tr>
<tr>
<td>One9</td>
<td>1991 1995 2000 2001 2004</td>
<td>0.191</td>
</tr>
<tr>
<td>197</td>
<td>1991 1995 2000 2001 2004</td>
<td>0.741</td>
</tr>
<tr>
<td>408</td>
<td>2000 2001 2004</td>
<td>0.675</td>
</tr>
<tr>
<td>404</td>
<td>2000 2001 2004</td>
<td>0.833</td>
</tr>
<tr>
<td>401</td>
<td>2000 2001 2004</td>
<td>0.775</td>
</tr>
</tbody>
</table>

*Year of return of anadromous adults

Table 3.9 Total exclusionary power of combined loci (for assignment of the first and second parent). The nine loci refer to the uppermost nine in Table 3.8. Samples used in the estimations are the same sets employed in Table 3.8.

<table>
<thead>
<tr>
<th>Number of loci used</th>
<th>Total exclusionary power (first parent)</th>
<th>Total exclusionary power (second parent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0.999876</td>
<td>0.999998</td>
</tr>
<tr>
<td>12</td>
<td>0.999998</td>
<td>1.000000</td>
</tr>
</tbody>
</table>

These calculations suggest that 9 loci would be adequate to resolve parentage of the offspring of randomly mating, unrelated parents. However, the above calculations are based on a set of assumptions that are unlikely to be fully met in the Girnock Burn study. The sample sets are not derived from an ‘idealised’ population; not all sample sets are in Hardy Weinberg equilibrium, generations are not discrete, reproductive success of spawners is unlikely to be random, and the population is relatively small with a significant number of spawners likely to be very closely
related (full or half-sibs). Therefore in order to explore the probability of false inclusion further a number of empirical investigations were carried out.

### 3.2.2 False inclusion in a real population

During the course of this study seven distinct parental assignment sets (PAS) were considered, detailed in Table 3.10.

**Table 3.10 Details of parental assignment sets (PAS).**

<table>
<thead>
<tr>
<th>PAS #</th>
<th>Offspring sample (#)</th>
<th>Age of offspring (#)¹</th>
<th>Year of return of all possible parents</th>
<th>Year of return of collected parents</th>
<th># loci used²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1997 parr (274)</td>
<td>1 (262) 2 (12)</td>
<td>1995</td>
<td>1995</td>
<td>9 or 11³</td>
</tr>
<tr>
<td>2</td>
<td>2004 smolts (407)</td>
<td>1 (2) 2 (189) 3 (194) 4 (8)</td>
<td>2002 2001 2000 1999</td>
<td>12 12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1995 adult returns (116)</td>
<td>3 (14) 4 (59) 5 (27)</td>
<td>1991 1990 1989</td>
<td>9 or 11³</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1999 adult returns (10)</td>
<td>3 (9)</td>
<td>1995 1995</td>
<td>9 or 11³</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2000 adult returns (89)</td>
<td>3 (29) 4 (37) 5 (9)</td>
<td>1996 1995 1994</td>
<td>9 or 11³</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2001 adult returns (41)</td>
<td>3 (4) 4 (23) 5 (9)</td>
<td>1997 1996 1995</td>
<td>9 or 11³</td>
<td></td>
</tr>
</tbody>
</table>

¹ age estimated by scale readings. Individuals whose age was not estimated are excluded from this column. Age is in years and corresponds to the total age of fish. # sampled shown in parentheses

² Locus sets are detailed in Chapter 2

³ 11 loci where initial assignments with nine loci were ambiguous (i.e. two or more males or females could be assigned)

⁴ nine loci used were both sire and dam were assigned; 11 used where only dam assigned
Two main points are apparent from examination of this table. Firstly, different numbers of loci (with potentially different resolving powers) were employed. At least nine loci were screened, but where considered appropriate two extra loci were assayed (in order to resolve obvious ambiguous assignments or where higher confidence was justified, i.e. in assigning dams in PAS 3-6 where no anadromous male was also assigned). Towards the end of the project a 12th locus was added to the multiplex panel and employed in screening the final two datasets (PAS 2 and 7, Table 3.10). Secondly, not all the potential parents were DNA profiled. With regards to this, the PAS fall into two main categories. For the assignment of juvenile fish, parr sampled in 1997 and smolts sampled in 2004 (PAS 1 and 2), the vast majority of offspring were estimated to be of an age where the corresponding parents had all been genotyped. Conversely, for datasets requiring assignment of adult offspring to previous year classes (PAS 3-7), only one targeted parental year class was genotyped, whereas the offspring sample would have parents derived from three or more year classes.

Three different approaches were used to investigate potential false inclusion errors:

1) **Assessing assignment using different numbers of loci**

All individuals for two datasets (PAS 2 and 7) were genotyped for the full locus panel, i.e. at twelve loci. These datasets were reanalysed using the earlier, and potentially less discriminatory, reduced locus panel (i.e. nine or eleven loci). For reanalysis of PAS 2, eleven loci were employed only when assignment was ambiguous. For reanalysis of PAS 7, eleven loci were used in cases when only a dam was assigned using 9 loci (i.e. with no corresponding anadromous male). The outcome is detailed in Table 3.11.
Table 3.11 Comparison of the numbers of assignments achieved with full and reduced locus panels for two of the parental assignment sets.

<table>
<thead>
<tr>
<th>PAS</th>
<th>Dam Only assigned</th>
<th>Dam and Sire assigned</th>
<th>Neither assigned/ Ambiguous assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12 loci</td>
<td>9/11 loci</td>
<td>Diff^4</td>
</tr>
<tr>
<td></td>
<td>12 loci</td>
<td>9/11 loci</td>
<td>Diff^4</td>
</tr>
<tr>
<td></td>
<td>12 loci</td>
<td>9/11 loci</td>
<td>Diff^4</td>
</tr>
<tr>
<td>#2</td>
<td>162</td>
<td>162</td>
<td>0</td>
</tr>
<tr>
<td>#7</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

^1 Number of cases where parental assignments differed

Little or no reduction in assignment resolution was observed with the application of the reduced locus set.

2) Assessing assignment using parental datasets that were known non-parentals

An attempt was made to assign offspring from PAS 1 and 2 (Table 3.10; 681 progeny in total) to the largest adult dataset available (2004 spawning run; n =117). The latter set could not possibly have been parents of these offspring. Nine loci were employed in the assignment of PAS1 offspring, and twelve loci in the assignment of PAS2 offspring. For PAS1 offspring, four (false) dam assignments and zero dam-sire pair assignments were observed, equivalent to a false inclusion rate of 1.5%. For PAS2 offspring, only one (false) dam assignment was observed and no dam-sire pairs were assigned, equivalent to a false inclusion rate of 0.2%. This is slightly higher than that predicted by Cervus, as might be expected as the stock in Gurnock Burn is not an ‘idealised’ population.

3) Assessing assignment where random errors were introduced to offspring datasets

In order to mimic a genotyping error or mutation, datasets were edited by hand to simulate a mistyping. In each case one locus genotype was selected and one allele score changed to that of the most common allele for that locus. This was carried out
for four samples in each set of offspring in PAS 1-7 (Table 3.10) and assignments performed. The outcomes were either that the initial female was still assigned or that a mismatch was created. None of the changes resulted in the assignment of an incorrect dam or dam-sire combination.

### 3.2.3 Conclusion - power of parentage analysis

When both dam and sire are positively assigned the probability of false inclusion tends to be extremely low, given the greater exclusionary power associated with assigning the second parent. As evidenced from both the theoretical calculations and empirical studies conducted here, it is unlikely that dam-sire combinations identified in this study are the result of false inclusion. There is a greater chance that a small proportion of dam only assignments may be false assignments. Due to such possible occurrences (albeit likely to be very few in number), only dam-sire assignments were used when assessing scale reading error.

### 3.3 Ageing salmon from scale readings

The age of Atlantic salmon can be estimated from the reading of ‘rings’ (circuli) that are seen in their scales. Rings are closer together when a fish is growing more slowly, i.e. during the winter months and when in freshwater. Thus, careful and practised scale examination can reveal the age of a fish, when it migrated to sea, and in some instances when, and how often a fish spawned. The determination of age of Atlantic salmon from scale readings is particularly useful in research on wild stocks. It has been applied in a vast number of studies (e.g. Summers 1996; Whalen et al. 2000; Jones et al. 2002; Stewart et al. 2002; Baum et al. 2004). However, the process is not error-free. One problem is that the ring pattern caused by winter growth can be mistaken for a ‘summer check’- when growth is slow during the summer (ICES 1984). In some cases the first annulus may be missed (Jensen and
Data quality and accuracy

Chapter 3

Johnson 1982), although this is generally not a problem with salmon from Girnock Burn (I. McLaren, personal communication). Another issue is the erosion of scale edges and sometimes the scale surface, as a result of re-absorption of material during the freshwater period prior to spawning in anadromous fish (ICES 1984). This is predominant in male fish due to the development of secondary sexual characteristics and tends to affect the estimation of sea-age (I. McLaren, personal communication). Tagging of fish for genetic analysis can be employed to determine the accuracy of scale readings, although there is very little information available on the results of such research. However, Lund and Hansen (1991) found 99% concordance between estimates of the sea-age of Atlantic salmon obtained from tagging and scale readings.

In this study fish age determined by scale reading (performed in all cases by FRS staff) and parentage analysis were compared. This was carried out in the following groups of samples:

- Group 1: 1+ parr sampled in 1997
- Group 2: 2 and 3 year old smolts sampled in 2004
- Group 3: Adult returns with 2-4 years spent in freshwater and 1-3 in sea-water.

It was possible that small numbers of samples in groups 1 were not aged 1+, and for small numbers in group 2 not to be aged 2 or 3 years.

3.3.1 1+ Parr sampled in 1997

Parr electrofished from the length of Girnock Burn in 1997 were selected by length to be presumed 1+ parr. Scale samples were also taken in order to verify their age. Of 152 parr assigned to both parents, three had not been aged as 1+ from scale readings. However, repeat scale readings confirmed that original readings were at
fault. Overall there was a 2% discrepancy (3/152) between age determined by genetic analysis and initial scale readings.

### 3.3.2 Smolts sampled in 2004

Smolts that were sampled as they left the traps in 2004 were mostly 2 or 3 years old. Of those that had been identified as 3 years old from scale readings and were assigned to both parents, 8% (10/130) could not be assigned to adults of the appropriate spawning year i.e. 2000. Of those that had been identified as 2 years from scale readings and were assigned to both parents, 3% (3/96) were not assigned to adults of the appropriate spawning year i.e. 2001. Smolts that had not been aged with confidence from scales, for example classified as 2?, were discounted from this analysis.

### 3.3.3 Adult sea returns of 1995, 1999, 2000, 2001 and 2004

Over the course of the study 54 returning spawners could be confidently assigned to parents of a previous spawning year (only individuals where both parents were assigned were considered). Thus the total ages of these individuals were known and compared to scale reading estimates (Table 3.12). All discrepancies involved an age estimate difference of one year only.
Table 3.12 Number of discrepancies between ages obtained from scale reading and inferred from parentage assignment.

<table>
<thead>
<tr>
<th>Year of return of parents tested</th>
<th>Year of return of offspring</th>
<th>Total number of offspring assigned</th>
<th>% (number) of all assigned with discrepancies in total age</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000 2001</td>
<td>2004</td>
<td>11</td>
<td>9.1 (1)</td>
</tr>
<tr>
<td>1991</td>
<td>1995</td>
<td>11</td>
<td>0.2 (0)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>54</td>
<td>7.4 (4)</td>
</tr>
</tbody>
</table>

1 Only offspring assigned to a female AND an anadromous male were included.

2 One sample was incorrectly aged on a spreadsheet, but was correctly aged on a scale packet. As the error was in data input as opposed to scale reading it was ignored here but included later when the error associated with using data from spreadsheets was required.

3.3.4 Discussion

Error rates ranged from 2% in estimation of parr aged 1+, to 7% in sea-returns and 8% in fish estimated as 3 year old smolts. These are higher than the 1% error rate found by Lund and Hansen (1991). Lund and Hansen (1991) only estimated the sea-age for wild fish, so it may be that errors in this study of anadromous fish were largely due to differences in the estimation of the number of years spent in freshwater. This was definitely the case in one of the four discrepancies found in this study for anadromous fish, but it was not possible to determine this for the remaining three samples. It may also be that erosion was not a problem in the stock studied by Lund and Hansen (1991), or that they sampled fish near to the river mouth, prior to its occurrence. In this study three of the four anadromous fish that were aged incorrectly were males. Although numbers are very small, since erosion occurs to a greater extent in males it does not rule out the possibility that this was a major contributory factor. Other discrepancies may have been due to human error in data input (one such error was identified, Table 3.12), inadvertently swapping scales during sampling or in reading large numbers of scales. In conclusion, there does not seem to be a particular cause for the errors found, but a range of uncertainties involved in estimating the age of Atlantic salmon from scales are likely to be
implicated. The errors identified in age estimation from scales in this stock may not necessarily apply to those that may occur in other stocks, due to possible differences in life histories.
3.4 Overall discussion- applications to other studies

The previous sections considered a range of areas where accuracy is important and ways in which they can be tested. To avoid errors in genotyping it was considered important to optimise the type of loci (avoiding di-nucleotides), PCR conditions and the width of bins to reflect the accuracy of size estimation and proximity of alleles. To test the accuracy of parentage analyses in this case it was useful to employ empirical tests, which could also be the case for other wild populations. A number of factors mean that estimating the age of Atlantic salmon from scale readings is not a clear-cut process and errors may be dependent on the study population and time of sampling.
Investigating natural and quasi-natural spawning

Chapter 4

4 Investigating natural and quasi-natural spawning

4.1 Introduction

4.1.1 Reproductive success and mating system in the wild

Spawning of wild Atlantic salmon takes place during autumn and winter.

Anadromous males compete with each other for access to females that are about to spawn, while mature parr adopt a ‘sneaking’ strategy. The female constructs a ‘nest’ or ‘egg pocket’ by moving away gravel to form a depression. The female then deposits eggs while milt is simultaneously released by the male(s). The eggs are subsequently covered with gravel, through movements made by the female. Another nest may be constructed in the depression left by the first nest. More nests may also be made in this way, with those in close proximity forming a ‘redd’. Initial observations suggested a ratio of one redd to one female (Hay et al. 1987, cited in Barlaup et al. 1994). However, it is now known that females may construct more than one redd, and that multiple females may contribute to a single redd (Barlaup et al. 1994; Taggart et al. 2001). No after-spawning care of eggs or offspring occurs. Although a wealth of information on the phenotypic plasticity of Atlantic salmon has been collected over the years (reviewed by Klemetsen et al. 2003), details of the mating system in Atlantic salmon have only recently been revealed. Initially there was a popular belief that anadromous fish remained monogamous throughout the spawning period (Belding 1934), although early observations of spawning in the wild suggested this was not the case as paired males were seen to be displaced by other males. Belding (1934) described it thus: ‘the role of [the] favored male is not secure, since he may be supplanted at any time during the season, day or hour’. In addition, Webb and Hawkins (1989) observed anadromous males moving between one or more females, in a study that involved tracking of radio-tagged anadromous
fish. The application of molecular markers has allowed a more detailed and definitive study of the reproductive success of spawners. Thompson et al. (1998) and Martinez et al. (2000), using genetic profiling techniques, deduced from analyses of offspring taken from discrete redds, that multiple males had fertilised eggs from a single female in the wild. However, in these studies they were not able to distinguish between mature parr and anadromous males as the parental fish had not themselves been genotyped. Martinez et al. (2000) also found females mating with more than one anadromous male, but in an artificial situation. Subsequent studies on wild stocks, also based on genetic profiling analyses, revealed that females mating with multiple anadromous males and males mating with multiple females were relatively common occurrences (Garant et al. 2001; Taggart et al. 2001). Garant et al. (2001) found that the reproductive success of males was positively correlated with the number of females they spawned with, although they stated that more research was required to confirm this observation.

Taggart et al. (2001) studied the spawning success of Atlantic salmon in Girnock Burn. This was investigated by determining the parentage of eggs that had been removed from naturally spawned redds. Attempts were made to sample eggs from all redds that were constructed, although there was no means of determining the extent to which this had actually been achieved. One of the three spawning years included in this survey was 1995. In this year 25 (of a total of 66) female spawners were not represented in the eggs that were sampled. Possible reasons for this finding include: the females had not spawned, redds had not been identified and so were not sampled, or superimposition led to eggs from some females not being sampled. One aim of the current study was to gain further insight into the reasons for the lack of representation of these females. This was achieved by sampling offspring from
outwith the confines of the redd, therefore not relying on redd identification for
sampling to take place. This also allowed an insight into the effectiveness of
sampling of the two schemes.

4.1.2 Supplementation schemes

The manipulation of wild stocks of salmonids has a long history. They have been
widely introduced into the southern hemisphere (Allendorf and Waples 1996),
although they are native to the northern hemisphere. For example brown trout
(Salmo trutta) have been introduced to Australia, South America and Africa (Laikre
1999), although they originate in Europe. Another example is the introduction of
Atlantic salmon from Canada to Australia in the 1960s (Reilly et al. 1999).

However, many populations and species have suffered declines, with 22 species of
salmonids being included in the 1994 IUCN Red List of Threatened Animals
(Allendorf and Waples 1996). Declines have been attributed to, inter alia,
overfishing, pollution, construction of hydroelectric schemes and other losses of
habitat.

Instead of treating the root cause of the decline, in numerous cases supplementation
of wild fish with fish reared in hatcheries has been used in an attempt to maintain
numbers. Unfortunately fishery management was generally not conducted with the
aim of conserving stocks prior to the 1970s (Utter 2004).

Monitoring of such programs was not usually implemented. With the advent of
genetic and biochemical markers it became possible to analyse the success (or
failure) of hatchery programs and their effect on native stocks. The first of these
programs involved using allozymes with alleles that differed in frequency between
hatchery and wild populations (e.g. Krueger and Menzel 1979 in brook trout
Salvelinus fontinalis). In later cases hatchery fish were bred to be homozygous at
particular allozyme loci (that differed to that found in the wild) thus allowing their identification following introduction (e.g. Taggart and Ferguson, 1984 in brown trout *Salmo trutta*; Chilcote *et al.* 1986 in steelhead trout *Oncorhynchus mykiss*).

Subsequent analyses have employed microsatellites in combination with more sophisticated analytical methods. These have included determining the origin of individuals through assignment tests whereby samples are assigned to wild or hatchery stocks based on multi-locus genotypes (e.g. Hansen *et al.* 2001a; Nielsen *et al.* 2001; Ruzzante *et al.* 2001). These studies have reported the full scope of potential outcomes, ranging from no or limited introgression, through some introgression with wild fish, to significant displacement of the wild stock (Hindar *et al.* 1991; Laikre 1999). However, in general, domesticated fish (either intentional hatchery releases or unintentional escapees from farming) have been found to be less fit than wild stocks in the natural environment (Hindar *et al.* 1991; Miller *et al.* 2004; Ruzzante *et al.* 2004; McGinnity *et al.* 2004).

With increasing recognition of the importance of conserving local adaptations, it was generally recommended that in order to maintain the genetic integrity of populations native fish should be used as broodstock in restoration programs (e.g. Hindar *et al.* 1991; Cuenco *et al.* 1993; Youngson *et al.* 2003; Ferguson 2004). Such programs, which involve native fish with the aim of conserving local populations, are known as ‘supportive’ (e.g. Tessier *et al.* 1997; Palm *et al.* 2003a; Ferguson 2004; Ruzzante *et al.* 2004) and will be referred to as such in this work. However, precise definitions can differ between authors (Goodman 2005).

Nevertheless, supportive breeding programs still have the potential to have negative impacts on the wild stock. Possible effects of supportive breeding are:
• A reduction in the genetic variation (due to inappropriate structuring of the breeding program)

• A reduction in the adaptation of the stock to the environment (through the inadvertent selection of characters beneficial in the artificial environment).

One particular downfall of supportive breeding schemes has been referred to as the ‘Ryman-Laikre effect’ (Waples and Do 1994; Tessier et al. 1997; Koljonen 2001). Under circumstances where the survival of offspring reared in a protective, captive environment is higher than that of offspring reared in the natural environment, the reproductive success of a segment of the population is increased (Ryman 1991). This results in an increase in the overall variance in reproductive success of breeders. Therefore, although supportive breeding can increase the census number of breeders in the short term, it may also result in a counter-productive decrease in effective population size ($N_e$), which is particularly apparent when few individuals are used as broodstock in the captive breeding (Ryman 1991).

Some empirical analyses of supportive breeding schemes have taken place. Tessier et al. (1997) estimated the effect of supportive breeding on $N_e$, taking into account unequal sex ratio and differences in outputs of spawners in the wild compared to those used as broodstock. In this study offspring were released as smolts. They showed that $N_e$ was more similar to the total number of actual breeders when the proportion of offspring from wild and hatchery breeders was more even. However, they did not take into account the variance in reproductive success of individual breeders.

Hansen et al. (2000) found strong indications of reduced effective population size in two populations of brown trout undergoing supportive breeding. The reason for the lowered $N_e$ was thought to be due to the pooling of milt from several males prior to
addition to eggs. Previous studies on hatchery procedures whereby milt from several males were mixed prior to addition to eggs showed that sires were not equally successful in fertilising eggs, in pink salmon (*Oncorhynchus gorbuscha*) (Garrat and Shirley 1985), chinook salmon (*O. tshawytscha*) (Withler 1988) and rainbow trout (*O. mykiss*) (Gile and Ferguson 1995). Thus in order to obtain a more even representation of males used as broodstock, a better strategy is to split eggs into batches and fertilise each batch with milt from a single male.

Tessier *et al.* (1997) and Hansen *et al.* (2000) used different methods to investigate effective population size; demographic and genetic approaches, respectively. Ardren and Kapuscinski (2003) used a demographic method to estimate the effective population size of a wild population of steelhead trout, accounting for unequal sex ratio, variable reproductive success of spawners and fluctuating population size. Approaches to estimating $N_e$ are described in more detail in Chapter 8.

In addition to possible reductions in effective population size as a result of supportive breeding, domestication selection may inadvertently take place. For example, Heath *et al.* (2003) found that egg size of chinook salmon continually declined over a period of approximately 10 years, during rearing in a hatchery environment. In two other populations of wild chinook salmon supplemented with hatchery reared fry, a decrease was also seen (Heath *et al.* 2003). McGinnity *et al.* (2004) found characteristics of an Atlantic salmon stock had changed after going through 18 generations of rearing in a hatchery (the broodstock initially originated from the native stock, but for the vast majority of spawning years broodstock consisted entirely of returning adults that originated in the hatchery). These characteristics included sex ratio (an excess of female returns), time of return to freshwater (later in the year) and the proportion of mature parr (an increase). There
may be inadvertent selection for early spawners, which may arise if all space in the
hatchery is taken by the time final spawners arrive. In the long term this can pose
problems because early spawners may be disadvantaged in the wild (Waples 1991).
Myers (2004) described how the lack of imprinting to natal sites led to greater
straying rates in hatchery salmon. Differences have been seen in the sensory
lamellae structures of hatchery and wild smolts of Atlantic salmon, those of wild
smolts being more developed than those of hatchery origin, which suggests this may
be implicated in differences in straying rates (CEFAS 2003). Hedrick et al. (2000a)
describe a supportive breeding program for chinook salmon (O. tshawytscha) that
was curtailed partly due to the hatchery fish returning largely to the site of the
hatchery rather than the natural spawning location of the stock.
An indirect consequence of hatchery releases of coho salmon in Oregon, which led
to an increase in numbers, was its de-listing from the Endangered Species Act in
this region (Myers et al. 2004). Clearly this was not a factor that aided recovery.
In order to reduce possible domestication selection it is favourable for the fish to
spend a minimal amount of time in the artificial environment. However, this needs
to be weighed against the gain in survival when offspring are reared for longer in
the hatcheries (Cuenco et al. 1993).
Further difficulties with supportive breeding have occurred when the best approach
to management is unclear. For example Primmer et al. (1999) reported on a stock of
Arctic charr in southeast Finland, where very few wild spawners could be caught for
broodstock. The decision had to be made to either increase numbers of broodstock
used by introducing fish from a different source, or to maintain the genetic integrity
and use those available. The latter option was taken. A similar problem arose in
Virginia (USA) where a restoration program for the American shad (Alosa
**Investigating natural and quasi-natural spawning**

* *apidissima* was being carried out. In this case no broodstock had been caught in the native river, leading the managers to resort to using broodstock from a nearby river (Brown *et al.* 2000).

Monitoring of supportive breeding programs that release offspring that are unable to be physically tagged (such as eggs or fry) is hindered as there should be little genetic difference between wild and released fish. In these cases it is necessary to determine the productivity of parents used in artificial spawning by using more intensive (and expensive) DNA profiling techniques to trace the origin of offspring.

### 4.1.3 Supportive breeding at Girnock Burn- Quasi-natural spawning

Due to consistent low returns at Girnock Burn in the late 1990s, natural spawning has been supported by or, more recently, replaced with ‘quasi-natural’ spawning since 2000. The process of quasi-natural spawning, as implemented at the Girnock Burn, involves stripping fish of eggs and milt, and then setting up crosses whereby eggs from each female are fertilised with milt from each male. Mature male parr are also used in order to mimic the natural situation (of fertilisation by sneaker males) at Girnock Burn, although the numbers of parr used are likely to be far fewer than those that contribute to natural spawning. A few minutes after adding milt, extra ‘insurance’ milt from a random male is added (the same male is used for each female). This is to protect against the unintentional use of sterile males in the first instance resulting in unfertilised eggs. Eggs from each female are then re-combined and layered between stones in baskets. The baskets are placed in a secured, covered ‘incubator’ in Girnock Burn. The following spring eggs are removed from the incubators and counted. Eggs from all families are mixed and planted out at an approximately even density throughout the burn, avoiding areas known to be especially prone to groundwater, which has shown to impair embryo development.
Investigating natural and quasi-natural spawning

(Youngson et al. 2004). Eggs are planted out in small containers (perforated with holes), containing small stones, in the stream bed. This approach minimises the period of time offspring spend in an artificial environment, in order to reduce selection for characteristics that may be beneficial in these circumstances, but not in the wild. Planting out as eggs also means that homing should not be impaired as homing capability is believed to be mainly developed during smoltification (when offspring of quasi-natural spawning would be in the natural environment) in a phase known as the sensory period for olfactory imprinting (Morin and Doving 1992). One theory behind imprinting is that the portion of the population in freshwater provide the chemical cue that is imprinted and later homed towards (Nordeng 1971; Morin and Doving 1992). Crosses were set up with the intention of having a relatively even output from each male and female used in spawning, thus maintaining or possibly increasing the effective population size in comparison to natural spawning. Another possible benefit of quasi-natural spawning was that regions of the burn inaccessible to spawning adults, but adequate as nursery ground, could be utilised as areas in which to plant out eggs, thus potentially increasing the carrying capacity of the burn.

4.1.4 Objectives

The two main aspects that were considered in this chapter were:

- An investigation into the spawning success of anadromous returns in 1995 (when only natural spawning occurred). This was explored by sampling 1+ parr offspring from outwith the confines of the redd and determining parentage. Findings were compared to those of Taggart et al. (2001), which were based on sampling eggs from within redds.
A comparison of the success of quasi-natural and natural spawning, in terms of the output and representation of spawners. Natural spawning that occurred in 1995, 2000 and 2001 was compared to quasi-natural spawning (that took place alongside natural spawning) in 2000 and 2001.

The success of the supportive scheme was investigated by utilizing preserved fin biopsies from smolts that were collected in 2004 as they passed through the traps at Girnock Burn on their spring migration. The vast majority of smolts were offspring of adults that spawned in 2000 and 2001. (Adult returns in 2000 and 2001 had also been sampled, although unfortunately the mature parr utilised in quasi-natural spawning had not). In both these years the majority of fish that returned to the burn were used in quasi-natural spawning, the remainder being placed above the traps to spawn naturally. This presented the opportunity to give an initial insight into the success of the scheme and possible problems associated with the approach. The monitoring should be useful in planning future management practices to maximise overall effective population size and in planning future sampling for monitoring purposes.
4.2 Materials and Methods

4.2.1 Samples

4.2.1.1 Natural spawning in 1995

The spawning event in 1995 involved the largest run of salmon entering the burn during the 1990s. A total of 128 fish returned to the traps, with 109 being placed above the traps to spawn naturally (a few kelts and diseased or dying fish were excluded). Of the 128 returns, tissue samples were retained and screened from 116 samples (including the 109 placed above the traps to spawn naturally), as described in Table 2.1. The main spawning run was completed in seven days, between the 25th and 31st October.

Sampling of eggs: During the spawning season (October-December) the burn was surveyed for redds every 1-2 days. An attempt was made to sample eggs from redds within 48 hours of spawning. In addition, photographs were taken and a marker stone placed on the bank to assist in locating the redd at a later date. The following spring redds from which no eggs had been obtained were re-excavated (this time more extensively) in an effort to obtain offspring samples. The eggs sampled in 1995 were reared in a hatchery and preserved (in ethanol) as alevins. It was intended that 20 eggs should be sampled from each redd, although in some cases fewer were found. DNA from samples (of eggs or alevins) was extracted using a modified phenol-chloroform method (Taggart et al. 2001).

Sampling of parr: During September 1997 a total of 274 presumably 1+ parr were sampled from the length of the burn. One sample was taken approximately every 30m, beginning at the traps and ending at the head of the burn. Adipose fin clips and
scale samples were taken from each sample and then the fish were returned to the burn.

All results that involved the representation of spawners in egg samples were obtained from a study carried out by Taggart et al. (2001).

4.2.1.2 Quasi-natural spawning alongside natural spawning (2000 and 2001)

Spawning events in 2000 and 2001: Due to consistent low returns in 1996-1999, FRS (Fisheries Research Services) decided to introduce a quasi-natural spawning regime for the stock in Girnock Burn in order to optimise the use of the eggs that were input into the burn. In 2000 a total of 89 fish returned to the traps. Of these, 76 were potential contributors to spawning above the traps, either through quasi-natural or natural spawning (although tissue samples were retained and screened from all 89 individuals, as described in Table 2.1). The run began on the 21st October and the last fish was caught in the trap on the 12th November. Of these, 14 females and 32 males were used in quasi-natural spawning while eight females along with 22 males were placed above the traps (on the same day) to spawn naturally (Table 4.1). In 2000 two sets of crosses were constructed during quasi-natural spawning (entitled set A and set B). In set A, nine females were utilised. The eggs from each female were divided into 21 roughly equal portions and each portion mixed with milt from a single male. Fourteen portions were each mixed with milt from different anadromous males, and seven portions with milt from different mature parr. Set B consisted of five females, each crossed with 18 anadromous males and three mature parr. Crosses were constructed in the same manner as for set A. Following the construction of crosses milt from an extra ‘insurance’ male was added. The insurance male was ‘likely’ to be a mature parr (no written record of the type of
Investigating natural and quasi-natural spawning

Chapter 4

male was made, although this information was obtained from a person present at the time).

In 2001 only 41 fish returned to the traps. The run began on the 8<sup>th</sup> October and finished on the 11<sup>th</sup> November. Seven anadromous males and 12 females were retained for quasi-natural spawning, and nine females together with eight males were released above the traps (over a three day period) to spawn naturally (Table 4.1). In this year only one set of crosses was constructed; eggs from each of the 12 females were divided into 11 roughly equal portions and each portion mixed with milt from a single male. Seven anadromous males and four mature parr were used. Following the construction of crosses, milt from an extra ‘insurance’ male was added. The insurance male was a mature male parr.

Table 4.1 Number of males and females that underwent quasi-natural and natural spawning in 2000 and 2001. The number of full-sib families in each set of quasi-natural spawning is shown. The number of females that were noted as part-spent are shown in brackets.

<table>
<thead>
<tr>
<th></th>
<th>Quasi-natural spawning</th>
<th>Natural spawning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000 set A&lt;sup&gt;1&lt;/sup&gt;</td>
<td>2000 set B&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>Number of females</td>
<td>9</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Number of anadromous males</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Number of mature parr</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Number of families (QN only)</td>
<td>189</td>
<td>168</td>
</tr>
</tbody>
</table>

<sup>1</sup>Two separate sets of crosses were set up in 2000.
<sup>2</sup>Records were incomplete for one female, the males used to fertilise eggs were not known. QN-quasi-natural.

Offspring of spawning in 2000 and 2001: Offspring were sampled as smolts caught in the downstream trap as they exited the burn, between the 3<sup>rd</sup> February and the 6<sup>th</sup> May in 2004. The adipose fin from one in five smolts was preserved in ethanol. Scale samples were also taken from these same biopsied fish for aging. A total of 410 adipose fin clips were preserved.
4.2.2 Genotyping

Egg samples and samples from adult returns in 1995 were screened with ten single locus minisatellite probes (Taggart *et al.* 2001). These data were made available for this project.

Parr samples and samples from adult returns in 1995 were typed at 9-11 microsatellite loci. Smolts sampled in 2004 and adult returns sampled in 2000 and 2001 were typed at 12 microsatellite loci (see Materials and Methods, Chapter 2).

4.2.3 Analysis

4.2.3.1 Parentage Analysis

Eggs sampled in 1995 and 1996 were assigned to adult returns of 1995 (as described by Taggart *et al.* 2001).

Parr that were sampled in 1997 were assigned to adult returns of 1995 using DAMFIND and SIREFIND (described in Materials and Methods, Chapter 2). Parentage of 407 smolts that were sampled in spring (February to May) 2004 was also determined using DAMFIND and SIREFIND (described in Materials and Methods, Chapter 2). Smolts were assigned to adult returns of 2000 and 2001.

4.2.3.2 Reproductive skew of natural and quasi-natural spawning

The reproductive success of adult spawners was estimated from the number of offspring (parr or smolts sampled) that they contributed to. A comparison was made between the representation of adult spawners in quasi-natural and natural spawning. The number of females mating with each male (and the number of males mating with each female) was also estimated for the two spawning regimes.
In addition, the representation of spawners was compared by using a measure of skew, the B index (Nonacs 2000, 2003). A review of 21 different measures of reproductive skew (prior to the introduction of the B index) showed that no single measure of skew was satisfactory when comparing groups of differing size or with different outputs (Kokko et al. 1999). In addition, some were difficult to interpret because the indices calculated could not easily be compared to a value corresponding to a random distribution (of offspring among parents)(Kokko et al. 1999). However, Nonacs (2000) devised the B index and compared it to seven other indices. It was found to be the single measure that was robust to different group size and productivity, and generally had the most positive features and was the most consistently reliable (Nonacs 2003). It was proposed as a benchmark measure (Nonacs 2000, 2003). To my knowledge no opposing views to this have been published. The B index has since been used to measure reproductive skew in the Argentine ant Linepithema humile (Fournier and Keller 2001), the social wasp Polistes carolina (Seppä et al. 2002), the Mediterranean ant Pheidole pallidula (Fournier et al. 2004), the rhesus macaque Macaca mulatta (Widdig et al. 2004), the tropical hover wasp Parischnogaster melyi (Fanelli et al. 2005), the mountain gorilla Gorilla beringei beringei (Bradley et al. 2005) and the Australian allodapine bee Exoneura robusta (Langer et al. 2006).

The value of the B index that corresponds to a random distribution (of offspring between parents) is zero. It is negative for a more even (than random) distribution and positive for a more skewed distribution. The Skew Calculator version 1.2 was used to estimate the B index, 95% confidence intervals of the index and a probability value for deviation of the distribution (of offspring among parents) from random (available from www.obeec.ucla.edu/Faculty/Nonacs, Nonacs 2003). The
confidence intervals assume a geometric distribution and therefore each group was tested for this using a Kolmogorov-Smirnov test for a geometric distribution implemented in STATISTICA v. 7.1.

4.2.3.3 Influence of natural and quasi-natural spawning on the effective number of breeders

The rate at which genetic diversity is lost is related to the effective population size \(N_e\), which is a key parameter in conservation (Palm et al. 2003b). Despite the actual value of \(N_e\) being difficult to estimate, it was still possible to investigate the effects of factors known to influence the effective population size. Although many factors affect the effective population size, those that are likely to differ between natural and quasi-natural spawning were of interest in this case. These were the variance in reproductive success along with the number of breeders and sex ratio, which could be drastically altered by the number of mature parr contributing to spawning.

The demographic effect on the effective number of breeders in a year was estimated using equations taken from Ardren and Kapuscinski (2003) (Equation 4.1-Equation 4.3). These estimate the influence of variance in reproductive output for males (Equation 4.1) and females (Equation 4.2), and the combined effect of these and unequal sex ratio (Equation 4.3).

\[
N_{\text{e,male}} = \left( N_{\text{m}} - \frac{1}{\bar{k}_{\text{m}}} \right) \left[ \frac{1}{\bar{k}_{\text{m}}} + \left( \frac{\sigma_{\bar{k}_{\text{m}}}}{\bar{k}_{\text{m}}} \right)^2 \right]^{-1}
\]

Equation 4.1Effective number of male breeders in a brood year taking into account variable reproductive success (from Ardren and Kapuscinski 2003).
Equation 4.2 Effective number of female breeders in a brood year taking into account variable reproductive success (from Ardren and Kapuscinski 2003).

\[ N_{bf} = \frac{\left( \bar{k}_f - 1 \right)}{\left[ \bar{k}_f + \left( \frac{\sigma^2_f}{\bar{k}_f} \right) - 1 \right]} \]

Equation 4.3 Effective number of breeders in a brood year taking into account unequal sex ratio and variable reproductive success of females and males (from Ardren and Kapuscinski 2003).

\[ N_{bf(Demo)} = 4 \left( \frac{N_{bf} \times N_{bm}}{N_{bf} + N_{bm}} \right) \]

Where:

- \( N_{bm} \) = Effective number of male breeders in a brood year taking into account variable reproductive success
- \( N_m \) = Census number of males
- \( \bar{k}_m \) = Mean output per male breeder
- \( \sigma^2_{km} \) = Variance in output of male breeders
- \( N_{bf(Demo)} \) = Effective number of breeders in a brood year, taking into account unequal sex ratio and variable reproductive success.

When \( f \) was substituted for \( m \), the corresponding value for female (rather than male) breeders was used.

The effective number of breeders in a brood year, accounting for unequal sex ratio and variable reproductive success was estimated for:

1. natural spawning in 1995
2. quasi-natural and natural spawning in 2000 and 2001

The reproductive success of spawners in 1995 was measured separately from egg and parr data. The reproductive success of spawners in 2000 and 2001 was estimated from smolts sampled in 2004.
The mean and variance of reproductive success employed were those estimated from the anadromous spawners (this was likely to differ when the reproductive success of mature parr was also taken into account, but this information was unavailable as samples of mature parr were not taken). The number of mature parr that contributed to natural spawning was estimated (as 10, 800, 1500 and 4000, for reasons described below) and included in $N_m$, the census number of males. The number of mature male parr used in quasi-natural spawning was also included in the calculations where appropriate.

The actual number of mature parr contributing to natural spawning was not known. In order to model the effects that different numbers of mature parr may have, a range of values was used. The four values used were:

1. an arbitrary low estimate of 10
2. an intermediate estimate of 800 (obtained by multiplying the number of redds identified in 1995, 109 (Taggart et al. 2001), by the maximum number of parr, seven, thought to have contributed to a single redd, based on limited sampling)
3. an intermediate estimate of 1500 (obtained by assuming 80% of male parr mature, that they do not mature until aged 1+, that half the parr in the burn were males, that half of the mature parr contributed to spawning, and mortality of parr between summer and autumn is 12.5%)
4. an upper limit of the number of mature male parr that might have been in the burn, of 4000 (assuming half the parr in the burn were male, that they do not mature until aged 1+, and mortality of parr between summer and autumn is 12.5%)

§ These estimates were obtained by using data from electrofishing that took place in Girnock Burn in the summers of 1969-1986 and 1999-2005. Initially, the
number of parr caught was extrapolated to the whole area of the burn above the traps (giving 9000 parr aged 1+ or above), then the assumptions above applied to give maximum and intermediate estimates. Data were supplied by Phil Bacon (personal communication).

In an additional, exploratory, analysis the effective number of breeders, taking into account variable reproductive success and unequal sex ratio (i.e. $N_{b(demo)}$), was estimated for spawners that underwent quasi-natural spawning in 2001, but under the hypothetical situation that these fish had spawned naturally. Thus the mean and variance of reproductive success estimated from natural spawning that took place in 1995, were applied to the spawners in 2001 that underwent quasi-natural spawning. The numbers of mature parr were estimated as 10, 800, 1500 and 4000, as previously for natural spawning.

One drawback of the estimates of variance in reproductive success for naturally spawning males was that they were based only on data from anadromous males. Input from mature parr would be likely to increase the variance. Therefore the effect of doubling the variance in reproductive success of males was investigated for natural spawning.

4.2.3.4 Reproductive Success in terms of numbers of smolts produced

The total number of smolts sampled that originated from quasi-natural and natural spawning (disregarding the contribution from each spawner) were compared in two ways. First, the average output of smolts per female that underwent quasi-natural and natural spawning was obtained. Secondly, the output of smolts from quasi-natural and natural spawning, taking into account the number of eggs initially used, was analysed. The number of eggs used in natural spawning was estimated by using
data from females employed in quasi-natural spawning. The length of each female used in quasi-natural spawning and the number of eggs it produced was known and used to construct an appropriate predictive model that estimated the number of eggs produced by any given female, given its length.

4.2.3.5 Survival of offspring of quasi-natural and natural spawning

Adult returns of 2004 aged 2.1, that had been spawned above the traps were assigned to spawners in 2000. A total of 13 sea returns of 2004 were assigned to spawners in 2000.
4.3 Results

4.3.1 Natural spawning in 1995

A total of 66 females were potential spawners in 1995. The majority of females identified were represented in both egg and parr samples (Figure 4.1 and Figure 4.2).

![Venn diagram showing the number of female spawners represented in egg and parr samples from natural spawning in 1995.](image)

Figure 4.1 Venn diagram showing the number of female spawners represented in egg and parr samples from natural spawning in 1995.

A limited number of females were represented only in egg or parr samples (three and six, respectively). However, there was a general consensus between the two sampling methods, with 19 females not being represented by either sample (Figure 4.2 and Figure 4.3).
Figure 4.2 The output of each anadromous female that was placed above the traps in 1995. Offspring were sampled as eggs (top) and 1+ parr (bottom). Date of placement of fish above traps is indicated by arrows (fish to the right side of arrows were placed above the traps on this date).
FEMALE SPAWNERS THAT WERE NOT DETECTED

Figure 4.3 Venn diagram showing the number of female spawners not represented in egg and 1+ parr samples from natural spawning in 1995.

Slightly more females were represented in the parr sample (44 compared to 41).

This may have been due to the sampling technique employed, which did not rely on the identification of redds, or a particular portion of the redd being sampled. For at least two females (#103 and #111; Figure 4.2) that were represented only in the parr sample, it seems likely that the redd(s) they created were overlooked. All 23 parr assigned to these two females were found within 870m of the traps, yet the most downstream redd identified was located c.1050m from the traps.

Three female spawners identified from the egg sampling in 1995 were not represented in the parr sampled in 1997. Surprisingly, one of these females (#88) had been assigned to the second highest number of eggs (Figure 4.3). A number of factors may account for this observation:

- the number of eggs assigned to a female was not representative of the actual number (perhaps due to sampling only a portion of redds),
- the mortality of offspring from this female was particularly high,
- the sampling of parr was not adequately extensive to have obtained samples of offspring from this female.

Four females were noted as ‘part-spent’ (#20, #106 and #107) or ‘fully-spent’ (#93) on entry to the traps. Despite this, three still produced offspring that were sampled,
showing at least some ‘part-spent’ fish and the female noted as ‘fully-spent’ spawned successfully when placed above the traps.

A total of 43 anadromous males were potential spawners in 1995 i.e. were put above the traps. The general consensus between females represented in egg and parr samples was also borne out with anadromous male spawners, the majority of males were either represented in both egg and parr samples (22 males) or in neither sample (15 males)(Figure 4.4, Figure 4.5 and Figure 4.6).

**DETECTED MALE SPAWNERS**

![Venn diagram](image)

Figure 4.4 Venn diagram showing the number of male spawners represented in egg and 1+ parr samples from natural spawning in 1995.
Output (number of eggs sampled) from anadromous males that were placed above the traps to spawn in 1995

Output (number of parr sampled) from anadromous males that were placed above the traps to spawn in 1995

Figure 4.5 The output of each anadromous male that was placed above the traps in 1995. Offspring were sampled as eggs (top) or 1+ parr (bottom). Date of placement of fish above traps is indicated by arrows (fish to the right side of arrows were placed above the traps on or after this date, but prior to the next date).
MALE SPAWNERS THAT WERE NOT DETECTED

![Venn diagram showing the number of male spawners not represented in egg and 1+ parr samples from natural spawning in 1995.](image)

That the majority of male and female spawners were represented in both parr and egg samples, or represented in neither sample, suggests that sampling was adequate to detect all but the minority of spawners that gave rise to offspring. This is supported by the sea returns (of 2000) that were offspring of spawners in 1995. All were offspring of male or female spawners in 1995 that had already been represented in egg or parr samples, or in both.

A relatively high proportion of potential spawners (29% and 35% for females and males, respectively) were not detected in either offspring sample. These individuals may not have spawned, or may have produced relatively few offspring, which were not represented with sampling. Two of the females not represented were later found as kelts, indicating they had indeed spawned above the traps. Their offspring may either not have survived, or not have been sampled. It is probably more likely that the parr sample would have given a better measure of the reproductive success of individual spawners as it did not rely on the identification of redds or sampling from a particular area within the redd.

In an attempt to identify potential factors determining the reproductive success of spawners, the length, age and time of return to the traps were compared between identified spawners and apparent non-spawners.
There was no obvious link between run time and successful spawning in females, since successful spawners were found early, mid and late in the spawning season (Figure 4.2). There was also no apparent difference in age (93% of spawners and 81% of apparent non-spawners were multi sea-winter fish, aged either 2.2 or 3.2 from scale reading). No significant difference in length (Mann Whitney Z=-0.90 P=0.37, Figure 4.7) between the groups was found.

There was also no apparent difference between successful and apparently unsuccessful males in terms of age (both consisted of returns aged 2.1, 2.2, 3.1 and 3.2) and there was no significant difference in length (T-test, t=0.028 df=41 P=0.98) (Figure 4.8). The final seven males to be placed above the traps (on or after the 12th November; Figure 4.5) were poorly represented in both samples; only two were positively identified as successful sires.

Figure 4.7 Length of anadromous females from the 1995 spawning run: detected vs. undetected spawners.
Despite the general concurrence in the representation of particular male and female spawners in egg and parr samples, they differed in the parental-pairings detected. Only 34 pairings were common to both egg and parr samples, with a total of 73 pairs being found in only one of the two samples (46 in the egg and 27 in the parr sample). The identification of more families from the egg sampling analysis may be directly related to sample sizes (n >1000 for eggs vs. 269 for parr sample). Clearly the combined datasets give a more complete picture in this regard.

Figure 4.9 Venn diagram showing the numbers of different parental-pairs that were represented in either parr or egg samples, or both.
Taggart et al. (2001) found evidence of size assortative mating in two of the three spawning years studied but not for the 1995 spawning year. It was therefore of interest to determine if extra data on pair-matings from the parr sampling altered this conclusion. No significant correlation (Pearson’s $r=0.036$, $P=0.71$) was found between the lengths of females and males that paired in 1995 (a total of 107 pairings were identified from samples of eggs taken in 1995 and from parr sampled in 1997) (Figure 4.10).

![Figure 4.10 Scatterplot showing the length of anadromous males and females, for all 107 pairs identified for spawning in 1995 (linear trendline included).](image)

There was a significant positive correlation between the extent of multiple mating (both dams and sires) and the number of offspring (parr) sampled (dams, Spearman’s $r=0.54$, $P=0.005$, Figure 4.11; sires, Spearman’s $r=0.55$, $P<0.001$, Figure 4.12).
Figure 4.11 Scatterplot (with linear trendline) of the number of offspring (1+ parr) sampled from each anadromous male spawner against the number of different female mates detected. The number of different females each anadromous male paired with was estimated from the parr sample of 1997.

Figure 4.12 Scatterplot (with linear trendline) of the number of offspring (1+parr) sampled from each anadromous female spawner against the number of different male mates detected. The number of different males each anadromous female paired with was obtained from the parr sample of 1997.
4.3.2 Representation of spawners in quasi-natural and natural spawning

Central to evaluating the effectiveness of quasi-natural spawning was demonstrating how the representation of spawners among resultant offspring compared to that of anadromous returns that spawned naturally. The representation of anadromous males and females that spawned naturally and quasi-naturally are depicted numerically in Table 4.2 and graphically in Figure 4.13 to Figure 4.16.

Table 4.2 Representation of anadromous male and female spawners that spawned quasi-naturally and naturally. Dates of samples refer to the year that anadromous fish returned to Girnock Burn. Offspring of returns of 1995 were sampled from eggs taken in 1995/1996, and parr sampled in 1997 (shown separately and combined). Offspring from returns of 2000 and 2001 were sampled as smolts in 2004. The representation of females, excluding females that were noted as ‘part-spent’ or ‘nearly kelt’, are shown in brackets.

<table>
<thead>
<tr>
<th></th>
<th>Quasi-Natural Spawning</th>
<th>Natural Spawning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% represented</td>
<td>Number represented/total</td>
</tr>
<tr>
<td>FEMALES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995 (parr)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1995 (egg samples)¹</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1995 (combined parr and egg)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2000</td>
<td>100 (100)</td>
<td>14/14 (11/11)</td>
</tr>
<tr>
<td>2001</td>
<td>100 (100)</td>
<td>12/12 (12/12)</td>
</tr>
<tr>
<td>MALES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995 (parr)</td>
<td>58</td>
<td>25/43</td>
</tr>
<tr>
<td>1995 (egg samples)</td>
<td>58</td>
<td>25/43</td>
</tr>
<tr>
<td>1995 (combined parr and egg samples)</td>
<td>65</td>
<td>28/43</td>
</tr>
<tr>
<td>2000A (first set quasi-natural spawning)</td>
<td>100</td>
<td>14/14</td>
</tr>
<tr>
<td>2000B (second set quasi-natural spawning)</td>
<td>83</td>
<td>15/18</td>
</tr>
<tr>
<td>2001</td>
<td>100</td>
<td>7/7</td>
</tr>
</tbody>
</table>

¹Results from Taggart et al. (2001).
Figure 4.13 The proportion of detected (shaded dark) and undetected female spawners. The proportion of females detected in natural spawning was taken from the parr sampled in 1997. Data for 2000 and 2001 are derived from smolts sampled in 2004. n=total number of females that were potential spawners (excluding those noted as ‘part-spent’ or ‘nearly kelt’).
Investigating natural and quasi-natural spawning

Chapter 4

Figure 4.14 The relative output of detected female spawners. Each shade represents a different female. The output of females in natural spawning in 1995 was taken from the parr output. The output of females that spawned in 2000 and 2001 was derived from the smolts sampled in 2004. n = number of detected female spawners (excluding those noted as ‘part-spent’ or ‘nearly kelt’).
Figure 4.15 The proportion of detected (shaded dark) and undetected male spawners. The proportion of males detected in natural spawning was taken from the parr sampled in 1997. Data for 2000 and 2001 are derived from smolts sampled in 2004. n=total number of anadromous males that were potential spawners.
Figure 4.16 The relative output of detected male spawners. Each shade represents a different male. The output of males in natural spawning in 1995 was taken from the parr output. The output of females that spawned in 2000 and 2001 was derived from the smolts sampled in 2004. n=total number of anadromous males detected.

Table 4.2 and Figure 4.13 show that the representation of spawners in quasi-natural spawning (up to the smolt stage) was more complete than in natural spawning, for both female and male spawners. That nearly all spawners used in quasi-natural
spawning were represented shows that the supportive breeding scheme was successful. It also reveals that the ‘insurance male’ used when setting up the crosses did not monopolise fertilisations. Not all anadromous males used in quasi-natural spawning in 2000 were represented; three males used in set B (Table 4.1) were not assigned to any smolts sampled. A number of factors acting either singly or in combination may account for their absence, e.g. issues of reduced fertility or sterility, differential survival / behaviour leading to under-representation in the spring smolt run, or sampling error due to the relatively large number of males (18) and few females (five, three of which were part-spent) used in this set. A smaller proportion of anadromous males that were placed above the traps to spawn naturally were represented in 2001 than in 1995 or 2000. This is likely due to the large number of males (22) in comparison to females (eight) that were potential (natural) spawners in this year.

The representation of successful female and male spawners in 1995 was apparently much more uniform compared to natural spawning in 2000 and 2001 (Figure 4.14 and Figure 4.16). A number of factors may account for this, including differences in sampling regimes (differential survival / behaviour may have led to under-representation in the spring smolt run), differences in sex ratios and absolute number of anadromous fish used in natural spawning in these years, or the occurrence of quasi-natural spawning alongside natural spawning in 2000 and 2001 (whereby naturally spawned eggs may have been disturbed by planting out eggs from quasi-natural spawning).

In order to statistically compare relative success of spawners, the B index and 95% confidence intervals of the index were estimated for the different groups of spawners (Figure 4.17 for females and Figure 4.18 for males). The distribution of
offspring among parents of each group was significantly different from random, except for a single group: offspring of females that underwent quasi-natural spawning in 2000. Groups that did not significantly differ from a geometric distribution are highlighted in Figure 4.17 and Figure 4.18. The 95% confidence intervals of these groups should be reliable, whereas they may be too broad for other groups (P. Nonacs, pers. comm.).
Figure 4.17 Reproductive skew of females, measured by Nonac's B index. Year refers to year of spawning. Eggs and parr are both samples of offspring of natural spawning. Upper and lower 95% confidence intervals are shown.

QN- quasi-natural spawning
Nat- natural spawning
*- distribution did not significantly differ from a geometric distribution
B- B index for the group
Equal B- value of B that corresponds to an equal distribution of offspring between spawners
Investigating natural and quasi-natural spawning

Figure 4.18 Reproductive skew of males, measured by Nonac's B index. Year refers to year of spawning. Eggs and parr are both samples of offspring of natural spawning. Upper and lower 95% confidence intervals are shown.
- QN- quasi-natural spawning
- Nat- natural spawning
- *- distribution did not significantly differ from a geometric distribution
- B- B index for the group
- Equal B- value of B that corresponds to an equal distribution of offspring between spawners

According to the B index, the reproductive success of both female and male spawners was significantly more skewed in natural spawning in 2000 and 2001, than in quasi-natural spawning for these years. However, it also showed that natural spawning that took place in 1995 (in the absence of quasi-natural spawning) resulted in a reproductive skew according to the B index, which was similar to that observed in quasi-natural spawning. This is surprising considering the difference in
representation of spawners between the types of spawning (i.e. a relatively high proportion of spawners were unsuccessful in natural spawning, while nearly all spawners employed in quasi-natural spawning were represented). Although the B index values are similar for these groups, the difference in absolute success (Figure 4.13 and Figure 4.15) may still be biologically important. B index values obtained for natural spawning that took place in 2000 and 2001 suggested more skewed distributions than values gained for the reproductive output of natural spawning in 1995 (when no quasi-natural spawning took place). This concurs with the observations above from Figure 4.14 and Figure 4.16. A number of reasons may be the cause, as suggested above (e.g. differences in sampling regimes (differential survival / behaviour may have led to under-representation in the spring smolt run), differences in sex ratios and absolute number of anadromous fish used in natural spawning in these years, or the occurrence of quasi-natural spawning alongside natural spawning in 2000 and 2001).

An additional element of spawning is the number of different individuals each spawner mates with. This aspect is not measured by the B index, but could be evolutionarily important as the greater the number of pairs that are formed, the greater the number of different genetic combinations, which increases the chances of advantageous combinations arising. The number of mates may also be linked to the distribution of kin throughout the burn. Increasing the number of pairs that are formed may allow an individual’s offspring to be more widely spread throughout the burn, which could be seen as a ‘bet-hedging’ strategy in the light of the variable environment. Therefore the number of males paired per female and the number of females paired per male was compared for quasi-natural and natural spawning (Table 4.3).
Table 4.3 The average number of anadromous males paired per female and the average number of females paired per male. Year relates to year of spawning. Average was taken from males and females that were represented in samples i.e., males and females had been paired with at least one other male or female. (Maximum number of females or males paired is shown in brackets). Females that were ‘part-spent’ or ‘nearly kelt’ were excluded from the calculation of the average number of males paired per female.

<table>
<thead>
<tr>
<th>Year</th>
<th>Quasi-natural spawning</th>
<th>Natural spawning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average number males paired per female</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (parr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (eggs) ¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (combined)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000 &amp; 2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average number females paired per male</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (parr)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (eggs) ¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995 (combined)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000 &amp; 2001</td>
</tr>
</tbody>
</table>

¹ Results from a previous study by Taggart et al. (2001).

The number of males paired per female was higher for quasi-natural spawning than in natural spawning in all cases (Table 4.3). This may be expected as in quasi-natural spawning each female’s eggs were fertilised by a minimum number of seven anadromous males. Of interest is the similarity between the number of females paired per male in quasi-natural spawning (in 2000 and 2001 combined) and natural spawning that took place in 1995. This suggests that setting up artificial crosses did not increase the number of females paired per male compared to that found in natural spawning in 1995, at least as measured by representation of spawners in offspring samples. A possible explanation is the low numbers of eggs each male
fertilised in set B of quasi-natural spawning in 2000, and the low numbers of females involved in this particular cross.

4.3.3 Influence on the effective number of breeders per year

In order to further investigate the effects of quasi-natural spawning in comparison to natural spawning, the effective number of breeders per year was estimated for the various sample sets. The mean and variance in reproductive success of different sets of spawners are shown below (Table 4.4). It is these values that influence $N_{b(demo)}$ (as demonstrated by Equation 4.1 to Equation 4.3). The standardized variance ($\text{variance}/(\text{mean})^2$) is also shown. This allows a comparison to be made between the groups, since it is largely independent of the mean (Nunney 1996). Nunney (1996) used this value to describe data on the variation in female fecundity in different species of mammals, insects and birds.

Table 4.4 shows that the standardized variance was lower for quasi-natural spawning than natural spawning, as would be expected. For quasi-natural spawning the standardized variance might be expected to be higher in females than males because of the different numbers of eggs produced by each female, whereas crosses were set up such that each male should have fertilised equal numbers of eggs. This was not the case in 2000, where anadromous males in set A fertilised c.1730 eggs each and anadromous males in set B fertilised c. 750 eggs each. This probably accounts for the comparatively high standardised variance in male output compared to females in 2000.

<table>
<thead>
<tr>
<th>Spawning year</th>
<th>Gender of spawner</th>
<th>Mean output</th>
<th>Variance</th>
<th>Standardized variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995 eggs</td>
<td>f</td>
<td>16.29</td>
<td>377.16</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>12.56</td>
<td>330.40</td>
<td>2.09</td>
</tr>
<tr>
<td>1995 parr</td>
<td>f</td>
<td>4.11</td>
<td>24.40</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>3.51</td>
<td>17.30</td>
<td>1.40</td>
</tr>
<tr>
<td>2000QN</td>
<td>f</td>
<td>11.73</td>
<td>21.62</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>3.19</td>
<td>4.93</td>
<td>0.49</td>
</tr>
<tr>
<td>2000NAT</td>
<td>f</td>
<td>5.83</td>
<td>53.37</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>1.36</td>
<td>17.10</td>
<td>9.20</td>
</tr>
<tr>
<td>2001QN</td>
<td>f</td>
<td>10.67</td>
<td>32.42</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>10.43</td>
<td>34.62</td>
<td>0.32</td>
</tr>
<tr>
<td>2001NAT</td>
<td>f</td>
<td>10.00</td>
<td>184.86</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>4.25</td>
<td>38.21</td>
<td>2.12</td>
</tr>
</tbody>
</table>

f- female  
m- male  
NAT- natural spawning  
QN- quasi-natural spawning  
1 output is number of smolts for 2000 and 2001, eggs or parr for 1995  
2 Standardized variance= (variance/(mean)^2)

The standardised variance was particularly high for natural spawning in anadromous males in 2000, which could be explained by the highly skewed sex ratio in this year (22 anadromous males and eight females were placed above the traps to spawn naturally in this year). In comparison, in 2001 eight anadromous males and nine females were placed above the trap to spawn naturally.

The estimates are likely to be more accurate for females because of the greater sample size of offspring, fewer offspring were identified from anadromous males due to the input of mature parr.

Table 4.5 shows the effective number of breeders per year ($N_{b(demo)}$), taking into account variable reproductive success and unequal sex ratio. Comparing the quasi-natural spawning that took place in 2000 and 2001, Table 4.5 shows that the effective number of breeders was a lower percentage of the initial number of
breeders for quasi-natural spawning that took place in 2000. This was probably due to the set-up of two separate sets of crosses in this year, with 18 anadromous males and 2 females (plus three females which had few eggs left), the other with 14 anadromous males and nine females. This resulted in a greater standardised variance in reproductive output for anadromous males in 2000 (compared to anadromous males in 2001) in addition to a more unequal sex ratio (this was nearly equal in 2001, when 12 females and 11 males were employed in crosses).

For natural spawning, the effect of contributions of differing numbers of mature parr was investigated. When the number of mature parr contributing was varied between 800 and 4000, there was little effect on the effective number of breeders. However, the effective number of breeders more than doubled when 800 rather than 10 mature parr contributed.
Table 4.5 Number of breeders and effective number of breeders (taking into account both unequal sex ratio and variable reproductive success) for quasi-natural and natural spawning in different years. Estimates of the number of mature male parr involved in natural spawning were varied (10, 800, 1500 or 4000).

<table>
<thead>
<tr>
<th>Type of spawning</th>
<th>Year of spawning</th>
<th>Mature parr used (U) or estimated</th>
<th>N (P)</th>
<th>N_{b(demo)}</th>
<th>N_{b(demo)}×100/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>QN</td>
<td>2000</td>
<td>10&lt;sup&gt;U&lt;/sup&gt;</td>
<td>53</td>
<td>28</td>
<td>53</td>
</tr>
<tr>
<td>QN</td>
<td>2001</td>
<td>4&lt;sup&gt;U&lt;/sup&gt;</td>
<td>23</td>
<td>19</td>
<td>83</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (E)</td>
<td>10</td>
<td>119</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (P)</td>
<td>10</td>
<td>119</td>
<td>54</td>
<td>45</td>
</tr>
<tr>
<td>Natural</td>
<td>2000</td>
<td>10</td>
<td>38</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>Natural</td>
<td>2001</td>
<td>10</td>
<td>26</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (E)</td>
<td>800</td>
<td>909</td>
<td>102</td>
<td>11</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (P)</td>
<td>800</td>
<td>909</td>
<td>111</td>
<td>12</td>
</tr>
<tr>
<td>Natural</td>
<td>2000</td>
<td>800</td>
<td>828</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Natural</td>
<td>2001</td>
<td>800</td>
<td>816</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (E)</td>
<td>1500</td>
<td>1609</td>
<td>106</td>
<td>7</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (P)</td>
<td>1500</td>
<td>1609</td>
<td>115</td>
<td>7</td>
</tr>
<tr>
<td>Natural</td>
<td>2000</td>
<td>1500</td>
<td>1528</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Natural</td>
<td>2001</td>
<td>1500</td>
<td>1516</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (E)</td>
<td>4000</td>
<td>4109</td>
<td>109</td>
<td>3</td>
</tr>
<tr>
<td>Natural</td>
<td>1995 (P)</td>
<td>4000</td>
<td>4109</td>
<td>118</td>
<td>3</td>
</tr>
<tr>
<td>Natural</td>
<td>2000</td>
<td>4000</td>
<td>4028</td>
<td>10</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Natural</td>
<td>2001</td>
<td>4000</td>
<td>4016</td>
<td>11</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

E-offspring sampled as eggs
P-offspring sampled as parr
QN- quasi-natural
N_{b(demo)}×100/N - the effective number of breeders as a percentage of the initial number of breeders

<sup>1</sup> Note this is the effective number of breeders per year, as opposed to the effective population size per generation, which is estimated in Chapter 8.

Estimates of the effective number of breeders for 1995 based on parr or egg data were similar. Comparing quasi-natural and natural spawning shows that as a
percentage of the original number of breeders (N), \( N_{b(demo)} \) was higher for quasi-natural spawning.

A comparison of the two methods of spawning was also approached in a different way, by estimating \( N_{b(demo)} \) for quasi-natural spawning in 2001 and then estimating \( N_{b(demo)} \) for the same numbers of anadromous fish, under the hypothetical situation of natural spawning. Thus, in order to mimic the effect of natural spawning on fish that actually underwent quasi-natural spawning, the mean and variance of reproductive success of natural spawning that took place in 1995 were applied to the set-up of quasi-natural spawning in 2001 (Table 4.6).

Table 4.6 The effective number of breeders in hypothetical situations and an actual set up. The number of breeders used in quasi natural spawning in 2001 were re-analysed using parameters (mean and variance of reproductive success) obtained from natural spawning that took place in 1995. The effective number of breeders estimated takes into account both variable reproductive success and unequal sex ratio (\( N_{b(demo)} \)).

<table>
<thead>
<tr>
<th>Hypothetical/actual situation</th>
<th>Origin of mean and variance of reproductive success</th>
<th>Number of anadromous males</th>
<th>Number of mature parr</th>
<th>Number of females</th>
<th>N</th>
<th>( N_{b(demo)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>QN</td>
<td>7</td>
<td>4</td>
<td>12</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Hypothetical Natural</td>
<td>7</td>
<td>10</td>
<td>12</td>
<td>29</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Hypothetical Natural</td>
<td>7</td>
<td>800</td>
<td>12</td>
<td>819</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Hypothetical Natural</td>
<td>7</td>
<td>1500</td>
<td>12</td>
<td>1519</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Hypothetical Natural</td>
<td>7</td>
<td>4000</td>
<td>12</td>
<td>4019</td>
<td>21</td>
<td>21</td>
</tr>
</tbody>
</table>

QN- quasi-natural

The number of effective breeders was lower for the hypothetical natural spawning. This was true even when 4000 mature parr were assumed to have contributed.

One of the drawbacks with the estimates of \( N_{b(demo)} \) was that the variance in reproductive success did not take into account the input from mature parr. It is likely that mature parr would individually have sired fewer offspring than anadromous males, which would have resulted in an increased variance in the output of males.
Investigating natural and quasi-natural spawning

To model the effect of increased variance in reproductive output the variance was doubled and $N_{b(demo)}$ re-estimated for natural spawning (Table 4.7).

Table 4.7 Number of effective breeders ($N_{b(demo)}$), when modelling the effect of increased variance in reproductive output, by doubling the variance. Estimated for natural spawning only.

<table>
<thead>
<tr>
<th>Year of spawning</th>
<th>Estimated number of mature parr</th>
<th>N</th>
<th>$N_{b(demo)}$</th>
<th>$N_{b(demo)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$1 \times \sigma^2_k$</td>
<td>$2 \times \sigma^2_k$</td>
</tr>
<tr>
<td>1995(E)</td>
<td>10</td>
<td>119</td>
<td>43</td>
<td>26</td>
</tr>
<tr>
<td>1995 (P)</td>
<td>119</td>
<td></td>
<td>54</td>
<td>33</td>
</tr>
<tr>
<td>1995(E)</td>
<td>800</td>
<td>909</td>
<td>102</td>
<td>63</td>
</tr>
<tr>
<td>1995 (P)</td>
<td>909</td>
<td></td>
<td>111</td>
<td>67</td>
</tr>
<tr>
<td>1995(E)</td>
<td>1500</td>
<td>1609</td>
<td>106</td>
<td>66</td>
</tr>
<tr>
<td>1995 (P)</td>
<td>1609</td>
<td></td>
<td>115</td>
<td>69</td>
</tr>
<tr>
<td>1995(E)</td>
<td>4000</td>
<td>4109</td>
<td>109</td>
<td>68</td>
</tr>
<tr>
<td>1995 (P)</td>
<td>4109</td>
<td></td>
<td>118</td>
<td>71</td>
</tr>
</tbody>
</table>

E-offspring sampled as eggs
P-offspring sampled as parr

$1 \times \sigma^2_k$ - Original estimate of variance in reproductive success for anadromous males and females used in calculating $N_{b(demo)}$.

$2 \times \sigma^2_k$ - The original estimate of variance in reproductive success for anadromous males and females was doubled in calculating $N_{b(demo)}$.

Table 4.7 shows that doubling the variance in reproductive output results in a decrease of $N_{b(demo)}$ of c. 40%.

4.3.4 Reproductive success in terms of smolt numbers

For spawning in 2000 and 2001 it was possible to compare the absolute output of quasi-natural and natural spawning. Smolt output per female (Table 4.8) was significantly higher (Mann-Whitney Z= -2.150 P= 0.032) for quasi-natural spawning c.f. natural spawning (median 10 vs. 4). However, females used in quasi-natural spawning were longer (Mann-Whitney test for fork length, Z=-2.845 P=0.004).
Table 4.8 Numbers of smolts sampled that were produced by each female in quasi-natural and natural spawning (2000 and 2001 combined, in order to maximise sample size). Three females were discounted from natural spawning as they had been noted as part-spent, and three from quasi-natural spawning for the same reason.

<table>
<thead>
<tr>
<th>Quasi-natural spawning</th>
<th>Natural spawning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>41</td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

Thus it was considered more prudent to compare smolt output against potential female fecundity. Based on data obtained from females used in quasi-natural spawning, two linear regressions were used to predict the numbers of eggs from grilse and multi-sea winter females (separately, as the relationship was non-linear when values were combined). Plots of the data (Figure 4.19 and Figure 4.20) and equations (Equation 4.4 and Equation 4.5) used to predict the number of eggs from females that spawned naturally, given their length, are depicted below.
Investigating natural and quasi-natural spawning

Chapter 4

Figure 4.19 Length of MSW females used in quasi-natural spawning and the number of eggs obtained from each.

Figure 4.20 Length of female grilse used in quasi-natural spawning and the number of eggs obtained from each.

\[ E = 19.507x - 8596.287 \]

Equation 4.4 Equation used to predict egg number (\( E \)) of MSW fish that spawned naturally, given their length in cm (\( x \)).

\[ E = 15.008x - 5732.297 \]

Equation 4.5 Equation used to predict egg number (\( E \)) of grilse that spawned naturally, given their length in cm (\( x \)).
Therefore the total number of eggs that were spawned naturally in 2000 and 2001 were estimated (Table 4.9). Taking into account the number of eggs used in quasi-natural and natural spawning, there was no difference in the output of smolts, since 30% of eggs were spawned naturally, and 30% of the smolts sampled also originated from natural spawning (Table 4.9). Thus the minimal impact approach, of rearing only up to the egg stage in an artificial environment, appears not to have given the offspring of quasi-natural spawning a great advantage over those from natural spawning.

### Table 4.9 Comparison of the output (in numbers of smolts) from quasi-natural and natural spawning, taking into account the number of eggs used. The number of eggs in natural spawning was estimated from their length. (Numbers in brackets relate to the number of anadromous females that were discounted as they had been noted as part-spent.)

<table>
<thead>
<tr>
<th>Year of spawning</th>
<th>Number of anadromous females used in spawning</th>
<th>Approximate initial number of eggs</th>
<th>Proportion of initial eggs (NAT/total)</th>
<th>Number of smolts represented</th>
<th>Proportion of smolts represented (NAT/total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000 QN</td>
<td>14</td>
<td>52051</td>
<td>0.28</td>
<td>139</td>
<td>0.20</td>
</tr>
<tr>
<td>2000 NAT</td>
<td>6 (2)</td>
<td>19943</td>
<td></td>
<td>35</td>
<td>0.20</td>
</tr>
<tr>
<td>2001 QN</td>
<td>12</td>
<td>56296</td>
<td>0.33</td>
<td>128</td>
<td>0.38</td>
</tr>
<tr>
<td>2001 NAT</td>
<td>8 (1)</td>
<td>27359</td>
<td></td>
<td>80</td>
<td>0.38</td>
</tr>
<tr>
<td>2000&amp;2001 QN</td>
<td>26</td>
<td>108401</td>
<td>0.30</td>
<td>267</td>
<td>0.30</td>
</tr>
<tr>
<td>2000&amp;2001 NAT</td>
<td>14 (3)</td>
<td>47302</td>
<td></td>
<td>115</td>
<td>0.30</td>
</tr>
</tbody>
</table>

**NAT**-natural spawning.

**QN**-quasi-natural spawning.
4.3.5 Marine survival of offspring of quasi-natural and natural spawning

The last spawning run sampled in this project was in 2004. It was possible that some of these fish represented the 2000 year class (i.e. fish that were aged 2.1). Parentage assignment identified 13 individuals that returned to Girnock Burn in 2004 as offspring of anadromous adults that spawned in 2000 (Table 4.10). These were all classified by age as grilse, as would be expected. Two returns from natural spawning were full sibs, while no returns that originated from quasi-natural spawning were identified as full sibs.

Table 4.10 The origin of adult returns that could be assigned to spawners in 2000, and the number of female and male spawners represented in the offspring.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Number of offspring assigned</th>
<th>Number of female spawners represented</th>
<th>Number of male spawners represented</th>
</tr>
</thead>
<tbody>
<tr>
<td>QN</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Nat</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

QN-Quasi-natural spawning
Nat-Natural spawning

Although numbers were too small for meaningful statistical analysis, offspring from both quasi-natural and natural spawning were represented. Thus there was no obvious indication that survival and homing of fish was influenced to any great extent by the type of spawning that produced them.
4.4 Discussion

4.4.1 Mating system in the wild (when no quasi-natural spawning took place)

There was a general concurrence between egg and parr samples in the particular males and females that were found to be detected and undetected spawners, largely confirming the findings of Taggart et al. (2001). However, it was likely that at least one redd (including two nests) had not been identified in the study carried out by Taggart et al. (2001). It cannot be determined whether or not the ‘unsuccessful’ spawners were in fact completely unsuccessful or had numbers of offspring too low to be detected. The finding of two females as kelts, that were apparently unsuccessful spawners suggests eggs were deposited, but whether any of these survived is not known. Three females that, according to field notes, were either part or fully spent were detected in the parr sample. This suggests that females with few eggs are still able to reproduce successfully, and the practise of putting them above the traps was valid. However, it may be that the fish that was noted as fully spent was incorrectly recorded or that it is difficult to determine when a female is fully spent.

The lack of representation of a notable proportion of putative male spawners ties in with observations made by Webb and Hawkins (1989) after radio tracking spawners at Girnock Burn in 1986. They found that some males showed little evidence of spawning success (defined as being observed to contribute to spawning).

In a study examining the parentage of fry from the natural spawning of anadromous male and female Atlantic salmon, Garant et al. (2001) found that 100% (41/41) of males were represented and that only 9% (3/35) of females were not represented, although the variance in success was high. However, it is very difficult to compare these findings to the present study due to the differing field conditions and methods.
employed. Sampling may have accounted for this better representation, since offspring were sampled beyond the confines of the redd (as fry), although fewer samples were taken: 650 fry in total. Another aspect that differed was the length of rivers available for spawning; in the present study 66 females were released into a stretch of approximately 8km, whereas 35 females were allowed to spawn in approximately 19km in the study by Garant *et al.* (2001). It may have been that a greater area of spawning habitat was available in the study by Garant *et al.* (2001), such that redd superimposition was less likely to have occurred, although this can only be speculated. Another relevant aspect is the timing of release. Anadromous males and females were all released into the study site at the end of August in the study conducted by Garant *et al.* (2001), although spawning takes place at a later stage, during October and November. This is in contrast to Girnock Burn, where the entrance of adult spawners to the burn is dictated largely by the water flow and occurs very near to the time of spawning. Other reasons for the lower representation of anadromous spawners could be washout of deposited eggs occurring during times of spate, groundwater effects on eggs that were deposited and predation (of the anadromous fish).

Although size assortative mating is common among other salmonid species (Foote, 1988) there was less than convincing evidence for this from data collected from Girnock Burn. Taggart *et al.* (2001) reported evidence of a positive correlation between male and female length for identified parental pairs from spawning in 1991. However this was not the case for 1992 or 1995 spawning data. The additional data provided by the current study for spawning in 1995 did not alter this finding.
The number of anadromous mates was correlated with reproductive success, for both males and females. This was also found by Garant et al. (2001). However they state that their finding could be due to not identifying all pairs formed. This is also likely to be the case in the current study, especially considering the large number of pairs (73) only represented by parr or egg samples. If not all pairs were detected it could mean that individuals that had mated with large numbers of females, but had low reproductive success may not have been detected. Thus this finding still requires more thorough sampling in order to determine the actual relationship.

A similar mating outcome to that identified here was found by Seamons et al. (2004), in steelhead trout. They found high numbers of potential spawners were not represented (43% and 23% of males and females, respectively), and a range of number of mates (0-10 for males and 0-5 for females). They suggested that the flexibility and complexity of the mating system might be advantageous in variable environments, which could well be the case in this (and other) stock(s) of Atlantic salmon. Ways in which multiple mating could be beneficial are; in producing a greater variety of genetic combinations, some of which may be useful in prevailing environmental conditions, in utilising different regions of the burn for offspring development (a ‘bet-hedging’ strategy), and avoidance of inbreeding (Avise et al. 2002). A flexible mating system may also be advantageous in allowing for variances in water flow during spawning (Taggart et al. 2001).

4.4.2 Representation of spawners in quasi-natural and natural spawning

The vast majority of spawners employed in quasi-natural spawning were represented in the smolts sampled. The only case when not all spawners were represented was when the sex ratio was heavily skewed towards males, and this may be due to chance sampling error among offspring. This is in contrast to natural
Investigating natural and quasi-natural spawning

Chapter 4

spawning, when complete representation of anadromous spawners was not found. This suggests that the method of quasi-natural spawning (involving crossing each male in a set with each female and distributing eggs from families throughout the burn) has been successful, in that the representation of anadromous spawners was more even and therefore increased the chances of their representation in the future gene pool.

Nonac’s B index (Nonacs 2000) was used as a way of statistically comparing the reproductive skew of different groups. This showed the natural spawning that occurred alongside quasi-natural spawning was significantly more skewed than quasi-natural spawning or natural spawning that occurred in 1995. Thus when managing the burn, it may be more beneficial to either have all fish spawned quasi-naturally, have greater numbers spawning naturally, or, when there is an excess of anadromous males, to place them (a long way) below the trap as they may encounter more females than are above the traps and possibly be more likely to reproduce.

Quasi-natural spawning resulted in more males being paired per female (but not more females being paired per male) than was detected to occur naturally. This ties in with observations from radio tracking made by Webb and Hawkins (1989) at Girnock Burn. They tracked anadromous males and females during the spawning season, and found that anadromous males tended to move further and more often than anadromous females, the implication being that they mated with different females in different parts of the river. Thus quasi-natural spawning may be a way of obtaining a higher number of pairings of males per female, but not necessarily females per male as this was shown already to be high in the natural spawning that occurred in 1995.
The numbers of females paired per male and vice versa were lower in this study (for both types of spawning) than those found by Garant *et al.* (2001), also in Atlantic salmon. Garant *et al.* (2001) found, on average, each male spawned with 6.4 females and each female spawned with 7.5 males. This may reflect different sampling regimes of the two studies, different mating systems in the two stocks, the shorter run time at Girnock Burn, the lower area of spawning habitat available at Girnock Burn, and possible effects of washout and groundwater on deposited eggs at Girnock Burn.

The actual reproductive success (as determined from smolt samples) of adults that spawned in 2000 and 2001 may have differed to that observed because the smolts sampled in the spring of 2004 only represented a portion of these offspring. Firstly, offspring of the same year class would have passed through the traps in the autumn prior to sampling. Secondly, in order to sample two year old smolts originating from spawners in 2000 and three year old smolts originating from spawners in 2001, samples of autumn migrants from 2002 and 2004 in addition to spring migrants from 2003 and 2005 would have been required. Therefore if a particular spawner only gave rise to offspring that undertook a particular life history tactic, such as smolting at three years, and no three year old smolts were sampled, the spawner would not have been represented. The extent to which this may have biased results is unknown.

The parr sample was not limited in this way, in that the whole cohort was available for sampling. However, the smolt sample was taken when fish were older, so had survived for longer in the freshwater environment. It was therefore potentially a slightly better indication of lifetime reproductive success of the spawners, although this must be considered in light of the probable bias in this latter sample. The egg
sample was limited in that only portions of the redd were sampled, although
sampling redds did allow for the position of spawning and therefore the ranging of
spawners to be determined. With hindsight a larger sample of 1+ parr would have
allowed for a more accurate estimation of the reproductive output of spawners in
1995 (particularly for males), and for the detection of additional mating pairs.
However, these samples had already been collected at the beginning of the current
project, and the sample size had been limited to minimise the amount of disturbance
caued.

When considering the design of studies comparing the reproductive success of
hatchery and natural-born spawners, Hinrichsen (2003) pointed out the importance
of having the appropriate baseline in order to provide a worthwhile comparison.
When this principle is applied to the present study, the 1995 dataset provided ‘an
appropriate baseline’ for natural spawning when comparing aspects of natural and
quasi-natural spawning, as no quasi-natural spawning took place in this year.
However, natural spawning that took place alongside quasi-natural spawning (in
2000 and 2001) allowed the success of the two methods to be measured under the
particular conditions that occurred in these years- for example water flow,
temperature and abundance of predators may have varied among years.

4.4.3 The effective number of breeders per year
Comparing the quasi-natural spawning that took place in 2000 and 2001, Table 4.5,
shows that the $N_{b(demo)}$ as a percentage of a percentage of the initial number of
breeders, was greater for quasi-natural spawning in 2001 than in 2000. As described
previously, this was likely due to the set-up of two separate sets of crosses in 2000,
one with 18 anadromous males and two females (plus three females which had few
eggs left), the other with 14 anadromous males and nine females. A more even sex ratio would have been preferable. However, the choice of parents was restricted by the sea-returns available on the day when crosses were conducted.

The main flaw in the estimation of $N_{b(demo)}$ was that the variance of reproductive success for males was calculated only from that of anadromous males. For natural spawning the variance in reproductive success may have increased if mature parr had been taken into account, since they are likely to individually have a lower reproductive success than anadromous males. An increase in variance in reproductive success was modelled, which resulted in a c.40% decrease in the $N_{b(demo)}$. In order to examine this further, the reproductive success of individual parr would be required in order to provide a more accurate estimate of the variance in reproductive success for males.

Overall, these results suggest that when a given number of breeders are used in quasi-natural spawning (as implemented at Girnock Burn) the resultant effective number of breeders will be greater than if they had spawned naturally. However, there may be more combinations of genetic types in the offspring of natural spawning if large numbers of mature parr contribute. It should be noted, however, that mature parr have not undergone selection in the marine environment. It may therefore be that the alleles carried by anadromous males are of more ‘value’ than those of mature parr. This is discussed further in Chapter 7.

In the long term it may be that attempting to maintain the effective number of breeders through quasi-natural spawning will have little effect on the effective population size, as it may be overridden by fluctuations in population size over time. Frankham (1995) found that temporal change in population size was the most influential factor on $N_e$, in a review of $N_e$ in 102 species, while Ardren and
Kapuscinski (2003) found that it was the second most influential factor on the effective population size of steelhead trout (the most influential being variable reproductive success). While management programs can help improve spawning success and freshwater survival rates, influencing marine survival would impossible or far more difficult to achieve.

4.4.4 Reproductive success in terms of numbers of smolts

When the number of eggs used in natural and quasi-natural spawning was taken into account there was no significant difference in the output of smolts from the two schemes. In some ways this was surprising as the survival of eggs in the burn over winter (for natural spawning) was likely to have been lower than that for quasi-natural spawning. In addition, eggs from quasi-natural spawning were planted out in areas that avoided regions subject to high levels of groundwater, which has been linked to impaired embryo development (Youngson et al. 2004). However, according to Cuenco et al. (1993) planting out offspring as eggs is not likely to result in an increase in survival, as the highest level of mortality occurs after emergence from the gravel. That there was no significant difference was beneficial as it meant that offspring from natural spawning were not at a (measurable) disadvantage in terms of survival to smolt stage. A well documented problem with artificial spawning programs is the Ryman-Laikre effect, whereby a small portion of the total broodstock are employed in the supportive breeding program, which are then over-represented in the offspring. The equality in survival and the set-up of crosses in quasi-natural spawning at Girnock Burn, that results in a representation of anadromous spawners that is no more skewed than natural spawning, suggests this is not occurring.
To a limited extent, quasi-natural spawning may provide a slight increase in the output of the burn. Quasi-natural spawning allows the use of regions of the burn suitable for egg deposition that would otherwise be either inaccessible to spawners due to low water levels, or do not tend to be used.

The distribution of kin throughout the burn was likely to differ for quasi-natural and natural spawning. Eggs from families were mixed prior to planting out in quasi-natural spawning, whereas females tend to nest in a few places in the burn during natural spawning. The roughly equal survival of offspring of natural and quasi-natural spawning shows that beneficial kin interactions, if they occur, are either still able to take place, or do not influence greatly the survival of offspring up to the smolt stage.

4.4.5 Marine survival of offspring of quasi-natural and natural spawning

No apparent differences in marine survival were observed between offspring of quasi-natural and natural spawning. This suggests that the survival and homing of fish that were produced through quasi-natural spawning has not been greatly impaired. This may be due in part to the minimal amount of time the offspring spent under artificial conditions (being planted out as eggs), which still allowed for imprinting. It also meant that offspring were present in freshwater at times that might be important if it is the salmon themselves that provide the chemical cues that are the basis of imprinting. This is in contrast to supportive breeding schemes where offspring are planted out as smolts, as a lack of imprinting to the natal stream can lead to higher straying rates in hatchery fish (Myers et al. 2004).

As such a low proportion of smolts actually return to Girnock Burn as adults to spawn, the value of attempting to decrease the variance in output of breeders may be questionable. However, for the small numbers of returns (13) in 2004 that could be
assigned to spawners in 2000, only two females were represented in offspring of natural spawning (two returns were also full sibs), whereas six different females were represented in offspring of quasi-natural spawning (none were identified as full sibs). This suggests that a lower variance in reproductive success of spawners to the smolt stage may also translate to later life stages (namely anadromous adults). However, it must be noted that the natural spawning that took place in 2000 initially had a limited number of females, so is not perhaps a comparable baseline.

Whether or not the initial representation of spawners is carried through from eggs to smolts to sea returns depends on how survival in the freshwater and marine environment is determined. If groups of migratory fish travel together, it may just be groups that follow a beneficial route that survive. Alternatively it may be those with favourable genetic make-ups that survive, in which case individuals that are more closely related will tend to survive anyway.

### 4.4.6 Future directions

This chapter describes the first attempt to quantify the genetic consequences of implementing quasi-natural spawning in Girnock Burn. Initially, this was implemented in relative haste and was not conducted with subsequent detailed genetic analysis foremost in mind. The study has been very enlightening, particularly in contrasting the genetic effects of quasi-natural and natural spawning. Nevertheless, in conducting the analyses, a number of weaknesses (particularly in sampling methodologies and design) became apparent, which should be addressed in future monitoring of quasi-natural spawning:

- it would be advantageous to sample the mature parr used in quasi-natural spawning. This would enable their individual reproductive success to be monitored, in addition to the anadromous males employed.
• details of the 'insurance' male used during crosses should also be retained in order that sterility / low fertility issues can be detected and quantified.

• it would also be beneficial to sample much more widely from each particular year class to provide a potentially much less biased sample. This would include sampling from the autumn migration in addition to the spring emigration (only the spring emigrants were sampled in this study) and sampling over multiple years.

• Larger sample sizes would provide a more accurate estimate of the reproductive success of individual anadromous spawners, particularly males. However, typing more samples would involve more time and expense. In addition, it would complicate routine sampling already in place, whereby one in every five smolts are examined in more detail (length is measured and scale samples are taken). In order for these data to tie in with tissue samples it is necessary to place the fin clips into individual wells filled with ethanol. This is a tedious task when in the field, in often inclement or severe weather conditions. An alternative method would be to collect all adipose fin clips (all Atlantic salmon captured in the downstream trap have the adipose fin removed) and place them in a single container of ethanol. This would be less time consuming but other data on the fish sampled would be lost.

All fish that returned to the burn in 2005 went through quasi-natural spawning. This was the first year that there was the capacity to do so and may be continued in future years. Positive aspects to this approach include:

• the representation of all anadromous returns should be more even. This should maximise $N_e$ for the population (provided adequate numbers of mature parr are involved in the crosses).
Investigating natural and quasi-natural spawning

- all potential parents can be sampled / genotyped. Without the input from anonymous parr, parentage assignment should be much simpler, potentially requiring fewer loci to allocate parents with high confidence.

Negative aspects of the scheme include;

- The risk of losing all eggs if they become infected or some unforeseen problem occurs with the incubator. (There was 30% mortality of the eggs from the crosses set up in 2005, far higher than in previous years. The reason for this is not known).

- Lack of selection during mating could cause fish that would otherwise not spawn to produce offspring. (However, there was no suggestion that strayers, identified using different methods, naturally had a substantially lower spawning success in Girdock Burn- see chapter 6).

- Fewer mature parr are likely to contribute to spawning, which is likely to differ from the natural situation.
5 The spatial distribution of juvenile kin and implications for sampling

5.1 Introduction

One potentially serious problem with sampling juveniles in order to determine population structure of salmonids was noted by Allendorf and Phelps (1981). They described, through simulation studies, how ‘it is dangerous to draw conclusions’ about the reproductive isolation of salmonid stocks, based on allele frequency data derived from juvenile samples, as such samples may originate from a small number of breeders. Despite this warning, juvenile samples have still been used in population genetic analyses in salmonids, probably due to the practicalities of sampling. With the advent of multiallelic molecular markers, and the ability to determine kinship of individuals from within samples in the current study, this problem has been revisited.

Empirical studies based on microsatellite and minisatellite data have found clustering of juvenile kin in at least some situations. Hansen et al. (1997) sampled brown trout aged 0+, 1 or 2 years, from two rivers in Denmark, the Tjaerbaek and Skibelund. In both rivers samples were obtained from a 200m stretch, although the majority of fish and all 0+ fish were obtained from a 20m stretch in the Skibelund. The authors found significant differences in mtDNA haplotypes between age classes within both rivers, and suggested that sampling of families may have contributed. The difference was particularly pronounced for the Skibelund samples, and further analysis with six microsatellite loci showed that the 0+ sample was likely to have comprised only three full sib groups. Accordingly Hansen et al. (1997) recommended the sampling of adult fish, or if utilizing juveniles aged 0+, sampling from river stretches greater than 100m in length and not to include large numbers of individuals caught at the same site. Mjølnerød et al. (1999) studied the distribution
of related Atlantic salmon juveniles along a 300m stretch of the River Jiesjokka in Norway. Using multilocus minisatellites they estimated the relatedness (based on band-sharing) of pairs and their geographic separation. The authors found a significant (but not strong) negative association between geographic separation and relatedness in pairs born in the same year. Mjølnerød et al. (1999) therefore recommended sampling from more than one age class when employing juveniles for population studies. A similar study comparing relatedness and geographic separation was conducted by Carlsson et al. (2004) in two streams, one in Scotland the other in Denmark. In the Findhu Glen Burn, Scotland, samples of 0+ brown trout were obtained from a 50m section and samples of fish older than 0+ from a 215m section. In the Dollerup Mollebaek, Denmark, samples of trout aged 1+ or older were obtained from a 125m section. The authors reported that samples found closer together also had a higher relatedness (based on 14 microsatellites) in the Findhu Burn, although this was not found in the Dollerup Mollebaek. On a smaller scale, Fontaine and Dodson (1999) evaluated whether Atlantic salmon juveniles occupying adjacent territories tended to be related. Fish aged 0+, 1+ and 2+ were sampled from areas of a large river (Trinite River) measuring 5 × 20m. The relatedness between pairs was estimated using four microsatellite loci and the geographic position recorded. The authors did not find a relationship between geographic separation and relatedness, of any age group sampled.

Where spatial association of juvenile kin has been found it is likely to be due, at least in part, to spawning in redds which leads to the hatching of related individuals in close proximity to each other. Beall et al. (1994) examined the dispersal of Atlantic salmon juveniles in the Lapitxuri Brook in France, a tributary where no Atlantic salmon occurred naturally. Following planting of eggs in the Brook, the
authors studied the distribution of 0+ fish by electrofishing. Approximately three months after hatching, 0+ fish were found to have mainly dispersed downstream; 89.2% were found within 1600m downstream from the artificial nest, while 3.7% were found within 750m upstream of the nest. The dispersal of Atlantic salmon from artificial nests was also examined by Webb et al. (2001) in Baddoch Burn (a tributary of the Dee, Scotland). Using single locus minisatellites, the authors determined the parentage of 0+ fish (aged 17 weeks) originating from known crosses and known nest sites. It was found that detected dispersal tended to be downstream and less than 1km from initial nest sites.

The implications of the distribution of kin for population analyses were considered in the current study; specifically how sampling strategy may affect results. Previous sampling of parr from the length of Girnock Burn (upstream of the fish traps) allowed this to be investigated by determining the distribution of kin through parentage analysis and comparing the genetic differentiation of sub-samples.
5.2 Methods

Samples were obtained from anadromous returns in 1995 that were caught in the upstream trap and 1+ parr (progeny of these returns in 1995) sampled from the burn in 1997. In 1995 113 anadromous returns were released above the traps to spawn, after an adipose fin clip (or sliver of adipose fin if the fin had previously been removed) was sampled from each. A total of 274 1+ parr were sampled by electrofishing from the length of Girnock Burn (above the fish traps) in 1997. These were selected by length to be 1+ parr. One fish was taken every c.30m, an adipose fin clip and a scale sample were obtained, and the fish replaced. Extraction of DNA from adipose fin clips is described in the Materials and Methods (Chapter 2). Nine to eleven loci were examined for each sample (Table 5.1) and a parentage analysis conducted, as described in the Materials and Methods (Chapter 2). Data had already been collected on the spawning position of anadromous returns in 1995, which had been examined by determining parentage of eggs sampled from redds (described in Taggart et al. 2001).

<table>
<thead>
<tr>
<th>Sample year</th>
<th>Life stage of sample</th>
<th>Relationship</th>
<th>Loci typed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>Anadromous return</td>
<td>Putative parents</td>
<td>2210 1605 2201 410 One9 407 202 171 197 (404 408)¹</td>
</tr>
<tr>
<td>1997</td>
<td>Parr</td>
<td>Putative Offspring</td>
<td>2210 1605 2201 410 One9 407 202 171 197 (404 408)¹</td>
</tr>
</tbody>
</table>

¹See material and methods for details

In order to determine how sampling strategy could affect the results of population genetic analyses when sampling juveniles of a single age class, the parr were divided into five equal groups using two methods. The first method was based on their distribution (as sampled) in the burn. The first 250 samples obtained were
divided into five groups of 50 samples, corresponding to samples from (adjacent) stretches of river approximately 1.47 km in length. For example, group one consisted of 50 parr sampled from the first 1.47 km above the traps and group two of 50 parr sampled between 1.5 km and 2.97 km above the traps. Hereafter these are referred to as ‘spatially discrete’ samples. The second method was designed to simulate sampling over a wider area. The 250 parr (the same individuals as used in method one) were split into groups of 50 by randomly sampling without replacement (using a sampling routine executed in SPSS v.14). These samples are referred to as ‘random samples’.

The groups of samples were then analysed using standard population genetics procedures, using the initial nine loci listed in Table 5.1, which were screened in all individuals. Hardy Weinberg equilibrium and differences in allele frequencies were tested for using exact tests and $F_{IS}$ was calculated using GENEPOP 3.4 (Raymond and Rousset 1995). For high precision exact tests were run with the following parameters: 20000 dememorization steps, 200 batches and 6000 iterations per batch. The standard error of the probability was always << 0.01. Pairwise values of $F_{ST}$ were calculated and the significance determined using ARLEQUIN 3.01 (Excoffier et al. 2005), with 20000 permutations. Again the standard error of the probability was always << 0.01.

The distribution of juvenile kin in the burn was examined by plotting the position from which groups of both full sibs and maternal half sibs were sampled. In cases where eggs had been assigned (in the previous study by Taggart et al. 2001) to corresponding anadromous sires and dams (in the case of full sibs) or the corresponding anadromous female (in the case of maternal half sibs), the position of the redd from which eggs had been sampled was also plotted.
5.3 Results

5.3.1 Parentage analysis
A total of 269 parr were assigned to parents in 1995. Twenty full sib groups with three or more progeny were identified and 11 full sib groups with two progeny. 17 maternal half sib groups with six or more members were identified, four groups with five members, eight groups with four members, and 15 groups with three or fewer members.

5.3.2 Population genetic analyses
Significant departures from Hardy Weinberg equilibrium were detected at individual loci in four of the five groups of ‘spatially discrete’ samples (without sequential Bonferroni correction, Table 5.2). However, combining over all loci revealed significant departures in only two groups of the five (both prior to and following sequential Bonferroni correction). In the random samples, significant departures from Hardy Weinberg equilibrium were detected at individual loci (without sequential Bonferroni correction) and for all loci combined (both prior to and following sequential Bonferroni correction), in two samples of five. Overall Fis values (calculated over all loci) were negative, corresponding to an excess of heterozygotes, in the spatially discrete samples. Both negative and positive values were found in the random samples.
Table 5.2 Significance of tests for Hardy Weinberg equilibrium (HW p), $F_{IS}$, and observed ($H_o$) and expected ($H_e$) number of heterozygotes for groups obtained from 1.47km lengths from the burn or randomly sampled from 7.47km.

<table>
<thead>
<tr>
<th>Group #</th>
<th>Locus</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2211</td>
</tr>
<tr>
<td>1</td>
<td>HW p</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td>-0.132</td>
</tr>
<tr>
<td></td>
<td>$H_o$</td>
<td>41.6</td>
</tr>
<tr>
<td></td>
<td>$H_e$</td>
<td>47.4</td>
</tr>
<tr>
<td>2</td>
<td>HW p</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td>$H_o$</td>
<td>38.7</td>
</tr>
<tr>
<td></td>
<td>$H_e$</td>
<td>41.1</td>
</tr>
<tr>
<td>3</td>
<td>HW p</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td>-0.071</td>
</tr>
<tr>
<td></td>
<td>$H_o$</td>
<td>41.1</td>
</tr>
<tr>
<td></td>
<td>$H_e$</td>
<td>44.1</td>
</tr>
<tr>
<td>4</td>
<td>HW p</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>$H_o$</td>
<td>43.1</td>
</tr>
<tr>
<td></td>
<td>$H_e$</td>
<td>44.4</td>
</tr>
<tr>
<td>5</td>
<td>HW p</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>$F_{IS}$</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>$H_o$</td>
<td>41.1</td>
</tr>
<tr>
<td></td>
<td>$H_e$</td>
<td>43.3</td>
</tr>
</tbody>
</table>

*significant at the 0.05 level following sequential Bonferroni correction (overall results for Hardy Weinberg equilibrium only). Initial $\alpha=0.05/5$
Tests for allelic differentiation revealed statistically significant differences in nine of ten pairwise comparisons of spatially discrete samples (Table 5.3). For random samples, no pairwise comparison was statistically significant (Table 5.3). Similar results were obtained when comparing the significance of $F_{st}$ values (Table 5.4); there were statistically significant differences in eight out of ten pairwise comparisons (following Bonferroni correction) of spatially discrete samples, while random samples were genetically homogenous (Table 5.4). In addition, for the few cases where pairwise comparisons of spatially discrete samples were not statistically significant (following sequential Bonferroni correction), the comparisons were between adjacent samples (sample numbers two and three, and three and four).

These results suggest that observed genetic differentiation between pairs of samples obtained over stretches of the burn c. 1.47km in length was due to sampling bias, since no genetic differentiation was observed when samples were taken randomly over the entire sampling area.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>0.95</td>
<td>0.12</td>
<td>0.36</td>
<td>0.94</td>
</tr>
<tr>
<td>2</td>
<td>&lt;0.01**</td>
<td>-</td>
<td>0.12</td>
<td>0.49</td>
<td>0.64</td>
</tr>
<tr>
<td>3</td>
<td>&lt;0.01**</td>
<td>0.65</td>
<td>-</td>
<td>0.16</td>
<td>0.34</td>
</tr>
<tr>
<td>4</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>0.01*</td>
<td>-</td>
<td>0.09</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>-</td>
</tr>
</tbody>
</table>

**significant at 0.01 level after sequential Bonferroni correction, initial $\alpha=0.01/10$

*significant at 0.05 level after sequential Bonferroni correction, initial $\alpha=0.05/10$
Table 5.4 Significance of Fst values for comparisons between pairs of samples. Above diagonal-random samples. Below diagonal- spatially discrete samples (groups relate to adjacent samples). Results are shown prior to sequential Bonferroni correction.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>0.91</td>
<td>0.15</td>
<td>0.48</td>
<td>0.87</td>
</tr>
<tr>
<td>2</td>
<td>&lt;0.01**</td>
<td>-</td>
<td>0.23</td>
<td>0.42</td>
<td>0.71</td>
</tr>
<tr>
<td>3</td>
<td>&lt;0.01**</td>
<td>0.48</td>
<td>-</td>
<td>0.39</td>
<td>0.53</td>
</tr>
<tr>
<td>4</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>0.04</td>
<td>-</td>
<td>0.11</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>&lt;0.01**</td>
<td>-</td>
</tr>
</tbody>
</table>

**significant at 0.01 level after sequential Bonferroni correction, initial α=0.01/10

5.3.3 Distribution of kin

In order to explore possible reasons for sample bias the distribution of identified full sib and maternal half sib groups was depicted graphically (Figure 5.1, Figure 5.2, respectively). Clustering of kin was evident in both cases. For example, in Figure 5.1 full sib groups 6 and 7 were only found within the first 1500m, while full sib groups 3 and 4 were only found upstream of 7000m. For groups of maternal half sibs, offspring of females 55, 103 and 111 were represented only in the first 1500m, while females 12 and 29 were represented only in samples taken upstream of 5000m (Figure 5.2).

Comparing the position of putative redds with the position of parr, suggested that dispersal from putative redds had occurred both upstream and downstream. In some cases families appeared to move in a certain direction. For example, full sib group number one appeared to have largely moved upstream from the putative redd, while members of full sib group number four appeared in general to have moved downstream (Figure 5.1). In contrast members of sib group two were not found to have moved in a particular direction; equal numbers were found upstream and downstream of the putative redd (Figure 5.1).
Figure 5.1 Distribution of full sibs (parr). Each diamond represents the position of a single parr, diamonds on a single horizontal line represent a full sib family. Only full sib groups consisting of three or more individuals are shown. Centres of circles represent the position of redds from which eggs were assigned to the corresponding dam and sire i.e. were members of the corresponding full sib group. 100 parr are depicted, of a total of 269.
Figure 5.2 Distribution of parr sampled that were offspring of particular female spawners. Each diamond represents the position of a single parr, diamonds on a single horizontal line represent the offspring of a particular dam. Data are only shown for female spawners where six or more offspring were detected. Centres of circles represent the position of redds from which eggs were assigned to the corresponding dam. 192 parr are depicted, of a total of 269.
5.3.4 Discussion

Analysis of allele frequencies from groups of individuals is a popular means of describing population structure of fishes (and other organisms). However, this presents a dilemma, as knowledge of population structure is necessary to inform unbiased sampling, whereas unbiased samples are required for the analysis of population structure.

Overall, the results show that sampling strategy did affect the results of population genetic analyses applied to 1+ parr in Girnock Burn in this study. Groups of 50 samples obtained over 1.47 km lengths of the burn showed statistically significant genetic differentiation, although this was not apparent when groups of 50 parr were randomly sampled over 7.47 km. This could be attributed to sampling of family groups, as these were shown to cluster, in some cases within stretches of the burn one to two km in length. Evidently, the recommendation by Hansen (1997) for brown trout that sampling should be conducted in stretches above 100m in length is not appropriate for 1+ Atlantic salmon parr in Girnock Burn. Sampling over larger regions, perhaps above 4-5km would be more appropriate when sampling a single age class of juveniles.

However, this may be logistically difficult and expensive to achieve. In order to avoid sampling a limited number of families and to obtain a better representation of the population as a whole, a more suitable method may be to sample more than one age class, as recommended by Mjolnerold et al. (1999). However, the distribution of families in Girnock Burn, a relatively small tributary, may differ to that in main rivers. Fontaine and Dodson (1999) describe the conditions in tributaries as high density (of fish) and low dispersal, in contrast to rivers, which they classified as having low density.
and high dispersal. Thus clustering of kin may not occur to the same extent in large rivers, or indeed in other tributaries where factors affecting dispersal and or spawning habits differ.

In this study there were some differences in the outcomes of tests for Hardy Weinberg equilibrium between spatially discrete and random samples (when focusing on individual loci prior to sequential Bonferroni correction). However, differences were relatively small, which suggests that results of Hardy Weinberg equilibrium tests may not necessarily be a useful predictor of such biased sampling.

A number of studies have now examined the temporal stability of genetic relationships among populations of salmonids. Problems may be encountered when using juvenile fish for such studies. Comparing families could lead to the observed degree of temporal genetic variation being higher than its actual level. Such studies have generally included samples from only adult fish (Palm et al. 2003b; Heath et al. 2002; Tessier and Bernatchez 1999), or mainly adult fish (Hansen et al. 2002; Nielsen et al. 1999a).

In some cases where juveniles have been used, comparatively little information regarding sampling protocols, such as the age of individuals or area sampled, has been given. For example Nielsen et al. (1999a) described the juveniles sampled from one river as parr sampled by electrofishing, with no indication as to the age class(es) included or the area electrofished. Hansen et al. (2002) stated the likely age composition of juveniles, but gave no details of the size of sampling area. Østergaard et al. (2003) conducted a study of the temporal and spatial genetic variation of populations of brown trout, based on samples consisting largely of juvenile fish.

Although the authors stated that these were mainly composed of two age classes and
that sampling had been carried out at ‘several points’ to avoid sampling a few families, there was no indication of the distance between sampling points. Garant et al. (2000) conducted a study on the temporal and spatial genetic variation of Atlantic salmon, by sampling 46-50 fry from regions of river greater than 200m$^2$. The authors found that the variation due to temporal instability or ‘random sampling error’ (i.e. biased sampling) was three times greater than the spatial component of variation. In the current study clustering of families had an impact on the allele frequencies of samples of 1+ parr taken from stretches of river c. 1.47 km in length, which corresponds to approximately 9700m$^2$ (based on a mean width of 6.6m). Had the fry in the study by Garant et al. (2000) also been clustered in a similar way, the observed allele frequencies may have been biased.

Webb et al. (2001) studied the dispersal of 0+ Atlantic salmon from artificial nests and found that this was generally downstream and within one km of the initial nest. Parr aged 1+ in the current study (a year older than the fish sampled by Webb et al. 2001) were found at greater distances from putative redds, but still remained clustered. In the families observed with larger numbers of detected progeny and cases where the position of putative redds had been identified, there appeared to be movement both upstream and downstream. However, it should be noted that not all redds may have been detected, and thus these observations should be treated with some caution. Beall et al. (1994) observed the dispersal of Atlantic salmon from an artificial nest in the year following hatching and found that the proportion of fish found upstream of the nest increased with time. In the current study, particular families appeared to move from redds in different directions. This is similar to the findings of Webb et al. (2001), who
reported that when the eggs from three families were planted out in the same nest, the pattern of movement downstream differed for each family. Thus there is the suggestion that juvenile salmon dispersal is not random, particularly among younger age classes. It is also likely that, without due diligence, electrofished samples are likely to be biased towards these most numerous individuals. Therefore much greater emphasis should be placed on devising and implementing suitably robust sampling procedures. Further work is warranted in order to determine the distribution of juvenile kin in other rivers. In the current study sampling of only a single age class was investigated. It may be that sampling single age classes is more of a problem at lower latitudes, where smoltification occurs at lower ages – one year old smolts are found in southern Europe (Klemetsen et al. 2003). In contrast, at higher latitudes a greater age range of juveniles will be present due to the higher age at smolting, for example, up to eight years in northern Quebec (Klemetsen et al. 2003).
6 Trap leakage and the extent of straying

6.1 Introduction

The reproduction of animals at the same ‘site’ as their parents is known as philopatry and is a common feature of salmonids. This strategy is of evolutionary significance as it can potentially drive the process of local adaptation, leading to speciation. Anadromous salmonids are renowned for their ability to home to natal streams to spawn, following often long and distant migrations to marine feeding grounds (e.g. Groot and Margolis 1991; Youngson and Hay 1996; Thedinga et al. 2000). However, homing is not perfect and straying has been recorded, to some degree, in the vast majority of studies on anadromous salmonids to date (for example appendix 1 in Hendry and Stearns 2004). Straying (or dispersal) can be defined as

‘the interpopulation movement between the natal area and the area where breeding first takes place’ (Fraser et al. 2004),

in contrast to gene flow which may be described as

‘the outcome of successful reproduction after dispersal’ (Fraser et al. 2004).

These definitions will also be applied in the current study i.e. ‘strayers’ may not necessarily reproduce successfully in the recipient population.

6.1.1.1 Implications of straying and gene flow

In salmonids high fidelity of homing theoretically leads to relatively restricted gene flow between populations, thus allowing potentially different adaptive characteristics to evolve (Taylor 1991; Dittman and Quinn 1996; Jonsson et al. 2003; McLean et al. 2004). However, it is difficult to prove conclusively that candidate characteristics are adaptive (Hansen et al. 2002). The degree of gene flow affects the geographical scale over which local adaptation is likely to occur. High
rates of gene flow will tend to act against localized adaptation (Adkison 1995). On a longer timescale gene flow among populations can prevent them from diverging into different species. Morjan and Rieseberg (2004) demonstrated that even a small level of gene flow could allow species to evolve as cohesive units. The authors reviewed estimates of gene flow in a wide range of species (including species of plants, fungi and animals) and found that the level of gene flow was theoretically too low to prevent divergence in many species. However, the authors concluded that they might be prevented from diverging by advantageous alleles under strong selection that would spread rapidly throughout a species despite low levels of gene flow.

While homing is a well documented feature of many salmonid species, straying has also played (and continues to play) a pivotal role in the success of this group of fishes. Straying is an essential attribute for the colonization of new habitat. Until 8000-15000 years ago most salmonids were confined to a restricted number of isolated refugia due to glaciation. The current distribution (of naturally established populations) is therefore a result of straying from these glacial refugia (Kinnison and Hendry 2004). There are records of the very recent colonization by coho salmon *Oncorhynchus kisutch*, of streams in Alaska that developed following glacial recession. The most recently colonised river was estimated to have formed less than 50 years ago (Milner *et al.* 2000). Straying has also been witnessed following human interference. Ayllon *et al.* (2006) attributed the presence of brown trout in a small stream that had been affected by damming, to strayers from nearby populations (the original population was lost during dam construction as the river was diverted). Straying may also be important for the persistence of small populations. Strayers to small populations are likely to maintain (or at least reduce
the rate of loss of) genetic diversity and so relieve or ameliorate negative effects of inbreeding (Consuegra et al. 2005). Primmer et al. (2006) found evidence that suggested populations of Atlantic salmon in tributaries of the Varzuga River in northwest Russia that were furthest from the river mouth received fewer strayers, and so might be more susceptible to the negative effects of declines in population size.

Gene flow also has implications for the management of salmonids with respect to their conservation. Conserving a particular population at the expense of others may result in the loss of gene flow to that (population) being conserved, which may in turn have negative effects (Hansen et al. 2002). In addition, supportive breeding programs may alter levels of gene flow through indiscriminately using sea-returns as broodstock, thus increasing the reproductive success of strayers.

6.1.1.2 Identification of strayers and the estimation of the effective number of migrants

Levels of straying may be estimated directly through physical tagging studies or, as developed more recently, through genetic tagging-profiling. The effective rate of straying (or the effective number of migrants) may be estimated indirectly through the analysis of allele frequencies.

Tagging studies may focus on the level of straying to a particular river or the level of straying from it. Offspring originating in a particular river may be tagged and, in the first instance, the numbers of tagged and untagged fish that return to the river determined, or in the second instance the locations of returning tagged fish may be identified (i.e the original river or alternative locations). Conventional tagging
involves attaching a physical marker to fish (such as coded wire tags), although Quinn et al. (1987) used the prevalence of parasites to infer the rate of straying in populations of sockeye salmon (*O. nerka*). An additional method is marking through temperature induced banding on otoliths (Mortensen et al. 2002).

A relatively recent technique is to explore the level of straying directly through the analysis of multilocus genotypes of individuals to determine their origin. This method is akin to physical tagging of individuals, except that the tag is inherent to the fish’s genetic makeup, so is never lost and is present at all life stages. Hansen et al. (2001b) reviewed the use of multilocus genotypes to assign individuals to the population of origin. The technique can be applied in a number of ways. For example Primmer et al. (2000) reported that the origin of a fish (Atlantic salmon) caught in an angling contest was more likely to be a fish market than the lake where the contest was held. Nielsen et al. (2001) applied assignment tests to determine whether current Danish populations of Atlantic salmon in five rivers were more likely to be derived from exogenous stocked fish or native fish, and found that a significant number of individuals from two of the populations were from native sources. Fraser et al. (2004) classified individuals in a population of brook charr (*Salvelinus fontinalis*), as strayers and non-strayers, also using an assignment method to determine the population of origin. A program commonly applied to such studies is GENECLASS 2 (Piry et al. 2004). This has previously been used numerous times to assign individuals to populations, for example in the diamondback terrapin *Malaclemys terrapin* (Hauswaldt and Glenn 2005), whitefly *Bemisia tabaci* (De Barro 2005), a pioneer tree species *Antirhea borbonica* (Litrico et al. 2005) and the three-spined stickleback *Gasterosteus aculeatus* (Taylor et al. 2004).
2006). It has also been used to identify first generation immigrants in the Mallorcan midwife toad *Alytes muletensis* (Kraaijeveld-Smit *et al.* 2005).

Estimating the *effective* number of migrants is obviously not possible from physical tagging studies, which merely indicate presence, rather than reproductive success of individuals. Estimating the effective number of migrants, indirectly, through genetic analysis is based on a theory introduced by Wright (1931). This states that

\[ N_e = 0.25(1/F_{ST} - 1) \]

where

\[ N_e = \text{effective number of migrants per generation} \]

\[ F_{ST} = \text{proportion of total genetic variation that is caused by differences among populations} \]

However, it is based on a number of assumptions that are rarely met in the real world, such as equal numbers of migrants between subpopulations, subpopulations of equal size and randomness with respect to dispersal (for example, that females are just as likely to disperse as males) (Whitlock and McCauley 1999). Other similar measures of differentiation have been developed, for example *R* _S_T_ by Slatkin (1995) for use with microsatellites based on a stepwise mutation model. These can be used to derive measures of gene flow, but still rely on the same island model assumptions (Whitlock and McCauley 1999). However, as these assumptions are likely to be violated, Whitlock and McCauley (1999) advised against using values of *F* _S_T_ (or other similar measures) to estimate the effective number of migrants. Other approaches have since been developed that use maximum likelihood methods based on coalescence theory such as that described by Beerli and Felsenstein (2001). These allow for asymmetric gene flow and unequal subpopulation sizes.
6.1.1.3  Factors that affect straying rate

Straying in salmonids may be influenced by environmental conditions and there are a number of theories explaining how selection might favour straying or philopatry (Hendry et al. 2004b). Straying rate in salmonids has been found to vary between populations of the same species (e.g. Unwin and Quinn (1993) compared to Hard and Heard (1999) in chinook salmon), over time for the same population (Jonsson et al. 2003) and between populations of different species inhabiting a common environment (Shapovalov and Taft 1954, cited in Quinn 1993). However, in general, straying tends to be between local rather than distant populations (Hendry et al. 2004b). Factors that may influence straying are water flow (Unwin and Quinn 1993), size of population (Jonsson et al. 2003), and the age (Hard and Heard 1999, Jonsson et al. 2003) and sex of the individual (Hard and Heard 1999; Jonsson et al. 2003).

Measures of straying from a river can be biased by variances in sampling intensity in different rivers. In addition tags may induce changes in the degree of straying. Thedinga et al. (2000) found elevated rates of straying in pink salmon that were coded wire tagged compared to pelvic fin clipped, and Habicht et al. (1998) found the ability of pink salmon to home could be affected by the position of coded wire tags.

6.1.1.4  Gene flow

Gene flow among populations is dependent on the (lifetime) reproductive success of strayers. The survival and reproductive success of non-native fish in native populations has received particular attention. This is due to the potential for non-native fish to interbreed with native stocks as a result of human activity. Such non-
native fish could be escaped farmed fish or releases of hatchery-reared fish for the purposes of conservation or fishing. Released hatchery-reared fish may be progeny of non-native broodstock, or be genetically dissimilar to native fish as a result of adaptation to the hatchery environment. Such non-native fish (whether escapees or purposely released) bear similarities to strayers in the wild although 1) they may be present in much greater numbers, and 2) they may be less well adapted to the natural environment as a result of domestication. Studies on the survival and reproductive success up to varying life stages of non-native fish and farmed fish in the wild, compared to those of native fish, have generally shown a higher survival of native fish (e.g. Skaala et al. 1996 in Salmo trutta comparing non-native hatchery fish and native fish; McGinnity et al. 2003 in Salmo salar comparing farmed and native fish; McGinnity et al. 2004, in Salmo salar comparing non-native and native fish) or a higher reproductive success of native fish (Fleming et al. 2000, in Salmo salar for lifetime reproductive success of farmed vs. native fish; McLean et al. 2004, in O.mykiss up to smolt stage comparing non-native hatchery fish and native fish). However, differences between hatchery or farmed fish and wild fish are likely to be greater than those between natural strayers and fish in the recipient population. Thus the interaction of wild strayers with recipient populations may therefore be more subtle, although it has been suggested that, in general, the level of gene flow among populations in the wild is lower than the level of straying (for example Tallman and Healey 1994; Hendry et al. 2004b).
6.1.2 Trap leakage and straying rate at Girnock Burn

An accurate estimate of straying rate to a region (or the degree of homing), based on tagged returns, requires that all juveniles leaving the tributary or river in question are tagged. At Girnock Burn a smolt trap based on a Wolf trap (Wolf 1951) is designed to catch all smolts and parr travelling downstream. All downstream migrating parr and smolts that are caught have the adipose fin removed and are microtagged before being placed downstream. The mark is retained in the returning adults, thus allowing the identification of anadromous adults that originated above the traps. Sea returns marked with an adipose fin clip are considered philopatric, as microtags that have been recovered from returning adults have only ever been traced back to Girnock Burn. Adult returns are recorded as clipped (having had the adipose fin removed) or unclipped. This allows an estimate of the degree of homing to the trap. For example, Youngson et al. (1994) investigated the level of homing to Girnock Burn (more specifically, the level of homing to the region above the traps in Girnock Burn). The authors found that 45% of returns to Girnock Burn (of an age expected to bear fin clips, between 1988 and 1990) were fin clipped. However, they indicated that the actual level of homing might in reality have been higher, as it was possible that trap leakage occurred (i.e. juveniles may have ‘leaked’ over the trap without being clipped). Although the downstream trap is designed to catch all smolts travelling downstream, the trap may overflow in times of spate and when grids become clogged with debris, such as fallen leaves (particularly in autumn). In these situations water and presumably fish are able to pass over the traps.

In a direct comparison of the reproductive success of philopatric and presumably non-philopatric returns, Taggart et al. (2001) found no significant difference in the reproductive success (up to the egg stage) of the two groups of returns in any of
three years studied (spawning in 1991, 1992, and 1995). However it was also possible in this study that a portion of non-philopatric returns were actually philopatric (i.e. due to trap leakage smolts may have evaded capture and not been marked).

Between 1990 and 2004 33% of adult returns caught at the Girnock traps were adipose fin clipped. The proportion of clipped and non-clipped fish varied greatly during this period (eg. from 0% clipped returns in 1997 to 64% in 1990, Figure 6.1).

![Graph showing numbers of clipped and unclipped returns to Girnock burn from 1990-2004](image)

**Figure 6.1 Number of sea returns to Girnock Burn that were clipped and unclipped between 1990 and 2004.**

The question then arises as to the origin of the unclipped returns. It may be that unclipped returning adults merely passed over the smolt trap (i.e. were due to trap
leakage), and so were actually philopatric. Alternatively they may have originated below the traps. Youngson and Hay (1996) describe the region towards which adult returns home as a ‘homing unit’. Since the traps provide an artificial barrier, they are not likely to coincide with the true boundary of the homing unit. Thus adults returning to the traps at Girnock Burn, but whose natal sites were below the traps, may still be homing to the same ‘unit,’ and be considered philopatric. Conversely, they may originate in other homing units. The origin of unclipped returns is considered in more detail in this chapter.

This chapter explores the extent of trap leakage and its implications at Girnock Burn. The following points are addressed:

1. Estimation of the extent of trap leakage. DNA profiles of unclipped (presumed non-philopatric) adult returns were compared with genotypes of adults that returned to Girnock Burn in previous years in order to identify possible philopatric individuals (i.e. fish that had evaded adipose fin clipping on the downstream migration).

2. Re-evaluation of homing rates in light of the information obtained from point 1.

3. Identification of adult returns that were likely to originate in different populations, through the analysis of multi-locus genotypes.

4. Comparison of the reproductive success of philopatric and non-philopatric adult returns.
6.2 Methods

6.2.1 Samples employed

In order to identify philopatric individuals and estimate the extent of trap leakage, three parent-offspring datasets were used (Table 6.1).

<table>
<thead>
<tr>
<th>Set</th>
<th>Parents</th>
<th>Potential Offspring</th>
<th># loci used</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year</td>
<td>Age assigned to parents</td>
<td># correct age for assignment</td>
</tr>
<tr>
<td>1</td>
<td>1991 95</td>
<td>1995 116 2.1</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>1995 116</td>
<td>1999 10^2 2.1</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>2000 89</td>
<td>2000 41 2.3 or 3.2</td>
<td>9</td>
</tr>
</tbody>
</table>

^1 Number of potential offspring that were positively identified (from scale reading) to be the correct age for assignment to parents. Clipped and unclipped returns are included.

^2 Only adult returns in 1999 aged, from scale reading, as ‘uncertain’ or 2.1 were employed

In order to compare the reproductive success of philopatric and non-philopatric individuals (including results from parentage assignments that identified unclipped individuals as philopatric), datasets shown in Table 6.2 were used.
### Table 6.2 Datasets employed in comparing the reproductive success of philopatric and non-philopatric returns.

<table>
<thead>
<tr>
<th>Set</th>
<th>Parents</th>
<th>Potential Offspring</th>
<th># loci used</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Year</td>
<td>#</td>
<td>Year</td>
</tr>
<tr>
<td>RS1</td>
<td>1995</td>
<td>116</td>
<td>1997</td>
</tr>
<tr>
<td>RS2</td>
<td>2000</td>
<td>30</td>
<td>2004</td>
</tr>
<tr>
<td>SRS1</td>
<td>1995</td>
<td>116</td>
<td>1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td>SRS2</td>
<td>2000</td>
<td>30</td>
<td>2004</td>
</tr>
</tbody>
</table>

1 RS- reproductive success
2 Number of potential offspring that were positively identified (from scale reading) to be the correct age for assignment to parents. Clipped and unclipped returns are included.
3 includes only adult returns that were released above the traps to spawn naturally (excludes adult returns used in quasi-natural spawning)
4 SRS-reproductive success up to sea return stage

### 6.2.2 Parentage assignment

The assignment of female parents to offspring was accomplished using DAMFIND, and the assignment of male parents to offspring by using SIREFIND, employing the criteria outlined in Chapter 2, section 2.5.2. Both dam only assignments (i.e. where the sire was assumed to be a mature parr), and anadromous parental pair assignments were included in the analyses.

### 6.2.3 Identification of strayers from population genotype data

GENECCLASS 2 was used to determine ‘migrants’ (or ‘strayers’) through the analysis of genotypes. GENECLASS 2 can be employed in two ways, firstly to assign individuals to populations, secondly to detect first generation immigrants. It allows the analyses to be carried out in situations when only a single population is
available (such as this), and was therefore employed in preference to other programs (i.e. IMMANC and STRUCTURE), which require more than one population (Maudet et al. 2002; Rannala and Mountain 1997). However, GENECLASS 2 assumes Hardy Weinberg equilibrium. As this assumption was violated in all years but 1991, the results may be affected.

The Rannala and Mountain (1997) Bayesian approach to estimate the likelihood of individuals being resident in a population or being a first generation migrant was employed in this study. Cornuet et al. (1999) and Maudet et al. (2002) found this approach gave more accurate results than other methods available in GENECLASS 2.

In both instances the Monte-Carlo resampling method put forward by Paetkau et al. (2004) was used to obtain a threshold value beyond which an individual was classified as non-resident or a first generation migrant. This was shown to be superior to other methods of resampling (with a reduced type I error), and differs in that it resamples a set of individuals of the same number to that in the actual sample, and samples gametes rather than alleles (Paetkau et al. 2004). A type one error rate of 0.001 was used, and 1000 individuals were resampled.

As many adults as possible were used in the analyses, since alleles should be representative of the population (Primmer et al. 2000), and with overlapping generations, this was the most certain way of representing the majority of alleles in the population.

Since 12 loci were used to type returns in 2000, 2001 and 2004, and only nine of these loci were used to type returns in 1991 and 1995, two separate groups of genotypes were analysed. One included adult returns in 1991, 1995, 2000, 2001 and
2004, and used nine loci. The other consisted of adult returns in 2000, 2001 and 2004 and included 12 loci.
6.3 Results

6.3.1Extent of trap leakage

Data for an initial estimate of trap leakage are presented in Table 6.3. Taking set 3 as an example: of 117 progeny screened (Table 6.1), seven individuals were assigned to parents in the 2000 spawning year class (these individuals would be aged 2.1). Of these seven individuals, one was unclipped. When offspring of female parents noted as ‘part-spent’ or ‘nearly kelt’ were included, 13 individuals could be assigned to the 2000 spawning year class, of which five were unclipped.

Table 6.3 Estimate of the level of trap leakage. Numbers of returns that could be assigned to salmon that spawned in Girnock Burn in 1991, 1995 and 2000 and the number of these that were unclipped—presumed to have leaked over the trap.

<table>
<thead>
<tr>
<th>Set</th>
<th>Year of return parents</th>
<th>Year of return of offspring</th>
<th># unclipped / # assigned</th>
<th>% unclipped</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2000</td>
<td>2004</td>
<td>1/7(5/13)</td>
<td>14 (38)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>23/68 (29/77)</td>
<td>34 (38)</td>
</tr>
</tbody>
</table>

\(^1\text{Numbers in parentheses include offspring of females that were noted as \textquote{part-spent} or \textquote{nearly kelt} on entry to the traps, otherwise these offspring are excluded.}\)

In datasets where female parents noted as ‘part-spent’ or ‘nearly kelt’ (in set two and set three) were identified, higher percentages of unclipped individuals were observed when offspring of these females were included. Successful spawning below the trap (as opposed to trap leakage) may account for this. Thus the estimate of mean trap leakage (of 34%) based on data excluding offspring of these female spawners is likely to be more reliable.
In an attempt to further purge the data of such spawners, another estimate of the extent of trap leakage was made, utilizing information obtained from the assignment of eggs to spawners in 1995 (from the study by Taggart et al. 2001). The location of redds in which females had spawned had been determined, and the maximum distance between redds sampled in which the same female had spawned was 3.02km. In this more conservative estimate of trap leakage only three groups of offspring were included, those that originated from:

- females that were noted as ‘unspent’ on entry to the traps
- a parental pair that had previously been identified in the egg survey, based on the assumption that few parental pairs spawn in different locations. (Four pairs of a total of 80 identified pairs were found to spawn in different locations according to the egg survey conducted by Taggart et al. (2001), on eggs spawned in 1995. The distance between redds was less than c.105m apart for each of three pairs of redds, while one was c.720m apart.)
- females that were found only to spawn more than 3km upstream from the traps (the maximum distance between spawning sites for a single female in 1995 was 3.02km, according to sampling carried out by Taggart et al. 2001).

Using only this subset of results, the extent of trap leakage was estimated as 37% (11 unclipped progeny / 30 assigned) c.f. 46% from the full dataset.

It was plausible that the level of trap leakage differed between years, due to varying periods of high water flow and varying levels of monitoring of the trap. In order to assess whether the level of leakage differed between years, a more detailed analysis of ‘set 2’ (see Table 6.3) was made, as this dataset included the highest number of offspring samples. The age, year of return and year of leaving freshwater for these
trap leakage and the extent of straying

Chapter 6

fish are shown in Table 6.4. The year in which offspring left freshwater was obtained either from the year in which it was spawned (from the assignment of parents to anadromous returns aged 2.1) or was estimated based on scale readings (for example, a return of total age four years may have spent two or three years in fresh water, which would have required data obtained from scale readings to discern). It was not possible to determine precisely when adult returns would originally have passed through (or over) the traps, on the downstream migration. It could either have been in the autumn as parr or in the spring as smolts. Fish passing through the traps in autumn as parr one year leave freshwater with smolts that pass through the traps in the following spring. Thus the rate of trap leakage estimated for smolts leaving freshwater in 1998 was related to conditions that may affect trap efficiency (such as water flow and the level of monitoring of the trap) that occurred in the autumn of 1997 and the spring of 1998.

Table 6.4 Age of offspring of spawners in 1995 used to compare the level of trap leakage in autumn 1997/spring 1998 to that in autumn 1998/spring 1999.

<table>
<thead>
<tr>
<th>Year left freshwater as smolt</th>
<th>Year returned to Girnock Burn</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>1999</td>
<td>2.1</td>
</tr>
<tr>
<td>1999</td>
<td>2000</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>2001</td>
<td>2.3</td>
</tr>
</tbody>
</table>

1 Includes parr passing through the traps in autumn 1997 and smolts passing through the traps in spring 1998

2 Includes parr passing through the traps in autumn 1998 and smolts passing through the traps in spring 1999

Estimates of the extent of trap leakage in the two periods are presented in Table 6.5. The freshwater age of 13 returns was uncertain; these were discounted. Offspring of
females that were noted as ‘part-spent’ or ‘nearly kelt’ were also discounted (the age of one of these was also uncertain).

Table 6.5 Estimates of the extent of trap leakage for smolts that left freshwater in 1998 and 1999.

<table>
<thead>
<tr>
<th>Year left freshwater as smolt</th>
<th># unclipped/ # returns assigned to parents</th>
<th>% age of total unclipped</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>5/7</td>
<td>71</td>
</tr>
<tr>
<td>1999</td>
<td>5/29</td>
<td>17</td>
</tr>
<tr>
<td>1998 &amp; 1999</td>
<td>10/36</td>
<td>28</td>
</tr>
</tbody>
</table>

1 according to time spent in freshwater as indicated by scale readings

As before, another estimate was calculated based only on those offspring that were thought likely to have originated above the traps (by applying the same criteria for selection of offspring as previously) (Table 6.6). As in Table 6.5, the freshwater age of 13 returns was uncertain (these were discounted) and offspring of females that were noted as ‘part-spent’ or ‘nearly kelt’ were also discounted.

Table 6.6 Estimates of the rate of trap leakage for smolts that left freshwater in 1998 and 1999. Only offspring that were thought likely to have originated above the traps were included.

<table>
<thead>
<tr>
<th>Year left freshwater as smolt</th>
<th># unclipped/ # returns assigned to parents</th>
<th>% age of total unclipped</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>2/4</td>
<td>50</td>
</tr>
<tr>
<td>1999</td>
<td>3/17</td>
<td>18</td>
</tr>
<tr>
<td>1998 &amp; 1999</td>
<td>5/21</td>
<td>24</td>
</tr>
</tbody>
</table>

1 according to time spent in freshwater as indicated by scale readings

The extent of trap leakage was significantly different between years in the initial estimate (exact test for contingency tables, P= 0.01), but not significantly different in the more conservative estimate using only females that were unlikely to have spawned below the traps (exact test for contingency tables, P=0.23). Thus further investigation would be required to determine whether trap leakage varies between years.
However, although numbers were low, the estimate of the extent of trap leakage was higher in 1998 than in 1999, in both estimates. The results suggest that the level of trap leakage differs between years.

### 6.3.2 Revised rate of homing

A revised estimate of the degree of homing was calculated. Previous indications of the level of homing to Girnock Burn, were obtained by classifying ‘homers’ as those returns that had been clipped. In this revised estimate homers also included unclipped individuals that had been assigned to previous returns to Girnock Burn i.e. they were either spawned above the trap and leaked over it without being clipped or were offspring of females that had spawned below the traps prior to entering the upstream trap. The numbers of individuals that would ordinarily be identified as homers (i.e. clipped offspring) were compared to numbers of individuals that could be assigned as being philopatric by DNA profiling (Table 6.7, summarised in Table 6.8). Only returns that were identified to be the correct age for assignment (according to scale reading) were included in the analysis.
Trap leakage and the extent of straying

Chapter 6

Table 6.7 Assignment of adult returns to adult spawners in previous years. Only offspring of a given age that were suitable for assignment, according to scale reading, were included.

<table>
<thead>
<tr>
<th>Year of return of parents</th>
<th>Year of return of offspring</th>
<th>Age of offspring</th>
<th>Result of assignment</th>
<th># offspring clipped</th>
<th># offspring not clipped</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>1995</td>
<td>2.1</td>
<td>Assigned</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not assigned</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1995</td>
<td>1999</td>
<td>2.1</td>
<td>Assigned</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not assigned</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>1995</td>
<td>2000</td>
<td>3.1 or 2.2</td>
<td>Assigned</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not assigned</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>1995</td>
<td>2001</td>
<td>3.2 or 2.3</td>
<td>Assigned</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not assigned</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>2004</td>
<td>2.1</td>
<td>Assigned</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not assigned</td>
<td>0</td>
<td>11</td>
</tr>
</tbody>
</table>

1 According to scale reading

Table 6.8 Summary of the number of returns that were clipped and unclipped, and the number of each that could be assigned to females that returned to Girnock Burn.

<table>
<thead>
<tr>
<th></th>
<th>Total # offspring clipped</th>
<th>Total # offspring unclipped</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total assigned</td>
<td>39</td>
<td>22</td>
<td>61</td>
</tr>
<tr>
<td>Total not assigned</td>
<td>5</td>
<td>26</td>
<td>31</td>
</tr>
<tr>
<td>Totals</td>
<td>44</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

The level of homing according to clipped and unclipped returns was 48% (44/92), while the estimate based solely on assignment was 66% (61/92). This latter estimate did not include five individuals that were clipped (and thus presumably philopatric) but nevertheless were not assigned by DNA profiling. This is most likely explained by incorrect scale reading. In Chapter 3, section 3.3, the error rate associated with aging adult returns by scale reading was estimated at 9%, not significantly different from that identified here. Thus accounting for scale reading error, by excluding eight individuals from the total (i.e. c.9% of 92) provides an adjusted level of homing of c.73% (61/84). Therefore the percentage of returns to the burn that would be considered ‘strays’ was approximately 27%, defining
strayers as returns that were not offspring of females that had entered the traps. Thus parentage analysis of sea returns allowed a more detailed analysis of the origin of strayers than was possible solely with clipping data. This has enabled a more biologically relevant estimate of straying to be estimated, rather than one based on an artificial (trap) boundary.

### 6.3.3 Identification of potential strayers from population genotype data.

Individuals that were likely to have strayed from different ‘populations’ were identified using assignment methods and resampling techniques implemented in GENECLASS 2. Outputs of these assignment analyses are presented in Table 6.9. Only individuals identified as strayers with a stringent probability of a false positive ≤0.001 are included.

<table>
<thead>
<tr>
<th>Spawning years included</th>
<th># loci utilized</th>
<th># samples</th>
<th>First generation migrants</th>
<th>Non-assignment to Girnock stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td></td>
<td></td>
<td>g00-013 P&lt;0.001</td>
<td>g00-013 P=0.001</td>
</tr>
<tr>
<td>1995</td>
<td></td>
<td></td>
<td>g00-065 P&lt;0.001</td>
<td>g00-065 P&lt;0.001</td>
</tr>
<tr>
<td>2000</td>
<td>9</td>
<td>458</td>
<td>g04-088 P&lt;0.001</td>
<td>g04-088 P&lt;0.001</td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td>g04-101 P=0.001</td>
<td>g04-101 P=0.001</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td>g00-013 P&lt;0.001</td>
<td>g00-013 P&lt;0.001</td>
</tr>
<tr>
<td>2000</td>
<td>12</td>
<td>247</td>
<td>g00-013 P&lt;0.001</td>
<td>g00-013 P&lt;0.001</td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td>g00-065 P&lt;0.001</td>
<td>g00-065 P&lt;0.001</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td>g00-070 P&lt;0.001</td>
<td>g00-070 P=0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>g04-088 P&lt;0.001</td>
<td>g04-088 P=0.001</td>
</tr>
</tbody>
</table>

1 Individual identification, g- Girnock, followed by the year of sampling (e.g. 00, in 2000). Final three digit number is the individual sample number.

No adult returns of 1991, 1995 or 2001 were identified as potential strayers under the criteria used. As one dataset was restricted in terms of the number of loci
examined (to nine rather than twelve loci) and the other dataset was restricted in
terms of the year classes examined (it did not include samples from 1991 or 1995),
only individuals that were identified (as ‘migrant’ or ‘not self-assigned’) in both
datasets were examined further. These comprised three individuals: g00-013, g00-
065 and g04-088. These three individuals were all males and were not fin clipped.
According to scale reading they were aged 3.1, 2.1, and 2.1 years, respectively.
Parentage analysis of g00-013 did identify a potential female parent for this
individual from the 1995 year class. Inspection of the composite genotype for g00-
013 revealed scoring for a unique allele at locus Ome9. (This allele had not been
found in any other adult return in 1991, 1995, 2000, 2001 or 2004). The individual
also possessed one rare allele at each of three other loci. It is possible that mature
parr harbour alleles that are rare or not present in the adult samples (although
perhaps unlikely). It is also possible that the identified strayer was spawned below
the traps, the sire being an anadromous strayer. Similar parentage analyses of g00-
065 and g04-088 did not identify possible dams / sires. However, in the case of g00-
065 (presumed age 2.1) this would not necessarily provide evidence for the fish
being a strayer, since its probable parental spawning year class (1996) was not
screened.
Overall the data suggest that there is not a great influx of strayers from populations
that differ widely at the loci tested.

6.3.4 Reproductive success of strayers identified by clipping and parentage
assignment
Sufficient numbers of offspring were identified among the datasets to allow a
comparison of the reproductive success between clipped and unclipped fish as well
as between philopatric fish and ‘strayers’ up to smolt and parr stage. Results are depicted in Table 6.10 and Table 6.11.

Table 6.10 Comparison of the reproductive success of clipped and unclipped spawners in three spawning years. Data analysed by exact test for contingency tables. (Female spawners noted as part or fully-spent on entry to traps were not included)

<table>
<thead>
<tr>
<th>Year of return of spawners</th>
<th>Sex of spawner</th>
<th>Offspring sample</th>
<th># spawners</th>
<th># offspring assigned to clipped and unclipped spawners</th>
<th>Exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Clipped</td>
<td>Unclipped</td>
<td>Clipped</td>
</tr>
<tr>
<td>1995 F</td>
<td>1997 parr</td>
<td>22</td>
<td>40</td>
<td>87</td>
<td>177</td>
</tr>
<tr>
<td>1995 M</td>
<td>1997 parr</td>
<td>21</td>
<td>22</td>
<td>84</td>
<td>68</td>
</tr>
<tr>
<td>2000 F</td>
<td>2004 smolts</td>
<td>6</td>
<td>0</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>2000 M</td>
<td>2004 smolts</td>
<td>9</td>
<td>13</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>2001 F</td>
<td>2004 smolts</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>74</td>
</tr>
<tr>
<td>2001 M</td>
<td>2004 smolts</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>34</td>
</tr>
</tbody>
</table>

In general there was no significant difference in the relative numbers of offspring from clipped and unclipped parents. However, this was not the case for female spawners in 2001, which was due to the particularly high output (41 smolts) for a single (unclipped, multi-sea-winter) female. However the p-value of 0.03 was not significant following sequential Bonferroni correction for multiple tests. This
female was identified as philopatric from assignment tests. There was no significant
difference in the relative numbers of offspring from philopatric and non-philopatric
spawners. This suggests the reproductive success of philopatric and non-philopatric
individuals did not differ, at least in the dataset studied.

A measure of the lifetime reproductive success of spawners would need to take into
account survival at sea. Reproductive success up to the stage of adult spawners
returning to Girnock Burn was explored between unclipped and clipped spawners
and philopatric and non-philopatric spawners. The parentage of sea returns was
determined and the success of clipped / unclipped spawners and philopatric/non-
philopatric spawners was compared. The sexes were examined separately.

Data for reproductive success up to adult return of clipped vs. unclipped spawners
are given in Table 6.12 and Table 6.13. The relative numbers of offspring from
clipped & unclipped parents returning to the burn did not differ significantly.

Table 6.12 Comparison of the reproductive success (up to adult return stage) of clipped and
unclipped female spawners. Data analysed by exact test for contingency tables. Female
spawners that were noted as part or fully-spent on entry to the traps were excluded.

<table>
<thead>
<tr>
<th>Year of return of spawners</th>
<th>Year of return of offspring</th>
<th># spawners</th>
<th># offspring assigned to clipped and unclipped spawners</th>
<th>Exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>1995</td>
<td>12</td>
<td>2</td>
<td>1.00</td>
</tr>
<tr>
<td>1995</td>
<td>1999 2000 2001</td>
<td>22</td>
<td>13</td>
<td>0.41</td>
</tr>
<tr>
<td>2000</td>
<td>2004</td>
<td>6</td>
<td>1</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Table 6.13 Comparison of the reproductive success (up to adult return stage) of clipped and unclipped male spawners. Data analysed by exact test for contingency tables.

<table>
<thead>
<tr>
<th>Year of return of spawners</th>
<th>Year of return of offspring</th>
<th># spawners</th>
<th># offspring assigned to clipped and unclipped spawners</th>
<th>Exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Clipped</td>
<td>Unclipped</td>
<td>Clipped</td>
</tr>
<tr>
<td>1991</td>
<td>1995</td>
<td>11</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>1995</td>
<td>1999 2000 2001</td>
<td>21</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>2000</td>
<td>2004</td>
<td>9</td>
<td>13</td>
<td>1</td>
</tr>
</tbody>
</table>

A similar analysis was undertaken between philopatric and non-philopatric spawners (identified by genotyping/parentage analysis, Table 6.14). Unfortunately this was only possible for male spawners. Only a small subset of spawners in 1995 was of the correct age (2.1) to be assigned to previous spawners (in 1991). Of those assigned only one was female, and this was clipped. In 2000, all females involved in natural spawning were clipped. Again there were no obvious indications of differential reproductive success (up to the stage of anadromous return) of philopatric/non-philopatric spawners.

Table 6.14 Comparison of the reproductive success (up to adult return stage) of philopatric and non-philopatric male spawners. Data analysed by exact test for contingency tables.

<table>
<thead>
<tr>
<th>Year of return of spawners</th>
<th>Year of return of offspring</th>
<th># spawners</th>
<th># offspring assigned to philopatric and non-philopatric spawners</th>
<th>Exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Philopatric</td>
<td>Non-philopatric</td>
<td>Philopatric</td>
</tr>
<tr>
<td>1995</td>
<td>1999 2000 2001</td>
<td>26</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>2000</td>
<td>2004</td>
<td>12</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>

6.3.5 Reproductive success of strayers identified by the analysis of genotypes

The previous estimates of the reproductive success of strayers did not make any distinction between strayers that may have come from just below the traps and those
that may have originated in different populations. The reproductive success of individuals identified as strayers using GENECLASS 2, through the analysis of multilocus genotypes, was therefore assessed.

Of the three individuals that were identified as potential strayers by GENECLASS 2 analysis, one was a return in 2004; no offspring were sampled from this year class. The remaining two returned to Girnock Burn in 2000 and only one of these was placed above the traps to spawn naturally. This successfully reproduced and it was represented in the smolt sample of 2004.
6.4 **Discussion**

6.4.1 **Extent of trap leakage**

Previous studies that have evaluated the efficacy of smolt traps have involved capture mark-recapture methods. Cheng and Gallinat (2004) estimated the efficiency of a rotary screw trap (designed only to capture a sample of downstream migrating smolts) and Cunjak and Therrien (1998) estimated the efficiency of a counting-fence (designed to trap all downstream migrating smolts). Cunjak and Therrien (1998) estimated the trap efficiency to be between 70% and 87% for smolts and between 65% and 82% for 1+ parr (maximum and minimum values over five years, for Atlantic salmon). These values were obtained by capturing a sample of smolts or parr at the counting fence, marking and releasing them above the counting fence, followed by re-capture (of a percentage) at the counting fence. Individuals not re-caught were assumed to have bypassed the trap. The values obtained were higher than those found in this study (a mean level of trap efficiency of 63%) for smolts and parr combined. This was probably due to the different river system and trapping facilities involved. A limitation of the capture-mark-recapture method is that the destination of the marked uncaptured fish is unknown- they might not actually reach the trap. In addition it only provides an estimate of the efficiency of the trap for the duration of the experiment, whereas the method employed in the current study provides an estimate of efficiency over the entire period of use.

Potential limitations of the genetic profiling method used in the current study were 1) the inability to identify reliably offspring that may have resulted from spawning below the trap and 2) the number of adult returns is limited in comparison with the number of smolts or parr available in a particular year (capture mark-recapture methods would involve smolts or parr, whereas the DNA profiling method relied on
adult returns). In addition a low degree of error may have been introduced through genotyping error or mutation (see Chapter 3). Particular attention was paid to females that may have spawned below the traps. An estimate of the extent of trap leakage was made excluding female spawners that had been noted as part-spent. A second estimate was made based on a subset of families where known spawning was confined to the upper reaches of the burn only, which should have further purged the dataset of offspring spawned below the traps. The values of leakage obtained were similar, 34% (obtained from data from three spawning years) in the former case and 37% (obtained from data from spawning in 1995) in the latter.

During the period of study the traps were cleared of debris by staff based approximately two hours from the traps at Girnock Burn. The traps are presently monitored by a person living in closer proximity to the traps and are cleared more often. However, the downstream trap is not cleared throughout the night. At Girnock Burn factors that tend to promote downstream migration are darkness and high flow (Youngson et al. 1994). Unfortunately these are also the conditions under which the downstream trap is more likely to clog. Therefore the estimate of trap leakage derived in this study may still reflect current levels of leakage.

The results suggested that the extent of trap leakage might differ between fish that left freshwater in 1998 and 1999, although the numbers involved were too low to determine this for the more conservative estimate of trap leakage (that involved solely offspring of females that were likely to have spawned only above the traps). Possible reasons for such differences include varying degrees of monitoring of the trap and varying periods of high water flow. However, the estimate for the rate of trap leakage for fish that left freshwater in 1998 was based on fish that spent two years in freshwater, and that for 1999 was based on fish that spent three years in
Trap leakage and the extent of straying

Chapter 6

freshwater. Although both years may have included fish that passed through (or over) the traps as parr in the autumn migration, those that passed through as smolts were likely to have differed in size, with 3 year old smolts being slightly larger on average (based on data obtained from smolts that passed through the traps in 2004, (Figure 6.2).

![Figure 6.2 Length of 2 and 3 year old smolts that passed through the traps in spring 2004.](image)

Smolts passing through the traps in 1998 would have been two years old and probably slightly smaller on average than those passing through as three year old smolts in 1999. It is possible that the difference in size of smolts may have affected trap efficiency. The extent of trap leakage was higher in 1998, when two-year old (smaller) smolts were sampled. Cheng and Gallinat (2004) suggested that larger fish might have been more manoeuvrable than small fish under strong water flows and so were less likely to be caught by a rotary screw trap. It may be that fish ‘prefer’ to
travel down into the Wolf trap than over it in high water flow, which could result in larger fish being more prone to capture. However, more data are needed to determine whether there is a difference in the level of trap leakage between years, what factors influence it and in which direction.

One of the difficulties involved in estimating trap leakage was the possibility that females had spawned prior to entering the traps. In 2005 no natural spawning took place above the traps. All offspring of this year class that originate above the traps underwent quasi-natural spawning, and so all crosses were of known pairings. Thus offspring of 2005 that return to the burn unclipped and are assigned to pairings known not to have been set up in quasi-natural spawning must originate below the traps. With no natural spawning above the traps the identification of unclipped returns that have leaked over the trap should therefore be slightly improved.

### 6.4.2 Revised rate of homing

The revised rate of homing was c.73%, which took into account fish that were unclipped but could be assigned to females that had reached the traps. This clearly shows that presence of an intact adipose fin is not a reliable indicator of straying in Girnock Burn, since only 48% of returns were clipped. The high rate of homing concurs with findings of sub-structuring of populations of Atlantic salmon (that have not been subjected to stocking) within rivers by, for example, Garant *et al.* (2001) and Primmer *et al.* (2006).

A question then arises concerning the origin of the apparent strayers to the region of the burn above the trap. If the boundary of the homing unit were the tributary, and the 27% of returns that were strayers originated in Girnock Burn below the traps, they would have been spawned in 900m of tributary. Considering the length of the burn above the traps is 8km this seems unlikely. However, it is dependent on the
productivity of the habitat within the 900m below the burn, which is difficult to predict as although there is more spawning substrate per km, due to shading from trees the growth rate of fry and parr is reduced (Iain McLaren, pers. comm.).

Alternative origins of the strays are the main stem of the river, other tributaries of the Dee, or even further afield. In order to estimate the relative contributions from each of the above, much more extensive sampling and genetic screening of the adult returns in these regions would be required followed by assignment tests to identify origins of putative strayers.

Leider (1989) studied the effect of the eruption of a volcano (Mount St Helens, Washington State) on the straying of steelhead trout. It was found that adult sea returns originating in tributaries severely affected by the eruption tended to return to other tributaries upstream, rather than downstream, of their natal river. If the same behaviour were applicable to Atlantic salmon in the Dee, then strays that enter Girnock Burn because they are prevented from entering their natal tributary (as might occur if water flow is low) may originate downstream rather than upstream of Girnock Burn.

Although it was not possible to determine the origin of all unclipped returns, parentage analysis of sea returns has allowed a more detailed examination of their origin than was possible solely with clipping data.

### 6.4.3 Identification of strayers through the analysis of population genotype data

Three adult returns (in five years of returns) were identified as unlikely to have originated from the population in Girnock Burn, based on genotype data. All were unclipped, as expected and were male grilse. Previous studies on straying in salmonid fishes have found male biased straying (Jonsson et al. 2003; Hard and
Trap leakage and the extent of straying  

Heard 1999), although others have found no detectable difference between the sexes (Thedinga et al. 2000; Hansen et al. 2001b). The influence of sea-age on straying has also been found to differ among studies, being more likely for individuals spending more time at sea (Jonsson et al. 2003 in Atlantic salmon), or less likely (Hard and Heard, 1999 in chinook salmon).

A better way of identifying strayers from different populations would be to sample other regions from where fish might stray, such as other tributaries and rivers. This is because there is more power in determining first generation migrants when all source populations are sampled (Piry et al. 2004). However, in reality this would be extremely costly.

6.4.4 Reproductive success of strayers that were identified through clipping, parentage assignment and the analysis of multilocus genotypes

Comparing the reproductive success (up to parr and smolt stage) of clipped and unclipped returns showed there was no significant difference between the groups. This is in concordance with the findings of Taggart et al. (2001). In order to distinguish between unclipped returns that had leaked over the trap or were spawned by females prior to entry to the traps, and unclipped returns that were not offspring of returns to the burn (specifically returns to the trap), additional comparisons were carried out. These treated unclipped returns that had been assigned to previous returns to the trap as ‘homers’ (i.e. philopatric). No significant differences in the reproductive success between the groups were found. However, not all age classes of unclipped returns were tested for philopatry through parentage analysis. For example unclipped returns in 1995 aged 3.2 would have been spawned in 1989, and would therefore need to be checked against returns in this year before philopatric state could be determined. Although the comparisons went some way towards
accommodating trap leakage and females that returned to the trap but spawned prior to entry, not all cases were likely to have been identified.

Clipped and unclipped returns, philopatric and non-philopatric returns were successful in producing offspring that returned as anadromous adults. This suggests that at least some of the offspring of strayers were adequately adapted to the environmental conditions to survive and return to spawn.

Previous studies have suggested that strayers in the wild tend to have a lower reproductive success than philopatric individuals (Hendry et al. 2004b; Tallman and Healey 1994). In the comparisons above the artificial boundary delineated by the traps was used to distinguish strayers and homers (with some degree of overlap introduced by identifying offspring of individuals that may have spawned below the traps prior to entry). In order to examine the reproductive success of strayers that were more likely to be from further afield, and therefore biologically defined strayers, individuals were identified that were not likely to have originated in the population in Girnock Burn (through the analysis of multilocus genotypes). Few strayers were identified in this manner, but one individual that was identified reproduced by natural spawning and its offspring survived (at least) to the smolt stage. This indicates that gene flow to the stock in Girnock Burn has occurred. This supposed gene flow may be important in maintaining genetic diversity in the Girnock Burn stock, though it could be argued to be disadvantageous in possibly disrupting adaptive allele complexes. The impact is dependent on the adaptive similarities of the strayer to those in Girnock Burn, the selection pressures experienced, and on the effective population size of the recipient population. In a smaller population straying may be more important in maintaining genetic diversity.
6.4.5 Implications for conservation

At Girnock Burn an important aspect of the current management of the stock is quasi-natural spawning. To date anadromous returns have been chosen for broodstock irrespective of their straying status- clipped or unclipped, genetically similar or dissimilar to the remaining stock. This would eliminate pre-zygotic mechanisms that might prevent strayers from reproducing. In addition the representation of returns used in quasi-natural spawning was found to be higher than that in natural spawning (for example, 100% of female spawners employed in quasi-natural spawning in 2000 were represented in offspring, while only 71% of female spawners that underwent natural spawning in 1995 were represented in offspring (Chapter 4). Thus it is possible that the reproductive success of strayers may be changed through quasi-natural spawning, although it is not known whether this would have any impact on the long-term survival of the stock. The effect of quasi-natural spawning will be difficult to examine due to the lack of knowledge on the origin of the anadromous adults that return to the burn and on the level of gene flow that occurs under natural circumstances. However, a very low number of returns were identified (through the analysis of population genotype data) as being unlikely to have originated from the Girnock Burn, and one of these reproduced naturally. Thus it is possible that quasi-natural spawning will not have a large impact on the reproductive success of ‘strayers’.

Consuegra et al. (2005) and Fraser et al. (2004) investigated the direction and magnitude of gene flow among populations of brook charr (Salvelinus fontinalis) and Atlantic salmon, respectively. Both studies uncovered complex systems, with asymmetric gene flow. In a wider context, these studies bring to attention the interactions of fish stocks within a watershed. For example if a management scheme
were to increase the output of a particular population or part of a population (albeit using native broodstock) would this tend to increase the number of strayers to other populations? Would this result in the disruption of adaptive allele complexes in the recipient populations? Answering these questions was beyond the scope of this study. However, an initial examination of the extent of homing (accounting for trap leakage) and possible genetic input of strayers was performed.
7 Mature parr; their reproductive success and potential impact on adaptation.

7.1 Introduction

7.1.1 Mature male parr

Alternative reproductive tactics (ARTs) in fish are common and include variation in mate choice and parental care (Avise et al. 2002). One such ART is ‘cuckoldry’ (also described as parasitism or sneaking), whereby fertilisations by secondary males are attempted by darting from hidden positions near or within the nest. In this way fertilisations may be ‘stolen’ from ‘fighter’ or ‘bourgeois’ males that fight for access to females or defend a territory. In the European bitterling (Rhodeus sericeus) males exhibit these two ARTs, either defending a territory– the ‘bourgeois’ male, or sneaking fertilisations. However, they are neither genetically nor morphologically different (Smith and Reichard 2005) and territorial males may also sneak fertilisations (Reichard et al. 2004). In other species of fish the distinction is more clear; in the ocellated wrasse (Symphodus ocellatus) and the bluegill sunfish (Lepomis macrochirus) the sneaking tactic is also exhibited, but performed by males that are smaller than the bourgeois males (Avise et al. 2002; Neff 2004).

In Atlantic salmon the ‘sneaking’ tactic is also adopted by males that are considerably smaller than bourgeois males; these are often less than a hundredth of the weight (Fleming 1996) and take the form of mature male parr. These are also known as precocious parr. In sockeye and coho salmon sneakers also occur, although they are termed ‘jacks’ (Fleming and Reynolds 2004).

Thus in Atlantic salmon, mature parr have the potential to fertilise eggs, without having undergone smoltification or experiencing the marine environment. The proportion of
Atlantic salmon parr that mature sexually has been found to vary with latitude. Valiente et al. (2005) reported that the proportion of mature parr decreased with increasing latitude, in populations on both the east and west of the Atlantic, over the entire (natural) range. A likely reason for the higher proportion of parr maturing in southern populations is the higher temperature, resulting in a higher growth opportunity (Martinez et al. 2000; Garcia-Vazquez et al. 2001). The impact of mature parr may be of particular importance in the future in relation to predicted climate change. Increases in freshwater temperature have been predicted across the United States (Mohseni et al. 1999). In the UK the annual average temperature has been predicted to increase by 2-3.5°C by 2080 (Hulme et al. 2002), and there is generally a close relationship between air and surface water temperature (Morrill et al. 2005). In Girnock Burn a rise in maximum temperature during spring and winter of approximately 2°C was seen between 1968 and 1997 (Langan et al. 2001). Considering the observed relationship between latitude and the proportion of parr that mature, it can be tentatively predicted that the proportion of parr maturing sexually may rise as the temperature and, with it, growth opportunity increase. Knowledge of the impacts of mature parr on populations is therefore required in order to predict some of the possible impacts of climate change on wild Atlantic salmon populations.

Much interest has been focused on the factors causing individual male parr to mature prematurely in freshwater. A number of studies have established a link between the physiology of parr and those that go on to mature as parr; early maturity has been associated with (increased) mesenteric fat levels (Rowe et al. 1991), growth rate
Mature parr; their reproductive success and potential impact on adaptation

(Berglund 1992) and size (Dalley et al. 1983; Prévost et al. 1992; Aubin-Horth and Dodson 2004).

Development of sexual maturity as parr can be thought of as a tactic adopted as a result of a conditional strategy, whereby above a certain threshold individual parr will mature in freshwater (Gross and Repka 1998; Aubin-Horth and Dodson 2004). Findings of Aubin-Horth and Dodson (2004) and Baum et al. (2004) suggest that the threshold varies even between sites within the same river system and Baum et al. (2005) proposed that the threshold may be modified by the environment.

7.1.2 Spawning success of mature parr

By comparison with the level of attention awarded to the factors influencing parr maturity, fewer studies have investigated the collective reproductive success of mature parr in the wild. Difficulties are encountered in studying spawning success through observation, particularly in the wild. However, the development of molecular markers has, comparatively recently, enabled the quantification of spawning success to be attempted. Early allozyme work by Hutchings and Myers (1988) revealed that mature parr were capable of fertilising eggs, albeit in an artificial setting (tanks). Later work by Jordan and Youngson (1992), also with allozymes, showed that mature parr contributed to spawning under more natural conditions, although the precise contribution was not quantified due to the low polymorphism of the marker used. Factors that may affect the proportion of eggs fertilised by mature parr include availability of mature parr, sperm competition, female mate choice, heterogeneity of substrate and the sex ratio of anadromous spawners. Thomaz (1995) found that in an artificial setting containing one anadromous female, one anadromous male and between
one and 12 mature parr, the proportion of eggs sired by mature parr was lower in the
second nest built. This was thought to be due to milt exhaustion of the mature parr, and
it was suggested that this may not relate to the situation in the wild, where there is
likely to be an over supply of mature parr. However, as the proportion of parr that
mature in populations varies greatly, perhaps there are situations where few males
mature as parr and thus the total proportion of eggs they fertilise may be limited by milt
exhaustion.

Gage et al. (1995) conducted an investigation, which suggested that the properties of
sperm produced by mature parr might influence fertilisation success. After comparing
various characteristics of sperm from male mature parr and anadromous males, the
authors reported that the mature male parr sperm were more motile and lived longer
than sperm from anadromous males, which could therefore be advantageous in gaining
fertilisations.

Another factor that may influence the reproductive success of mature parr is female
mate choice. Female mate choice is exhibited in sockeye salmon, Oncorhynchus nerka
(Burgner 1991) and at spawning time female coho salmon (Oncorhynchus kisutch)
show aggression towards males other than the dominant male (Sandercock 1991).

Maekawa et al. (1993) reported that anadromous female Dolly Varden (Salvelinus
malma) showed aggression towards sneaker (stream-resident) males but never towards
anadromous males. They also found that attempts from stream-resident males to
fertilise eggs tended to result in a decrease in the number of times that the female
released eggs. One reason for such behaviour might have been the consumption of eggs
by stream-resident males, which had been observed in the closely related Miyabe charr
(Salvelinus malma miyabbei). However, this had not been observed in the population of Dolly Varden under study, leading the authors to suppose that this was not a factor. Maekawa et al. (1993) suggested that females might exhibit such behaviour in order to promote anadromy and fitness in offspring, although this would be dependent on anadromy and fitness being heritable characters. In Atlantic salmon Thomaz (1995) reported that anadromous females were aggressive towards mature male parr and it was previously thought that females would not spawn in the presence of mature parr without an attending anadromous male (Caswell et al. 1984), although this has since been observed (e.g. Jones and Hutchings 2001). It may be that the degree of aggression directed towards mature parr by females affects their collective reproductive success. Jones and Hutchings (2001) postulated that with a more heterogeneous substrate mature parr would be more likely to be able to sneak fertilisations un-noticed. Thus mature parr may be more successful in obtaining fertilisations in some natural situations rather than artificial environments with less debris within which to hide. Jones and Hutchings (2001) found that the collective reproductive success of mature parr decreased with increasing competition from anadromous males. In a comparison of different studies described in the literature (see Table 7.1) this relationship was not apparent. However, many different factors could be influencing the collective success of mature parr so the studies are unlikely to be directly comparable. Nevertheless, Table 7.1 illustrates the paucity of information available on the proportion of offspring sired by mature parr in a natural setting, the low numbers of studies that examined large numbers of anadromous spawners and the narrow range of the degree of competition from anadromous males that has been observed.
Table 7.1 Comparison of studies on Atlantic salmon spawning that have determined the proportion of offspring sired by mature parr.

<table>
<thead>
<tr>
<th>Type of setting</th>
<th># anadromous females</th>
<th># anadromous males</th>
<th>% potential anadromous spawners that were male</th>
<th># mature parr</th>
<th>% offspring sired by mature parr</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>100</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>1</td>
<td>33</td>
<td>10-23</td>
<td>52.5</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>1</td>
<td>33</td>
<td>7</td>
<td>75.8</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>N</td>
<td>66</td>
<td>43</td>
<td>39</td>
<td>?</td>
<td>50</td>
<td>J Taggart, pers. comm.</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>4</td>
<td>40</td>
<td>?</td>
<td>~11</td>
<td>Jordan and Youngson 1992</td>
</tr>
<tr>
<td>N</td>
<td>55</td>
<td>39</td>
<td>41</td>
<td>?</td>
<td>51</td>
<td>J Taggart, pers. comm.</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>5</td>
<td>4.7</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>1</td>
<td>~5</td>
<td>Hutchings and Myers 1988</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>20</td>
<td>~20</td>
<td>Hutchings and Myers 1988</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>20</td>
<td>23</td>
<td>Jones and Hutchings 2002</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>1</td>
<td>25</td>
<td>Moran et al. 1996</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>10</td>
<td>25.9</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>2</td>
<td>26</td>
<td>Thomaz et al. 1997</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>20</td>
<td>29.9</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>20</td>
<td>37.1</td>
<td>Jones and Hutchings 2002</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>12</td>
<td>39</td>
<td>Moran et al. 1996</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>12</td>
<td>40</td>
<td>Thomaz et al. 1997</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>10</td>
<td>55</td>
<td>Jones and Hutchings 2001</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>6</td>
<td>89</td>
<td>Moran et al. 1996</td>
</tr>
<tr>
<td>N</td>
<td>34</td>
<td>38</td>
<td>53</td>
<td>?</td>
<td>52</td>
<td>J Taggart, pers. comm.</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>7</td>
<td>54</td>
<td>~100</td>
<td>65.1</td>
<td>Martinez et al. 2000</td>
</tr>
</tbody>
</table>

1 A- artificial, N- natural setting
2 updated from Taggart et al. 2001

Jones and Hutchings (2001) concluded that a greater collective reproductive success of mature parr was associated with lower competition from anadromous males. However, this was based on an experimental study investigating only three levels of competition,
involving; i) one female, one anadromous male and 5 to 20 mature parr, ii) two females, one anadromous male and 7 to 23 mature parr, iii) one female, no anadromous male and five mature parr. Without the inclusion data from this third situation (involving no anadromous male) the relationship was not significant. The study was also conducted under artificial conditions. The relationship was not tested, and therefore may not hold, under more complex natural situations.

7.1.3 Genetic diversity and mature parr
Given the potential success of mature parr in fertilising eggs, considerable attention has been directed towards their influence on effective population size. L’Abée-Lund (1989) gave an estimate of the effective population size of a small population of Atlantic salmon in Norway, excluding and including mature parr. However, these estimates were rather crude as they were based solely on the numbers of anadromous males, anadromous females and mature male parr in the system. The authors were able to count the fish in the system as it had been treated with rotenone, in order to kill the salmon and eliminate *Gyrodactylus salaris*. The estimates of the effective population size took into account unequal sex ratio and overlapping generations, but neither variance in reproductive success nor fluctuations in population size were considered. They concluded that mature parr *theoretically* increased the effective population size above the recommended minimum of 50 in the population under study. Moran and Garcia-Vazquez (1998) found multiple paternity of eggs from a single female. Eggs were sampled from a redd (constructed under natural conditions) and parentage was inferred using single-locus minisatellite markers. Interpretation of the egg genotypes showed they all could have originated from a single female, and a minimum of six
males. They suggested this multiple paternity could be a mechanism for increasing the effective population size and thus maintaining genetic diversity, particularly in small populations. However, the authors were not able to determine the identity of males (whether anadromous males or mature parr). More thorough analyses of the impact of mature parr on the effective population size have been reported by Martinez et al. (2000), Garcia-Vazques et al. (2001), Jones and Hutchings (2001) and Jones and Hutchings (2002), again using genetic analyses. Garcia-Vazquez et al. (2001) demonstrated in a semi-artificial setting (a 650m long section of a stream containing mature male parr into which seven anadromous males and six anadromous females were released) that mature parr contributed to spawning and related this to the survival of small populations in southern Europe. They estimated the sex ratio and proportion of between cohort pairings in Atlantic salmon from five rivers in southern Europe under two scenarios, i) 0% of embryos sired by mature male parr ii) 60% of embryos fertilised by mature parr. In the populations under study the sex ratios of anadromous fish were heavily biased towards females, thus the mature parr input resulted in an evening of the sex ratio. A lower rate of between-cohort pairing was also predicted, which the authors suggested would result in a decrease in inbreeding. The authors concluded that as a result of these mechanisms, the effective population size would be increased. Jones and Hutchings (2001) quantified the effect of mature parr input on effective population size by identifying individual mature parr and accounting for variance in reproductive success. This was conducted in an artificial setting, which therefore enabled samples to be obtained from the mature parr involved, and subsequent identification of parentage using microsatellite markers. They concluded...
that the increase in effective population size due to mature parr would be most
pronounced when numbers of anadromous males was low, because they found an
increased proportion of eggs fertilised by mature parr when the intensity of competition
from anadromous males was reduced.
Valiente et al. (2005) found no relationship between conservation status and genetic
diversity. Conservation status was based on the number of adult individuals, including
for example a category for stable populations and another for declining populations.
Valiente et al. (2005) thought the relationship might be attributed to mature parr
maintaining genetic diversity in populations with a low conservation status. Consuegra
et al. (2005) also found genetic diversity to be higher than expected in small
populations, but attributed this to gene flow (due to anadromous adults rather than
mature parr).
The effective population size is concerned with the rate of loss of genetic diversity due
to stochastic forces, rather than systematic forces such as selection (Wang 2005). The
impact of mature parr on effective population size has already been examined, as
described above. However, the environment experienced by mature male parr and
anadromous males differs and mature parr are able to reproduce without having
experienced the marine environment. To my knowledge the effect of reproductively
successful mature parr on the propagation of alleles under selection has not been
investigated.

7.1.4 Consequences of the mature parr ART for adaptation to the environment
Under certain circumstances it may be that reproductively successful mature parr
propagate alleles that confer adaptation to the freshwater environment, but which are
Mature parr; their reproductive success and potential impact on adaptation

not favourable for anadromy. This hypothesis is based on the ability of mature parr to reproduce without having experienced the marine environment, smoltification, or the return migration. The life stages involved are depicted in Figure 7.1.

Figure 7.1 Life history of Atlantic salmon illustrating the two main ARTs of males. Dark arrows show where alleles conferring adaptation to the freshwater environment may be propagated. Not all stages will be present in all populations.

There is more likely to be a difference in adaptation to the conditions encountered by mature parr and anadromous males if the selection intensity is strong and differs between the environments encountered, if the different tactics have a heritable component and mature parr do not later migrate to sea and return as anadromous males. Female spawning may act to prevent the propagation of alleles conferring an advantage
in freshwater, as in the vast majority of populations they undergo migration to sea prior to spawning; females very rarely mature as parr (Esteve 2005). However, if selection intensity is high then females carrying alleles that confer a disadvantage for anadromy may have a higher mortality. Alternatively, alleles may be propagated mainly in the male lineage if they are located near to the sex-determining region, or may only be expressed in males.

One of the factors making differential adaptation more likely is heritability of parr maturity. A heritable component to parr maturity in Atlantic salmon was found by Thorpe et al. (1983), Gjerde (1984) and Duston et al. (2005). Thorpe et al. (1983) described how the incidence of male parr maturity was significantly higher among offspring of mature parr aged 1+, than among offspring of anadromous males aged a total of two or three years. Gjerde (1984) constructed six artificial crosses between fish of different ages, including females aged 4 or 5 years, anadromous males aged 4 or 5 years and 1 year old mature male parr. Ten males (of a particular age) and 10 females (of a particular age) were used in each cross. Among the resultant offspring mature male parr only developed in the progeny of crosses that were initially constructed with mature male parr. Duston et al. (2005) found that male offspring of mature parr had higher rates of maturity as parr than male offspring of anadromous males. In this study the anadromous males were MSW fish, although their precise age was not given.

Gjerde (1984) stated that the heritability of parr maturity had not been estimated, and to the writer’s knowledge this is still the case. Faster growth of offspring of mature parr (compared to offspring of anadromous males) was found prior to exogenous feeding by Garant et al. (2002). Thus, it is plausible that there may be a genetic distinction
between mature parr and anadromous males, allowing selection for different characteristics.

Another factor that makes differential adaptation more likely is strong selection pressures that differ between anadromous and freshwater-resident fish. Anadromous fish are subject to factors that cause mortality on the migration to and from the sea, and during the period in the marine environment, to which mature parr are not subjected. Stresses encountered include osmoregulation during the transition between fresh and saltwater (and back to freshwater), energy expenditure associated with migration (throughout the journey, but perhaps particularly on the upstream migration when no feeding occurs and delays may transpire due to environmental conditions) and predation (specifically for anadromous fish as they may encounter more predators and be less able to avoid them) (Hendry et al. 2004a). In addition, a different array of pathogens will be encountered in the marine environment. Mature parr resident in freshwater are free from these additional stresses and therefore might be less constrained in evolving adaptations specific to the freshwater environment. Indeed, differences have been reported between sympatric populations of nonanadromous and anadromous fish of the same species, which were likely to be adaptations to the different life histories (anadromy and nonanadromy). For example, Birt et al. (1991) reared progeny of sympatric anadromous and non-anadromous Atlantic salmon in a common environment. They found differences in parameters associated with parr-smolt transformation, including; seasonal levels of body silvering, gill Na^+-K^+ ATPase activity and size and number of gill chloride cells. These differences were suggested to have a genetic basis as the fish were reared in a common environment and would be
consistent with the nonanadromous fish being in the process of dispensing ‘with smolting changes associated with a switch to hypoosmoregularity mechanisms’. Wood and Foote (1996) found morphological and developmental differences between two sympatric populations of sockeye salmon *Oncorhynchus nerka*, one nonanadromous and the other anadromous. The authors found differences in gill raker number (a heritable character) and growth rate when reared under controlled conditions.

A specific factor that may cause mortality, and therefore be a selection pressure, in anadromous Atlantic salmon is high river temperature on the return migration. High temperatures and low flow have been identified as a problem in Devon for Atlantic salmon entering estuaries, associated with delayed or failed entry (Solomon and Sambrook 2004). In the Aberdeenshire Dee, Smith and Hawkins (1995, cited in Solomon and Sambrook 2004), found that migrating salmon might be subject to a difference in temperature of up to 10°C over a few metres (due to stratification of saline and fresh water of different temperatures). In addition sea surface temperature in spring has been linked to survival and growth rate of Atlantic salmon post-smolts, both in Europe and North America (Friedland *et al.* 2000, 2003, 2005) although the relationship is complex and not necessarily direct. Another factor that may affect mortality is growth rate. Handeland *et al.* (2003) found that fish that had been selected for a higher growth rate were less able to tolerate seawater.

When considering what role mature parr might play in the adaptive response of the population, it is genetically determined traits under selection that are important. Unwin *et al.* (2003) found family-specific survival rates at sea of chinook salmon, suggesting a ‘considerable genetic influence over survival’.
7.1.5 Issues investigated

The content of this chapter is partly concerned with exploring the role of sexually mature parr in the adaptive response of populations to the environment. Much attention has been directed towards the influence of mature parr on effective population size (and this aspect was covered in Chapter 4). However, to the writer’s knowledge the way in which mature male parr may affect (specifically) the propagation of alleles that are under selection has not been considered. Mature parr are able to reproduce after only experiencing freshwater conditions, which may lead to the propagation of alleles that confer adaptation to freshwater conditions but not necessarily conditions experienced by anadromous fish.

An additional aspect of male parr maturity investigated was the relationship between the sex ratio of anadromous spawners and the collective reproductive success of mature parr. Although there have been a large number of studies on the proportion of parr that mature as parr (e.g. Baum et al. 2004; Valiente et al. 2005), fewer have investigated the proportion of eggs sired by mature male parr, particularly in the wild. From a semi-natural setting Jones and Hutchings (2001) concluded that a greater collective reproductive success of mature parr was associated with lower competition from anadromous males. One aim of the research described in this chapter was to provide an insight into the way in which the sex ratio of anadromous spawners may affect the collective reproductive success of mature parr in the wild.

7.1.6 Mature parr at Girnock Burn

Girnock Burn was the study site for all the issues examined. The proportion of male parr that mature in freshwater (as parr) in Girnock Burn has not been quantified.
However mature male parr are present in large numbers in the system. They have been recorded after being trapped (in the downstream trap) during the autumn. Buck and Youngson (1982) suggest their emigration at this time could be a means of maximising the chances of encountering anadromous females. Comparing data collected over an eleven year period, the authors found that the date on which mature male parr were first found in the trap (in autumn) was significantly correlated with (and generally succeeded) the first date that anadromous females were placed above the traps. In addition, in 1978 no anadromous fish were placed above the traps and in this year substantially more mature parr were trapped in the downstream trap than previously.

Buck and Youngson (1982) suggest that in ‘normal’ years the presence of anadromous returns acts to prevent emigration of mature parr, but that the presence of adult females may also result in an increase in general activity, giving rise to the observed correlation between the onset of mature parr emigration and first date of female entry to the burn.

Male parr may initially mature in freshwater aged 1+. An added complexity is that males that do mature in freshwater may go on to smolt and return to the burn as anadromous males. This has been deduced from microtagged returns (Iain McLaren, personal communication). However, the prevalence of this has not been quantified.

Previous studies found the proportion of eggs presumed to be fertilised by mature parr to be between 50 and 52% (based on data in Taggart et al. 2001), and 10.8% (95% CI 5.2-22.5%) (Jordan and Youngson 1992). The latter value may not have been a true representation of the situation throughout the burn as only a small section of the burn was studied (80mX6m); a total of six female and four anadromous males were released into this section (and prevented from moving upstream or downstream), while parr
Mature parr; their reproductive success and potential impact on adaptation

4. Mature parr; their reproductive success and potential impact on adaptation

Chapter 7

were able move into and out of the study area on the downstream side (but not the upstream side). In addition, the estimate was not precise as only a single diallelic allozyme marker was used. This allowed positive identification of progeny that were sired by mature parr in some cases (depending on the genotype), but sire type could not be distinguished in others and so was estimated based on the frequency of one of the alleles in mature male parr sampled from the study site.

7.1.7 Objectives

Four main questions were investigated. Although it is very difficult to identify selection signals, particularly in a complex natural environment, relevant data were available from the present study and therefore explored. All issues were investigated by studying the stock of Atlantic salmon in Girnock Burn, as outlined below:

1. Does the sex ratio of anadromous adults influence the collective reproductive success of mature parr?

With larger numbers of anadromous males (compared to the number of anadromous females) the collective reproductive success of mature parr may be lowered due to fewer sneaking opportunities brought about by higher levels of aggression from anadromous males.

This issue was approached by comparing the reproductive success of mature parr in years with different sex ratios of anadromous adults. Both new and existing data were employed. Data were already available on the sex ratio of anadromous returns and subsequent reproductive success of mature parr for natural spawning that occurred in 1991, 1992 and 1995. In addition new data were obtained on the reproductive success
of mature parr for natural spawning that occurred in 2000 and 2001. In these five
spawning years the sex ratio of potential anadromous spawners (male:female) ranged
from 0.65-2.75.

2. Do mature parr propagate alleles that are advantageous in the freshwater
   environment? Specifically, is the survival of offspring of mature parr greater than
   the survival of offspring of anadromous males in the freshwater environment?
Mature parr are able to complete their life cycle in the freshwater environment leading
to the possibility that they propagate alleles that are advantageous in the freshwater
environment. Anadromous males must survive both marine and freshwater
environments, in addition to upstream and downstream migrations. Thus anadromous
males may tend to propagate alleles that do not confer such an advantage in the
freshwater environment as those propagated by mature parr.
The survival of offspring of the two sire types in freshwater was compared by sampling
cohorts at two stages in freshwater. Paternity was determined (mature parr or
anadromous male) and the proportion sired by mature parr in the two stages (of each
cohort) compared.

3. Do mature parr propagate alleles that are disadvantageous in the marine
   environment? Specifically, in the marine environment, is the survival of offspring
   of mature parr lower than the survival of offspring of anadromous males?
Mature parr are able to reproduce without experiencing the marine environment, or
downstream and upstream migrations. This may release the selection pressure on
certain alleles and lead to the propagation of alleles that do not confer as great a level of adaptation to the marine environment as those propagated by anadromous males.

An initial investigation into the survival of offspring of the two types of sire in the marine environment was made by comparing the proportion of smolts to the proportion of adult sea returns that were sired by mature parr.

4. Over the lifespan of anadromous fish, is there any evidence for differential mortality of mature parr offspring vs. anadromous parr offspring?

A preliminary exploration into the overall survival (to anadromous return stage) of mature parr offspring and anadromous male offspring was made. The proportion of eggs sired by mature parr was compared to the proportion of anadromous returns sired by mature parr.
7.2 Methods

7.2.1 Sampling

Details of the samples utilised are provided in Table 7.2, including the year of sampling, number and life stage sampled, and the loci examined.

Table 7.2 Details of samples studied.

<table>
<thead>
<tr>
<th>Year of sampling</th>
<th># samples</th>
<th># loci examined</th>
<th>Life stage</th>
<th>Loci examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>95</td>
<td>9-11(^1)</td>
<td>Adult return</td>
<td>Set I</td>
</tr>
<tr>
<td>1995</td>
<td>116</td>
<td>9-11(^1)</td>
<td>Adult return</td>
<td>Set I</td>
</tr>
<tr>
<td>1999</td>
<td>10</td>
<td>12</td>
<td>Adult return</td>
<td>Set II</td>
</tr>
<tr>
<td>2000</td>
<td>89</td>
<td>12</td>
<td>Adult return</td>
<td>Set II</td>
</tr>
<tr>
<td>2001</td>
<td>41</td>
<td>12</td>
<td>Adult return</td>
<td>Set II</td>
</tr>
<tr>
<td>2004</td>
<td>117</td>
<td>12</td>
<td>Adult return</td>
<td>Set II</td>
</tr>
<tr>
<td>1997</td>
<td>274</td>
<td>9-11(^1)</td>
<td>Parr</td>
<td>Set II</td>
</tr>
<tr>
<td>2004</td>
<td>407</td>
<td>12</td>
<td>Smolts</td>
<td>Set II</td>
</tr>
</tbody>
</table>

\(^1\) see Material and Methods (Chapter 2) for details

\(^2\) Set I = 2210 1605 2201 410 One9 407 202 171 197 (404 408)

Set II = 2210 1605 2201 410 One9 407 202 171 197 404 408 401

Adults were sampled after being caught in the upstream trap. Parr were sampled from the length of Girnock Burn in September 1997 by electrofishing, as detailed in Chapter 2. Smolts were sampled between February and May 2004 in the downstream trap, which is designed to trap all fish moving downstream. A tissue sample was taken from one in every five smolts caught. Tissue samples for extraction of DNA consisted of adipose fin clips (parr and smolts) and a punch or sliver of tissue removed from the adipose fin of anadromous returns. Details of DNA extraction, genotyping and parentage analysis, in addition to a diagram illustrating the relationship between samples are provided in Chapter 2. However, a summary of the parentage analyses
performed is given in Table 7.3, which shows five sets of putative parents and offspring.

Table 7.3 Study samples relating parent and offspring relationships.

<table>
<thead>
<tr>
<th>Putative Parents</th>
<th>Putative Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Year sampled</td>
<td>Life Stage</td>
</tr>
<tr>
<td>1991</td>
<td>Adult return</td>
</tr>
<tr>
<td>1995</td>
<td>Adult return</td>
</tr>
<tr>
<td>1995</td>
<td>Adult return</td>
</tr>
<tr>
<td>2000</td>
<td>Adult return</td>
</tr>
<tr>
<td>2001</td>
<td>Adult return</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data were also included that were already available from a previous study by Taggart et al. (2001). Samples that were involved are outlined in Table 7.4

Table 7.4 Samples included from a previous study. Details of DNA extraction and genotyping were described in Taggart et al. (2001).

<table>
<thead>
<tr>
<th>Year of sampling</th>
<th># samples</th>
<th># minisatellite loci examined</th>
<th>Life stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>94</td>
<td>7</td>
<td>Adult return</td>
</tr>
<tr>
<td>1992</td>
<td>72^1</td>
<td>7</td>
<td>Adult return</td>
</tr>
<tr>
<td>1995</td>
<td>109^1</td>
<td>10</td>
<td>Adult return</td>
</tr>
<tr>
<td>1991</td>
<td>574</td>
<td>7</td>
<td>Eggs</td>
</tr>
<tr>
<td>1992</td>
<td>469</td>
<td>7</td>
<td>Eggs</td>
</tr>
<tr>
<td>1995</td>
<td>1075</td>
<td>10</td>
<td>Eggs</td>
</tr>
</tbody>
</table>

^1 number of adult returns placed above trap, as opposed to the number initially caught in the upstream trap
7.2.2 Spawning success of mature parr; effects of differing sex ratios of anadromous returns

The first aim was to investigate the possibility of a link between the sex ratio of anadromous adults and the collective reproductive success of mature parr. It was hypothesized that with fewer anadromous males compared to the number of anadromous females, the higher the collective reproductive success of mature parr may be (assuming similar numbers of mature male parr in each year).

The sex ratio of potential anadromous spawners (male:female) ranged from 0.65 to 2.75 in the years under study (Table 7.5). Natural spawning that took place in 1991, 1992, 1995, 2000 and 2001 was examined.

Table 7.5 Datasets used to investigate spawning success of mature parr under differing sex ratios of anadromous spawners.

<table>
<thead>
<tr>
<th>Year of spawning</th>
<th>Sex ratio (male:female)</th>
<th># anadromous males</th>
<th># anadromous females</th>
<th>Offspring sampled (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.71</td>
<td>39</td>
<td>55</td>
<td>Eggs (574)</td>
</tr>
<tr>
<td>1992</td>
<td>1.12</td>
<td>38</td>
<td>34</td>
<td>Eggs (469)</td>
</tr>
<tr>
<td>1995</td>
<td>0.65</td>
<td>43</td>
<td>66</td>
<td>Eggs (1075)</td>
</tr>
<tr>
<td>2000</td>
<td>2.75</td>
<td>22</td>
<td>8</td>
<td>Smolt (407)</td>
</tr>
<tr>
<td>2001</td>
<td>0.89</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

1 # potential anadromous male spawners; # potential anadromous female spawners

2 Total number of offspring sampled, not all were the correct age to be assigned to spawners in 1995, 2000 and 2001

Ideally the spawning success of mature parr would have been estimated from the proportion of eggs they fertilised. However, only smolt samples were available from the (natural) spawning that occurred in 2000 and 2001.

Linear regression was used to test for a relationship between the sex ratio of anadromous spawners and the percentage of offspring sired by mature parr.
7.2.3 Survival of offspring from alternative sire types in freshwater

In order to compare the survival of mature parr offspring in freshwater to the survival of anadromous male offspring in freshwater, estimates of the proportion of fish sired by each type were obtained at two points in time. This was examined for a total of three cohorts. The first cohort examined consisted of progeny of fish that spawned in 1995, which all spawned naturally. An estimate of the proportion of eggs produced by each sire type was available from the study conducted by Taggart et al. (2001), from a sample of 1075 eggs. A subsequent estimate was made of the proportion of parr aged 1+ (sampled in September 1997) that were sired by mature parr and anadromous males.

Two further cohorts were examined, which were offspring of quasi-natural spawning. The initial proportion of eggs fertilised by mature parr and anadromous males was known from the crosses that had been set up. In 2000 two sets of crosses were constructed. Set A consisted of nine females. The eggs from each female were divided into 21 roughly equal portions and each portion mixed with milt from a single male. Fourteen portions were each mixed with milt from different anadromous males, and 7 portions with milt from different mature parr. Set B consisted of five females, each crossed with 18 anadromous males and three mature parr. Crosses were constructed in the same manner as for set A. The final cohort examined consisted of offspring of spawners in 2001. In this year only one set of crosses was constructed; eggs from each of 12 females were divided into 11 roughly equal portions and each portion mixed with milt from a single male. Seven anadromous males and four mature parr were used.

Following the construction of crosses in 2000 and 2001 milt from an extra ‘insurance’ male was added. In 2001 the insurance male was a mature male parr, in 2000 this was also ‘likely’ to be the case (no written record of the type of male was made, although
Mature parr; their reproductive success and potential impact on adaptation

this information was obtained from a person present at the time). A subsequent estimate of the proportion of smolts that were sired by mature parr was made.

In order to compare the proportions of eggs and parr or smolts sired by mature parr, 95% confidence intervals of the proportions were calculated. If the 95% confidence intervals overlapped then proportions were considered not to be significantly different.

The standard error of each proportion was calculated as:

\[ SEP = \sqrt{(pq)/n} \]

SEP = standard error of a proportion
p = proportion of offspring sired by mature parr
q = 1 - proportion of offspring sired by mature parr
n = number of offspring

The 95% confidence interval was then calculated as the proportion ±1.96 X SEP. In situations with low numbers of samples (where nXp or nXq were lower than five), confidence intervals were calculated using the exact probabilities of the Binomial distribution, as suggested by Bland (2000). These were calculated using an MSDOS program (biconf), obtained from [www.sghms.ac.uk/depts/phs/staff/jmb/](http://www.sghms.ac.uk/depts/phs/staff/jmb/).

An exact test for contingency tables implemented in ‘struc’ (Raymond and Rousset 1995) was employed in some cases in order to clarify (or otherwise) results.

7.2.4 Survival of offspring from alternative sire types in the marine environment

A preliminary investigation into the survival in the marine environment of progeny of mature parr compared to the survival of progeny of anadromous males was conducted. This was achieved by determining the sire type (anadromous male or mature parr) of smolts in addition to the sire type of sea returns in a single cohort. The progeny of fish
Mature parr; their reproductive success and potential impact on adaptation

that spawned in 2000 were investigated. This included offspring originating from both quasi-natural and natural spawning. Smolts and adult returns were sampled in 2004 (those assigned to spawners in 2000 were aged three years in total; either three year old smolts or 2.1 grilse). It was therefore assumed that offspring of anadromous males and mature parr were equally likely to spend three years in freshwater, or to spend two years in freshwater and one at sea (as described in the discussion, this may not be a valid assumption).

Proportions of offspring sired by mature parr that were sampled as smolts, and as adult returns, were compared as in the previous section (7.2.3).

7.2.5 Overall Survival

The overall survival, from egg stage to anadromous return, of mature parr offspring and anadromous male offspring was compared. The sire type of eggs and anadromous returns of two cohorts were examined (Table 7.6).

<table>
<thead>
<tr>
<th>Spawning year</th>
<th># Eggs sampled</th>
<th>Year of return of anadromous offspring sampled (#) (^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>574</td>
<td>1995 (116)</td>
</tr>
<tr>
<td>1995</td>
<td>1075</td>
<td>1999 (10), 2000 (89), 2001 (41)</td>
</tr>
</tbody>
</table>

\(^1\)Numbers of anadromous offspring assigned to spawners in 1991 and 1995 are provided in results

The proportion of anadromous returns sired by mature parr was compared to the proportion of eggs sired by mature parr as in section 7.2.3.
7.3 **Results**

In all the following results when paternity was determined, anadromous males were positively assigned. In cases where no anadromous male was assigned (but an anadromous female was assigned), paternity was attributed to mature parr. A summary of the outcomes of parentage analyses performed is given in Table 7.7, which shows the five sets of putative parents and offspring.

<table>
<thead>
<tr>
<th>Year sampled</th>
<th>Life Stage</th>
<th># sampled</th>
<th>Year sampled</th>
<th>Life Stage</th>
<th># assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>Adult return</td>
<td>95</td>
<td>1995</td>
<td>Adult return</td>
<td>13</td>
</tr>
<tr>
<td>1995</td>
<td>Adult return</td>
<td>116</td>
<td>1997</td>
<td>Parr</td>
<td>269</td>
</tr>
<tr>
<td>1999</td>
<td>Adult return</td>
<td>6</td>
<td>2000</td>
<td>Adult return</td>
<td>35</td>
</tr>
<tr>
<td>2000</td>
<td>Adult return</td>
<td>89</td>
<td>2001</td>
<td>Adult return</td>
<td>10</td>
</tr>
<tr>
<td>2001</td>
<td>Adult return</td>
<td>41</td>
<td>2004</td>
<td>Adult return</td>
<td>13</td>
</tr>
</tbody>
</table>

### 7.3.1 Spawning success of mature parr; effects of differing sex ratios of anadromous returns

It was postulated that the collective reproductive success of mature parr might depend on the sex ratio of anadromous males to females, with more anadromous males per female resulting in a decrease in the success of mature parr. No significant relationship was seen between sex ratio of anadromous spawners and the percentage of offspring sired by mature parr (Pearson’s $r=-0.75$, $P=0.14$)(Table 7.8, Figure 7.2).
Table 7.8 Reproductive success of mature parr with differing sex ratios of anadromous spawners.

<table>
<thead>
<tr>
<th>Year of spawning</th>
<th>Sex ratio (male:female)</th>
<th># anadromous male spawners</th>
<th># female spawners</th>
<th>% offspring sired by mature parr</th>
<th># sired by mature parr</th>
<th># sired by anadromous males</th>
<th>Offspring sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.71</td>
<td>39</td>
<td>55</td>
<td>51</td>
<td>293(^1)</td>
<td>281(^1)</td>
<td>eggs</td>
</tr>
<tr>
<td>1992</td>
<td>1.12</td>
<td>38</td>
<td>34</td>
<td>52(^1)</td>
<td>243(^1)</td>
<td>226(^1)</td>
<td>eggs</td>
</tr>
<tr>
<td>1995</td>
<td>0.65</td>
<td>43</td>
<td>66</td>
<td>50(^1)</td>
<td>533(^1)</td>
<td>542(^1)</td>
<td>eggs</td>
</tr>
<tr>
<td>2000</td>
<td>2.75</td>
<td>22</td>
<td>8</td>
<td>43</td>
<td>23</td>
<td>30</td>
<td>smolt</td>
</tr>
<tr>
<td>2001</td>
<td>0.89</td>
<td>8</td>
<td>9</td>
<td>58</td>
<td>46</td>
<td>34</td>
<td>smolt</td>
</tr>
</tbody>
</table>

\(^1\) Data from J Taggart, personal communication

Figure 7.2 The percentage of offspring sired by mature parr compared to the sex ratio of potential anadromous spawners. Data represents different spawning years, indicated by labels.

The year in which the highest percentage of offspring was sired by mature parr was 2001, when there was also the lowest absolute number of anadromous males. In 2000...
the lowest percentage of offspring were sired by mature parr, and this was also the year with the highest sex ratio (male:female) of anadromous spawners.

The reproductive success of mature parr ranged between 43% and 58%. This suggests that the sneaking tactic of mature parr was successful even in the presence of large numbers of anadromous males.

7.3.2 Survival of offspring from alternative sire types in freshwater

Comparison of the proportion of eggs and subsequent life stages (parr or smolts) sired by mature parr in different years and in different sets of quasi-natural spawning showed that there was no statistically significant difference between the proportions (Figure 7.3). Results are only comparable within years and for spawning in 2000, only within set A and within set B. Details are shown in Table 7.9.
Mature parr; their reproductive success and potential impact on adaptation

Figure 7.3 A comparison of the proportion of offspring sired by mature parr at different freshwater life stage; eggs compared to parr and eggs compared to smolts. Offspring of three spawning years were examined; 1995, 2000 and 2001. 95% confidence intervals are shown.

* Two sets of crosses were constructed in quasi-natural spawning in 2000, described in 7.2.3, 2000A refers to set A, 2000B refers to set B.
Table 7.9 Initial percentage of eggs found to be sired by mature parr, and the percentage of subsequent life stages (parr or smolts) that were sired by mature parr. Eggs and offspring from one female in 2001 were discounted as records on the crosses constructed with it were incomplete.

<table>
<thead>
<tr>
<th>Year of spawning</th>
<th>%age eggs sired by mature parr</th>
<th># eggs sired by mature male parr</th>
<th># eggs sired by anadromous males</th>
<th>Subsequent life stage sampled</th>
<th>%age subsequent life stage sired by mature parr</th>
<th># subsequent life stage sired by mature parr</th>
<th># subsequent life stage sired by anadromous males</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>50</td>
<td>533&lt;sup&gt;1&lt;/sup&gt;</td>
<td>542&lt;sup&gt;1&lt;/sup&gt;</td>
<td>parr</td>
<td>43</td>
<td>117</td>
<td>152</td>
</tr>
<tr>
<td>2000 set A</td>
<td>33</td>
<td>c. 11658&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>c. 23315&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>smolts</td>
<td>35</td>
<td>34</td>
<td>64</td>
</tr>
<tr>
<td>2000 set B</td>
<td>14</td>
<td>c. 2225&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>c. 13349&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>smolts</td>
<td>7</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>2001</td>
<td>36</td>
<td>c. 19731&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>c. 34528&lt;sup&gt;φ&lt;/sup&gt;</td>
<td>smolts</td>
<td>32</td>
<td>35</td>
<td>73</td>
</tr>
</tbody>
</table>

<sup>1</sup> Results from Taggart et al. (2001)

<sup>φ</sup> Estimated from the number of eggs used in quasi-natural spawning. Eggs from each female were counted; approximately 33% in 2000 set A, 14% in 2000 set B and 36% in 2001 were fertilised with milt from mature parr.

7.3.3 Survival of offspring from alternative sire types in the marine environment

Numbers of smolts and anadromous returns assigned to spawners in 2000 are given in Table 7.10

Table 7.10 Outcome of parentage analysis for offspring from quasi-natural and natural spawning in 2000.

<table>
<thead>
<tr>
<th>Spawning year&lt;sup&gt;1&lt;/sup&gt;</th>
<th># smolts assigned</th>
<th># anadromous returns from 2004 assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000A</td>
<td>98</td>
<td>6</td>
</tr>
<tr>
<td>2000B</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>2000N</td>
<td>53</td>
<td>6</td>
</tr>
</tbody>
</table>

<sup>1</sup> 2000A- set A of quasi-natural spawning in 2000

2000B- set B of quasi-natural spawning in 2000

2000N- natural spawning in 2000
There was no significant difference between the proportion of anadromous returns that were sired by mature parr and the proportion of smolts sired by mature parr in either natural spawning, or quasi-natural spawning that that occurred in 2000 (Figure 7.4). Set B of quasi-natural spawning was not depicted graphically due to the very low number of anadromous returns assigned (1).

![Proportion of offspring sired by mature parr](image)

**Figure 7.4** Comparison of the proportion of offspring sired by mature parr prior and subsequent to a period in the marine environment; smolts compared to anadromous returns. 95% confidence intervals are shown. Both female and male offspring are included.

2000A - offspring of ‘set A’ of quasi-natural spawning that was constructed in 2000
2000N - offspring of natural spawning that occurred in 2000

Numbers of smolts and anadromous adults assigned to each category are shown in Table 7.11, where the data are further broken down into male and female anadromous adults.
Table 7.11 Proportion (with numbers in parentheses) of female and male sea returns (that could be assigned to previous spawners at Girnock Burn) that were sired by mature parr. The percentage of earlier life stages that were sired by mature parr is also given.

<table>
<thead>
<tr>
<th>Details of spawners (parents)</th>
<th>Details of offspring (smolts and anadromous returns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category of spawning in 2000</td>
<td>Proportion of smolts sired by mature parr (#)</td>
</tr>
<tr>
<td></td>
<td>Proportion anadromous returns sired by mature parr (#)</td>
</tr>
<tr>
<td>2000A</td>
<td>0.35(34/98)</td>
</tr>
<tr>
<td>2000B</td>
<td>0.80(3/41)</td>
</tr>
<tr>
<td>2000N</td>
<td>0.43(23/53)</td>
</tr>
</tbody>
</table>


2000B- offspring of set B of quasi-natural spawning in 2000

2000N- offspring of natural spawning in 2000
7.3.4 Overall survival of offspring from alternative sire types in the marine environment

Numbers of eggs and anadromous returns that could be assigned to spawners in 1991 and 1995 are given in Table 7.12.

Table 7.12 Outcome of parentage analysis.

<table>
<thead>
<tr>
<th>Spawning year</th>
<th># eggs assigned</th>
<th>Year of return of anadromous offspring assigned (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>Eggs (574)</td>
<td>1995 (13)</td>
</tr>
</tbody>
</table>

There was a significant difference in the proportion of eggs and anadromous returns that were sired by mature parr in 1991, but not in 1995 (Figure 7.5, details in Table 7.13). This was clarified by an exact test for contingency tables, with a significant difference for 1991 (P=0.01), but not for 1995 (P=0.09). However, both point estimates of the proportion of anadromous returns sired by mature parr were lower than the estimates of the proportion of eggs sired by mature parr.
Figure 7.5 Proportion of eggs sired by mature parr in comparison to the proportion of anadromous adults sired by mature parr.

*‘adult’ refers to the life stage of the offspring sampled, i.e anadromous adult return. For example the stage classified ‘1991 adult’ relates to offspring of natural spawning that occurred in 1991 that have survived one year in the marine environment and returned to Girnock Burn in 1995 as anadromous adults.
Table 7.13 Proportion (with numbers in parentheses) of female and male sea returns (that could be assigned to previous spawners at Girnock Burn) that were sired by mature parr. The proportion of earlier life stages that were sired by mature parr is also given.

<table>
<thead>
<tr>
<th>Spawning year</th>
<th>Proportion of eggs sired by mature parr (#)</th>
<th>Year of return of anadromous offspring</th>
<th>Proportion female returns sired by mature parr (#)</th>
<th>Proportion male returns sired by mature parr (#)</th>
<th>Proportion male and female returns sired by mature parr (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.51(293/574)</td>
<td>1995</td>
<td>0.00(0/1)</td>
<td>0.17(2/12)</td>
<td>0.15(2/13)</td>
</tr>
<tr>
<td>1995</td>
<td>0.50(533/1075)</td>
<td>1999, 2000, 2001</td>
<td>0.46(11/24)</td>
<td>0.30(8/27)</td>
<td>0.37(19/51)</td>
</tr>
</tbody>
</table>
Chapter 7

7.4 Discussion

The current study was conducted in the setting of the complex natural environment. As such only strong selection signals were likely to have been detected. This was particularly true when exploring survival in the marine environment and the ‘overall’ survival, due to limited numbers of anadromous returns that could be assigned to the specific year classes investigated.

7.4.1 Reproductive success of mature parr with differing sex ratios of anadromous spawners

Jones and Hutchings (2001) found that the overall reproductive success of mature parr was higher when there was less competition from anadromous males i.e. the sex ratio of anadromous spawners (male:female) was higher. However, this was based on findings from an artificial set up involving a limited number of spawners; a single anadromous male, between one and two anadromous females and five to 23 mature male parr. In addition, the relationship was not statistically significant when the data point corresponding to 100% fertilisation of eggs by mature parr (when no anadromous male was present) was omitted. Despite this, it does seem theoretically plausible that mature parr spawning success would be reduced by aggression from anadromous males, and that such aggression may be more apparent when more anadromous males are present.

In the current study, conducted in a natural setting, no statistically significant relationship was found between the sex ratio of anadromous spawners and the reproductive success of mature parr. It may be that such a relationship does exist but was not identified due to the small number of data points (five). Alternatively, it may be that in the wild, the situation is more complex and the sex ratio of anadromous spawners is perhaps not a good proxy for the intensity of anadromous male
Mature parr; their reproductive success and potential impact on adaptation

Chapter 7

competition. The actual sex ratio on spawning grounds, or the operational sex ratio may have differed somewhat to the sex ratio calculated from the total number of anadromous spawners released above the traps in a single spawning year. Females generally spawn within 7-10 days, while anadromous males may remain active on the spawning grounds for a month or more (Fleming 1996), therefore the operational sex ratio was likely to have been biased more towards males. However, it might have been that when the number of anadromous returns was low the operational sex ratio was closer to that calculated from the overall number of anadromous returns placed above the traps. Fewer spawners would have limited the maximum number of males per female and perhaps made it more unlikely for the longer active reproductive period of males to alter the operational sex ratio due to a lower density throughout the burn. In 2001 the number of anadromous spawners released above the traps to spawn naturally was low, amounting to eight males and nine females. The reproductive success of mature parr was also highest in this year. Perhaps in 2001 the operational and calculated sex ratio were closer than in other years with larger numbers of returns. The numbers of returns allowed to spawn naturally in this year were manipulated; the remaining returns were utilised in quasi-natural spawning. However, such low numbers of returns do occur in the wild. For example, the number of female returns to the river Mandeo in northwest Spain varied between two and thirty-two between 1984 and 1996 (Martinez et al. 2000). It may therefore be that the sex ratio of anadromous spawners has more influence on the collective reproductive success of mature parr in populations such as this, where the number of anadromous returns is low. Thus the relationship
reported by Jones and Hutchings (2001) is perhaps only applicable to populations with low numbers of returns.

There is indirect evidence that the relationship may also be applicable to steelhead trout (*Oncorhynchus mykiss*). Ardren and Kapuscinski (2003) found that the ratio of effective population size (estimated using a temporal method) to the number of anadromous spawners was higher in years with low numbers of returns, which they termed ‘genetic compensation’. Seamons *et al.* (2004) suggested this might be due to increased collective reproductive success of mature parr when numbers of returns were low. Assuming this to be the case, and considering the findings of Jones and Hutchings *et al.* (2001) (that increased competition from anadromous males results in a lower input from mature parr), genetic compensation would only be seen if females tended to outnumber males. Anadromous females outnumbered anadromous males in nine of the ten years that Ardren and Kapuscinski (2003) reported genetic compensation. Thus it is plausible that mature parr could be the cause of this ‘genetic compensation,’ although other factors such as reduced redd superimposition or a greater representation of all anadromous spawners in offspring could also have contributed. This could have implications in Girnock Burn if the numbers of anadromous adults returning to Girnock Burn were to fall dramatically. In years with sex ratios skewed towards females but with high overall numbers of anadromous returns, the operational sex ratio may remain skewed towards males. However, in years with sex ratios skewed towards females with *low* numbers of returns the operational sex ratio may also be skewed towards females possibly resulting in an increase in the collective reproductive success of mature parr in natural spawning (above that seen with high overall numbers of returns).
Another factor that may have further complicated the relationship between the operational and calculated sex ratios in the study described here was the release date of spawners: anadromous fish were generally not released above the traps on a single day. In 1995 they were released between the 16th October and 16th November and the sex ratio (male:female) was not constant throughout this period. For example between the 18th October and 25th October there were equal numbers of male and female anadromous spawners above the traps, although between the 25th October and 16th November the sex ratio varied but was always less than one. Thus the sex ratio calculated from the total numbers of anadromous adults released above the traps was really only applicable after the final anadromous spawner had been released above the traps. However, in 2001 they were released on a single day, which again may have resulted in the operational and calculated sex ratios being more similar in this year. It may also have been more similar in 2000 as all anadromous returns were released above the traps within a three-day period.

In order to confirm (or refute) these predictions regarding operational sex ratios and calculated sex ratios, additional studies on the collective reproductive success of mature parr with high and low numbers of anadromous returns and a range of sex ratios could be carried out. It would be beneficial in such studies to release all anadromous spawners simultaneously and to have equal numbers of mature parr.

In the current study, different life stages were sampled (eggs and smolts) in order to estimate the collective reproductive success of mature parr. Ideally eggs would have been sampled, but these were not available for spawning in 2000 or 2001. Utilizing smolts in these years could have been problematic had the proportion of smolts sired by
mature parr not reflected the proportion of eggs sired by mature parr (i.e. the spawning success). The proportion of eggs and smolts sired by mature parr for quasi-natural spawning in 2000 and 2001 was compared in section 7.3.2. Results indicated no significant difference in either year.

7.4.2 Survival of offspring from alternative sire types in freshwater

There was no evidence that the survival of mature parr offspring was greater than the survival of anadromous male offspring in this population in the years studied. These data suggest that, at least in the stock examined, the offspring of mature parr were not significantly better adapted to the freshwater environment than the offspring of anadromous males.

There were both advantages and disadvantages to sampling parr and smolts for respective estimates of progeny survival from 1995 and 2000/2001 spawners. Sampling parr enabled a cross section of the entire cohort to be sampled, but allowed the survival to be monitored over a shorter period of time. Conversely only a single age class of smolts were sampled from each year (two year old smolts from spawning in 2001 and three year old smolts from spawning in 2000), thus only a portion of the whole cohort was sampled. As such, samples of two or three year old smolts would not have included male parr that matured at two or three years, respectively. Since male parr maturity is known to have a heritable component (Thorpe et al. 1983; Gjerde 1984; Duston et al. 2005), this could have affected the proportion of smolts sampled that were sired by mature parr. As sampling of parr occurred prior to migration it may have provided a more accurate estimate of the proportion of the entire cohort that was sired by mature parr at this stage.
There are a number of reasons that could explain why no evidence was found suggesting that mature parr offspring had a greater survival rate than anadromous male offspring in this population in the years studied. Firstly, there may not be a genetic distinction between mature parr and anadromous males. The heritability of the mature parr tactic may be masked by environmental influences on its expression, such that there is little difference in the proportion of male offspring of mature parr and anadromous males that mature as parr. Secondly, mature parr are able to smolt in following years and return to spawn as anadromous males. The degree to which this occurs is unknown, but it may be to the extent that precludes detection of any genetic differences between mature parr and anadromous males. Thirdly, selection pressures on resident and anadromous fish may not be adequately divergent to result in selection for different characters. It may also be that there are many genes involved with the expression of characters that may be under selection, thus reducing the degree of selection on each. Another possibility is that relevant selection pressures are not constant through time.

Despite there being no suggestion that mature parr offspring survive better in freshwater than anadromous male offspring in the study population, this may not be applicable to other populations or other species that exhibit similar life history strategies.

### 7.4.3 Survival of offspring from alternative sire types in the marine environment

It was postulated that the offspring of mature parr might have a lower survival rate in the marine environment, than the offspring of mature parr, because mature parr reproduce without having undergone selection in the marine environment. No
significant difference in the proportion of smolts and anadromous returns sired by mature parr in 2000 was seen. However, in both comparable sets of samples where more than one anadromous fish was sampled, the point estimate of the proportion of adult returns sired by mature parr was lower than that for the proportion of smolts sired by mature parr. Considering the low numbers of returns examined, this could warrant further study.

A potential problem with the comparison made was that both smolts and anadromous returns were sampled in 2004. Although originating from the same spawning year (2000), they adopted different life history tactics. The smolts sampled were aged three years, whereas the adult returns underwent smoltification aged two years. It may be that the proportion of three year old smolts sired by mature parr differed to the proportion of two year old smolts sired by mature parr—due to maturation of male parr and smoltification (in a particular year) being mutually exclusive events. If this were the case, the samples would not be strictly comparable.

An additional problem is that, had the proportion of anadromous adults sired by mature parr been significantly lower than the proportion of smolts sired by mature parr, it would not have been possible to determine the cause of this. As hypothesised, it could be due to the expression of alleles, in mature parr offspring, that conferred a disadvantage in the marine environment. However, it could also be attributed to male parr maturity being a heritable character, resulting in male offspring of mature parr tending to also become mature parr. Previous studies on Atlantic salmon have identified lower survival of male parr that matured as parr, both in freshwater (Myers 1984, from interpreting numbers of mature male parr, female parr, immature parr, male
smolts and female smolts in different years), and in the marine environment (Lundqvist et al. 1994, from tagging). In order to distinguish between the two scenarios, knowledge on the life histories of anadromous returns (whether or not they matured as parr) would be required.

Although no significant difference was found in the current investigation, numbers of anadromous returns were low which reduced the power of the study. It may therefore be worth investigating this in a different population with higher numbers of returns in each year. It would be advantageous to obtain data on the sire type of a particular set (or sets) of fish prior to and following a period in the sea in order to eliminate the problem of different life history tactics being followed. For example, two year old smolts could be sampled, and subsequently, after a period at sea the same set could be sampled as anadromous fish.

7.4.4 Overall survival of offspring from alternative sire types in the marine environment

A significant difference in the proportion of eggs and anadromous returns sired by mature parr was seen in 1995, but not 1991. However, in both years, the point estimate of the proportion of adult returns sired by mature parr was lower than that for the proportion of eggs sired by mature parr.

In this comparison, as with previous comparisons, it was not possible to distinguish the effect of possible mortality due to maturation as parr and possible mortality due to proposed selection acting on alleles propagated by the two sire types. However, the effect of either factor on the samples studied was not dramatic as not every anadromous return was sired by an anadromous male; mature parr did successfully sire anadromous returns.
7.4.5 Future directions

Knowledge of the impact of mature parr spawning and subsequent recruitment on the success of anadromy is important in predicting the consequences of using mature parr in quasi-natural spawning, and the possible effects of envisaged global warming. However, there are also other important issues that need to be considered in predicting the effect of current and future changes, as a range of scenarios can be predicted at the present time.

One reason for the decrease in abundance of Atlantic salmon is thought to be a decline in marine survival (Hawkins 2000; Potter and Crozier 2000; Middlemas et al. 2003). It can be postulated that a reduction in marine survival could result in an increase in the benefits of parr maturity. In addition, if freshwater temperature rises, the proportion of parr that mature in a given population is likely to increase. Perhaps Atlantic salmon will come to follow life history strategies similar to that seen in some sea trout populations, whereby males are primarily freshwater resident, while the females are anadromous (Klemetsen et al. 2003). Other possible outcomes are that growth opportunity in freshwater will increase to allow maturation of females in freshwater, resulting in ‘landlocked’ populations, or that anadromy will become too costly and populations die out due to a lack of anadromous females. At the present time these scenarios are highly speculative. Key questions that would assist in determining the role of mature parr in the future of Atlantic salmon populations are:

1. Would an increase in freshwater temperature result in an increase in the proportion of parr maturing?
2. Has there been an increase in the proportion of mature parr, as a result of increased mortality in the marine environment?
3. Does an increased proportion of mature parr translate into an increase in the collective reproductive success of mature parr?

The first question has been addressed, to a certain extent, by Baum et al. (2005) in a study on Atlantic salmon. The authors determined the proportion of male parr that matured in the wild above and below a point where discharged warm water (from a whisky distillery) resulted in an average mean increase in temperature of 2.7°C. They found no significant increase in the proportion of male mature parr below the discharge point. This finding contrasted with the results of increasing the temperature in a laboratory study involving rainbow trout, where an increase in the proportion of early maturing males was seen (Crandall and Gall 1993).

With regard to the second question, an increase in the proportion of residuals (a form of sockeye remaining in freshwater to mature) among adult males in populations of sockeye salmon in two lakes in Russia (from 13 to 82% and from 26 to 92%) over 40 years was seen (Hendry et al. 2004a). This coincided with declines in sea returns, so a decline in survival at sea (due to fishing pressure) may have been the cause.

The third question has not been addressed in wild populations. Therefore there is a requirement for more research before the effects of declining marine survival and climate change on wild Atlantic salmon populations, mediated by mature parr are predicted.
8 Genetic diversity over time and estimation of the effective population size

8.1 Introduction

8.1.1 Genetic diversity over time

Genetic analyses have been used in the context of conservation to detect bottlenecks (eg. Jones et al. 2004 in Tasmanian devils *Sarcophilus harrisii*), to estimate effective population size, $N_e$ (eg. Aspi et al. 2006 of Finnish wolves *Canis lupus*) and to evaluate the success of restoration efforts (eg Brown et al. 2000 in the American Shad, *Alosa sapidissima*). Often investigations take the form of temporal studies on genetic diversity - via one or more samples taken at two or more intervals in time. Archived samples have proven particularly useful in such approaches, particularly with the advent of PCR screening techniques. For example Johnson et al. (2004) used tissue samples from greater prairie-chickens dating back to 1951 to estimate the effective population size of populations that had undergone a drastic reduction in census population size. In salmonids, archived scale samples have been used to determine the effects of human impacts such as hatchery breeding schemes (e.g. Hansen et al. 2000, in brown trout, Säisä et al. 2003, in Atlantic salmon, Guinard et al. 2003 in lake trout *Salvelinus namaycush*), and changes to habitat (e.g. Shrimpton and Heath 2003, in chinook salmon *Oncorhynchus tshawytscha*). Temporal studies have also been employed to investigate the degree of stability of the spatial structure of populations (e.g. in Atlantic salmon, Nielsen et al. 1999a; Tessier and Bernatchez 1999).

One of the main aims in conservation programs is the maintenance of genetic diversity (Consuegra et al. 2005). The loss of genetic diversity is thought to lessen the ability of a population to adapt to new challenges, though there is little empirical evidence for this (Amos and Balmford 2001). However, in one notable study
Frankham et al. (1999) showed that inbred *Drosophila* were less able to adapt to NaCl challenges than outbred lines.

One measure of genetic diversity increasingly used in conservation-based studies is allelic richness, which can be defined as the mean number of alleles per locus in a sample (Leberg 2002). Samples of different size can be compared by standardising the number of alleles to the smallest sample through rarefaction. This is statistically more powerful than other methods of comparing samples of unequal size such as reducing all samples to the size of the smallest sample or eliminating small samples (Leberg 2002). Reductions in allele richness between historical and contemporary samples have been observed, for example, in a population of grizzly bears (*Ursus arctos*) that was isolated from other populations (Miller and Waits 2003) and in greater prairie chickens that underwent a reduction in census population size (Bellinger et al. 2003). In salmonids, reductions in allele richness have also been observed in populations subject to human impacts; in Atlantic salmon (Säisä et al. 2003; Lage and Kornfield 2006), and in lake trout (Guinand et al. 2003).

In addition, comparisons of the genetic composition of temporally spaced samples have been conducted in order to monitor conservation efforts and to gain further insights into the dynamics of populations. In Atlantic salmon, Säisä et al. (2003) investigated the effects of captive breeding in two rivers in Finland, which were ‘closed’ due to the construction of power plants. The authors compared samples obtained from the original wild stock to samples taken once artificial breeding had commenced, over a maximum period of seven generations and using seven microsatellite loci. The authors found that statistically significant differences in allele frequencies among years were common, in contrast to a large naturally reproducing stock (in the river Teno, which forms the border of northern Finland...
and Norway) where differences were generally not statistically significant
(Bonferroni correction for multiple testing was not applied, although significant
results would have remained so if it had been employed). This was attributed partly
to the breeding program and partly to the original wild populations (prior to
hatchery breeding) being smaller than the contemporary wild population in the
River Teno. A similar study to elucidate the effect of hatchery stocking in a
population of steelhead trout (\textit{Onchorynchus mykiss}) in the Kitimat River, British
Columbia, was conducted by Heggenes \textit{et al.} (2006), using ten microsatellite loci
over a period of 27 years. The authors detected no statistically significant genetic
differentiation between samples obtained prior to and following the implementation
of the hatchery scheme (following sequential Bonferroni correction for multiple
testing), probably due in part to the hatchery practices involving large numbers of
wild fish and multiple year classes (Heggenes \textit{et al.} 2006).

In populations with overlapping generations, fluctuations in allele frequencies over
time are theoretically more complex than in populations with discrete generations
(Waples 1990; Waples and Teel 1990; Ryman 1997). Based on simulations
informed by salmonid population data, Waples and Teel (1990) described how
changes in allele frequencies over short time periods (1-5 years) were only slightly
lower than those seen when comparing samples taken many more years apart (10-25
years). Ryman (1997) found, also from models, that changes in allele frequencies
between cohorts were higher than between years (where a cohort is composed of
individuals born in the same year and a sample from a particular year is composed
of parts of different cohorts- as in the adult sea returns for a single year). Palm \textit{et al.}
(2003b) similarly observed this in brown trout; allele frequency heterogeneity was
greater among cohorts than among samples obtained in different years.
A number of studies involving salmonids have now examined the temporal stability of genetic population structure. Stability of genetic population structure has been found in Atlantic salmon- among three populations sampled over a maximum of 76 years (Nielsen et al. 1999a) and among four populations sampled over a maximum period of 24 years (Tessier and Bernatchez 1999)- and in Brown trout- among five populations sampled over a maximum period of 88 years (Hansen et al. 2002) and in two populations sampled over a 20 year period (Palm et al. 2003b). However, contrary results have also been reported. Heath et al. (2002) described temporal instability in spatial structure in a study on steelhead trout (Onchorhynchus mykiss) among three populations sampled over a maximum of 39 years, while Østergaard et al. (2003) did not find temporal stability in populations of brown trout that inhabited an unstable environment, among four rivers sampled a maximum of 53 years apart.

Samples obtained at a single point in time can also be used to infer the history of a population, namely to test for a recent bottleneck. One method is based on the rate of loss of heterozygosity compared to the number of alleles; during a bottleneck the rate of loss of heterozygosity is expected to be lower than the rate of reduction in the number of alleles. This is because more low frequency alleles are expected to be lost than high frequency alleles (Williamson-Natesan 2005). A test based on this is implemented in the program BOTTLENECK (Cornuet and Luikart 1996; Piry et al. 1999). An alternative approach is based on the range and number of alleles at microsatellite loci. During a bottleneck the number of alleles is expected to decline more rapidly than the range of alleles (since only the loss of very large or very small alleles contribute to a reduction in size range). Garza and Williamson (2001) developed the M ratio based on this theory, and a statistical means of evaluating the significance of M given the number of loci, number of samples, mutation mode and
a value derived from the effective population size and mutation rate (for details see section 8.2.3). Comparing these two methods for detecting bottlenecks, Wiliamson-Natesan (2005) found that under conditions where a bottleneck had lasted several generations, initial population size was high (or mutation rates were high) and the population had some time to recover, the M ratio method was more likely to detect a bottleneck correctly. Under circumstances where the bottleneck was less severe, or very recent, the test based on heterozygosity (and assuming a two phase model of mutation- for microsatellites) was more likely to correctly detect a bottleneck (William-Natesan 2005). Both methods have been widely applied, for example in studies of the black rhinoceros *Diceros bicornis* (Harley *et al.* 2005), house finch *Carpodacus mexicanus* (Hawley *et al.* 2006) and damselfly *Coenagrion mercuriale* (Watts *et al.* 2006), where evidence for bottlenecks was provided by one or both approaches. However, both methods were designed to detect dramatic and recent changes (Williamson-Natesan 2005) - defined as occurring approximately in the past $2N_e-4N_e$ generations for the approach based on expected heterozygosity (Piry *et al.* 1999), but probably longer than this with the M ratio (Garza and Williamson 2001). In order to detect more long term, gradual changes in population size an alternative method proposed by Beaumont (1999) may be employed (Garza and Williamson 2001; Williamson-Natesan 2005). Beaumont (2003) latterly developed a method to allow the estimation of recent changes in population size, although at least two temporally separated samples are required, and it is the change in size in the period between the two samples that is examined. Hansen *et al.* (2006) simulated reductions in size of a population of North Sea Houting, *Coregonus oxyrhynchus*, and found this method to be more powerful than the M ratio in
detecting the decline, although a greater degree of sampling is required for the procedure developed by Beaumont (2003).

### 8.1.2 Estimation of effective population size

Wright (1931) introduced the concept of the effective population size. The effective population size (or effective population number) is the size of an ideal population (one which has for example random mating, constant size and random variation in reproductive success) in which the magnitude of some genetic property is equal to that of the actual population (Crow and Kimura 1970). There are a number of different forms of \( N_e \), based on different genetic properties, although the most widely used approaches are the inbreeding effective size and variance effective size. The former is based on the rate of decrease of heterozygosity in a population due to inbreeding, the latter on the degree of drift (Leberg 2005; Wang 2005). Negative effects of reduced effective population size include loss of variability, which may result in a reduction in fitness and / or loss of evolutionary potential, inbreeding depression (an increase in the occurrence of the expression of deleterious recessive alleles and loss of heterozygosity), and mutational meltdown (a decrease in the purging of deleterious alleles) (Amos and Balmford 2001).

There has been great interest in estimating the effective population size of populations and different approaches to estimating \( N_e \) have recently been reviewed by Leberg (2005) and Wang (2005). The first distinction in approaches is the use of genetic or demographic data. Estimating \( N_e \) using demographic data requires detailed information on the characteristics of the population in question, such as variance in reproductive success and sex ratios of breeders, in addition to the census size (Caballero 1994). Unfortunately, these data are rarely available. Using genetic
data, $N_e$ may be estimated indirectly in a variety of ways based on, for example, the
degree of heterozygote excess or linkage disequilibrium.

A method of particular use in populations with overlapping generations, such as
Atlantic salmon, is the temporal method. This is based on measuring the change in
allele frequencies between two or more samples taken at different times, and
estimating $N_e$ assuming the observed changes are solely the result of genetic drift.
The majority of temporal methods have been developed for populations with
discrete generations, but if samples are taken at least one generation apart (although
more generations are preferable) and sampling is representative, these methods
‘should apply approximately’ to populations with overlapping generations (Wang
2005). Two exceptions, which were developed specifically for overlapping
generations, are i) a method put forward by Jorde and Ryman (1995) for species
with iteroparous life histories and ii) another introduced by Waples (1990) for
species with semelparous life histories. However, Wang (2005) describes
difficulties in applying the method proposed by Jorde and Ryman (1995), due to
uncertainty in the optimal grouping of data. Both these methods are moment
estimators. Alternative techniques (to moment estimators) have been developed that
are reported to be more precise and accurate estimators of $N_e$ (Wang 2001; Berthier
et al. 2002; Tallmon et al. 2004). However, the performance of these estimators in
populations with overlapping generations has not been specifically addressed
(although Berthier et al. 2002 included a population of otters (*Lutra lutra*) in their
comparison of likelihood and moment based estimators, the true value of $N_e$ was not
known). The alternative techniques are pseudo-likelihood based (for example that
developed by Wang 2001 and implemented in the program MLNE), likelihood-
based (for example that detailed by Anderson et al. 2000 and applied in
MCLEEPS), or coalescent based, involving Bayesian statistics (such as that implemented in the program TMVP, Beaumont 2003 or in the program CoNe, Anderson 2005).

Unfortunately the majority of the above methods assume an isolated population and immigration can ‘substantially bias estimates of \( N_e \)’ (Wang and Whitlock 2003). However, a method developed by Wang and Whitlock (2003) allows the joint estimation of immigration rate and \( N_e \), although samples from source populations are required.

Aspi *et al.* (2006), Lippé *et al.* (2006) and Poulsen *et al.* (2006) have applied different temporal methods to the same datasets in real populations (of wolves *Canis lupus*, the copper redhorse *Moxostoma hubbsi* and cod *Gadus morhua*, respectively).

Aspi *et al.* (2006) used four different estimators and found all gave similar results (\( N_e \) ranged from 38 to 43) with relatively narrow confidence limits (the widest interval of a single estimate was 19-99). The most imprecise estimate was from the moment estimator of Waples (1989). The other methods employed were the coalescent-based method of Beaumont (2003) implemented in TMVP, the likelihood-based method of Anderson *et al.* (2000) implemented in MCLEEPS, and the pesudo-likelihood method of Wang and Whitlock (2003) implemented in MNE.

Lippé *et al.* (2006), obtained results from three temporal methods, although only quoted confidence intervals for one. The point estimate of \( N_e \) obtained from the moment estimator of Waples (1989) was approximately half that obtained from the likelihood estimator of Anderson *et al.* (2000), while \( N_e \) estimated using the likelihood method of Wang (2001) implemented in MLNE was in between (and had 95% confidence intervals encompassing the other estimates). Poulsen *et al.* (2006)
concluded that the results gained from the moment estimator of Waples (1989) and the pseudo-likelihood estimator of Wang (2001) were reasonably similar, although one was approximately double the other (1068 compared to 2067) and confidence limits were wide (reaching infinity in one estimate). They also used the method of Berthier *et al.* (2002) implemented in TM3, but found the results from one population to be dependent on the maximum $N_e$ given as prior information.

In Atlantic salmon, $N_e$ has been estimated in recent years using temporal methods and microsatellite data by Lage and Kornfield (2006), who employed the likelihood method of Wang (2001), by Säisä *et al.* (2003) using the moment estimator of Waples (1990) developed for populations with overlapping generations, by Consuegra *et al.* (2005) who also used the moment estimator of Waples (1990) as well as the pseudo-likelihood method of Wang and Whitlock (2003), and by Elliot and Reilly (2003) using the coalescent-based estimator developed by Berthier *et al.* (2002).
8.1.3 Objectives

One of the objectives of this study was to monitor, using genetic methods, the stock of Atlantic salmon in Girnock Burn. This was of particular importance given the introduction of quasi-natural spawning in 2000 and low numbers of female returns between 1997 and 2003. The samples available also provided the opportunity to estimate $N_e$ using temporal methods. Specifically, the aims were to:

- Compare allele frequency distributions between pairs of years and determine whether there was evidence for differentiation between temporally spaced samples
- Compare allele frequency distributions between cohorts (as far as was possible with the current dataset) and compare this to results seen from comparisons among years
- Determine whether there was any evidence for recent bottlenecks according to the M ratio or the method implemented in BOTTLENECK
- Apply different temporal methods to the stock of Atlantic salmon in Girnock Burn, particularly the moment estimator developed for populations with overlapping generations by Waples (1990), and other likelihood and coalescent based estimators developed for populations with discrete generations
- Estimate $N_e$ from demographic data
- Compare the results and consider the utility of estimating $N_e$. 
8.2 Methods

8.2.1 Samples

Tissue samples were obtained from adult returns to Girnock Burn in 1991, 1995, 2000, 2001 and 2004. Protocols for DNA extraction and genotyping are described in Chapter 2. Samples were typed at nine to twelve microsatellite loci (Table 7.2). Fish were aged using scale reading.

Table 8.1 Details of samples studied.

<table>
<thead>
<tr>
<th>Year of return</th>
<th># samples</th>
<th># samples aged 4 years</th>
<th># loci examined</th>
<th>Loci examined a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>95</td>
<td>34</td>
<td>9</td>
<td>Set A</td>
</tr>
<tr>
<td>1995</td>
<td>116</td>
<td>59</td>
<td>9</td>
<td>Set A</td>
</tr>
<tr>
<td>2000</td>
<td>89</td>
<td>37</td>
<td>12</td>
<td>Set B</td>
</tr>
<tr>
<td>2001</td>
<td>41</td>
<td>23</td>
<td>12</td>
<td>Set B</td>
</tr>
<tr>
<td>2004</td>
<td>117</td>
<td>76</td>
<td>12</td>
<td>Set B</td>
</tr>
</tbody>
</table>

a Set A- 2210 1605 2201 410 One9 407 202 171 197
Set B- 2210 1605 2201 410 One9 407 202 171 197 404 408 401

8.2.2 Analysis of temporal stability of genetic diversity between 1991 and 2004

Allele frequencies in different years were depicted graphically for each locus following the method employed by Tessier and Bernatchez (1999). Differences in allelic frequencies between pairs of years were tested using an exact test implemented in GENEPOP 3.4 (Raymond and Rousset 1995). For high precision the tests were run with the following parameters: 20000 dememorization steps, 200 batches and 6000 iterations per batch. Under these parameters the standard error of the probability (P) value for each locus was always <0.006. The P- value for genic differentiation for each sample pair across all nine loci (Fisher’s method) was then computed. A second, related measure of differentiation, pairwise Fst values (between pairs of years) were also computed with significance being tested by a
permutation approach (20000 permutations) using ARLEQUIN 3.01 (Excoffier et al. 2005).

In addition, the previous tests were used to compare allele frequencies between cohorts where possible. Unfortunately complete cohorts were not available in the dataset, therefore parts of cohorts were compared; these consisted of anadromous returns of total age four years from each of the years sampled (1991, 1995, 2000, 2001 and 2004, which correspond to parts of the cohorts spawned in 1986, 1990, 1995, 1996, and 1999, respectively). Where appropriate sequential Bonferroni corrections were employed to account for multiple testing (Rice 1989).

Hedrick (2005) developed a standardised genetic differentiation measure, which was calculated between pairs of years. FSTAT 2.9.3 (Goudet 2001) was employed to obtain Fst values, and subsequently RECODEDATA (Meirmans 2006) was used to obtain the maximum value possible given the variation present. The original value was divided by the maximum value to obtain the standardised measure.

Pearson’s correlation test was used to determine whether there was a correlation between allele frequencies in the samples that were temporally most distant (adult returns of 1991 and 2004), an approach taken from Tessier and Bernatchez (1999). Allelic richness per locus in each year was estimated with FSTAT 2.9.3 (Goudet 2001), which implements the rarefaction procedure of Petit et al. (1998).

8.2.3 Testing for a recent bottleneck

To test for a recent bottleneck the program BOTTLENECK version 1.2.02 was employed (Cornuet and Luikart 1996; Piry et al. 1999). A two phase model of mutation, with 95% stepwise mutations and 5% multi-step mutations was used, as recommended by Piry et al. (1999) for microsatellites. Results from the Wilcoxon test are reported, as this is considered to be the most robust and powerful when
using fewer than 20 loci (Piry et al. 1999). BOTTLENECK version 1.2.02 was also employed to conduct a qualitative test for a recent bottleneck, based on the allele frequency distribution (if this is ‘L-shaped’ it suggests a recent bottleneck has not occurred, Luikart et al. 1998). For both tests all years were investigated (9 loci were examined for 1991 and 1995 while 12 loci were examined for 2000, 2001 and 2004).

The M ratio was also used to test for the occurrence of a recent bottleneck. The M ratio was calculated as

\[ M = \frac{k}{r} \]

Where

\( k \) = number of microsatellite alleles

\( r \) = range of microsatellite alleles (in number of repeat units)

Critical values of M (\( M_c \)) were estimated using the program Critical_M (Garza and Wiliamson 2001). Values below \( M_c \) indicate that there has been a recent reduction in population size, assuming the mutation model used in estimating \( M_c \) was applicable. The following parameters were employed:

\[ \Delta_g = 3.5 \]

\( p_s = 0.9 \)

\( \theta = 0.12, 0.24, 1.2 \) or 12

Where

\( \Delta_g \) = average size of non single step mutations

\( p_s \) = proportion of single step mutations

\( \theta = 4N_e \mu \)
Genetic diversity over time and estimation of the effective population size  Chapter 8

\[ N_e = \text{effective population size} \]

\[ \mu = \text{mutation rate} \]

Values of \( \Delta_g \) and \( p_s \) were recommended by Garza and Williamson (2001). \( \theta \) could correspond to (for example) an \( N_e \) of 50, 100, 500 or 5000 and \( \mu \) of \( 6 \times 10^{-4} \), or \( N_e \) of 5, 10, 50 and 500 and \( \mu \) of \( 6 \times 10^{-3} \). A mutation rate of \( 6 \times 10^{-4} \) is the average of that obtained by O’Reilly \textit{et al.} (1998) and Norris \textit{et al.} (2000) for microsatellites in Atlantic salmon.

8.2.4 Estimating the effective population size

Four temporal methods were used to estimate the effective population size.

Temporal methods are dependent on samples taken at two or more stages in time, and all require that the number of generations between samples is known. The generation length in populations with overlapping generations is taken to be the average age of spawners (Hill 1979; Miller and Kapuscinski 1997). Two estimates of generation length were made for this study, assuming;

i) Only anadromous adults contributed to spawning. The average age of anadromous spawners was estimated by obtaining the average age of returns (according to scale readings) between 1991 and 2004 (inclusive). This gave a generation length of 4.1 years.

ii) Mature parr also contributed to spawning and that half of all progeny were sired by mature parr aged 2 years. This resulted in a generation length of 3.6 years.

In addition, all estimates were made with two sets of sampling years. Firstly, 1991 and 2004, which were preferable because they were the furthest apart. The corresponding number of generations between 1991 and 2004 were therefore:
i) 3.17 generations assuming no mature parr input (13 / 4.1)

ii) 3.61 generations assuming input from mature parr (13 / 3.6)

Unfortunately the sample obtained in 2004 may not have been representative (in terms of age classes, see section 8.3.1), thus estimates were also repeated using samples obtained in 1991 and 2000. The corresponding number of generations between 1991 and 2000 were therefore:

i) 2.20 generations assuming no mature parr input (9 / 4.1)

ii) 2.50 generations assuming input from mature parr (9 / 3.6)

The four methods used to estimate $N_e$ were:

1. The moment-based estimator of Waples (1990), which was specifically developed for a semelparous life history with overlapping generations. The following equation was utilized, which is applicable for sampling schemes whereby adults are sampled (as in the current situation):

$$N_b = \frac{b}{2(F-1) / \bar{S} + 1/N)}$$

Where

$N_b$=effective number breeders per year

$b$= an empirically derived parameter based on the number of years between samples and the average generation length

$F$=variance in allele frequency change

$\bar{S}$=harmonic mean of number of individuals in the two samples

$N$=number of breeders in earlier sample taken
And from Tajima (1992)

\[ N_e = N_b \hat{y} \]

Where

\( \hat{y} \) = average generation length

F was estimated using the program NeEstimator version 1.3 (Peel et al. 2004), which employs equation 9 from Waples (1989). The parameter \( b \) was estimated with the method described by Tajima (1992). \( N \) was taken to be 95. 95% confidence intervals of \( N_e \) were estimated by using modified values of \( F \) (Waples 1989, Shrimpton and Heath 2003) obtained using:

\[ nF / \chi^2_{0.025(n)} \text{ and } nF / \chi^2_{0.975(n)} \]

Where

\( n = \) sum of the number of independent alleles for each locus (where the number of independent alleles is the total number less one)

2. The pseudo-likelihood method of Wang (2001) as implemented in MLNE v1.1 (Wang and Whitlock 2003) was used, with maximum \( N_e \) set to 1000 and 10000. The number of generations between 1991 and 2004 was taken to be three, not accounting for mature parr, while it was taken to be four generations when accounting for mature parr (only integers were accommodated). The corresponding values for 1991 and 2000 were two and three generations.

3. A method developed by Beaumont (2003) and implemented in the program TMVP was used. The basis for this approach was a coalescent based method
introduced by Berthier et al. (2002) and employs a MCMC (Markov chain Monte Carlo) method with importance sampling. A thinning interval of 10 and size of proposal distribution of parameter updates of 0.3 was used. An importance sample of 10000 was employed. This method allows for more than two temporally spaced samples. However, in order for a fairer comparison with other methods of estimating $N_e$, the calculations were conducted while only using two samples. The mode and 95% upper and lower highest probability density (HPD) intervals were obtained with code provided by Mark Beaumont and implemented in R. The number of generations between 1991 and 2004 was taken to be 3.17, not accounting for mature parr, while it was taken to be 3.61 generations when accounting for mature parr. The corresponding values for 1991 and 2000 were 2.20 and 2.50 generations.

4. An alternative method also based on the coalescent approach of Berthier et al. (2002) was employed. This was developed by Anderson (2005) and uses an ‘improved’ importance sampling distribution compared to that utilised by Beaumont (2003). The method was implemented in the program CoNe (Anderson 2005). 90000 Monte Carlo repetitions were carried out. The number of generations between 1991 and 2004 was taken to be 3.17, not accounting for mature parr, while it was taken to be 3.61 generations when accounting for mature parr. The corresponding values for 1991 and 2000 were 2.20 and 2.50 generations.

In addition, estimates of the effective population size based on demographic factors were made. Methods described in Chapter 4, section 4.2.3.3 were used to take into
account the effect of variable reproductive success and unequal sex ratio on the
effective number of breeders per year ($N_{b(demo)}$). The mean output and variance in
spawning success obtained for spawning that occurred in 1995 (estimated from parr
sampled in 1997) was applied to spawning that occurred in the years from 1995-
2004. Estimates of $N_{b(demo)}$ were made for each year (1995-2004) assuming 10 or
800 mature parr were involved. Quasi-natural spawning (as well as natural
spawning) occurred in years from 2000 to 2004. Therefore another estimate of
$N_{b(demo)}$ in the years from 2000-2004 was made, assuming the mean output and
variance in reproductive success was equal to that calculated for quasi-natural
spawning that occurred in 2001 (as in Chapter 4). Estimates of $N_{b(demo)}$ were then
combined for the years from 1995 to 2004, to produce an estimate of $N_e$ per
generation; using equation 1 from Waples (2002):

$$N_e = g \bar{N}_b$$

Where

$g =$ average generation length

$= 4.1$ years (excluding input from mature parr)

$= 3.6$ years (including input from mature parr)

$\bar{N}_b =$ harmonic mean of number of breeders per year

Although originally developed for populations with discrete generations, in
simulations Waples (2002) found that the $N_e$ (per generation) of a semelparous
population with overlapping generations and fluctuating population size was as low
as, or lower than, this estimate.
8.3 Results

8.3.1 Genetic diversity between 1991 and 2004

Allele frequencies in each year are depicted in Figure 8.1 to Figure 8.12, the area of each circle represents the frequency of a particular allele in a specific year. Only adult returns in 2000, 2001 and 2004 were typed for loci Ssa404, Ssa408 and Ssa401.

![Figure 8.1](image_url)

*Figure 8.1 Allele frequencies of adult returns at locus Ssa2210 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).*
Figure 8.2 Allele frequencies of adult returns at locus *Ssa1605* in five spawning years (1991 $n=95$, 1995 $n=116$, 2000 $n=89$, 2001 $n=41$, 2004 $n=117$).

Figure 8.3 Allele frequencies of adult returns at locus *Ssa2201* in five spawning years (1991 $n=95$, 1995 $n=116$, 2000 $n=89$, 2001 $n=41$, 2004 $n=117$).
Figure 8.4 Allele frequencies of adult returns at locus Ssa407 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).

Figure 8.5 Allele frequencies of adult returns at locus Ssa410 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).
Figure 8.6 Allele frequencies of adult returns at locus One9 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).

Figure 8.7 Allele frequencies of adult returns at locus Ssa171 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).
Figure 8.8 Allele frequencies of adult returns at locus Ssa197 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).

Figure 8.9 Allele frequencies of adult returns at locus Ssa202 in five spawning years (1991 n=95, 1995 n=116, 2000 n=89, 2001 n=41, 2004 n=117).
Figure 8.10 Allele frequencies of adult returns at locus $Ssa408$ in three spawning years (2000 $n=89$, 2001 $n=41$, 2004 $n=117$).

Figure 8.11 Allele frequencies of adult returns at locus $Ssa404$ in three spawning years (2000 $n=89$, 2001 $n=41$, 2004 $n=117$).
Genetic diversity over time and estimation of the effective population size  Chapter 8

Figure 8.12 Allele frequencies of adult returns at locus Ssa401 in three spawning years (2000 n=89, 2001 n=41, 2004 n=117).

Visually the allele frequencies appeared to be broadly similar among years (Figure 8.1 to Figure 8.12). However, exact tests of allelic differentiation showed there were significant differences between six sets of years, reduced to three following sequential Bonferroni correction; 2004 and 1991, 2004 and 1995, 2004 and 2000 (Table 8.2). Results are shown over all (nine) loci.

Table 8.2 P-values of exact tests for allelic differentiation: pairwise comparison of all returns in different years. Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.286</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td><strong>0.015</strong></td>
<td>0.456</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.096</td>
<td>0.149</td>
<td><strong>0.045</strong></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td><strong>0.006</strong></td>
<td>&lt;<strong>0.001</strong></td>
<td><strong>&lt;0.001</strong></td>
<td><strong>0.041</strong></td>
<td>-</td>
</tr>
</tbody>
</table>

*significant following sequential Bonferroni correction, initial $\alpha=0.05/10$

** significant following sequential Bonferroni correction, initial $\alpha=0.01/10$
Fst values were low and no comparison was significant after sequential Bonferroni correction, although four comparisons were significant prior to sequential Bonferroni correction (Table 8.3).

Table 8.3 Estimates of multilocus pairwise Fst values between year classes (below diagonal).
Above diagonal: P value computed by permutation (using Arlequin). Values in bold are significant at 0.05 level prior to Bonferroni correction. No comparison was significant after sequential Bonferroni correction (initial $\alpha=0.05/10$)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>-</td>
<td>0.234</td>
<td></td>
<td>0.151</td>
<td>0.016</td>
</tr>
<tr>
<td>1995</td>
<td>0.0006</td>
<td>-</td>
<td>0.210</td>
<td>0.062</td>
<td>0.007</td>
</tr>
<tr>
<td>2000</td>
<td>0.0032</td>
<td>0.0007</td>
<td>-</td>
<td>0.061</td>
<td>0.020</td>
</tr>
<tr>
<td>2001</td>
<td>0.0018</td>
<td>0.0026</td>
<td>0.0029</td>
<td>-</td>
<td>0.081</td>
</tr>
<tr>
<td>2004</td>
<td>0.0023</td>
<td>0.0024</td>
<td>0.0023</td>
<td>0.0025</td>
<td>-</td>
</tr>
</tbody>
</table>

The standardised measure of genetic differentiation (Table 8.4) developed by Hedrick (2005) was higher than the non-standardised measure.

Table 8.4 Standardised measure of genetic differentiation among years, calculated using FSTAT and RECODEDATA.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.0042</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.0217</td>
<td>0.0049</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.0114</td>
<td>0.0185</td>
<td>0.0198</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>0.0155</td>
<td>0.0162</td>
<td>0.0154</td>
<td>0.0157</td>
<td>-</td>
</tr>
</tbody>
</table>

When comparing cohorts (as opposed to years), a greater number of comparisons were significant (Table 8.5, Table 8.6).
Table 8.5 P-values of exact tests for allelic differentiation: comparison of cohorts (anadromous returns aged 4 years.) Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991c</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995c</td>
<td>&lt;0.001**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000c</td>
<td>&lt;0.001**</td>
<td>0.664</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001c</td>
<td>0.004*</td>
<td>0.447</td>
<td>0.066</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2004c</td>
<td>&lt;0.001**</td>
<td>&lt;0.001**</td>
<td>&lt;0.001**</td>
<td>0.015</td>
<td>-</td>
</tr>
</tbody>
</table>

*significant following sequential Bonferroni correction, initial $\alpha=0.05/10$

** significant following sequential Bonferroni correction, initial $\alpha=0.01/10$

c- only anadromous fish aged 4 years were compared (the largest component year class in all samples analysed). Thus 1991c relates to the anadromous fish that returned in 1991 aged 4 years i.e. part of the cohort that were spawned in 1986.

Table 8.6 Estimates of multilocus pairwise Fst values between cohorts (anadromous returns aged 4 years), below diagonal. Above diagonal: P value computed by permutation (using Arlequin). Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991c</td>
<td>-</td>
<td>&lt;0.001**</td>
<td>&lt;0.001**</td>
<td>0.015</td>
<td>&lt;0.001**</td>
</tr>
<tr>
<td>1995c</td>
<td>0.00976</td>
<td>-</td>
<td>0.571</td>
<td>0.176</td>
<td>0.003*</td>
</tr>
<tr>
<td>2000c</td>
<td>0.01619</td>
<td>-0.00070</td>
<td>-</td>
<td>0.045</td>
<td>0.009</td>
</tr>
<tr>
<td>2001c</td>
<td>0.00950</td>
<td>0.00246</td>
<td>0.00649</td>
<td>-</td>
<td>0.026</td>
</tr>
<tr>
<td>2004c</td>
<td>0.01117</td>
<td>0.00476</td>
<td>0.00513</td>
<td>0.00610</td>
<td>-</td>
</tr>
</tbody>
</table>

*significant following sequential Bonferroni correction, initial $\alpha=0.05/10$

** significant following sequential Bonferroni correction, initial $\alpha=0.01/10$

c- only anadromous fish aged 4 years were compared. Thus 1991c relates to the anadromous fish that returned in 1991 aged 4 years i.e. part of the cohort that were spawned in 1986.

It is of note that the majority of statistically significant comparisons (either prior to or following sequential Bonferroni correction) in Table 8.2 and Table 8.3 involved returns in 2004. Due to the complex nature of fluctuations in allele frequencies in populations with overlapping generations, this observation was further investigated. Sea returns in 2004 included the highest percentage of a single cohort in comparison
with other years (Figure 8.13). Thus returns in this year may have been more representative of a particular cohort, than of the stock as a whole.

![Composition of sample](image)

**Figure 8.13** Cohorts represented in each year of adult returns.

This was investigated further by evening the representation of cohorts in the 2004 year class (by removing half of the returns aged 4 years, chosen at random). The composition of the manipulated year class is shown in Figure 8.14, identified as ‘2004i,’ which included 79 individuals.
Evening the representation of cohorts resulted in no significant difference in allele frequencies between any comparison involving 2004 following sequential Bonferroni correction, and fewer and less significant differences prior to Bonferroni correction (Table 8.7, Table 8.8).

**Table 8.7** P-values of exact tests for allelic differentiation: pairwise comparison of returns in different years when the 2004 year class was manipulated by evening cohorts (2004i). Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>0.294</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.016</td>
<td>0.452</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.092</td>
<td>0.142</td>
<td>0.046</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2004i</td>
<td>0.033</td>
<td>0.014</td>
<td>0.027</td>
<td>0.141</td>
<td>-</td>
</tr>
</tbody>
</table>

*significant following sequential Bonferroni correction, initial \(\alpha=0.05/10\)

** significant following sequential Bonferroni correction, initial \(\alpha=0.01/10\)
Table 8.8 Estimates of multilocus pairwise Fst values (below diagonal) between year classes when the 2004 year class was manipulated by evening cohorts (2004i). Above diagonal: P-value computed by permutation (using Arlequin). Values in bold are significant at 0.05 level prior to Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>-</td>
<td>0.232</td>
<td><strong>0.005</strong></td>
<td>0.147</td>
<td><strong>0.021</strong></td>
</tr>
<tr>
<td>1995</td>
<td>0.00059</td>
<td>-</td>
<td>0.203</td>
<td>0.059</td>
<td><strong>0.018</strong></td>
</tr>
<tr>
<td>2000</td>
<td>0.00316</td>
<td>0.00071</td>
<td>-</td>
<td>0.064</td>
<td>0.176</td>
</tr>
<tr>
<td>2001</td>
<td>0.0018</td>
<td>0.00258</td>
<td>0.00293</td>
<td>-</td>
<td>0.215</td>
</tr>
<tr>
<td>2004i</td>
<td>0.00273</td>
<td>0.00249</td>
<td>0.00123</td>
<td>0.0017</td>
<td>-</td>
</tr>
</tbody>
</table>

*significant following sequential Bonferroni correction, initial $\alpha=0.05/10$

**significant following sequential Bonferroni correction, initial $\alpha=0.01/10$

There was a significant correlation (P<0.001) between allele frequencies in 1991 and 2004 (Figure 8.15).

![Figure 8.15](image.png)  

Figure 8.15 Correlation between allele frequencies in adult returns sampled in 1991 and 2004. Alleles from all nine loci are represented.

There was no consistent change in allelic richness between 1991 and 2004 (Table 8.9).
Table 8.9 Allelic richness per locus for each year, and combined over loci for each year. Standardised to a minimum sample size of 41 (in year 2001) using rarefaction.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2210</td>
<td>11.3</td>
<td>10.1</td>
<td>9.3</td>
<td>10.0</td>
<td>11.4</td>
</tr>
<tr>
<td>1605</td>
<td>8.3</td>
<td>7.5</td>
<td>8.1</td>
<td>7.0</td>
<td>6.7</td>
</tr>
<tr>
<td>2201</td>
<td>19.1</td>
<td>18.2</td>
<td>19.9</td>
<td>18.0</td>
<td>20.0</td>
</tr>
<tr>
<td>410</td>
<td>20.7</td>
<td>21.7</td>
<td>23.8</td>
<td>21.0</td>
<td>22.4</td>
</tr>
<tr>
<td>One9</td>
<td>3.0</td>
<td>3.0</td>
<td>3.5</td>
<td>3.0</td>
<td>3.4</td>
</tr>
<tr>
<td>407</td>
<td>20.7</td>
<td>20.9</td>
<td>20.5</td>
<td>20.0</td>
<td>18.2</td>
</tr>
<tr>
<td>202</td>
<td>10.6</td>
<td>9.7</td>
<td>9.9</td>
<td>9.0</td>
<td>9.9</td>
</tr>
<tr>
<td>171</td>
<td>10.3</td>
<td>10.5</td>
<td>10.6</td>
<td>11.0</td>
<td>10.3</td>
</tr>
<tr>
<td>197</td>
<td>20.1</td>
<td>20.5</td>
<td>18.0</td>
<td>19.0</td>
<td>23.9</td>
</tr>
<tr>
<td>Total over all loci</td>
<td>124.1</td>
<td>122.1</td>
<td>123.7</td>
<td>118.0</td>
<td>126.1</td>
</tr>
</tbody>
</table>

8.3.2 Testing for a recent bottleneck

There was no evidence for a recent bottleneck according to tests implemented in BOTTLENECK (Cornuet and Luikart 1996; Piry et al. 1999), results of the quantitative test are shown in Table 8.10. The allele frequency distribution curves were all approximately ‘L-shaped’, suggesting no recent bottleneck had occurred.

Table 8.10 Results of test for bottleneck implemented in BOTTLENECK, assuming a two phase mutation model.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># loci used</td>
<td>9</td>
<td>9</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Probability</td>
<td>0.65</td>
<td>0.13</td>
<td>0.52</td>
<td>0.30</td>
<td>0.47</td>
</tr>
</tbody>
</table>

The indication of a bottleneck using the M ratio was dependent on the value of \( M_c \) used (Table 8.11). In 2000, 2001 and 2004 the significance of the M ratio was dependent on the value of \( \theta \) used to estimate \( M_c \). \( \theta \) is based on the effective population size and mutation rate. Neither the mutation rate nor \( N_e \) were known, although values of \( \theta \) equal to 0.12 or 0.24 do not seem unreasonable, as they could correspond to a mutation rate of \( 6\times10^{-4} \), the average of that obtained by O’Reilly et
al. (1998) and Norris et al. (2000) for microsatellites in Atlantic salmon and an $N_e$ of 50 or 100. Garza and Williamson (2001) suggest that M ratios below 0.68 should be taken as evidence of a recent bottleneck; all observed M ratios were higher than this. The M ratios in the more recent years (2000, 2001 and 2004) were lower than those in 1991 and 1995. This was still the case when the M ratio was calculated using the same 9 loci for all years (Table 8.11).

Table 8.11 M ratios for samples of anadromous fish obtained between 1991 and 2004, and corresponding estimates of $M_c$ with varying values of $\theta$. M ratios lower than $M_c$ indicate a recent bottleneck- such values are shown in bold.

<table>
<thead>
<tr>
<th>Year of sample</th>
<th>Number of loci</th>
<th>Number of individuals $^a$</th>
<th>Average M $^b$</th>
<th>$M_c$ $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\theta=0.12$</td>
<td>$\theta=0.24$</td>
</tr>
<tr>
<td>1991</td>
<td>9</td>
<td>95</td>
<td>0.91</td>
<td>0.84</td>
</tr>
<tr>
<td>1995</td>
<td>9</td>
<td>116</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>2000</td>
<td>12</td>
<td>89</td>
<td>0.79</td>
<td>0.85</td>
</tr>
<tr>
<td>2000</td>
<td>9</td>
<td>89</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>2001</td>
<td>12</td>
<td>41</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td>2001</td>
<td>9</td>
<td>41</td>
<td>0.85</td>
<td>0.84</td>
</tr>
<tr>
<td>2004</td>
<td>12</td>
<td>117</td>
<td>0.82</td>
<td>0.85</td>
</tr>
<tr>
<td>2004</td>
<td>9</td>
<td>117</td>
<td>0.85</td>
<td>0.84</td>
</tr>
</tbody>
</table>

$^a$ Number of diploid individuals (Critical_M treats each individual as two ‘samples’)

$^b$ Average over loci

$^c$ Estimate of effective population size

Effective population sizes estimated using temporal methods and samples obtained in 1991 and 2004, for different generation lengths, are shown in Table 8.12. The number of generations between 1991 and 2004 were higher (3.61 or 4 generations) when mature parr were included and lower (corresponding to 3.17 or 3 generations) when mature parr were excluded from the estimation of generation length.
## Table 8.12 Estimates of $N_e$ per generation using different methods, based on data from 1991 and 2004. Name of programs utilised are given in parentheses.

<table>
<thead>
<tr>
<th>Method (program)</th>
<th>Mature parr input included</th>
<th># generations between 1991 and 2004</th>
<th>$N_e$ max</th>
<th>$N_e$ estimate</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waples 1990, Tajima 1992</td>
<td>N</td>
<td>3.17</td>
<td>n/a</td>
<td>1410</td>
<td>1194-1666</td>
</tr>
<tr>
<td>Waples 1990, Tajima 1992</td>
<td>Y</td>
<td>3.61</td>
<td>n/a</td>
<td>1238</td>
<td>1048-1463</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>N</td>
<td>3</td>
<td>1000</td>
<td>916</td>
<td>373-1000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>N</td>
<td>3</td>
<td>10000</td>
<td>916</td>
<td>373-10000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>Y</td>
<td>4</td>
<td>1000</td>
<td>1000</td>
<td>477-1000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>Y</td>
<td>4</td>
<td>10000</td>
<td>1173</td>
<td>480-10000</td>
</tr>
<tr>
<td>Anderson 2005 (CoNe)</td>
<td>N</td>
<td>3.17</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anderson 2005 (CoNe)</td>
<td>Y</td>
<td>3.61</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>N</td>
<td>3.17</td>
<td>1000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>N</td>
<td>3.17</td>
<td>10000</td>
<td>1777</td>
<td>508-9542&lt;sup&gt;HPD&lt;/sup&gt;</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>Y</td>
<td>3.61</td>
<td>1000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>Y</td>
<td>3.61</td>
<td>10000</td>
<td>1992</td>
<td>666-9566&lt;sup&gt;HPD&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> (in estimate of generation length) Y = yes, N = no

<sup>b</sup> n/a = not required/ unable to manipulate in some methods

<sup>c</sup> - = none calculated, possibly due to too little differentiation between samples

<sup>HPD</sup> = 95% highest posterior density limits are quoted for the method of Beaumont (2003).

Estimates of effective population sizes using samples obtained in 1991 and 2000, for different generation lengths, are shown in Table 8.13. The number of generations between 1991 and 2000 were higher (2.5 or 3 generations) when mature parr were included and lower (corresponding to 2.2 or 2 generations) when mature parr were excluded from the estimation of generation length.
Table 8.13 Estimates of \( N_e \) per generation using different methods, based on data from 1991 and 2000. Name of programs utilised are given in parentheses.

<table>
<thead>
<tr>
<th>Method</th>
<th>Mature parr input included</th>
<th># generations between 1991 and 2004</th>
<th>( N_e ) max</th>
<th>( N_e ) estimate</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waples 1990, Tajima 1992</td>
<td>N</td>
<td>2.20</td>
<td>n/a</td>
<td>1876</td>
<td>1453-2444</td>
</tr>
<tr>
<td>Waples 1990, Tajima 1992</td>
<td>Y</td>
<td>2.50</td>
<td>n/a</td>
<td>1648</td>
<td>1276-2146</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>N</td>
<td>2</td>
<td>1000</td>
<td>595</td>
<td>237-1000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>N</td>
<td>2</td>
<td>10000</td>
<td>595</td>
<td>237-10000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>Y</td>
<td>3</td>
<td>1000</td>
<td>791</td>
<td>321-1000</td>
</tr>
<tr>
<td>Wang 2001 (MLNE V1.12)</td>
<td>Y</td>
<td>3</td>
<td>10000</td>
<td>791</td>
<td>321-10000</td>
</tr>
<tr>
<td>Anderson 2005 (CoNe)</td>
<td>N</td>
<td>2.20</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anderson 2005 (CoNe)</td>
<td>Y</td>
<td>2.50</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>N</td>
<td>2.20</td>
<td>1000</td>
<td>852</td>
<td>339-995(^{\text{HPD}})</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>N</td>
<td>2.20</td>
<td>10000</td>
<td>1267</td>
<td>157-9293(^{\text{HPD}})</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>Y</td>
<td>2.50</td>
<td>1000</td>
<td>868</td>
<td>369-995(^{\text{HPD}})</td>
</tr>
<tr>
<td>Beaumont 2003 (TMVP)</td>
<td>Y</td>
<td>2.50</td>
<td>10000</td>
<td>1672</td>
<td>260-9443(^{\text{HPD}})</td>
</tr>
</tbody>
</table>

\(^{a}\) (in estimate of generation length) Y = yes, N = no  
\(^{b}\) n/a = not required/ unable to manipulate in some methods  
\(^{c}\) - = none calculated, possibly due to too little differentiation between samples  
\(^{\text{HPD}}\) = 95\% highest posterior density limits are quoted for the method of Beaumont (2003).

For all estimates, confidence limits were wide and point estimates were high, ranging from 595-1992. Estimates based on samples obtained in 1991 and 2004 were generally higher than those based on samples obtained in 1991 and 2000.

Demographic estimates of \( N_e \) are shown in Table 8.14. Estimates ranged from 95 to 144, depending on the assumptions made. All were lower than any point estimates obtained using temporal methods.
Table 8.14 Demographic estimates of $N_e$, for the years 1995-2004.

<table>
<thead>
<tr>
<th>Modelled on QN/Nat spawning in 2000-2004</th>
<th>Assumed generation time (years)</th>
<th># mature parr $= 10^b$</th>
<th># mature parr $= 800^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nat</td>
<td>3.6</td>
<td>95</td>
<td>118</td>
</tr>
<tr>
<td>Nat</td>
<td>4.1</td>
<td>108</td>
<td>135</td>
</tr>
<tr>
<td>QN</td>
<td>3.6</td>
<td>121</td>
<td>127</td>
</tr>
<tr>
<td>QN</td>
<td>4.1</td>
<td>137</td>
<td>144</td>
</tr>
</tbody>
</table>

$^a$ for years 2000-2004 both natural (Nat) and quasi-natural (QN) spawning occurred. In these years $N_{b(demo)}$ estimates were made based on mean output and variance in reproductive success assuming natural spawning (Nat), or quasi-natural spawning (QN).

$^b$ in natural spawning.
8.4 Discussion

8.4.1 Genetic diversity through time

Over the 13 year period studied (between 1991 and 2004) there was no consistent change in allele richness and, in general, no observable change in genetic composition of the stock of Atlantic salmon in Girnock Burn. This suggests that low numbers of female returns between 1997 and 2000 have not drastically affected the observed genetic diversity, at least by the indicators used in this study. Further analysis of the genetic diversity in years subsequent to 2004 would be required to determine the effect of persistence of low numbers of returns in 2001-2003. Other temporal studies on the genetic composition of Atlantic salmon populations have also found generally non-significant changes in genetic diversity or non-significant genetic differentiation between samples of adults: Nielsen et al. (1999a) in samples taken 37 years apart, Tessier and Bernatchez (1999) in samples taken 13-24 years apart (approximately 3-5 generations), Säisä et al. (2003) in samples taken 31 years apart (approximately 6 generations). Significant differences were found by Säisä et al. (2003) within populations affected by stocking procedures.

Palm et al. (2003b) point out that significant differences in allele frequencies are more likely to be detected over longer time periods, but that due to overlapping generations significant differences may be seen between cohorts, or when samples do not consist of even representation of a cohort. In this study where significant differentiation (according to Fst values or comparison of allele frequencies) among years was found, it was generally between the sample obtained in 2004 and samples taken in previous years, which was likely due to the uneven representation of cohorts (specifically, the over-representation of a particular cohort in 2004). Palm et al. (2003b) found that the magnitude of difference in allele frequencies was higher.
among cohorts than among years, in two populations of brown trout. In the study described here, significant differences in allele frequencies were more common in comparisons of cohorts (although data were limited), than in comparisons of years. This could have implications in other studies comparing temporal variation in allele frequencies in species with overlapping generations. It may be advantageous to determine the composition (in terms of cohorts) of samples being compared prior to drawing conclusions about the stability of allele frequencies in the population under study, although in many cases this may not be practical. In addition, when estimating the effective population size of populations with overlapping generations using the temporal method, it would be advantageous to ensure that different cohorts are represented in the samples being compared. If this were not the case, there could be a downwards bias in the effective population size estimated, as observed fluctuations in allele frequencies may be greater than that in the population as a whole. Due to this potential problem the effective population size in this study was estimated over two different time periods, comparing returns in 2004 and 1991 in addition to returns in 2000 with 1991, since cohorts were represented more evenly in 2000 than in 2004.

The degree of correlation between allele frequencies in 1991 and 2004 was similar to that seen by Tessier and Bernatchez (1999). The authors obtained correlation coefficients of 0.93-0.98 for Atlantic salmon with 7 loci and 3-5 generations between samples while that observed in this study was 0.90. There was no evidence for a bottleneck according to two of the three tests used, and the results from the other test were dependent on the critical value of M employed. This suggests that the low numbers of returns between 1997 and 2000 did not significantly reduce the effective population size. However, the tests relied on
knowledge of the mutation mode of the loci used. The M ratio test had reduced power due to imprecision in the critical value of M. Positive aspects of these tests are the requirement for only a single sample and they are not computationally intensive (in contrast to the method of Beaumont (2003) for detecting changes in effective population size).

8.4.2 Estimating the effective population size

The point estimates of $N_e$ from the temporal methods appeared to be high, given that the highest number of adult returns in the years examined was 117, the unequal sex ratio of spawners, and the observed unequal spawning success of anadromous returns (Chapter 4). Such apparently high estimates may have been obtained because:

1. $N_e$ is indeed large
2. assumptions made in estimating $N_e$ were not valid
3. the methods themselves were problematic

In the first instance, $N_e$ could be larger than that initially presumed if the adults passing upstream through the traps are only a small part of the entire population i.e. the traps represent an artificial rather than natural boundary. In the second instance, assumptions that may not have been met were selective neutrality of markers, random sampling of the population, absence of mutation and absence of immigration (Leberg 2005). Since the markers used were microsatellites the majority (at least) were likely to have been effectively neutral (Jarne and Lagoda 1996). As the lower boundary of the population is unknown, it is difficult to determine whether sampling was random, although random sampling would have been more improbable had juveniles rather than sea returns been sampled. The degree of mutation over the time period examined should be negligible (Leberg
2005). Possibly the most important assumption that may not have been met was that of no immigration. By the nature of their life histories, Atlantic salmon returning from the sea may stray to sites other than their natal origin and possibly then breed there. Indeed, three male grilse caught at the traps in Girnock Burn in 2000 and 2004 were likely to have originated in a different population, based on their genotypes (Chapter 6). In addition, a proportion of returning adults are of unknown origin (approximately 25% from one estimate in Chapter 6, as indicated by the absence of fin clips and inability to assign parentage to previous spawners at Girnock Burn). Immigration can drastically affect the change in allele frequencies, causing an upwards or downwards bias in the estimate of $N_e$ (Wang and Whitlock 2003). An assumption specific to the moment estimator of Waples (1990) is that of semelparity. The extent to which mature parr at Girnock Burn re-mature as anadromous fish is unknown. However, Heath et al. (2002) applied this method (Waples’ 1990 moment estimator) to steelhead trout with a low degree of repeat spawning (<10%).

The third possibility was that the methods employed were flawed. The moment estimator can be biased when allele frequencies are close to 1 or 0 (Waples 1990; Turner et al. 2001), as was the case for the loci employed- there were many low frequency alleles. However, Turner et al. (2001) found that ‘binning’ low frequency alleles was ‘not warranted’ as it decreased precision of the estimate. In addition, Shrimpton and Heath (2003) found that binning or not binning alleles made little difference to their estimates, despite the occurrence of a number of alleles of low frequency. Tallmon et al. (2004) tested the accuracy of a range of temporal methods, including that by Wang (2001) and Beaumont (2003) used in this study. The authors found that, when applied to an experimental population of mosquito
fish, the 95% confidence (or credible intervals) of the two afore-mentioned methods did not include the true value of $N_e$. Another cause for concern is that temporal methods are more reliable for populations with small values of $N_e$ (Säisä et al. 2003; Leberg 2005). This is possibly why Aspi et al. (2006) found a range of methods gave very similar estimates for $N_e$ in a population of wolves, with relatively narrow confidence intervals.

It is difficult to determine whether the sample taken in 2004, due to its uneven composition in terms of cohorts, affected the results. Samples obtained in 1991/2004 were attractive because they spanned a longer time period, although the samples obtained in 1991/2000 were more representative of different cohorts. They estimated $N_e$ over a different time periods, so will not necessarily produce the same results, as highlighted by Waples (2005). The effect of different generation times (due to the inclusion or exclusion of mature parr) differed depending on the method, with the moment estimator predicting a lower $N_e$ with decreased generation time (corresponding to the inclusion of mature parr), whereas in all other estimates $N_e$ increased with decreased generation time (the inclusion of mature parr). The maximum value of $N_e$ affected the confidence intervals, but not the point estimate of $N_e$ in Wang’s (2001) method, whereas altering the maximum value of $N_e$ also affected the point estimate of $N_e$ obtained using Beaumont’s (2003) method.

Unfortunately it was not possible to compare results obtained with the method of Beaumont (2003) to those from Anderson (2005), due to the absence of results with the latter method- possibly because differentiation was too low between samples.

An important consideration when applying these methods is computational time. The lower this is, the more likely the method is to be thoroughly tested. The method of Anderson (2005) was considerably quicker than that of Beaumont (2003) (even
when values were obtained from Anderson’s method, results not shown). Applied to
data in the current study, the method of Beaumont (2003) took around 24 hours to
compute, whereas that of Anderson (2005) took one to two hours. In addition, the
method of Anderson (2005) provides a point estimate and confidence limits directly,
whereas further analysis of the output is required in the method of Beaumont
(2003).
Temporal methods are more reliable for populations with a small $N_e$. Therefore
populations that do have a large $N_e$ would require samples that were widely spaced
in time, in order for adequate genetic differentiation between samples and hence for
more reliable estimates to be made. Obtaining such samples may not always be
feasible. Clearly, the accuracy of the estimate will also be dependent on meeting the
assumptions made in the process, namely that of no immigration. A low degree of
genetic differentiation may also be the result of immigration (as opposed to high $N_e$)
and accounting for this would presumably give more reliable estimates. However,
for methods that do not take into account the level of immigration, it would be
useful to know the level of genetic differentiation that is required to obtain
meaningful estimates of $N_e$. However, it could be argued that the estimation of $N_e$ in
a conservation context may be of more interest for populations with small $N_e$.
The demographic estimates of $N_e$ were nearly an order of magnitude lower than any
of the point estimates given by the genetically based temporal methods. One of the
drawbacks of the demographic estimate is that it is based on the initial number of
potential spawners, which was unknown; spawners below the traps that were part of
the population unit were not included.
Temporal methods gave point estimates of around 600-2000 whereas the
demographic estimates were in the region of 100-150. Given that $N_e$ is on average
only 11% of the census size (Frankham et al. 2002), and that 33% of returns to Girnock Burn between 1990 and 2004 were identified as having originating above the traps (from adipose fin clips), the lower estimates may seem more likely. However, the rate of loss of genetic diversity is also dependent on immigration. If immigration occurs constantly over a long period, the rate of change of allele frequencies approaches that predicted by the effective size of the whole metapopulation, thus $N_e$ is overestimated using temporal methods if immigration is not accounted for (Wang and Whitlock 2003). The demographic method is not affected by immigration, but also reveals nothing about how immigration may or may not be affecting the rate of loss of genetic diversity. Therefore in order to understand how genetic diversity in the stock of Atlantic salmon in Girnock Burn may be affected in the future, it would be useful to know both the level of immigration and $N_e$.

Other studies that have estimated $N_e$ in Atlantic salmon populations have also used the methods employed in the current study. Lage and Kornfield (2006) used the method developed by Wang (2001) to estimate $N_e$ of an endangered population of wild Atlantic salmon, with one to nine generations separating samples and seven microsatellite markers. Their point estimates were low (48-202) by comparison with estimates using this method for Girnock Burn. Confidence intervals were wide (the widest from 40–1000, and 1000 was the maximum $N_e$ specified) if three or fewer generations separated samples, but narrower when a larger number of generations separated samples (for example 81-213 when sampling occurred nine generations apart). This suggests that if more generations had separated the samples from Girnock Burn, narrower confidence intervals may have been obtained. It may be that more differentiation would be seen over a longer period, resulting in more
precise estimates. Säisä et al. (2003) used the moment estimator of Waples (1990) (which accounts for overlapping generations) to estimate \( N_e \) in a large wild population of Atlantic salmon. Over the longest time period examined (56 years) the point estimate of \( N_e \) was 1412, with a 95\% confidence interval of \( 475-\infty \). This may be a reflection of the lower reliability of the temporal method for large populations and/or the occurrence of low frequency alleles (which were not pooled). It would be interesting to apply other temporal methods to this dataset. The authors used a different sampling scheme (using juveniles) to that in the current study, which required the use of a different equation from Waples (1990), obviating the need for the inclusion of ‘N’ (the number of breeders). This may be advantageous as estimating \( N \) undoubtedly introduces an additional source of error. Consuegra et al. (2005) estimated \( N_e \) for four small populations of Atlantic salmon, using Waples’ (1990) moment estimator, as well as the method of Wang and Whitlock (2003). For the latter separate estimates were calculated, while taking into account and ignoring immigration. Low estimates were obtained with all three methods, with confidence intervals that were narrower by comparison with those obtained in the current study. Estimates that were obtained while allowing for immigration were lower than all other estimates, although these did not take into account overlapping generations and were based on samples taken only 4 years apart. Elliot and Reilly (2003) estimated \( N_e \) for a hatchery population of Atlantic salmon in Tasmania, using the method of Berthier et al. (2002) implemented in TM3.

The most accurate means of determining \( N_e \) for Atlantic salmon that exhibit a semelparous life history with overlapping generations is still unclear. Simulations or studies on experimental populations with known \( N_e \) would be required to compare the accuracy and precision of Waples’ moment estimator specifically developed for
such a scenario, to that of other temporal methods (likelihood and coalescent based) developed for populations with discrete generations (but which should approximate $N_e$ values for populations with overlapping generations if a long interval occurs between samples). It would be useful for such studies to include loci with a range of numbers of alleles, and to be conducted for high and low values of $N_e$.

In order to obtain more accurate estimates of $N_e$ for Girnock Burn it may be necessary to sample over longer time periods, for example utilising scale samples, and/or to sample from populations that are possible sources of migrants. The latter would enable the estimation of immigration rate and $N_e$ jointly, using the method put forward by Wang and Whitlock (2003).

Wang (2005) drew attention to the limitations of calculating $N_e$ using genetic methods, given difficulties in its calculation and the lack of information it provides about the basis of its value. In contrast, Waples (2004) was optimistic, claiming it was now possible to obtain the ‘detailed demographic information necessary to directly estimate $N_e$ (and $N_e/N$) with high precision’. This may be the case under some circumstances, but estimating $N_e$ using the demographic method proved problematic in Girnock Burn, despite detailed demographic information. The writer’s view is that the limitations described by Wang (2005) cannot be stressed too greatly, although for small isolated populations it may be possible to calculate $N_e$ with a greater degree of accuracy. In a conservation context the interpretation of $N_e$ is also dependent on the level of immigration. For example Consuegra et al. (2005) found the effective population sizes of four populations of Atlantic salmon were very low, although they maintained genetic variation at levels similar to that seen in larger populations. The authors suggested this might be due to asymmetric gene flow. In addition, Fraser et al. (2004) postulated that sex-biased dispersal could
reduce the rate of loss of genetic diversity in populations with low $N_e$. In the long
term, immigration can result in the overestimation of $N_e$, as it can produce a
reduction in the rate of change of allele frequencies theoretically reaching that of the
effective size of the whole metapopulation (Wang and Whitlock 2003). Thus
estimates of $N_e$ should be interpreted with great care, due to the (in)accuracy of the
estimate and possible implications of immigration.
9 Concluding Remarks

The application of highly polymorphic molecular markers to archived samples and samples recently obtained from Atlantic salmon in Girnock Burn has allowed a greater insight into a range of aspects of this stock. These include:

- mating system
- reliability of molecular methodology
- spatial distribution of juvenile kin
- extent of homing
- the reproductive success of mature parr and potential impact on adaptation
- an evaluation of genetic diversity over time
- estimation of the effective population size
- effectiveness of a management strategy

The stock of Atlantic salmon in Girnock Burn consists of wild fish that have been subject to almost entirely natural conditions, particularly before 2000 when a form of spawning intervention was instigated. This has allowed a study of Atlantic salmon in close to ‘real world’ conditions. An ever–present concern with tank or experimental studies is the applicability of results to conditions in the field. However, studies undertaken in the wild face other, possibly more daunting challenges. Distinguishing the effect of ‘signal’ from the ‘noise’ of the natural environment is much more difficult to achieve, since the level of control that can be exercised over both study organism and environment is greatly restricted. In the current study, for example, the number of sea returns that entered the upstream trap limited the sample of sea returns available for
Concluding remarks

Chapter 9

analysis in a particular year, a factor not under the control of the researcher. This was also largely a retrospective study, which was advantageous in allowing a greater time span to be investigated than would otherwise have been allowed by the duration of the study. However, this also resulted in a lack of flexibility in sample collection. For example samples of mature male parr used in quasi-natural spawning would have been a valuable resource to characterise, but unfortunately these had not been retained. The current study was undertaken in a single, relatively small, spate tributary of one river system and has revealed novel insights into the complex dynamics of Atlantic salmon population genetics. Just how representative the findings of this study are in relation to other Atlantic salmon populations remains to be established. Similar types of study should certainly be conducted elsewhere.

An important finding of the study with regards to the stock management was that quasi-natural spawning was successful in giving a more complete and even representation (according to the standardised variance) of spawners than natural spawning (Chapter 4). However, this manipulation apparently also affected the output of natural spawning. The reproductive success of fish involved in natural spawning in 2000 and 2001 (i.e. when quasi-natural spawning was also practised) was significantly more skewed than that of natural spawning in 1995, when no quasi-natural spawning took place. This suggests that the current strategy, to use all anadromous returns in quasi-natural spawning with no natural spawning above the traps, should result in a better overall representation of (anadromous) spawners. However, this approach also introduces the risk of losing substantial numbers of eggs during incubation, which in turn may inadvertently impose selection for characters beneficial in the incubator, but not
necessarily so in the natural environment. Other more subtle, but potentially evolutionary significant, consequences may follow from assisted spawning. The practise eliminates potential sexual selection forces. Furthermore, planting out progeny in mixed family units may alter the distribution of juvenile kin in the burn. The consequences of doing this are unknown, although 1+ parr that were spawned naturally in Girnock Burn were found to be clustered in family groups (Chapter 5). In addition the occurrence of kin-biased behaviour in Atlantic salmon juveniles has been suggested by studies conducted under experimental conditions (Brown and Brown 1996; Griffiths and Armstrong 2002). For such interactions to occur requires that juveniles are able to recognise kin. Recently Rajakaruna et al. (2006) reported that juvenile Atlantic salmon were able to discriminate between kin and non-kin; such recognition being partially influenced by a major histocompatibility complex (MHC) class II gene haplotype.

The over-winter mortality of incubated eggs is generally low (averaging 4% from 2000-2002), although more recently, in the winter of 2005/2006 approximately 30% of eggs were lost, the reasons for which are still not apparent. Therefore the first potential disadvantage described above could pose a real problem. Domestication selection is probably not a great issue in Girnock Burn, as 1) incubation time is relatively short, 2) the incubator is situated within the burn, and 3) mortality is generally low. The higher mortality in 2005 may have been due to factors that would have also been encountered in the natural environment. It could present a problem, however, if mortality were specific to the conditions in the incubator. A catastrophic loss of progeny, due to incubator malfunction is always a risk, but could be managed by splitting resources among incubators or sites. Overall, the potential shortcomings of quasi-natural
spawning do not appear to have had drastic consequences (as far as could be monitored); anadromous returns in 2004 included offspring of quasi-natural spawning from 2000.

An aspect of quasi-natural spawning that was likely to differ substantially from natural spawning in Girnock Burn was the number of mature parr contributing to offspring. Although the number of mature parr that contribute successfully to natural spawning has not been quantified, it is apparent, from observations and a previous genetic investigation (Taggart et al. 2001) that large numbers, probably hundreds, successfully spawn. It is likely to be orders of magnitude greater than the number currently used in quasi-natural spawning. A future line of research could be to estimate the number of mature parr involved in natural spawning more accurately. It would probably only be feasible to conduct such a study under semi-natural conditions, perhaps a cordoned-off section of a spawning region, to allow for the collection of tissue samples from all putative spawners. Such a set-up would also allow further investigation into the success of mature parr with differing sex ratios of anadromous spawners. Although the spawning success of individual mature parr was investigated by Thomaz et al. (1997) and Jones and Hutchings (2001), limited numbers of mature parr (a maximum of 23) and anadromous spawners (a maximum of one anadromous male and two anadromous females) were used and the experiments were conducted in relatively small areas (the largest section measured 10m × 3m). Quasi-natural spawning allows the proportion of eggs fertilised by mature parr to be manipulated. It would be useful to know what the optimum level would be, in terms of the long-term survival of the stock, although this is perhaps an unattainable objective. It was postulated that mature parr have the
potential to propagate alleles that may confer a disadvantage in the marine environment, although an exploration of this did not reveal definitive evidence for such a scenario. Any disadvantage cannot be strong as offspring of mature parr were found to return to the burn as anadromous adults (Chapter 7). It may be of interest to examine this in other populations, perhaps where there are larger numbers of anadromous returns in each year and a low percentage of males that mature as parr.

An important consideration in the conservation of species is climate change. It was postulated that the predicted increase in temperature may result in a greater incidence of male parr maturation. As a result there may be a decline in the survival of anadromous males, since Lundqvist et al. (1994) reported a lower marine survival of males that had previously matured as parr. Leading from this, it may be speculated that the proportion of eggs fertilised by mature parr in the wild could increase. Under such circumstances the best course of action for quasi-natural spawning remains unclear. The level of reproductive success of mature parr in the wild, dictated by (changing) environmental conditions, may differ to that artificially imposed by quasi-natural spawning. An increase in the proportion of eggs fertilised by mature parr would be likely to result in a greater degree of isolation of the stock (as the scope for straying of anadromous males is higher than for parr) and a decrease in the average generation length. Generation length is one of the many factors affecting the effective population size. Theory predicts that the effect of juveniles reproducing is to decrease the generation length, which in turn acts to lower $N_e$ (e.g. Nunney 1993).

An attempt was made in the current study to estimate the effective population size using temporal methods based on allelic data and a method based on demographic data.
The estimate based on demographic data was lower than all point estimates obtained using temporal methods (based on genetic data). One possible reason for this disparity is the influence of immigration, which was not taken into account in the temporal methods explored. A future line of research could be to estimate the effective population size and effective number of migrants simultaneously using the method developed by Wang and Whitlock (2003). However, this approach requires samples to be obtained from populations that are the sources of migrants. As most migrants are likely to come from within the Dee catchment this should be feasible (though demanding on resources, time and finance), by extensive electrofishing of other tributaries supplemented with samples from angler caught fish. More precise estimates of $N_e$ might have been gained if the samples utilised had been obtained further apart in time. A potential resource that could be exploited is scale samples taken from anadromous returns to Girnock Burn dating back to 1966. However, although microsatellite loci have been successfully amplified in the past from old scale samples of Atlantic salmon (eg Nielsen et al. 1997; Ayllon et al. 2006b; Lage and Kornfield 2006), there are technical problems associated with this approach, e.g. poor quality and small amounts of DNA (Nielsen et al. 1999b). If it proved possible to exploit these samples (and thereby generate reliable genetic data) the information collected would also allow the comparison of allelic richness over a greater number of years than was explored in the current study.

One aspect of this investigation focused on the extent of trap leakage. An estimate of a mean level of c.34% was made (measured as the percentage of returns assigned to spawners at Girnock Burn, that were unclipped). Estimation of the extent of trap
leakage was hampered by the possibility that females had spawned below the traps prior to being placed above them (although measures were taken to exclude the offspring of such females from the analyses). With current management practices all Atlantic salmon originating above the traps are now from known crosses (as they are solely the product of quasi-natural spawning). Thus, future sample datasets should provide a means to improve the estimate of the extent of trap leakage, as offspring of pairings between anadromous spawners that were not constructed in quasi-natural spawning could be positively identified as having originated below the traps. Despite the demonstration that trap leakage is occurring, it is unlikely that the level will be reduced. Most escapement is likely to occur during periods of high water flow (flood conditions) with fish passing over the trap. Even if increased manpower were available to monitor and clear the downstream trap more frequently than at present (particularly through the night), during spate conditions this may not be a safe option and in some conditions it clogs more rapidly than it can be cleared. The identification of returns to Girnock Burn that were philopatric, but unclipped (presumably due to trap leakage) allowed a more refined estimate of the level of homing to the burn than was otherwise possible based solely on the ‘clipping’ status of adult returns. Olfaction is known to be essential for the accurate homing of anadromous salmonid species (Dukes et al. 2006). It is thought that imprinting on odours in the natal stream occurs during the parr-smolt transformation (PST) period, and that such odours are then used as ‘orientation cues’ during homing. Recent findings suggest that both environmental and pheromonal cues are involved (Dukes et al. 2004). At Girnock Burn, migratory fish leave as parr during the autumn or as smolts the following spring.
It would be of interest to determine whether fish that pass through the traps in autumn are more likely to stray than those passing through as smolts in spring, since autumn parr might have a lower degree of imprinting to the burn. McCormick et al. (1998) speculated that pollutants might interfere with the development of olfactory imprinting and the subsequent homing of adults. In salmonids, Saucier et al. (1991) reported that exposure of rainbow trout to a low level of copper (comparable to levels encountered in the ‘natural environment’ as a result of industrial activity) resulted in the impairment of olfactory discrimination, while Scholz et al. (2000) described how exposure to nominal levels of diazinon (an insecticide) disrupted the homing ability of chinook salmon.

Other studies have investigated the physiological effects of pollutants, particularly metals, on the olfactory system of fishes (reviewed in Scott and Sloman 2004). It can be postulated that homing, perhaps in a range of fish species, has been (and will continue to be) affected by the influx of pollutants into river systems and the marine environment as a result of anthropogenic influences.

Based on the current study (Chapters 5 and 8), a number of recommendations can be made for future genetic sampling strategies. For population genetic analyses, it was found that when sampling juvenile kin, large regions of Girnock Burn needed to be electrofished to provide an apparently unbiased sample. Preferably, more than one age class should be included when sampling. The applicability of this to other regions could be determined by future research. It was also found that when a cohort was over-represented in a year class of anadromous returns, allele frequencies were significantly affected. It may therefore be useful to obtain information on the age of anadromous fish
when sampled for population genetic analyses, in order to determine whether the representation of cohorts is heavily skewed.

The current study was dependent on accurately typing highly polymorphic microsatellite markers, using methodologies that allowed a relatively high throughput (multiplexing of multiple loci within single PCRs and capillary-based electrophoresis). This presented a greater degree of resolution than was previously attainable with the application of minisatellites, which are labour intensive to screen and require relatively large amounts of isolated DNA. However, despite the progress in methodologies, errors in scoring microsatellite alleles were identified. An effort should be made to identify possible errors in all allele-based data, irrespective of the application. Although errors may be apparent in parentage analyses (using assignment rather than maximum likelihood methods), whereby putative instances of mis-typing can be flagged and re-analysed, in population studies this is not the case. Therefore re-analysis of a sub-sample of individuals should be carried out in order to determine the reliability of data.

A problem that has been identified with microsatellites is the inability to easily standardise data across laboratories; Smith et al. (2003) reported that this was time-consuming and expensive. The marker of the future for intensely studied organisms may be the SNP. Smith et al. (2003) described how these were relatively inexpensive and rapid to assay in comparison to other markers such as microsatellites and allozymes. Although generally only bi-allelic (Schlötterer 2004), such that large numbers of loci must be examined to provide a level of information content comparable to that provided by more polymorphic markers, typing methodologies that negate the need for electrophoresis and have a high level of automation mean that assays are
relatively rapid and easily standardised across laboratories (Smith et al. 2003).

However, even with such sophisticated procedures, there is still the potential for error, the occurrence of which should be analysed as a matter of course.
Appendix

A1 Appendix

1.1 DNA extraction- modified chelex protocol

Based on a method by Estoup et al. (1996), modified by J Taggart and P Prodöhl.

Chelex solution: 10% chelex (Sigma) in autoclaved TE (10mM Tris, 1 mM EDTA, pH8.0), 0.1% SDS

1. Heat chelex mixture on a magnetic stirrer with stir bar to 60°C.
2. Remove a punch from the fin clip 2mm in diameter using a biopsy punch.
   Squeeze this dry with a tissue and place in a well of a 96-well plate.
3. Repeat for remainder of samples in the set to be processed
4. Add 3μl of proteinase K (10mg/ml) to each well.
5. Add 100μl of chelex solution to each well using a pipette tip with the end removed (to allow the chelex to be included)
6. Heat at 55°C deg in an oven for 3hours.
7. Spin the plate at 1000g for 1min.
8. Heat plate at 95°C for 10 min.
9. Spin plate at 1000g for 1min.
10. Store at -25°C.
11. Prior to use mix (shake while ensuring plastic cover is firmly attached) and spin at 1000g for 1min.
Appendix

1.2  DNA extraction using Wizard® SV Genomic DNA Purification System

(Promega, Cat. No. A2360)

Digestion Solution Master Mix
(total 275µl per sample, only 260µl used to allow for pipetting errors)

Solution Volume per sample

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei Lysis Solution</td>
<td>200µl</td>
</tr>
<tr>
<td>0.5M EDTA (pH8.0)</td>
<td>50µl</td>
</tr>
<tr>
<td>Proteinase K 20mg/ml</td>
<td>20µl</td>
</tr>
<tr>
<td>RNase A Solution 4mg/ml</td>
<td>5µl</td>
</tr>
</tbody>
</table>

1. Remove punch 3mm in diameter from sample using a biopsy needle. Squeeze dry with tissue and place in microtube.
2. Repeat with remainder of set of samples to be processed.
3. Add 260µl Digestion Solution Master Mix to each sample.
4. Shake until all samples are immersed in digestion solution.
5. Incubate at 55°C for 16 hours if adult tissue, and minimum of 3 hours if smolt tissue. Shake tubes half way through incubation.
6. Add 240µl of Wizard®SV Lysis Buffer to each microtube. Vortex.
7. Tip contents of each microtube into a Wizard® SV Minicolumn Assembly (one for each sample).
8. Spin at 13000g for 3 minutes.
9. Discard contents of lower part of Wizard® SV Minicolumn Assembly (the collection tube).
Appendix

10. Add 640µl Wizard®SV Wash Solution (with 95% ethanol added).

11. Spin at 13000g for 1min.

12. Remove contents of collection tube.

13. Repeat 3 times (a total of 4 washes).

14. Spin at 13000g for 3 min.

15. Transfer top of Wizard® SV Minicolumn Assembly to a microtube.

16. Add 100µl nuclease free water.

17. Leave for 2min at room temperature.

18. Spin at 13000g for 1min.

19. Add 50µl nuclease free water if adult sample, 30µl nuclease free water if smolt sample. Ensure all drops are in contact with bottom, rather than side, of the tube.

20. Leave for 2 min at room temperature.

21. Spin at 13000g for 3min.

22. Keep only the microtube with eluted DNA. Store at -25°C.
Appendix

1.3  *PCR components and thermocycling conditions*

Throughout the course of the study, PCR conditions were modified to take into account the changing semi-automated detection apparatus (i.e. ABI and Beckman fragment analyzers), different DNA extraction methodologies used and additions to the locus panel employed.

Non-fluorescent primers and ‘ABI’ fluorescent primers were ordered from MWG Biotech while Beckman fluorescent dyes were supplied by Proligo France SAS.

Other PCR components were supplied by ABGENE

i.e. Taq DNA Polymerase (5U / µL ) Cat No. AB-0192

- inc 10× Buffer II solution (100 mM Tris-HCl, pH8.3; 500 mM KCl) and separate 25 mM MgCl₂ solution

- dNTP set (high concentration 100 mM each) Cat No. AB-0315

The varying protocols are detailed below in Table A1 - A10. The products of PCRs described in Table A1- Table A4 were run on ABI fragment analyzers, products of PCRs described in Table A5-Table A10 were run on Beckman Coulter fragment analyzers. Template DNA for PCR conditions described in Table A1-A4 was extracted using either the phenol chloroform or chelex protocol. Template DNA for PCR conditions described in Table A5- Table A10 was extracted using a commercially available kit (the Wizard® SV Genomic DNA Purification System).

Unless otherwise stated, the concentration is the same for forward and reverse primers. Sequences of primers are detailed in Table 2.2 and 2.3.
Appendix

Table A1 PCR conditions for multiplex 1. Final reaction volume was 10µL. Steps 2-4 were repeated 24 times (25 cycles in total).

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1×</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td></td>
</tr>
<tr>
<td>MgCl₂</td>
<td>2.5mM</td>
<td></td>
</tr>
<tr>
<td>2201 (F&amp;R)</td>
<td>0.38µM</td>
<td>1 95 2min</td>
</tr>
<tr>
<td>1605 (F&amp;R)</td>
<td>0.08µM</td>
<td>2 95 30s</td>
</tr>
<tr>
<td>2210 (F&amp;R)</td>
<td>0.1µM</td>
<td>3 58 30s</td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td>4 72 50s</td>
</tr>
</tbody>
</table>

Table A2 PCR conditions for multiplex 2. Final reaction volume was 10µL. Steps 2-4 were repeated 24 times (a total of 25 cycles). Template DNA was extracted using either the phenol chloroform or chelex protocol.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1×</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td></td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td></td>
</tr>
<tr>
<td>410 (F&amp;R)</td>
<td>0.35µM</td>
<td>1 95 2min</td>
</tr>
<tr>
<td>407 (F&amp;R)</td>
<td>1.2µM</td>
<td>2 95 30s</td>
</tr>
<tr>
<td>One9 (F&amp;R)</td>
<td>0.25µM</td>
<td>3 58 30s</td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td>4 72 2min</td>
</tr>
</tbody>
</table>
Appendix

Table A3 PCR conditions for multiplex 3. Final reaction volume was 10µL. Steps 1-3 were repeated 7 times (total of 8 cycles). Steps 4-6 were then repeated 19 times (a total of 20 cycles). Template DNA was extracted using either the phenol chloroform or chelex protocol.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1×</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td>Step</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td>1</td>
</tr>
<tr>
<td>197 (F&amp;R)</td>
<td>0.15µM</td>
<td>2</td>
</tr>
<tr>
<td>202 (F&amp;R)</td>
<td>0.2µM</td>
<td>3</td>
</tr>
<tr>
<td>171 (F&amp;R)</td>
<td>0.25µM</td>
<td>4</td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

Table A4 PCR conditions for multiplex 4. Final reaction volume was 10µL. Steps 2-4 were repeated 24 times (a total of 25 cycles). Template DNA was extracted using either the phenol chloroform or chelex protocol.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1×</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td>Step</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td>1</td>
</tr>
<tr>
<td>404 (F&amp;R)</td>
<td>0.35µM</td>
<td>2</td>
</tr>
<tr>
<td>408 (F&amp;R)</td>
<td>0.6µM</td>
<td>3</td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>
Table A5 PCR conditions for multiplex 5, used for DNA from adults that returned in 2000 and 2001. Final reaction volume was 20µL. Steps 2 to 4 were repeated 29 times (total of 30 cycles) before steps 5 and 6 were carried out. 2µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td>buffer</td>
<td>1×</td>
<td>1</td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td>2</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td>3</td>
</tr>
<tr>
<td>2201 (F&amp;R)</td>
<td>0.38µM</td>
<td>4</td>
</tr>
<tr>
<td>1605 (F&amp;R)</td>
<td>0.2µM</td>
<td>5</td>
</tr>
<tr>
<td>2210 (F&amp;R)</td>
<td>0.12µM</td>
<td>6</td>
</tr>
<tr>
<td>410 (F&amp;R)</td>
<td>1µM</td>
<td></td>
</tr>
<tr>
<td>407 (F&amp;R)</td>
<td>1.2µM</td>
<td></td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td></td>
</tr>
</tbody>
</table>

Table A6 PCR conditions for multiplex 5, used for DNA from adults that returned in 2004 and for smolts that were collected in spring 2004. Final reaction volume was 20µL. Steps 2 to 4 were repeated 34 times (total of 35 cycles) before steps 5 and 6 were carried out. 2µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td>buffer</td>
<td>1×</td>
<td>1</td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td>2</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td>3</td>
</tr>
<tr>
<td>2201 (F&amp;R)</td>
<td>0.38µM</td>
<td>4</td>
</tr>
<tr>
<td>1605 (F&amp;R)</td>
<td>0.2µM</td>
<td>5</td>
</tr>
<tr>
<td>2210 (F&amp;R)</td>
<td>0.145µM</td>
<td>6</td>
</tr>
<tr>
<td>410 (F)</td>
<td>1µM</td>
<td></td>
</tr>
<tr>
<td>410 (R)</td>
<td>1.2µM</td>
<td></td>
</tr>
<tr>
<td>407 (F&amp;R)</td>
<td>1.2µM</td>
<td></td>
</tr>
<tr>
<td>Taq</td>
<td>0.045U/µL</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix

**Table A7 PCR conditions for multiplex 6 used for DNA from adults that returned in 2000, 2001, and 2004.** Final reaction volume was 20µL. Steps 2 to 4 were repeated 29 times (total of 30 cycles) before steps 5 and 6 were carried out. 2µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1x</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td></td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td></td>
</tr>
<tr>
<td>404 (F&amp;R)</td>
<td>1µM</td>
<td></td>
</tr>
<tr>
<td>408 (F&amp;R)</td>
<td>0.6µM</td>
<td></td>
</tr>
<tr>
<td>171 (F&amp;R)</td>
<td>0.3µM</td>
<td></td>
</tr>
<tr>
<td>One9 (F&amp;R)</td>
<td>0.15µM</td>
<td></td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µL</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

**Table A8 PCR conditions for multiplex 6 used for DNA extracted from smolts sampled in spring 2004.** Final reaction volume was 20µL. Steps 2 to 4 were repeated 34 times (total of 35 cycles) before steps 5 and 6 were carried out. 3µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>buffer</td>
<td>1x</td>
<td></td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td></td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td></td>
</tr>
<tr>
<td>404 (F&amp;R)</td>
<td>1µM</td>
<td></td>
</tr>
<tr>
<td>408 (F&amp;R)</td>
<td>0.5µM</td>
<td></td>
</tr>
<tr>
<td>171 (F&amp;R)</td>
<td>0.3µM</td>
<td></td>
</tr>
<tr>
<td>One9 (F&amp;R)</td>
<td>0.18µM</td>
<td></td>
</tr>
<tr>
<td>Taq</td>
<td>0.035U/µl</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>
Table A9 PCR conditions for multiplex 7 used for DNA extracted from adults sampled in 2000, 2001 and 2004. Final reaction volume was 20µL. Steps 1 to 3 were repeated 7 times (total of 8 cycles), then steps 4 to 6 were repeated 19 times (total of 20 cycles) before step 7 was carried out. 2µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td>buffer</td>
<td>×0.9</td>
<td>1</td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.2mM</td>
<td>2</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.5mM</td>
<td>3</td>
</tr>
<tr>
<td>197 (F&amp;R)</td>
<td>0.15µM</td>
<td>4</td>
</tr>
<tr>
<td>202 (F&amp;R)</td>
<td>0.36µM</td>
<td>5</td>
</tr>
<tr>
<td>401 (F)</td>
<td>0.45µM</td>
<td>6</td>
</tr>
<tr>
<td>401 (R)</td>
<td>0.5µM</td>
<td>7</td>
</tr>
<tr>
<td>Taq</td>
<td>0.025U/µl</td>
<td></td>
</tr>
</tbody>
</table>

Table A10 PCR conditions for multiplex 7 used for DNA extracted from smolts sampled in 2004. Final reaction volume was 20µL. Steps 2 to 4 were repeated 7 times (total of 8 cycles) then steps 5 to 7 were repeated 24 times (total of 25 cycles) before step 8 was carried out. 3µL of undiluted DNA was used.

<table>
<thead>
<tr>
<th>Component</th>
<th>Final concentration</th>
<th>Thermocycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Step</td>
</tr>
<tr>
<td>buffer</td>
<td>1×</td>
<td>1</td>
</tr>
<tr>
<td>each dNTP</td>
<td>0.23mM</td>
<td>2</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.7mM</td>
<td>3</td>
</tr>
<tr>
<td>197 (F&amp;R)</td>
<td>0.17µM</td>
<td>4</td>
</tr>
<tr>
<td>202 (F&amp;R)</td>
<td>0.44µM</td>
<td>5</td>
</tr>
<tr>
<td>401 (F&amp;R)</td>
<td>0.51µM</td>
<td>6</td>
</tr>
<tr>
<td>Taq</td>
<td>0.03U/µl</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>
Appendix

1.4 Example traces of multiplexed loci.

Example traces obtained for each multiplex are depicted in Figure A1 to Figure A13.

![Example trace of multiplex 1. From left to right, are loci Ssa2210, Ssa1605 and Ssa2201.](image)

Figure A1 Example trace of multiplex 1. From left to right, are loci Ssa2210, Ssa1605 and Ssa2201.
Appendix

Figure A2 Example trace of multiplex 2, showing all fluorors combined.

Figure A3 Example traces of multiplex 2, showing each fluor separately. Loci included are One9, Ssa410 and Ssa407 (from top to bottom).
Appendix

Figure A4 Example trace of multiplex 3, showing all fluors combined.

Figure A5 Example traces of multiplex 3, showing each fluor separately. Loci included are \textit{Ssa171}, \textit{Ssa202} and \textit{Ssa197} (from top to bottom).
Appendix

Figure A6 Example trace of multiplex 4, showing both fluors combined.

Figure A7 Example traces of multiplex 4, showing each fluor separately. Loci included are Ssa408 and Ssa404 (from top to bottom).
Appendix

**Figure A8** Example trace of multiplex 5, showing all fluors combined.

**Figure A9** Example traces of multiplex 5, showing each fluor separately. Loci included are, in top trace from left to right *Ssa2210, Ssa1605, Ssa2201*, middle trace *Ssa407*, lower trace *Ssa410*. 
Appendix

Figure A10 Example trace of multiplex 6, showing all fluors combined.

Figure A11 Example traces of multiplex 6, showing each fluor separately. Loci included are, in top trace from left to right One9 and Ssa171, middle trace Ssa404, lower trace Ssa408.
Appendix

Figure A12 Example trace of multiplex 7, showing all fluors combined.

Figure A13 Example traces of multiplex 7, showing each fluor separately. Loci included are Ssa401, Ssa197, and Ssa202 (from top to bottom).
Bibliography


Amos, W. and Balmford, A. 2001 When does conservation genetics matter? *Heredity* **87**: 257-265


Birt, T.P., Green, J.M. and Davidson, W.S. 1991 Contrasts in development and smolting of genetically distinct sympatric anadromous and nonanadromous Atlantic salmon,


Caballero, A. 1994 Developments in the prediction of effective population size. Heredity 73: 657-679

Cairney, M., Taggart, J.B. and Hoyheim, B. 2000 Characterization of microsatellite and minisatellite loci in Atlantic salmon (Salmo salar L.) and cross-species amplification in other salmonids. Molecular Ecology 9 (12): 2175-2178


CEFAS (The Centre for Environment, F.a.A.S. 2003 Research on migratory salmonids, eel


Dittman, A.H. and Quinn, T.P. 1996 Homing in Pacific salmon: mechanisms and ecological


Gage, M.J.G., Stockley, P. and Parker, G.A. 1995 Effects of alternative male mating


Gjerde, B. 1984 Response to individual selection for age at sexual maturity in Atlantic salmon. *Aquaculture* **38** (3): 229-240


Goudet, J. 1995 FSTAT (version 1.2) A computer program to calculate F-statistics. *Journal of Heredity* **86**: 485-486


London.


IUCN Red List of Threatened Species [www.iucnredlist.org](http://www.iucnredlist.org) Downloaded 10th November 2006


Lage, C. and Kornfield, I. 2006 Reduced genetic diversity and effective population size in an endangered Atlantic salmon (*Salmo salar*) population from Maine, USA. *Conservation Genetics* 7 (1): 91-104


Luikart, G., Allendorf, F.W., Cornuet, J-M. and Sherwin, W.B. 1998 Distortion of allele frequency distributions provides a test for recent population bottlenecks. The Journal of Heredity 89 238-247


Meirmans, P.G. 2006 Using the AMOVA framework to estimate a standardised genetic differentiation measure. *Evolution* **60** 2399-2402


Myers, R.A. 1984 Demographic consequences of precocious maturation of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences* **41** (9): 1349-


Skaala, O., Jørstad, K.E. and Borgstrom, R. 1996 Genetic impact on two wild brown trout (*Salmo trutta*) populations after release of non-indigenous hatchery spawners. *Canadian Journal of Fisheries and Aquatic Sciences* **53** (9): 2027-2035


Taggart, J.B. 2007 FAP: an exclusion-based parental assignment program with enhanced predictive functions. *Molecular Ecology Notes* **7** (3) 412-415


Taggart, J.B. and Ferguson, A. 1990b Hypervariable minisatellite DNA single locus probes for the Atlantic salmon, Salmo salar L. *Journal of Fish Biology* **37** (6): 991-993


Tessier, N., Bernatchez, L. and Wright, J.M. 1997 Population structure and impact of supportive breeding inferred from mitochondrial and microsatellite DNA analyses
in land-locked Atlantic salmon *Salmo salar* L. *Molecular Ecology* 6 (8): 735-750


Unwin, M.J., Kinnison, M.T., Boustead, N.C. and Quinn, T.P. 2003 Genetic control over survival in Pacific salmon (*Oncorhynchus Spp.): experimental evidence between and within populations of New Zealand chinook salmon (*O. tsawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 60 (1): 1-11


Wang, J.L. 2005 Estimation of effective population sizes from data on genetic markers. Philosophical Transactions of the Royal Society B-Biological Sciences 360 (1459): 1395-1409


Waples, R.S. 2002 Effective size of fluctuating salmon populations. Genetics 161 (2): 783-791


Waples, R.S. 2005 Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? Molecular Ecology 14 (11): 3335-3352


Watts, P.C., Saccheri, I.J., Kemp, S.J. and Thompson, D.J. 2006 Population structure and the impact of regional and local habitat isolation upon levels of genetic diversity of the endangered damselfly Coenagrion mercuriale (Odonata: Zygoptera). Freshwater Biology 51 (2): 193-205


Wolf, P. 1951 A trap for the capture of fish and other organisms moving downstream. Transactions of the American Fisheries Society 80: 41-45


Wright, S. 1931 Evolution in mendelian populations. Genetics 16 (2): 0097-0159


