# The adaptive diversity of Arctic charr (*Salvelinus alpinus*) in Scotland

Sam Fenton, BSc MRes

Submitted in fulfilment of requirements for the Degree of Doctor of Philosophy

School of Biodiversity, One Health & Veterinary Medicine

College of Medical, Veterinary & Life Sciences

University of Glasgow

February 2024

# Abstract

Intraspecific, within-species, diversity is critical to the continued existence of populations, species, and entire ecosystems and so the conservation of intraspecific diversity is increasingly valued in conservation policy. The Arctic charr (*Salvelinus alpinus*) represents one of the most highly diverse species in the world with high levels of phenotypic and genetic variation seen across its Holarctic range including the appearance of distinct trophic specialist populations. The emergence of such a vast range of phenotypes and genotypes makes the Arctic charr a key study system for exploring the process behind the emergence of biodiversity. The high diversity seen gives the species high conservation and scientific value but also makes defining important units of conservation to protect that diversity challenging, with questions about whether Arctic charr should even be considered a single species. Defining these units of conservation and understanding the factors underlying patterns of genetic diversity and differentiation requires large-scale genetic studies, currently lacking in the British Isles.

In this thesis, I investigated patterns of genetic differentiation and structuring across Scotland and the rest of the British Isles to define important units of conservation in Arctic charr. I highlighted instances of trophic specialist populations, known as ecotypes, as high priority for conservation action while questioning an existing taxonomy of endemic species which were delineated based on phenotypic characteristics. I investigated the factors the underlie the emergence of diversity highlighting ecosystem size as important in the extent of phenotypic and genetic diversity seen but also for underlying instances of trophic specialisation. Analysis of the colonisation history of Arctic charr revealed that the British Isles were colonised by multiple distinct ancestral sub-lineages and that glacial history had a major role in shaping patterns of genetic differentiation and diversity with regions still covered by ice during the colonisation period of the species. Finally, I investigated the genomics of key phenotypes that underlie repeated instances of trophic specialisation finding that only small numbers of SNPs were shared across replicates indicating the use of different genomic pathways to achieve similar phenotypes. This thesis represents an important body of work both for the conservation of Arctic char but more broadly our understanding of the emergence of diversity and the repeatability of evolution.

# Table of Contents

# List of Tables

# List of Figures

# Preface

**Chapter 2.** The tissue samples used in this chapter came from collaborators from the Universities of Bangor, Aberdeen, Highlands and Islands, the Marine Institute, and the Dee District Salmon Fishery Board along with existing samples at the University of Glasgow. Generation of genetic data including DNA extraction, library preparation, and subsequent processing were all performed by myself along with all analyses. The same genetic dataset was used in **Chapters 3 and 5**.

**Chapter 6.** This chapter centres around the re-analysis of an existing dataset of Arctic charr ecomorph pairs generated by Arne Jacobs at the University of Glasgow. Arne Jacobs generated the ddRADseq libraries, mapped them to the genome, and did the landmarking for head shape and body shape. All subsequent processing and analysis were conducted by me. The QTL database used in this chapter is based on a database originally collected by Arne and augmented to include more recent studies by myself using the same methodology before I mapped it to the *Salvelinus* sp. genome and ran all the subsequent analyses.

# Acknowledgements

I would first like to thank my supervisors, Colin Adams, Colin Bean, and Kathryn Elmer for their support throughout this PhD and under whom I have grown tremendously as a scientist. Without their wide range of knowledge, from Arctic charr to conservation policies, this PhD quite simply could not have been done. On a more personal level, their encouragement, kindness, and good humour have made what could have been an incredibly stressful 4 years into an incredibly rewarding and enjoyable experience. Doing a PhD has been long-standing dream of mine, and I will always be eternally grateful to them for giving me this opportunity.

I would like to thank my collaborators from across the British Isles for providing samples for the main body of work for the thesis. Collecting all these samples from all over Scotland and the rest of the British Isles simply would not have been achievable in the span of a single PhD. So, I would like to thank Sam Martin, Eric Verspoor, Ron Greer, Victoria Pritchard, Anthony Smith, Elvira de Eyto. and Sam Poultney for allowing this project to achieve everything that is has. I would also like to thank Kara Layton for her insight into Arctic charr over the last few years and for her help in running the genetic offset analysis.

I would also like to thank the various members of the roof labs who have been there during my PhD including Mark Sanda, Elizabeth Kilbride, Jenna Nichols, Nic Strowbridge, Peter Koene, Darren Hunter, Li Chao, Hongxin Xie, Kevin Schneider, Marco Crotti, Hans Recknagel, Molly Uzzell, John Smout, and many more I will have forgotten to mention. Not only have they all been a tremendous help with analyses and giving feedback on my work but also helped create a wonderful environment to work in and it was always a pleasure coming in each day. Special thanks must go to Arne Jacobs who really took me under his wing early on in my PhD and has remained a constant source of insight throughout the rest of my PhD. A similar thanks must go to Maria Capstick for teaching me all the ins and outs of my various lab work protocols and always being ready to help with my various lab-based queries. The roof labs simply would not run without you.

Thanks, must also go to my many long-standing friends located throughout the UK, both for the entertainment and laughs we have had down the years but also for their understanding and accommodation for my incredibly busy schedule. I have had a very work-focused mindset for the last few years and their understanding and support throughout that time period has been a huge help. Among them, special note must go to Liam German with

whom I have shared the experience of doing a PhD. Our weekly chats about all things PhD related and everything else have always been a highlight and I am extremely excited for the next opportunities for us both going forward.

Finally, I must thank my Mum and my sister Abby. As much as they might joke that I only come home to see the pets (Ruby in particular), I truly love anytime I get to see them and their support throughout these years has been tremendous. Mum, you always taught me to try my best and encouraged me to follow my dreams. I don't think I would be where I am today and have achieved all that I have without your dedication and love. Thank you for everything.

This thesis is dedicated to my late father, who passed away a few weeks after my initial submission. My dad was the reason I ever wanted to do science. He lit that fire in me and always encouraged me to follow my dreams. I never would have made it this far without everything he ever did for me. Love you always, Dad.

# Author's Declaration

This thesis, and all the work contained within it, was conducted between October 2019 and February 2024 by myself, unless otherwise stated. No part of this thesis has been submitted towards the fulfilment of any other degree.

Signature: _____

Printed name: Sam Fenton

# Definitions and Abbreviations

AAH: Ard Achadh
ABL: a'Bhaid-Luachraich
AGB: a'Garbh-bhaid Mor
AM: Ancient Migration
ARK: Arkaig
BCH: Bruicheach
BOD: Bodlyn
BHO: Braig Horrisdale
BIIS: British-Irish Ice Sheet
Bn: Benthivore
BRO: Brora
BSLMM: Bayesian sparse linear mixed model
BUN: Bunaveela
CLD: Calder
CON: Conniston
COL: Coulin
DAM: Damh
DBH: Dubh
ddRADseq: Double digest restriction-site associated DNA sequencing
DOI: Doine
DON: Doon
DTC: Duntelchaig
DUG: Dughaill
ERN: Earn
ERT: Ericht
ESU: Evolutionarily significant unit
FAH: Fannich
FIN: Finn
GAR: Garry
gLGM: global Last Glacial Maximum
GLM: Generalised linear model
GO: Gene ontology
GWAS: Genome-wide association study
HA: Hydrometric area
IM: Isolation with migration
IMchange: Isolation with change in migration rate
IMint: Isolation with migration and introgression
INS: Insh
KID: Kindrum
LAG: Laggan
LCR: Luichart
LCY: Lochy
LMM: Linear mixed model
LOC: Loch
LUB: Lubnaig
LVT: Langavat
LLS: Loch Lomond Stadial
MAR: Maree
MED: Meadie
MLT: Mealt
MOE: More
MRI: Morie
MRK: Merkland
mtDNA: Mitochondrial DNA
MU: Management unit
NAV: Naver
ND1: NADH dehydrogenase 1
NBR: Nam Brac
OSG: Osgaig
PAD: Padarn
Pl: Planktivore
Ps: Piscivore
QTL: Quantitative trait loci
RAN: Rannoch
SC: Secondary contact
SCint: Secondary contact with introgression
Si: Strict isolation
SEA: na Sealga
SEI: Seil
SHN: Shin
STK: Stack
SNP: Single nucleotide polymorphism
TEG: Treig
TRF: Tarff
TUM: Tummel
UAI: Uaine

# Chapter 1: General Introduction

## 1.1 Protection of biodiversity

Biological diversity, biodiversity, broadly refers to the variability among living organisms (Sepkoski, 1997; National Research Council, 1999). One of the key evolutionary processes behind the origin and maintenance of biodiversity is natural selection (Nosil, 2012). Adaptation to local environment through natural selection plays a major role in the generation of new phenotypes and genotypes and thus biodiversity. The processes that lead to the emergence of phenotypic and genetic diversity may also lead to the formation of new species, i.e. speciation (Seehausen et al., 2014). For example, divergent natural selection between populations in different environments may lead to selection against hybrids between said populations and the development of reproductive isolation, a key stage in speciation (Nosil et al., 2009). Crucially though the processes underlying the emergence of diversity are not always the same as those that lead to instances of speciation (Butlin et al., 2001). How the variation of phenotypes and genotypes is partitioned across a species can vary. In many species, for example those with continuous distributions, phenotypic and genetic variation is continuous across populations while in species with disjunct distributions this variation can be discontinuous across populations. This discontinuous variation can appear in the form of distinct populations that have adapted to different local environments, often referred to as ecotypes, which can represent an early stage to potential speciation (Ravinet et al., 2016).

Biodiversity exists across different temporal and hierarchical levels. Across its hierarchical continuum, biodiversity exists on levels as small as genes all the way up to whole biomes (Allendorf et al., 2012). Within this spectrum exists four major levels of biodiversity: genes, populations, species, and ecosystems (Hughes et al., 1997; Bowen, 1999; Luck et al., 2003) which all need to be protected for proper conservation of biodiversity to occur. While conservation across all four of these levels is important, protection of diversity on the population level may be the most critical (Luck et al., 2003). The emergence of diversity within and across populations, in the form of new phenotypes and genotypes, is crucial to the continued existence of theses populations but also the species itself. Intraspecific, within-species, variation is the fuel upon which evolution depends (Schindler et al., 2010; Le Coeur et al., 2021) and without it, species would live and die by their current form.

Conservation across all four of the major hierarchical levels of biodiversity is critical due to how they interact with one another. For example, each species in an ecosystem provides an important functional role for that ecosystem. Some species can even strongly influence community structure and ecosystem function by extension. These species are known as ecosystem engineers and either through their own physical structure or by transforming the state of living or non-living material (Jones et al., 1994; Hastings et al., 2007; Nicastro et al., 2020). Species of mussel (*Brachidontes*) modify their environment through their mussel beds creating habitats for a range of species, with the species inhabiting them varying depending on the environment created by each mussel species (Arribas et al., 2013; Nicastro et al., 2020). As such, the loss of a species can drastically impact the functioning of the ecosystem as a whole (Cardinale et al., 2000). Similarly, only protecting small amounts of suitable habitats for a species can lead to small population sizes, where genetic drift can hinder natural selection from purging deleterious alleles, which can threaten the survival of the species (Lynch et al., 1995; Robinson et al., 2023). Research has shown that intraspecific variation can have large ecological effects on primary productivity, species co-existence, and ecosystem resilience and recovery (Reusch et al., 2005; Des Roches et al., 2018). For example, variation in gill raker size and spacing in three-spine stickleback (*Gasterosteus aculeatus*), which indicates different ecological specialisations, affects prey community structuring and primary productivity (Harmon et al., 2009).

Biodiversity also exists across three major temporal components, past, present, and future (Sepkoski, 1997; Bowen, 1999), each of which are crucial to protect. Past biodiversity often refers to the different ancestral lineages that exist within and across species while contemporary/present diversity refers to current patterns of genetic and ecological variation and their influence on ecosystems (Allendorf et al., 2012). Future biodiversity, the ability of species and populations to adapt to future environments, is arguably the most important of the three. This is also the most difficult to identify as it requires the ability to predict evolution and future adaptive radiations (Bowen, 1998) which goes beyond just the current diversity in genes and includes the generation of novel diversity needed for future adaptation. This needs both the protection of populations and species but also the key ecological and evolutionary processes behind the emergence of diversity, both on level of genes and the external environmental factors that drive variation (Allendorf et al., 2012). This highlights why we need to work across all the major hierarchical levels in order to properly protect biodiversity.

Managing across the different hierarchical levels and temporal elements can be challenging as it requires an understanding of the genes and processes the underlie local adaptation, how populations are differentiated from one another, the ancestral lineages that make up a species, and the external environmental factors that drive the emergence of diversity which are all often complex and interlinked. These problems are magnified in highly diverse species, as while the high diversity allows them to have high ecological versatility, it also makes the identification of key population units to conserve difficult and the boundaries on what should be considered a distinct species complicated (Kocher, 2004; Allendorf et al., 2012).

## 1.2 Intraspecific conservation units

While much of the conservation focus is still placed at the species level (Coates et al., 2018), there is increasing effort to recognize major elements of intraspecific diversity for conservation action to properly protect species and their adaptive potential and by extent functioning ecosystems (Hughes et al., 1997; Funk et al., 2012). In recognition of this, conservation management now frequently involves the delineation of intraspecific conservation units to manage and protect intraspecific diversity (Fraser and Bernatchez, 2001; Bonin et al., 2007; Allendorf et al., 2012). This often requires the delineation of multiple genetically distinct populations/population units to protect (Hughes et al., 1997; Funk et al., 2012; Allendorf et al., 2012). Maintenance of multiple genetically distinct population across multiple distinct local environments can help protect a range of local adaptations, as well as the possibility of adaptation to future changes, thereby maintaining the processes behind biodiversity. How diversity is distributed in a species will have major implications for how it should be conserved, for example ecotype populations in disjunct distributions which are adapted to different local environments can make for clear population units for conservation action (Jacobs et al., 2020).

A number of different types of intraspecific conservation units have been defined and used to achieve the aim of protecting biodiversity, each with slightly different focuses (Moritz, 1994; Palsbøll et al., 2007; Minter et al., 2021). For example, gene conservation units (GCUs) focus on the preservation of genetic diversity in designated areas and are commonly used in species of flora, while evolutionarily significant units (ESUs) focus on the conservation of genetically distinct populations in-situ and are more typically used in species of fauna (De Guia and Saitoh, 2007; Minter et al., 2021). These different conservation units often have multiple different definitions which place differing emphasis

on the different elements of biodiversity (Allendorf et al., 2012). Some definitions of an ESU focus on the need to identify reciprocal monophyly which places most of the emphasis on past diversity in the form of identifying distinct lineages (Moritz, 1994) while others promote the idea of exchangeability, i.e., that individuals can be moved between populations and perform the same ecological role, where populations forming unique roles should be considered ESUs (Crandall et al., 2000).

The existence of multiple definitions highlights both that conservation efforts often have different focuses for which a single universal definition does not fit but also that each definition has their own limitations. For example, the idea of exchangeability is highly impractical for species where common garden or reciprocal transplant experiments cannot easily be performed to test exchangeability (Allendorf et al., 2012). Similarly, a focus on reciprocal monophyly can overemphasise small, isolated populations as evolutionarily significant due to high levels of genetic drift leading to them becoming monophyletic quickly (Paetkau, 1999; Allendorf et al., 2012). While ESU definitions still vary to an extent, many focus on identifying 1) reproductive isolation and 2) ecological or adaptive uniqueness to ensure the future evolutionary potential of the species with the latter usually requiring concordance across multiple data types, e.g., genetic, phenotypic, and geographic data (Waples, 1991; Allendorf et al., 2012). This consensus ensures recognition of historical distinctiveness while also having a key adaptive element, with the longer the isolation has been and the more different the environment is increasing the likelihood a population represents an ESU, while also not relying on a single criterion like reciprocal monophyly.

While this consensus is a reasonable and applicable approach, it places the emphasis greatly on strongly established diversity. That is to say, to be considered a conservation unit worth maintaining it must already show substantial genetic differentiation, often used as a proxy for reproductive isolation, and adaptive and ecological distinctiveness in concordance with other data types. Evolution and speciation exist on a continuum and as such there exists examples where phenotypic differences between populations have been established but this is not yet matched by notable genetic differentiation. For example, instances of recent ecotype divergence (Jacobs et al., 2020) may go undervalued under many conservation unit definitions despite their clear importance.

## 1.3 Ecological speciation and ecotype divergence

Understanding how populations diverge and new species arise is a key question in evolutionary biology (Seehausen et al., 2014). The maintenance of divergence between

populations and species is often dependent on the formation of reproductive barriers with the level of reproductive isolation an indicator of the extent of speciation (Nosil et al., 2009; Nosil, 2012). These barriers are either pre-zygotic (before mating) or post-zygotic (after mating) with pre-zygotic barriers and extrinsic post-zygotic barriers generally thought to evolve quicker than intrinsic post-zygotic barriers (Hendry et al., 2007; Egan et al., 2012; Seehausen et al., 2014). Pre-zygotic barriers generally involve individuals from different populations no longer considering each other mate candidates, for example through the sexual selection seen in many bird species (Edwards et al., 2005), or restrictions to the opportunities for mating between populations, due to the use of different habitats/spawning sites or different reproductive cycles (Feder et al., 1994; Fillatre et al., 2003; Nosil et al., 2006). Post-zygotic barriers can either be intrinsic, e.g., that mating forms inviable offspring through genetic incompatibilities (Presgraves, 2010), or extrinsic for example in the form strong selection against hybrids through divergent or disruptive selection (van der Sluijs et al., 2008).

Geographical isolation often acts as key reproductive barrier in many examples of speciation (Coyne and Allen Orr, 2004). Geographical barriers to gene flow can allow for populations to accumulate difference between them, either through natural selection or genetic drift, or allow for the maintenance of existing differentiation (Loretán et al., 2020). Instances where the differentiation of populations in geographical isolation leads to divergence into different species is referred to as allopatric speciation (Butlin et al., 2008) and is considered to be the most common mode of speciation (Coyne and Allen Orr, 2004). However, geographical isolation is not essential for speciation to occur, with examples of speciation occurring between populations with overlapping ranges, referred to as sympatric speciation, or those in close geographic proximity, parapatric speciation (Via, 2001; Niemiller et al., 2008). While there is evidence for cases of strict sympatric divergence, that is no geographic barriers to no gene flow at any time, many cases of supposed sympatric speciation involve secondary contact between distinct colonising lineages which initially diverged in allopatry, known as secondary contact divergences (Bolnick and Fitzpatrick, 2007; Rougeux et al., 2017; Poelstra et al., 2018).

While geography is often important in allowing the divergence of populations, the key evolutionary processes that lead to speciation include natural selection and sexual selection (Coyne and Allen Orr, 2004). The process by which adaptation to different ecological environments leads to the development of reproductive isolation is known as ecological speciation (Nosil, 2012). Populations living in different ecological environments will likely

experience divergent, ecological-based, natural selection due to the need for adaptation to different resources and habitats, for example different wing patterns in *Heliconius* butterflies or phenotypic differences between beach and river spawning sockeye salmon (*Oncorhynchus nerka*) (Jiggins, 2008; Hendry, 2009). Instances of these ecologically diverged populations existing within a species are often described as ecotypes. Hybrids between these populations may experience reduced fitness as they are not properly adapted to either environment (Hendry et al., 2007; Gow et al., 2007). This in turn would lead to reduced gene flow between populations and the development of reproductive barriers between them (Filchak et al., 2000; Garduño-Paz et al., 2010). Examples of ecological speciation have been documented in numerous species and is at the centre of many adaptive radiations, where a single ancestral lineage rapidly diversifies to a range of local environments evolving diverged traits to occupy those habitats (Rundell and Price, 2009). Ecological speciation occurs on a spectrum of divergence (Hendry, 2009) from continuous adaptive variation with no reproductive isolation (Bolnick and Lau, 2008) to clear adaptive differences with irreversible reproductive isolation, i.e., genetic incompatibilities for hybrids (Rogers and Bernatchez, 2007).

## 1.4 The genomic underpinnings of local adaptation, reproductive isolation, and ecological speciation

A major goal of evolutionary biology is identifying the loci that are responsible for local adaptation, reproductive isolation, and ecological speciation enabled by advancements in genomic techniques (Cruickshank and Hahn, 2014). In cases where populations diverged recently or are diverging under gene flow, for example in cases of sympatric or parapatric speciation, it is generally expected for the genome to show highly heterogenous patterns of genetic differentiation (Nadeau et al., 2012). In these instances, most of the genome will show low levels of genetic differentiation while a small number of regions show high levels of differentiation or fixation between populations which are referred to as genomic islands (Wolf and Ellegren, 2017; Hejase et al., 2020). These islands are thought to be associated with genes and regions underlying local adaptation and reproductive isolation and can arise by a number of mechanisms (Turner et al., 2005; Yeaman, 2013). Genomic islands can form through genomic hitchhiking where a key adaptive locus under divergent selection between populations reduces migration for surrounding loci resulting in higher genetic differentiation (Barton, 2000). Large-effect loci, loci which have a large effect on the related trait, have an increased chance of contributing to local adaptation in the face of gene flow as they resist the homogenising effect of gene flow (Yeaman and Whitlock, 2011). Genomic islands can

consist of multiple adaptive loci which form tight linkage to prevent recombination breaking up favourable combinations of alleles (Yeaman, 2013). Colocalization of adaptive loci, or quantitative trait loci (QTLs), has been shown in a number of species, for example in various cichlid fish, and acts as a key mechanism by which populations can rapidly diverge (Fruciano et al., 2016; Hahn et al., 2017). Genomic islands can also arise through reduced diversity and background selection in regions with low recombination rate or through introgression of adaptive alleles across populations or erosion of larger regions of high differentiation following secondary contact between lineages (Cruickshank and Hahn, 2014; Seehausen et al., 2014; Yeaman et al., 2016).

A number of different statistics can be used to measure divergence between populations across the genome in order to identify regions of high differentiation and the selective forces acting upon them (Cruickshank and Hahn, 2014). The extent of genetic differentiation across genome provides important information on the likely scenario under which the populations diverged, for example high genetic differentiation across much of the genome implies the populations initially diverged under allopatry, or involved secondary contact, while low genome-wide genetic differentiation indicates a recent sympatric divergence (Seehausen et al., 2014; Jacobs et al., 2020). Measures of divergence can be described as either measures of relative or absolute divergence (Cruickshank and Hahn, 2014). The most common measures of relative genetic divergence between populations are based on Wright's $F_{ST}$ (Wright, 1931) and use normalised difference in allele frequencies between populations to determine the extent of genetic differentiation across the genome and at specific loci. $F_{ST}$ and $F_{ST}$-like measures can be inflated when genetic diversity is low and are dependent on levels of within-population variation (Cruickshank and Hahn, 2014). Low genetic diversity can be the result of small population sizes, population bottlenecks, and strong founder's effects (Li and Roossinck, 2004) but regions of low recombination rate and/or high gene diversity can also reduce within-population variation and inflate values of $F_{ST}$ (Seehausen et al., 2014). These regions can occur due to background selection against deleterious mutations, selective sweeps, or the direct influence of recombination on genetic diversity (Charlesworth, 1994; Nielsen et al., 2005; Spencer et al., 2006; Hwang and Wang, 2023). To account for these factors, measures of absolute divergence, such as $D_{XY}$, use the average number of pairwise differences between sequences from two populations, excluding all comparisons within the same population, and can be corrected for local mutation rate (Nei, 1987; Cruickshank and Hahn, 2014). Consequently, this prevents islands of high differentiation in $D_{XY}$ forming due to linked selection although $D_{XY}$ can be influenced by ancestral diversity and substitution rate (Begun et al., 2007; Cruickshank and Hahn, 2014).

Most studies will employ multiple metrics such as $F_{ST}$ and $D_{XY,}$ as they are independent of one another, to untangle divergent selection and gene flow from neutral processes (Seehausen et al., 2014).

Studies looking into the regions underlying local adaptations suggest that the same genes or genomic regions can show repeated divergence between replicates populations or species experiencing similar environmental conditions (Elmer and Meyer, 2011) although how common this is across replicates remains unclear.

## 1.5 Parallel evolution and its genomic underpinnings

The repeated appearance of the same phenotypes across ecosystems suggests both that similar environments impose similar selective pressures and that only a few phenotypic solutions are favoured in response (Elmer and Meyer, 2011; Bolnick et al., 2018). This implicates an important role for natural selection in evolution as other mechanisms such as de novo mutations and genetic drift are unlikely to repeatably produce the same phenotypes (Conte et al., 2012). These repeated phenotypes are seen in numerous species and are often cited as examples of parallel evolution (Colosimo et al., 2005; Elmer et al., 2014; Soria-Carrasco et al., 2014; Auer et al., 2018). Commonly these repeated phenotypes take the form of ecotype populations which have specialised to similar environments such as the repeated divergence of benthic-limnetic ecotypes in a number of freshwater fish species (Baillie et al., 2016; Rivas et al., 2018; Jacobs et al., 2020). How similar the direction and magnitude of phenotypic change are across replicates can be strongly influenced by the degree of environmental and genetic similarity and as such variable patterns of phenotypic parallelism can be seen across replicates (Oke et al., 2017; Stuart et al., 2017).

While loci underlying local adaptation have been identified in a number of species (Fruciano et al., 2016), the degree to which the same genomic underpinnings are used across replicates and species to create those phenotypes is less known (McGirr and Martin, 2018; Magalhaes et al., 2021; Salisbury and Ruzzante, 2021). That is to say, do parallel phenotypes commonly have parallel genomic bases with selection acting upon the same genomic architecture, i.e., when parallel phenotypes result from the same nucleotide mutation in the same gene, or are parallel phenotypes arising through alternative genetic pathways, i.e., from different mutations in the same gene or through mutations in different genes resulting in the same phenotypes (Elmer and Meyer, 2011; Rivas et al., 2018). Examples from various species suggest variable patterns of shared genomic underpinnings with some replicates showing high levels of genetic parallelism while others indicate that only a few key loci are

shared across replicates and in some cases across multiple species (Holliday et al., 2016; Rivas et al., 2018; Salisbury and Ruzzante, 2021). Studies on three-spine stickleback show that loci underlying parallel evolution of marine and freshwater ecotypes in the Eastern Pacific were not shared in replicates found elsewhere in the species distribution while in *Littorina saxatilis* replicated instances of crab and wave ecotype divergences show little evidence of genetic parallelism despite being found less than 10 km apart (Ravinet et al., 2016; Fang et al., 2020).

The response to selection for each replicate depends on their respective genetic architecture, e.g., recombination and mutation rates and epigenetic modifications, which can vary among populations and across the genome (Bolnick et al., 2018). A shared genomic architecture may facilitate the repeated accumulation of differences in the same genomic regions (Seehausen et al., 2014). The probability of parallel genomic underpinnings of phenotypes declines with increasing time to the last common ancestor between the populations, lineages, or species in question (Conte et al., 2012). The degree of genomic parallelism will also depend on how polygenic the adaptive phenotype is, with highly polygenic phenotypes less likely to share the same underpinnings as multiple different genetic pathways can result in similar phenotypes (Elmer and Meyer, 2011; Salisbury and Ruzzante, 2021). Understanding the predictability of evolution requires thorough investigation of the environments each replicate inhabits, their genetic architecture, and their histories.

## 1.6 Colonisation and glacial history

A key contributor to patterns of genetic variation and the adaptive potential of a species is the distinct ancestral lineages that make up their contemporary distribution (Roberts and Hamann, 2015). While the ranges of some species seem to have arisen from a single colonising lineage, particularly in high isolated regions (Csapó et al., 2020; Woods et al., 2021), many contemporary distributions are made up of a number of ancestral lineages/sub-lineages (Bernatchez, 2001). These distinct lineages often diverged in allopatry to one another before acting as the sources of colonisation. This is particularly true for temperate species existing in regions that experienced ice coverage during previous glacial cycles (Clark et al., 2012). The harsh glacial conditions forced many temperate flora and fauna to exist only in isolated safe havens known as glacial refugia (Hewitt, 1999; Hewitt, 2000). For example, the Balkan, Italian, and Iberian peninsulas are well known as key glacial refugia in Europe for many temperate species while the Brittany peninsula has been

identified as a refugium for many marine and coastal species (Taberlet et al., 1998; Hoarau et al., 2007). The isolated nature of glacial refugia means that populations existing in different refugia can differentiate over time, due to the lack of gene flow, through natural selection and genetic drift (Hewitt, 1996; Baines et al., 2004; Angst et al., 2022). As such, each population may have gained key mutations that influence the adaptive potential of that lineage. At the end of glacial conditions, these differentiated populations then acted as the major lineages which colonised contemporary distributions.

When looking at colonisation history some species, particularly those with disjunct distributions, show evidence that the different lineages that make up their distribution remained geographically separate and colonised different parts of their range (Crotti et al., 2020; Crotti et al., 2021). Conversely other species show evidence that different colonising lineages came into contact with one another (Finnegan et al., 2013). A number of different outcomes can occur at these points of contacts. The two lineages may have mixed resulting in derived populations that show high genetic diversity, described as genetic diversity hotspots (Chiocchio et al., 2021). The two lineages may have remained separate to one another and diversified to different parts of the shared environment as seen in many examples of ecotype divergence (Jacobs et al., 2020). The identification of the key ancestral lineages that make up the contemporary distribution of a species is vital in helping us protect biodiversity and understand the future adaptive potential of species.

## 1.7 Arctic charr (*Salvelinus alpinus*)

Fish that inhabit recently glaciated systems, for example lakes that formed at the end of the last ice age, have often been used in studies on local adaptation, parallel evolution, and ecological speciation (Schluter, 1996; Jonsson and Jonsson, 2001). A common feature of postglacial lakes is the appearance of closely related ecologically and often genetically distinct ecotype/species pairs that have arisen due to rapid ecological speciation since the last ice age with these co-existing trophic specialists most often found in salmonid species (Rogers and Bernatchez, 2007; Gow et al., 2007; Jacobs et al., 2020). Commonly these trophic specialists are diverged along the benthic-limnetic axis with benthivorous ecotypes, that feed in the littoral or profundal habitats, and planktivorous and piscivorous ecotypes that feed on zooplankton and predominantly other fish species respectively (Schluter, 1996; Elmer, 2016). Of all the salmonid species, Arctic charr shows particular high phenotypic and genetic diversity, making them an ideal system for studying the genomic basis of adaptation

and ecological speciation, the drivers behind the emergence of diversity, and how to protect the diversity of a highly diverse species (Jonsson and Jonsson, 2001; Klemetsen, 2013).

The Arctic charr is the most northerly freshwater fish species in the world being found as far north as Ellesmere Island in Canada (Lehnherr et al., 2018). The species has a wide-ranging distribution across the Holarctic which spans North America, the Arctic, Europe, and Siberia (Reist et al., 2013). The species distribution consists of five major lineages (Acadian, Arctic, Atlantic, Bering, and Siberian) which correspond to discrete regions and diverged from each other during the Pleistocene glaciation (Brunner et al., 2001; Moore et al., 2015). More recent divergences are known to have occurred within the lineages during the last ice age with these distinct sub-lineages then acting as the sources for the colonisation after the end of the last ice age (Jacobsen et al., 2022). The northern part of the species distribution is a mix of anadromous and landlocked-resident (non-anadromous) populations with many populations consisting of both phenotypes (Jørgensen and Johnsen, 2014). This variation in phenotype is driven by a number of factors including the trade-off between freshwater productivity and migratory distance to the sea (Finstad and Hein, 2012). In the southern parts of the distribution, all populations are non-anadromous including the British Isles which my research focuses on.

Arctic charr shows incredibly high phenotypic and genetic variability across populations to the extent that it has been described as the most variable known vertebrate species in the world (Klemetsen, 2013). This includes variation in morphology, meristic characteristics, colouration, spawning, foraging, and diet with much of this variation seen in the British Isles (Maitland et al., 2007; Maitland and Adams, 2018). Patterns of genetic differentiation between populations are highly mosaic with nearby populations often showing similar levels of genetic differentiation to distant populations suggesting limited levels of gene flow between many populations particularly true for non-anadromous parts of the range (Wilson et al., 2004; Kapralova et al., 2011).

Many of the phenotypic and genetic differences seen are related to different trophic specialisations with ecotype divergences well known in the species (Jonsson and Jonsson, 2001; Doenz et al., 2019). The extent of phenotypic and genetic divergence between ecotypes varies considerably across replicates and is strongly influenced by ancestral history with sympatric ecotypes generally more similar to each other than ones that arose through secondary contact (Jacobs et al., 2020). Reproductive isolation between ecotypes is mostly

maintained through pre-zygotic barriers, such as differences in spawning time and habitat or size-assortative mating (Alekseyev et al., 2002; Garduño-Paz et al., 2012).

The high diversity seen in Arctic charr has also led to debates about whether this represents a single highly diverse species or a number of different species or a species complex (Klemetsen, 2010). Within the British Isles, the high phenotypic variation seen led to the classifications of a number of endemic species during the Victorian age (Adams and Maitland, 2007). Over time consensus became that a single highly diverse species exists in the British Isles before much of the old taxonomy was revived (Kottelat and Freyhof, 2007). The classification of different endemic species has a major impact for the protection of biodiversity with many of the endemic species being isolated to one or a small number of locations. The relationship between these phenotypic classifications and levels of genetic differentiation have yet to be determined and so the validity of this taxonomy still remains a question.

The high diversity seen gives Arctic charr, and the environments it inhabits, high conservation and evolutionary value and so increasing interest has been put into defining intraspecific conservation units, such as ESUs, within the species. However, the high diversity seen poses challenges on how to properly define these units and a lack of large-scale genetic studies limits our understanding of patterns of genetic differentiation and adaptive potential needed to assess how to define units.

## 1.8 Thesis outline

In this thesis, I will attempt to answer questions about how we handle the conservation of highly diverse species using the highly diverse salmonid species Arctic charr (*Salvelinus alpinus*) as a study species. I do this by: investigating the key genes that underlie local adaptation, identifying the key contemporary and historical processes that underlie patterns of diversity and differentiation, identifying the key historical lineages that make up the species, delineating appropriate intraspecific conservation units, and assessing the taxonomy of putative endemic species.

### *1.8.1 Chapter 2: Defining conservation units in a highly diverse species: A case study from Arctic charr*

The aim of the second chapter is to define important conservation units within a highly diverse species, the Arctic charr. We first analysed patterns of genetic differentiation and structuring across populations to determine a criterion for defining intraspecific

conservation units, in tandem with phenotypic and geographical data, and assess the existing taxonomy of endemic species in the British Isles. We then assessed the susceptibility of different populations to loss due to various threats including climate change and predation to determine which populations may be of higher priority for conservation action.

### *1.8.2 Chapter 3: Increased ecological opportunity leads to higher genetic diversity and probability of trophic specialisation in Arctic charr*

The third chapter focuses on identifying the extrinsic factors that drive the emergence of diversity, both phenotypic and genetic. We first analysed patterns of genetic and phenotypic diversity to determine how environmental factors such as ecosystem size influence levels of diversity. We then investigated the environmental factors that underlie instances of trophic specialisation by using data from all 187 lakes known to contain Arctic charr in Scotland. We then used these findings to create a predictive model to identify potential unknown ecotype divergences in the lesser studied populations in this range.

### *1.8.3 Chapter 4: How glaciation impacted evolutionary history and contemporary genetic diversity of flora and fauna in the British Isles*

The fourth chapter consists of a literature review which summarises known changes in ice coverage in the British Isles from the global Last Glacial Maximum (*ca.* 27 ka) through to the start of the Holocene (*ca.* 11.5 ka). We then discuss how ice coverage affected patterns of contemporary genetic diversity, through the formation of glacial refugia and distinct ancestral lineages, providing various case studies demonstrating its influence.

### *1.8.4 Chapter 5: The historical patterns that have shaped contemporary genetic differentiation across populations of Arctic charr in Scotland*

The fifth chapter focuses on understanding how patterns of historical colonisation and glacial history have influenced contemporary patterns of genetic differentiation in Arctic charr. We first investigated the colonisation history of Arctic charr in the British Isles using sequences from across the Holarctic to identify the key colonising lineages/sub-lineages. We then investigated contemporary patterns of genetic differentiation to determine how colonisation history and the varying patterns of glacial history have influenced contemporary patterns.

### *1.8.5 Chapter 6: Genomic underpinnings of head and body shape in Arctic charr ecomorph pairs*

The sixth chapter focuses on the genomic underpinnings behind phenotypic divergence and whether it is consistent across replicates and species. We identified loci

associated with variation in head and body shape morphology between benthivorous-planktivorous ecotype pairs of Arctic charr before determining whether these loci are shared consistently across replicates and whether they are under the same selective pressures in each pair. We then investigated whether the same genomic regions were important to phenotypic divergence across multiple salmonid species.

### *1.8.6 Chapter 7: General discussion*

The final chapter synthesises the findings of the previous chapters and discuss their contribution to our understanding of the genomics and drivers of diversity and ecological specialisation, the conservation of highly diverse species, and the impact of historical processes on contemporary differentiation.

# Chapter 2: Defining conservation units in a highly diverse species: A case study for Arctic charr

## 2.1 Abstract

Defining appropriate conservation units is crucial to the protection of populations, species, and biodiversity. Accurate delineation of conservation units is particularly difficult in highly diverse species like the Arctic charr (*Salvelinus alpinus*) which shows high levels of phenotypic and genetic variation within and across populations. The criteria used to define units of conservation can lead to inaccurate assessment of conservation priority with attempts to quantify important population units having led to the classification of a number of putative endemic species in the British Isles based solely on morphometric analysis. The delineation of appropriate units of conservation and identifying vulnerable populations within the species in Scotland has been limited by a lack of national-scale genetic studies of populations. To assess the taxonomy of putative endemic species and identify evolutionary significant and vulnerable populations for conservation, we conducted a national-scale genetic study of Arctic charr populations in Scotland (N=64 populations). Our analyses showed little support for the endemic species classifications, with these putative species found sporadically across our phylogenetic trees and patterns of differentiation within the range seen for the rest of Arctic charr. We found that most populations represented their own distinct genetic clusters through admixture analysis with low levels of gene flow resulting in high genetic differentiation. We then used our genetic data along with phenotypic and geographic information to identify populations of evolutionary significance. Finally, we assessed the vulnerability of populations to a range of threats including climate change and predation to identify populations to prioritise for conservation.

## 2.2 Introduction

The management and protection of biodiversity requires both the delineation of appropriate units of conservation and an assessment of their vulnerability to damage or loss. While conservation action is frequently directed at the species level, an increasing body of work has demonstrated the importance of intraspecific, within-species, diversity to ecological versatility, ecosystem recovery, and adaptation to changing environments (Reusch et al., 2005; Schindler et al., 2010; Des Roches et al., 2018; Le Coeur et al., 2021). For many species, particularly highly diverse ones, conservation management at a species level is insufficient for protecting vital diversity and it is becoming increasingly common practice to define intraspecific conservation units (Mee et al., 2015; Forester et al., 2022). However, there are many different functional definitions of what constitutes a biologically meaningful conservation unit and the criteria used to define them can drastically impact the assessment of conservation importance (Mee et al., 2015; Mussmann et al., 2020; Minter et al., 2021). Issues about the criteria used to define units of conservation can be seen in the Kottelat and Freyhof (2007) review of the freshwater fishes of Europe which often used the single criterion of phenotypic characteristics to delineate important evolutionary units leading to a number of highly debated species classifications (Adams and Maitland, 2007; Etheridge et al., 2012; Crotti et al., 2020; Barthelemy et al., 2023). Despite this, the delineation of intraspecific units is keystone to conservation policy.

With the aim of conserving the extant diversity and protecting the adaptive potential (future biodiversity) of a species, many different types of intraspecific conservation unit focus on identifying reproductively isolated groups, or groups showing high genetic differentiation as proxy for isolation (Allendorf et al., 2012). Two of the most commonly defined conservation units are Evolutionarily significant units (ESUs) and Management Units (MUs) (Moritz, 1994; Funk et al., 2012). Broadly an ESU can be defined as a reproductively isolated population or group of populations that shows notable adaptive differentiation or functional diversity, however a number of different definitions exist with some placing the emphasis on contemporary patterns of diversity while others focus more on distinct historical lineages (Waples, 1991; Moritz, 1994; De Guia and Saitoh, 2007; Casacci et al., 2014; Coates et al., 2018). In contrast, MUs are generally defined as demographically independent, reproductively isolated, units (Moritz, 1994; Funk et al., 2012) without the element of adaptive differentiation used for ESUs. Many studies define both these units in hierarchy with broader ESUs containing multiple MUs (Funk et al., 2012;

Shaney et al., 2020; Forester et al., 2022) but the applicability of this approach differs with species and context.

Many studies defining intraspecific conservation units have used a criterion based solely in genetic data (Robertson et al., 2014) even before the advent of large-scale genomic datasets, which has allowed for high precision delineation of conservation units utilising both selection neutral and adaptive SNP markers (Funk et al., 2012). How ESUs are determined can also determine on the amount of data used, for example genome-wide differentiation versus differentiation at single genes like greb1-like gene (Waples and Lindley, 2018; Waples et al., 2022). However, just using genetics can undervalue instances of recent divergences, for example when high morphological differentiation is not yet matched by high genetic differentiation (Apolônio Silva de Oliveira et al., 2017). We therefore argue that various data types, for example genetic, phenotypic, and geographic data, should be considered where possible when defining units of conservation. Firstly, this allows for information from different data types to provide important evidence that may confirm genetic information, with this idea of concordance a key part of the original definition of an ESU (Ryder, 1986). But secondly, because different types of data can provide an alternative insight on what should be deemed evolutionarily significant (Funk et al., 2012).

Beyond the definition of conservation units, another important requirement of conservation management is the need to identify supporting habitats and sites that may require urgent conservation action or monitoring because they are more vulnerable to loss (Campbell et al., 2017). This can involve investigating both the population and the environment it exists in. For example, small habitats with lower population carrying capacity will limit the size of a population and potentially its overall diversity and thus its ability to respond to change (Fischer and Lindenmayer, 2007). Increasing temperatures are a threat to many species, with increased water temperature shown to affect the development and overall morphology of freshwater species (Campbell et al., 2021). Interaction between these different factors can magnify the threat level, for example populations in smaller lakes may have less refuge from high water temperatures and are therefore more at risk (Campbell et al., 2017). Thus, using population genetics and environmental information to identify the level of the threat to a population from various factors can help identify vulnerable sites to prioritise for conservation action.

A species that demonstrates the complexity of defining units of conservation well is the salmonid fish species Arctic charr (*Salvelinus alpinus*). Across its Holarctic range, the species shows very high variation in phenotypic traits and high levels of genetic variation (Jonsson and Jonsson, 2001; Wilson et al., 2004; Klemetsen, 2010; Recknagel et al., 2017; Maitland and Adams, 2018). In the British Isles, unlike in the more northern parts of its range, the species is non-anadromous and remains in freshwater throughout the life cycle (Finstad and Hein, 2012; Jørgensen and Johnsen, 2014; Maitland and Adams, 2018) thereby inhibiting contemporary gene flow between populations in unconnected river catchments. Across its range, there are a number of lakes that contain multiple distinct populations, known as ecotypes, diverged to specialise in different trophic niches (Klemetsen, 2010; Doenz et al., 2019; Jacobs et al., 2020). These ecotypes are often distinguished by differences in their diet and foraging tactics but can be distinguished by other factors such as morphology and spawning (Malmquist et al., 1992; Adams et al., 1998; Garduño-Paz et al., 2012; Hooker et al., 2016). The extent of divergence between co-existing ecotypes varies by location and each represent important examples of the functional diversity of the species (Schluter, 1996; Adams and Huntingford, 2004; Adams et al., 2007a, et al., 2007; Moccetti et al., 2019; Fenton et al., 2023a) (Chapter 6). Many of these instances of ecotype divergence are the result of secondary contact and admixture of two lineages, whilst others represent recent sympatric divergences from a single lineage and therefore show low levels of genetic differentiation but show clear differentiation in phenotypes such as spawning time (Kettle-White, 2001; Garduño-Paz et al., 2012; Jacobs et al., 2020).

The high intraspecific diversity of Arctic charr has resulted in both the species and their environments being designated as having high conservation and scientific value. Therefore, there is particular interest in identifying evolutionary significant and/or vulnerable sites for Arctic charr for conservation purposes. Recent guidelines for selection of Sites of Special Scientific Interest (SSSIs) in Great Britain make a strong case for selection and protection of sites housing within-species diversity and seeks to preserve the habitats giving rise to it (Bean et al., 2018). There is also a strong movement towards the protection of genetic diversity in response to 2020 Aichi Targets which includes fish species (Hollingsworth et al., 2019). However, given that management resources are limited we need a way to define populations or groups of populations that are arguably more important than others to prioritise.

The diversity seen across populations of Arctic charr makes identifying these important populations and delineating appropriate units difficult. The Kottelat and Freyhof

(2007) review of the freshwater fishes of Europe defined many putative species which are currently found within the IUCN Red List (IUCN, 2023). In the context of Arctic charr, these were often Victorian classifications revived then added to other geographic locations based on available morphological data (Adams and Maitland, 2007). This resulted in a large increase from a single species of Arctic charr to include a further 20 which were endemic to the British Isles, some only found in a single lake, and the suggestion that many more species remained to be described (Kottelat and Freyhof, 2007). Since their delineation, there has been much debate about the validity of these putative endemic species and whether this represents a valid criterion for delineation of conservation units (Adams and Maitland, 2007). Primarily this is because the high phenotypic diversity within Arctic charr and lack of connectivity among lacustrine habitats means that morphological similarities between populations do not always reflect a shared ancestral history. Instead, they often result from local adaptation and convergence, or phenotypic plasticity, as is well described in freshwater fish species (Præbel et al., 2013). This criterion, if applied to the wider range, could lead to the definition of new species for each waterbody that contain Arctic charr which has massive ramifications for conservation (Savvaitova, 1995; Adams and Maitland, 2007). However, a proper assessment of these putative endemic species based on genetic data has been limited due to a lack of large-scale studies.

We conducted a national scale genetic study of Arctic charr in Scotland to identify units of conservation and vulnerable populations which may require specific action at an individual site level. We inferred the genetic distinctiveness of populations across Hydrometric Areas through genetic structuring and phylogenetic analyses using a genome-wide dataset of SNPs. We tested the phenotypic delineation of 14 putative endemic species against genomic and mitochondrial data to determine whether this represented a viable criterion for identifying units of conservation. We used our genomic analyses based on neutral and adaptive SNPs in tandem with phenotypic and ecological findings from previous studies and geographical information to delineate ESUs and MUs within the species and highlight populations of significance to protect. Finally, we generated a number of susceptibility metrics to evaluate the vulnerability of populations to range of threats, such as climate change and predation, to highlight populations we believed are particularly at risk to loss.

## 2.3 Material and Methods

### *2.3.1 Criterion for defining intraspecific conservation units*

We focused on the delineation of ESUs and MUs in our approach to defining intraspecific conservation units due to their frequent use in other fish species (De Guia and Saitoh, 2007). We defined ESUs as populations or groups of populations that: 1) showed high reproductive isolation to other groups using genetic differentiation as a proxy, and 2) showed substantial adaptive differentiation, important functional diversity or represented populations of geographic importance. Groups that showed evidence of reproductive isolation but did not show sufficient evidence for our second criteria were considered MUs. Given the lack of anadromy in the species, we expect to see pronounced divergence at neutral loci (Robertson et al., 2014) due to low levels of gene flow. As such, we choose not to use an ESU definition based in reciprocal monophyly (Moritz, 1994) as this would overvalue small, isolated populations that have high levels of genetic drift and low genetic diversity (Allendorf et al., 2012) which could increase the extinction risk of many populations and the species as a whole if considered the most significant (Weeks et al., 2016). Given the expected high levels of population differentiation, we felt that larger groupings for ESUs used in a hierarchical approach, where larger ESUs are defined which contain multiple MUs (Funk et al., 2012; Shaney et al., 2020), would only capture a small portion of variation between ESU groups and have high levels of variation within each group and therefore be inappropriate for the dataset. This approach would also undervalue the individual sites/populations that show substantial adaptive differentiation or functional diversity as they would only be valued as MUs.

### *2.3.2 Assessment of endemic species classifications*

Classification of Arctic charr populations as ‘endemic species’ was based on information in (Kottelat and Freyhof, 2007) (Table 2-1). We tested the validity of both the classification of these populations as putative species and the classification of populations into the same putative species through the use of mitochondrial and genome-wide datasets. We argued that these putative species, if they are to be considered as such, should show high levels of genetic differentiation relative to all populations not considered to be the same species at a level higher than that seen between other populations in our dataset. While complete lineage sorting is not a requirement, to be considered a distinct species, the populations representing putative species should drive the patterns of large-scale genetic structuring to some extent in our phylogenetic analyses or show distinct patterns. This could

be in the form of a distinct branch in a phylogenetic tree or a distinct haplotype in our mitochondrial haplotype network. There should be some evidence that these distinct patterns are shared across populations of the same endemic species where relevant. In instances where populations are considered to be multiple different putative species based on the localities in Kottelat and Freyhof, 2007, as seen in lochs Maree and Coulin (Lakes 34 and 36 on Figure 2-1), these populations had to show high levels of genetic diversity which would indicate the presence of multiple distinct populations along with evidence of split genetic structuring or multiple haplotypes.

**Table 2-1.** Populations in our dataset belonging to various endemic species as indicated by Kottelat and Freyhof. Original species descriptions are indicated. Several species have no Latin name, so the name indicated in Kottelat and Freyhof, 2007 is provided.

| **Species** | **Locations in our dataset** | **Original species description** |
|---|---|---|
| *S.killinensis* | Doine | (Gunther, 1865) |
| *S.mallochi* | Shin | (Regan, 1909) |
| *S.maxillaris* | Merkland, Coulin, Maree, Langavat | (Regan, 1909) |
| *S.perisii* | Padarn, Bodlyn | (Gunther, 1862) |
| *S.struanensis* | Rannoch (Pl) | (Maitland, 1881) |
| Rannoch benthivore | Rannoch (Bn) | (Kottelat and Freyhof, 2007) |
| Rannoch piscivore | Rannoch (Ps) | (Kottelat and Freyhof, 2007) |
| *S.youngeri* | Eck, Awe (Pl), Lee, Earn, Doon, Talla | (Friend, 1956) |
| Awe spring spawn | Awe (Bn) | (Kottelat and Freyhof, 2007) |
| Maree (large eye) | Maree | (Kottelat and Freyhof, 2007) |
| Maree (slender) | Maree, Uaine, Coulin | (Kottelat and Freyhof, 2007) |
| *S.inframundus* | Mealt | (Regan, 1909) |
| Windermere (Spring) | Conniston | (Kottelat and Freyhof, 2007) |
| *S.colii* | Finn | (Gunther, 1862) |

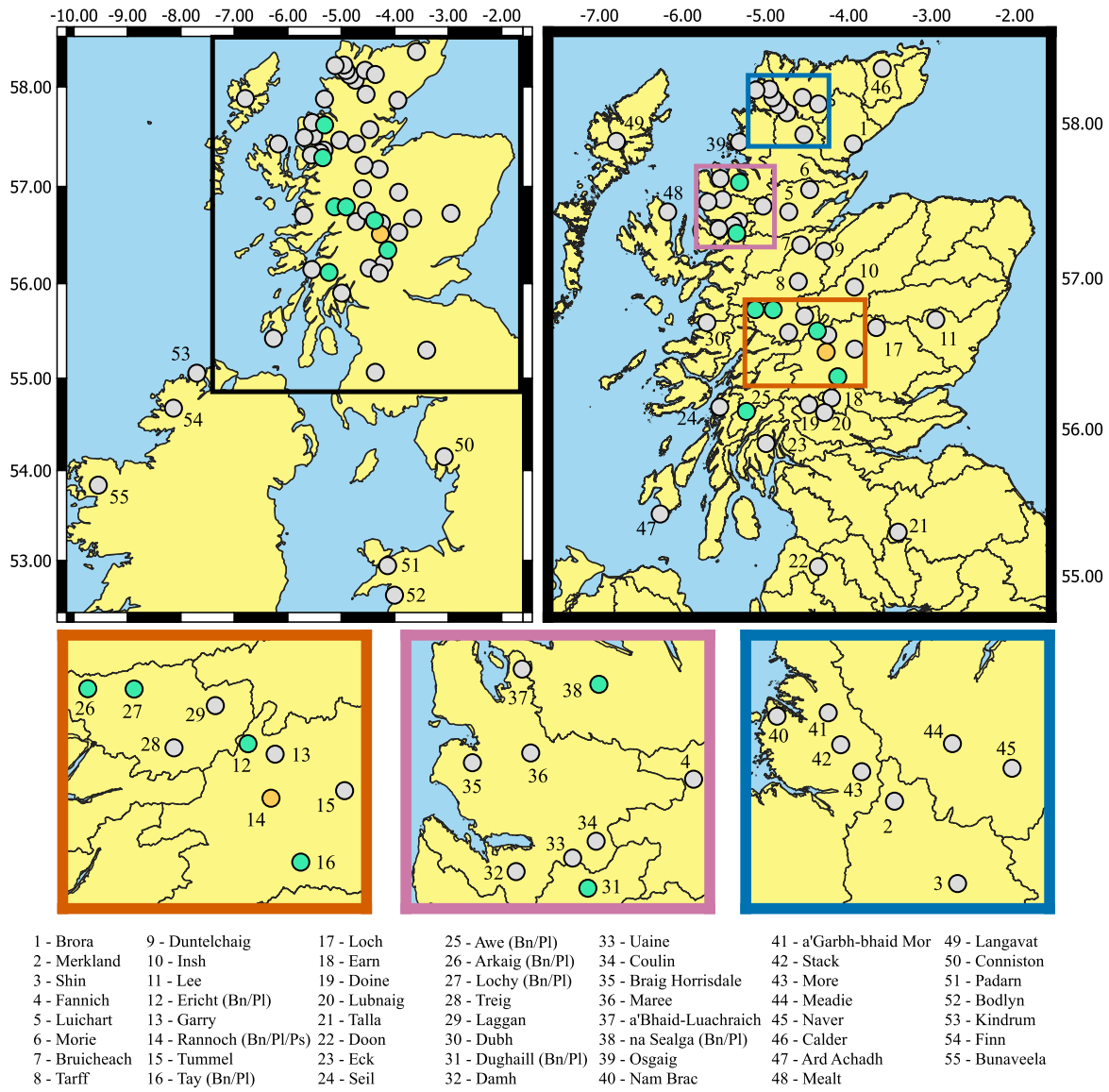


**Figure 2-1.** Dataset map of our populations across the British Isles. Lake origin is indicated on the map in ascending number order following Hydrometric Area order. Lakes with multiple ecotypes are indicated by Bn/Pl/Ps delineators in the legend and with green (bimodal) and oranges (trimodal) colours respectively. Bn refers to benthivore, Pl refers to planktivore and Ps refers to piscivore. Hydrometric area boundaries are indicated on the Scotland specific panel in the top right. Colour panels are used for regions with high density of populations.

### *2.3.3 Sample acquisition and collection*

Arctic charr were sampled from 49 Scottish lakes, three Irish lakes, two Welsh lakes, and one English lake (Figure 2-1) between 1997 and 2021 using 30m x 1.5 m Nordic multi-panel mesh survey gill nets (CEN, 2015). Sample numbers and locations for each lake are indicated in Table 2-1. Several of the sampled lakes in Scotland contain populations with multiple ecotypes (Fraser et al., 1998; Adams et al., 1998; Jacobs et al., 2020). We also have three instances of parapatric divergences (two neighbouring lakes from the same river system supporting two ecotypes), Shin-Merkland, Doine-Lubnaig, Stack-More (Adams et al., 2006; Adams et al., 2007b). Each ecotype was considered its own population in all steps of data generation and analysis unless otherwise specified. Lochs Maree and Stack are both known to contain multiple ecotypes (Adams et al., 2008) however these were not sampled explicitly and so each lake was considered to have one population for this study. In total, we had samples from 64 different Arctic charr populations covering 26 of the 27 Hydrometric Areas known to contain extant Arctic charr populations. This included 10 individuals per population for each unimodal populations and five individuals per ecotype in multiple ecotype lakes, wherever possible (Table 2-2). Hydrometric Areas are defined as integral river catchments or several contiguous river catchments which are designated across the British Isles (National River Flow Archive, 2014).

**Table 2-2.** Overview of each of our 64 populations with summary stats for ddRADseq data. N refers to number of individuals, Pa refers to number of private alleles, He to expected heterozygosity, Ho to observed heterozygosity, π to nucleotide diversity, Ar to allelic richness.

| Pop ID | Latitude | Longitude | Sampling Year | N | Pa | Ho | He | π | Ar |
|---|---|---|---|---|---|---|---|---|---|
| a'Bhaid-Luachraich | 57.81 | -5.55 | 2012 | 10 | 32 | 0.093 | 0.075 | 0.080 | 1.080 |
| a'Garbh-bhaid Mor | 58.39 | -4.95 | 2021 | 10 | 80 | 0.065 | 0.048 | 0.051 | 1.051 |
| Ard Achadh | 55.36 | -6.16 | 2021 | 10 | 164 | 0.048 | 0.039 | 0.042 | 1.042 |
| Arkaig (Benthivore) | 56.97 | -5.14 | 2007 | 4 | 0 | 0.116 | 0.094 | 0.110 | 1.110 |
| Arkaig (Planktivore) | 56.97 | -5.14 | 2007 | 5 | 0 | 0.116 | 0.098 | 0.110 | 1.110 |
| Awe (Benthivore) | 56.30 | -5.23 | 1997 | 5 | 0 | 0.119 | 0.108 | 0.125 | 1.124 |
| Awe (Planktivore) | 56.30 | -5.23 | 1997 | 5 | 0 | 0.121 | 0.103 | 0.119 | 1.119 |
| Bodlyn | 52.80 | -4.01 | 2002 | 2 | 0 | 0.112 | 0.089 | 0.128 | 1.128 |
| Braig horrisdale | 57.67 | -5.67 | 2016 | 9 | 75 | 0.067 | 0.056 | 0.060 | 1.060 |
| Brora | 58.05 | -3.95 | 2010 | 4 | 1 | 0.095 | 0.074 | 0.088 | 1.088 |
| Bruicheach | 57.39 | -4.58 | 2009 | 8 | 24 | 0.075 | 0.059 | 0.065 | 1.065 |
| Bunaveela | 54.02 | -9.54 | 2001 | 10 | 127 | 0.054 | 0.044 | 0.047 | 1.047 |
| Calder | 58.52 | -3.59 | 2021 | 3 | 0 | 0.108 | 0.084 | 0.102 | 1.102 |
| Conniston | 54.34 | -3.07 | 2005 | 1 | 0 | 0.129 | 0.065 | 0.129 | 1.129 |
| Coulin | 57.55 | -5.35 | 1997 | 9 | 0 | 0.079 | 0.066 | 0.071 | 1.071 |
| Damh | 57.50 | -5.57 | 2008 | 2 | 0 | 0.112 | 0.069 | 0.104 | 1.104 |
| Doine | 56.34 | -4.48 | 1997 | 4 | 0 | 0.118 | 0.088 | 0.106 | 1.106 |
| Doon | 55.24 | -4.37 | 1997 | 6 | 0 | 0.112 | 0.090 | 0.100 | 1.100 |
| Dubh | 56.88 | -5.71 | 2021 | 5 | 12 | 0.124 | 0.091 | 0.108 | 1.108 |
| Dughaill (Benthivore) | 57.47 | -5.35 | 2013/2014 | 5 | 2 | 0.102 | 0.087 | 0.098 | 1.098 |
| Dughaill (Planktivore) | 57.47 | -5.35 | 2013/2014 | 5 | 0 | 0.105 | 0.093 | 0.105 | 1.105 |
| Duntelchaig | 57.35 | -4.29 | 2010 | 4 | 0 | 0.122 | 0.089 | 0.113 | 1.113 |
| Earn | 56.38 | -4.23 | 2016 | 9 | 27 | 0.102 | 0.083 | 0.089 | 1.089 |
| Eck | 56.08 | -4.99 | 1997 | 10 | 59 | 0.060 | 0.049 | 0.057 | 1.057 |
| Ericht (Benthivore) | 56.82 | -4.39 | 2007 | 4 | 1 | 0.127 | 0.106 | 0.126 | 1.126 |
| Ericht (Planktivore) | 56.82 | -4.39 | 2007 | 5 | 0 | 0.145 | 0.121 | 0.136 | 1.136 |
| Fannich | 57.64 | -4.96 | 2007 | 9 | 2 | 0.068 | 0.056 | 0.060 | 1.060 |
| Finn | 54.86 | -8.14 | 2012 | 1 | 0 | 0.051 | 0.025 | 0.051 | 1.051 |
| Garry | 56.80 | -4.25 | 2007 | 8 | 2 | 0.115 | 0.095 | 0.103 | 1.103 |
| Insh | 57.12 | -3.93 | 2007 | 10 | 8 | 0.092 | 0.077 | 0.082 | 1.082 |
| Kindrum | 55.23 | -7.71 | 1997 | 7 | 43 | 0.072 | 0.058 | 0.065 | 1.065 |
| Laggan | 56.92 | -4.55 | 2007 | 6 | 0 | 0.115 | 0.099 | 0.110 | 1.110 |
| Langavat | 58.04 | -6.82 | 1997 | 10 | 81 | 0.087 | 0.074 | 0.078 | 1.078 |
| Lee | 56.90 | -2.95 | 1997 | 9 | 35 | 0.108 | 0.090 | 0.096 | 1.096 |
| Loch | 56.85 | -3.67 | 2007 | 10 | 83 | 0.037 | 0.029 | 0.031 | 1.031 |
| Lochy (Benthivore) | 56.96 | -4.92 | 2009 | 2 | 0 | 0.188 | 0.124 | 0.171 | 1.171 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Lochy (Planktivore) | 56.96 | -4.92 | 2009 | 5 | 0 | 0.123 | 0.104 | 0.116 | 1.116 |
| Lubnaig | 56.30 | -4.30 | 1997 | 10 | 2 | 0.103 | 0.078 | 0.100 | 1.100 |
| Luichart | 57.62 | -4.78 | 2009 | 10 | 29 | 0.102 | 0.086 | 0.092 | 1.092 |
| Maree | 57.71 | -5.53 | 2008 | 10 | 7 | 0.131 | 0.117 | 0.124 | 1.124 |
| Meadie | 58.33 | -4.56 | 2007 | 5 | 2 | 0.101 | 0.078 | 0.089 | 1.089 |
| Mealt | 57.61 | -6.18 | 2009 | 2 | 0 | 0.029 | 0.018 | 0.027 | 1.027 |
| Merkland | 58.23 | -4.73 | 1997 | 10 | 26 | 0.068 | 0.052 | 0.065 | 1.065 |
| More | 58.28 | -4.87 | 1998 | 4 | 0 | 0.121 | 0.091 | 0.113 | 1.113 |
| Morie | 57.75 | -4.47 | 2009 | 3 | 0 | 0.123 | 0.087 | 0.116 | 1.116 |
| Nam Brac | 58.38 | -5.12 | 2010 | 3 | 0 | 0.085 | 0.062 | 0.081 | 1.081 |
| naSealga (Benthivore) | 57.79 | -5.31 | 2015 | 5 | 0 | 0.089 | 0.071 | 0.083 | 1.083 |
| naSealga (Planktivore) | 57.79 | -5.31 | 2015 | 5 | 0 | 0.088 | 0.072 | 0.084 | 1.084 |
| Naver | 58.30 | -4.35 | 2011 | 9 | 27 | 0.118 | 0.100 | 0.107 | 1.107 |
| Osgaig | 58.05 | -5.32 | 2015 | 9 | 35 | 0.086 | 0.073 | 0.079 | 1.079 |
| Padarn | 53.13 | -4.13 | 2020 | 10 | 65 | 0.095 | 0.082 | 0.087 | 1.087 |
| Rannoch (Benthivore) | 56.70 | -4.23 | 2010 | 1 | 0 | 0.101 | 0.050 | 0.101 | 1.101 |
| Rannoch (Planktivore) | 56.70 | -4.23 | 2010 | 6 | 5 | 0.115 | 0.098 | 0.109 | 1.109 |
| Rannoch (Piscivore) | 56.70 | -4.23 | 2010 | 4 | 0 | 0.133 | 0.111 | 0.128 | 1.128 |
| Seil | 56.32 | -5.55 | 2021 | 9 | 77 | 0.046 | 0.039 | 0.042 | 1.042 |
| Shin | 58.12 | -4.57 | 1997 | 7 | 1 | 0.111 | 0.095 | 0.104 | 1.104 |
| Stack | 58.34 | -4.92 | 1997 | 5 | 0 | 0.115 | 0.102 | 0.118 | 1.118 |
| Talla | 55.48 | -3.40 | 1997 | 8 | 0 | 0.098 | 0.084 | 0.091 | 1.091 |
| Tarff | 57.15 | -4.61 | 2012 | 8 | 57 | 0.049 | 0.038 | 0.042 | 1.042 |
| Tay (Benthivore) | 56.52 | -4.14 | 1997 | 5 | 1 | 0.127 | 0.105 | 0.122 | 1.122 |
| Tay (Planktivore) | 56.52 | -4.14 | 1997 | 5 | 1 | 0.121 | 0.107 | 0.123 | 1.123 |
| Treig | 56.81 | -4.73 | 2007 | 7 | 0 | 0.127 | 0.124 | 0.135 | 1.135 |
| Tummel | 56.71 | -3.95 | 2007 | 9 | 18 | 0.126 | 0.105 | 0.113 | 1.113 |
| Uaine | 57.52 | -5.39 | 1997 | 10 | 0 | 0.067 | 0.060 | 0.064 | 1.064 |

### 2.3.4 DNA extraction and ddRADseq

DNA was extracted from fin clips, muscle, gill, and head kidney tissue using the Mackerey-Nagel NucleoSpin DNA Extraction Kit following manufacturer recommendations. DNA quality and quantity were assessed using agarose gel electrophoresis and the Qubit Fluorometer with Life Technologies dsDNA BR assay. Four libraries of 90 samples plus five technical replicates were prepared using a ddRADseq protocol for Illumina, modified from Recknagel et al., 2015, and following Jacobs et al., 2020. Double restriction enzyme digestion as performed with the *MspI* and *PstI* proteins. A size selection range of 145-295bp was used to ensure compatibility with pre-existing data using the same methodology. Final RAD-tag enrichment was performed over 10 PCR amplification cycles. Paired end 75bp sequencing was performed on the Illumina NextSeq500 platform at Glasgow Polyomics. Raw data (ddRADseq NCBI short read bioproject: PRJNA607173) previously generated for eight Scottish lakes (Awe, Dughaill, na Sealga, Tay, Lubnaig, Merkland, Uaine, Eck) (Jacobs et al., 2020) was downloaded from NCBI.

### 2.3.5 Mitochondrial ND1 amplification, sequencing and processing

NADH dehydrogenase 1 (ND1) gene was amplified from genomic DNA with PCR, using forward primer B1NDF 5'-TAAGGTGGCAGAGCCCGGTA-3' and reverse primer B1NDR 5'-TTGAACCCCTATTAGCCACGC-3' (Schenekar et al., 2014). PCR reagents and thermocycle profile can be found in Table A2-1. The PCR product was sequenced in both forward and reverse directions at Dundee Sequencing and Services (www.dnaseq.co.uk). Where possible, ND1 was generated for the same individuals that were present in the ddRADseq dataset (Table 2-3). Chromatograms for forward and reverse reads were checked using *MEGA v11* as quality control with any low-quality reads dropped (Tamura et al., 2021). *Uniprot UGENE v45* (Okonechnikov et al., 2012) was used to first trim sequences to a minimum quality of 38 before forming a consensus sequence from forward and reverse reads while mapping to a reference sequence from GenBank (Accession number: AF154851.1) (Doiron et al., 2002). Sequences for *Salvelinus fontinalis* (Accession number: MF621738.1) and *Salvelinus namaycush* (Accession number: OM736821.1) were used as outgroup species. Sequences were formed into a multisequence alignment using *MUSCLE* with a TCS haplotype network then made using *PopArt v1.7* (Leigh and Bryant, 2015). P-distances were calculated using *MEGA* and nucleotide diversity was calculated in *DNASP v6* (Rozas et al., 2017).

**Table 2-3.** Summary of mitochondrial ND1 data. Eight populations from the ddRADseq dataset do not have any mtDNA data so are not included here. Haplotypes are named as they are in Figure 2-7. Π refers to mtDNA nucleotide diversity.

| Population | NmtDNA | Haplotypes | π |
|---|---|---|---|
| a'Bhaid-Luachriach | 4 | Brit_1 | 0 |
| a'Garbh-bhaid Mor | 4 | Brit_1 | 0 |
| Ard Achadh | 4 | Brit_1 | 0 |
| Arkaig (Benthivore) | 3 | Brit_1 | 0 |
| Arkaig (Planktivore) | 3 | Brit_1 | 0 |
| Awe (Benthivore) | 3 | Brit_1 | 0 |
| Awe (Planktivore) | 4 | Brit_1, Scot_19 | 0.00158 |
| Braig Horrisdale | 4 | Brit_1, Scot_4 | 0.00106 |
| Bruicheach | 3 | Brit_1, Scot_10 | 0.00072 |
| Bunaveela | 3 | Ire_2 | 0 |
| Calder | 3 | Brit_1, Scot_15 | 0.00068 |
| Coulin | 7 | Scot_1 | 0 |
| Doine | 4 | Brit_1, Scot_8 | 0.00052 |
| Doon | 3 | Scot_11, Scot_12 | 0.00068 |
| Dubh | 5 | Brit_1, Scot_22 | 0.00068 |
| Dughaill (Benthivore) | 3 | Brit_1 | 0 |
| Dughaill (Planktivore) | 4 | Scot_1 | 0 |
| Earn | 5 | Brit_1, Scot_20 | 0.00087 |
| Eck | 9 | Brit_1 | 0 |
| Ericht (Benthivore) | 4 | Brit_1 | 0.00068 |
| Ericht (Planktivore) | 2 | Scot_16, Scot_21 | 0.00308 |
| Fannich | 3 | Scot_1 | 0 |
| Finn | 3 | Ire_1 | 0 |
| Garry | 4 | Scot_17, Scot_21 | 0.00294 |
| Insh | 3 | Brit_1, Scot_18 | 0.00068 |
| Kindrum | 2 | Brit_1 | 0 |
| Laggan | 2 | Brit_1 | 0 |
| Langavat | 4 | Brit_1 | 0 |
| Lee | 4 | Scot_5 | 0 |
| Loch | 5 | Brit_1 | 0 |
| Lochy (Benthivore) | 2 | Scot_6, Scot_10 | 0.00205 |
| Lochy (Planktivore) | 5 | Brit_1, Scot_13 | 0.00042 |
| Lubnaig | 3 | Brit_1 | 0.00072 |
| Luichart | 3 | Scot_14 | 0 |
| Maree | 3 | Brit_1, Scot_1 | 0.00072 |
| Meadie | 4 | Brit_1 | 0 |
| Merkland | 4 | Scot_1 | 0 |

| | | | |
|---|---|---|---|
| More | 3 | Brit_1 | 0 |
| na Sealga (Benthivore) | 2 | Brit_1 | 0 |
| na Sealga (Planktivore) | 3 | Brit_1, Scot_9 | 0.00113 |
| Naver | 3 | Brit_1 | 0 |
| Osgaig | 2 | Scot_14 | 0 |
| Padarn | 4 | Brit_1 | 0 |
| Rannoch (Benthivore) | 4 | Brit_1, Scot_7 | 0.00054 |
| Rannoch (Planktivore) | 4 | Scot_1, Scot_2, Scot_3, Scot_21 | 0.00054 |
| Rannoch (Piscivore) | 6 | Brit_1, Scot_7 | 0.00177 |
| Seil | 5 | Brit_1, Scot_15 | 0.00041 |
| Shin | 2 | Brit_1, Scot_1 | 0.00108 |
| Stack | 3 | Brit_1 | 0 |
| Talla | 5 | Scot_11 | 0 |
| Tarff | 3 | Scot_16 | 0 |
| Tay (Benthivore) | 5 | Brit_1 | 0 |
| Tay (Planktivore) | 5 | Brit_1, Scot_10 | 0.00042 |
| Treig | 4 | Brit_1, Scot_10 | 0.00052 |
| Tummel | 3 | Brit_1 | 0 |
| Uaine | 3 | Scot_1 | 0 |

### 2.3.6 Processing of ddRADseq data

Raw sequence quality was assessed using *FastQC v0.11.9* (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The *process_radtags* pipeline in *Stacks v2.60* was used for demultiplexing raw sequence data (Rochette et al., 2019). To prevent a drop in sequence quality affecting mapping, the first 6bp of the second read were trimmed off using *Trimmomatic v0.39* (Bolger et al., 2014). Mapping was performed to the annotated *Salvelinus sp.* genome (ASM291031v2) using *bwa mem v0.7.16* using a seed length of 20 (Li, 2013). Reads with a mapping quality < 20 were filtered out using *samtools v1.15* (Danecek et al., 2021). Any individuals with < 1M mapped reads were removed due to the high amounts of missing data. A total of 10 individuals were dropped due to low read counts. The *gstacks* module of *Stacks*, was run with the –gt-alpha set to 0.01 to ensure called genotypes were reliable when RAD-loci building. SNPs were retained in the populations module of *Stacks* if they met the following criteria: present in 66% of all individuals in each population and across all populations, a minimum minor allele frequency of 0.01, maximum observed heterozygosity of 0.5 with only the first SNP per locus retained. We used the script *filter_hwe_by_pop.pl* (https://github.com/jpuritz/dDocent) to filter any SNPs that were out of Hardy-Weinberg Equilibrium within populations, although no such SNPs were detected. *VCFtools v0.1.16* was used to identify and remove any SNPS with read depth <10x or >30x (Danecek et al., 2011).

### 2.3.7 Genotyping error

We used an R script published by (Mastretta-Yanes et al., 2015) to estimate the rate of genotyping error using 20 replicated individuals. Error rate was calculated to be 2.6%. The R script was then adapted to identify SNPs that showed a high rate of mismatching across replicates. SNPs that mismatched in more than 1 replicate pair (>5% genotyping error rate) were removed from the dataset for all individuals. Genotyping error was then recalculated resulting in a mean error rate of 1.4%. Replicate samples were then removed from the dataset, retaining the sample with the higher coverage. After filtering we had 24,878 SNPs and 410 individuals from 64 different populations.

### 2.3.8 Summary statistics

Pairwise $F_{ST}$, observed and expected heterozygosity, number of private alleles and FIS were generated for all populations and loci using the –fstats parameter in the *populations* module of *Stacks*. Allelic richness was estimated for each population using the *hierfstat v0.5-10* R package (Goudet, 2005) (R Core Team, 2021).

### 2.3.8 Identification of SNPs correlated to environmental and climatic variables

We conducted a Redundancy Analysis (RDA) using the *vegan v2.6.2* R package to identify SNPs associated with variation in environmental and climatic variables which represent our putative adaptive SNP dataset to analyse patterns of adaptive differentiation (Oksanen J et al., 2022). Environmental data for 19 variables were collected from WorldClim using the raster R package *v3.5-15* using latitude and longitude co-ordinates (resolution 2.5 arcmin) (Hijmans and van Etten, 2012; Fick and Hijmans, 2017). Bathymetric data on lake surface area, mean and max depth, and altitude was collected for each lake from published work (Murray and Pullar, 1910; Maitland and Adams, 2018). An additional bathymetric variable, littoral zone percentage (defined as the area of the lake shallower than 5 metres depth) was calculated using available bathymetric maps and *ImageJ v1.50i* (Murray and Pullar, 1910; Schneider et al., 2012).

To create a dataset of uncorrelated variables, as collinearity can affect genotype-environment patterns (Ratner, 2009), a collinearity matrix was made in R for all environmental and bathymetric variables. Variables with correlation coefficient ($\rho$) of $< 0.7$ were retained as independent variables and in cases where variables showed high correlation ($\rho > 0.7$) only one variable was retained. This left us with a set of 10 uncorrelated variables (Table A2-2). SNPs that were within three standard deviations of the mean of each constrained axis deemed to be significant based on ANOVA were considered to be putatively adaptive. In our case this was all 10 axes and thus we identified 1,104 SNPs as associated with environmental and bathymetric variables.

### 2.3.9 Genes and regions containing adaptive SNPs

To determine if our putative adaptive SNPs were related to adaptive processes. we investigated the genes and regions containing these SNPs. First, we compared the position of adaptive SNPs with a database of markers for quantitative trait loci (QTL) for various phenotypes from a range of salmonid species mapped to the *Salvelinus sp.* genome (Fenton et al., 2023a) (Chapter 6). Second, we compared SNP positions to all the annotated genes in the *Salvelinus sp.* genome (ASM291031v2) to identify genes containing associated SNPs. Both of these comparisons were carried out using *BEDtools v2.27.1* (Quinlan and Hall, 2010), with a limit of ± 100kbp for QTLs and ± 1kbp for genes to be considered to be near an adaptive SNP. We then ran GO term overrepresentation analysis (ORA) for any genes deemed to contain adaptive SNPs using the *TopGO v2.40.0* R package (Alexa and Rahnenfuhrer, 2020) using all genes containing any RAD loci from our ddRADseq data

(associated or otherwise) as the comparison dataset. Results were summarised using *REVIGO* (Supek et al., 2011).

### *2.3.10 Analysis of population structuring*

For analysis of population structuring, we removed any non-neutral SNPs as these can affect clustering (Funk et al., 2012). This included the putatively adaptive SNPs associated with environmental variables and SNPs in the top 5% of global $F_{ST}$ scores, inferred from *hierfstat* (Goudet, 2005). This left a dataset of 22,763 SNPs. Admixture modelling was run using *PopCluster* for K=2 to K=64 with medium scaling and 10 replicates per K (Wang, 2022). The FSTIS and DLK2 estimators were used to determine the optimum number of K (Figure A2-1). Principal component analyses were run for all SNPs and associated SNPs using the *adegenet R package v2.1.10* (Jombart, 2008). Distance-based phylogenetic relationship between populations were constructed using an unrooted Neighbour-Joining (NJ) tree in the *poppr R package v2.9.3* (Kamvar et al., 2015). Any SNPs with missing data were removed to prevent missing data affecting the phylogenetic trees. This left a dataset of 5,949 SNPs for phylogenetic analyses. The tree was run with 1,000 bootstraps with no root.

### *2.3.11 Calculating genetic offset*

Future climatic data was derived from the Coupled Model Intercomparison Project Phase 5 (CMIP5) under the Representative Concentration Pathways (RCP) scenario RCP4.5. The RCP 4.5 scenario assumes $CO_2$ emissions begin to decline by 2045. SNPs associated with climatic variables were determined using an RDA as described earlier. We removed any SNPs that showed 5 or less different minor allele frequency values across all populations, leaving us with a dataset of 235 SNPs. Genetic offset was not calculated for Lough Finn and Conniston Water due to insufficient numbers of individuals. To model current genotype-environment relationships, *gradientforest R package v0.1.32* was used with 250 trees (Ellis et al., 2012). This model was then used to transform current and future climatic data based on their importance in explaining genomic variation with the Euclidean distance between these current and future values representing genetic offset. Due to the way gradientforest models genotype-environment relationships, a single genetic offset score was calculated for each lake meaning that there was a single score for lakes containing multiple ecotypes rather than one per ecotype.

### 2.3.12 Population susceptibility

Sensitivity scores for each loch and information of fish communities were taken from Maitland and Adams, 2018. Sensitivity scores were calculated based on three bathymetric and location variables: Maximum depth of the loch, surface area of the loch, and loch altitude (Winfield et al., 2010). Deeper lakes were deemed to offer more refuge from high surface water temperatures and so were given a lower sensitivity score. Similarly, higher altitude was deemed to provide some protection from thermal stress so higher altitude was given a lower score. Small lakes will support smaller populations than larger ones and are impacted more by extreme climate events as a result (Fischer and Lindenmayer, 2007) so lakes with a smaller surface area were given a higher score. Scores for each variable were combined to give a score ranging from 3 to 11 with a score of 3 being a lake of least concern for population loss while 11 would be of highest concern (Table A2-3). The number of piscine predatory species in each lake was used to determine vulnerability of Arctic charr population to loss to predation. Brown trout (*Salmo trutta*), Northern pike (*Esox lucius*), European eel (*Anguilla anguilla*), European perch (*Perca fluviatilis*) were all counted as predatory species so scores for each lake could range from 0 to 4 (Table A2-4).

## 2.4 Results

### 2.4.1 Summary statistics

Across the genomic dataset (N = 24,878 SNPs) of 410 individuals from 64 populations of Arctic charr, the Loch Lochy benthivore population was the most genetically diverse with the highest observed heterozygosity and allelic richness ($H_O$ = 0.188, $A_R$ = 1.171) while the Loch Mealt population was the least genetically diverse ($H_O$ = 0.029, $A_R$ = 1.023) (Table 2-2). The Loch Ard Achadh ($P_A$ = 164), Lough Bunaveela ($P_A$ = 127), and Loch Loch populations ($P_A$ = 83) had the highest numbers of private alleles; with the number of private alleles negatively correlated with observed heterozygosity ($P < 0.001$, $R^2 = 0.329$).

### 2.4.2 Population structuring

A neighbour-joining tree separated populations by drainage divide (Figure 2-2). Populations in river systems that flow to the east coast of Scotland appear in one cluster, which is separate to populations in west-flowing river systems and outgroup populations from the rest of the British Isles. The populations representing the different putative endemic species and those representing the same putative species were found across all the major sections of the tree and did not form their own distinct clusters. In a principal component

analysis using all SNPs, individuals from lochs Coulin, Uaine, Fannich, Maree, and the Dughaill planktivore populations separate from all other populations along EV1 which explained 3.62% of the total variance (Figure 2-3A). The Dughaill benthivore population forms an intermediate group along the EV1 axis while EV2 (explaining 2.57% of the total variance) splits the Loch Loch population from all others. No clear groupings representing the different putative species were found across these or lower EV axes.

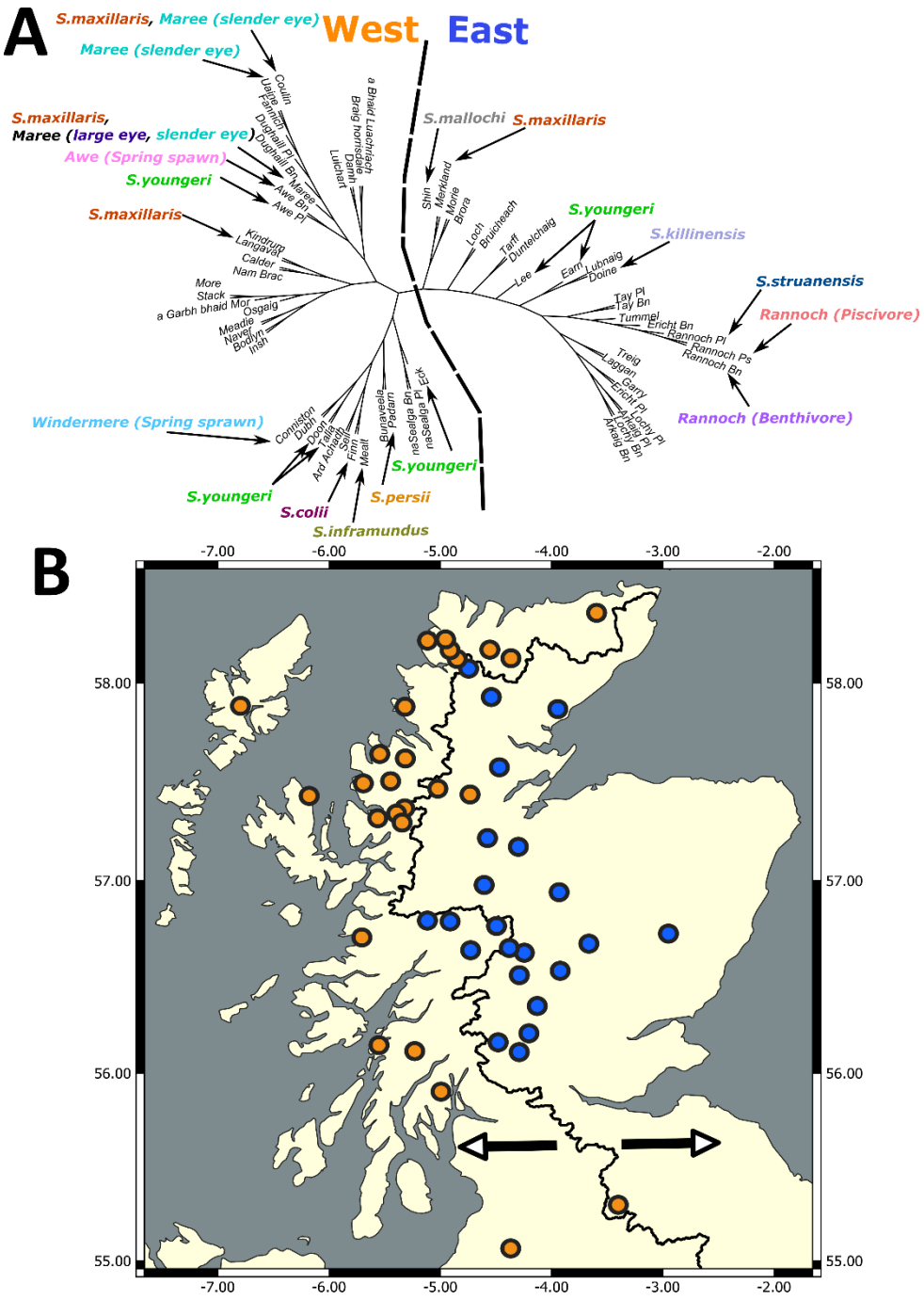


**Figure 2-2.** Neighbour-joining plot. NJ plot (A) is split into east and west flowing river system groups and is indicated by the dotted line Endemic species delineations are noted by coloured names and arrows for each relevant population. B) shows the geographical distribution of this divide with the black line showing where the drainage divide occurs, populations to the left-hand side appear in west draining river catchments and those to the right in east draining. Populations are coloured by which side of the neighbour-joining tree they appear in, orange for west and blue for east. All outgroup populations are found in the west flowing group.

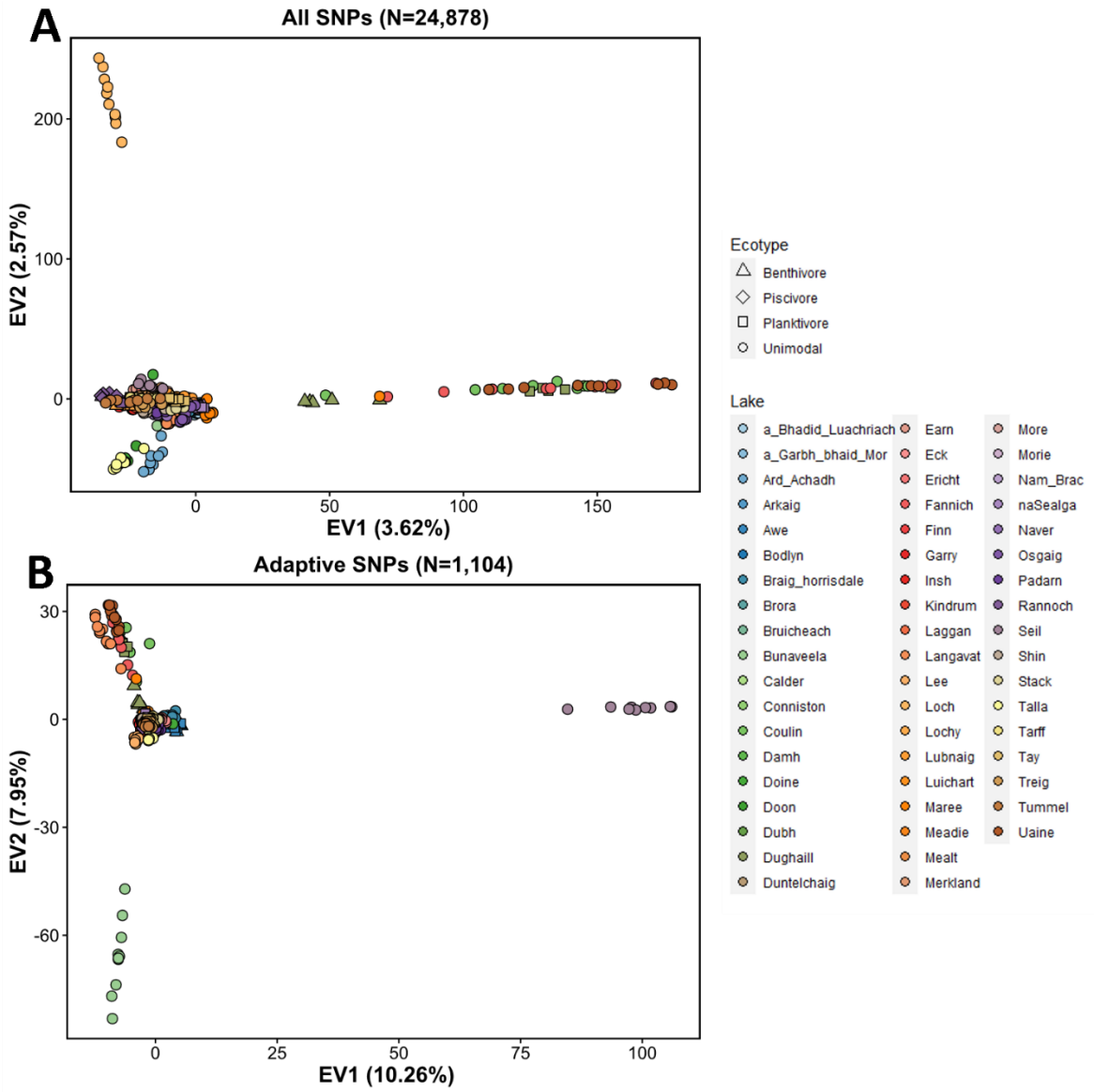


**Figure 2-3.** Principal components analyses based on all SNPs (A) (N=24,878) and adaptive SNPs (N=1,104) (B). Individual points are coloured by lake of origin indicated in the legend. Shape is based on population ecotype, triangle for benthivores, diamond for piscivores and squares for planktivores. Unimodal populations are given circles.

Genetic differentiation was high in most comparisons between populations with a mean pairwise $F_{ST}$ of 0.256 across all populations (Figure 2-4). The lowest pairwise $F_{ST}$ ($F_{ST}$= 0.045) was between Loch Doon and the Talla Reservoir population (which is a conservation refuge population derived from Loch Doon ancestors (Maitland et al., 2007)). The highest was between Lough Finn in Ireland and the Loch Rannoch benthivore population ($F_{ST}$ = 0.663), influenced by the small number of individuals we had for each population. Several populations such as Loch Loch (mean pairwise $F_{ST}$ = 0.384), Loch Seil (mean pairwise $F_{ST}$ = 0.371), and Loch Tarff (mean pairwise $F_{ST}$ = 0.357) showed particularly high mean pairwise $F_{ST}$ in all comparisons (all pairwise $F_{ST}$ > 0.2). Comparing levels of genetic differentiation between the different species, putative and *alpinus*, shows that many of the endemic species show levels of genetic differentiation that are on average lower than mean pairwise comparison across all populations (Figure 2-5).

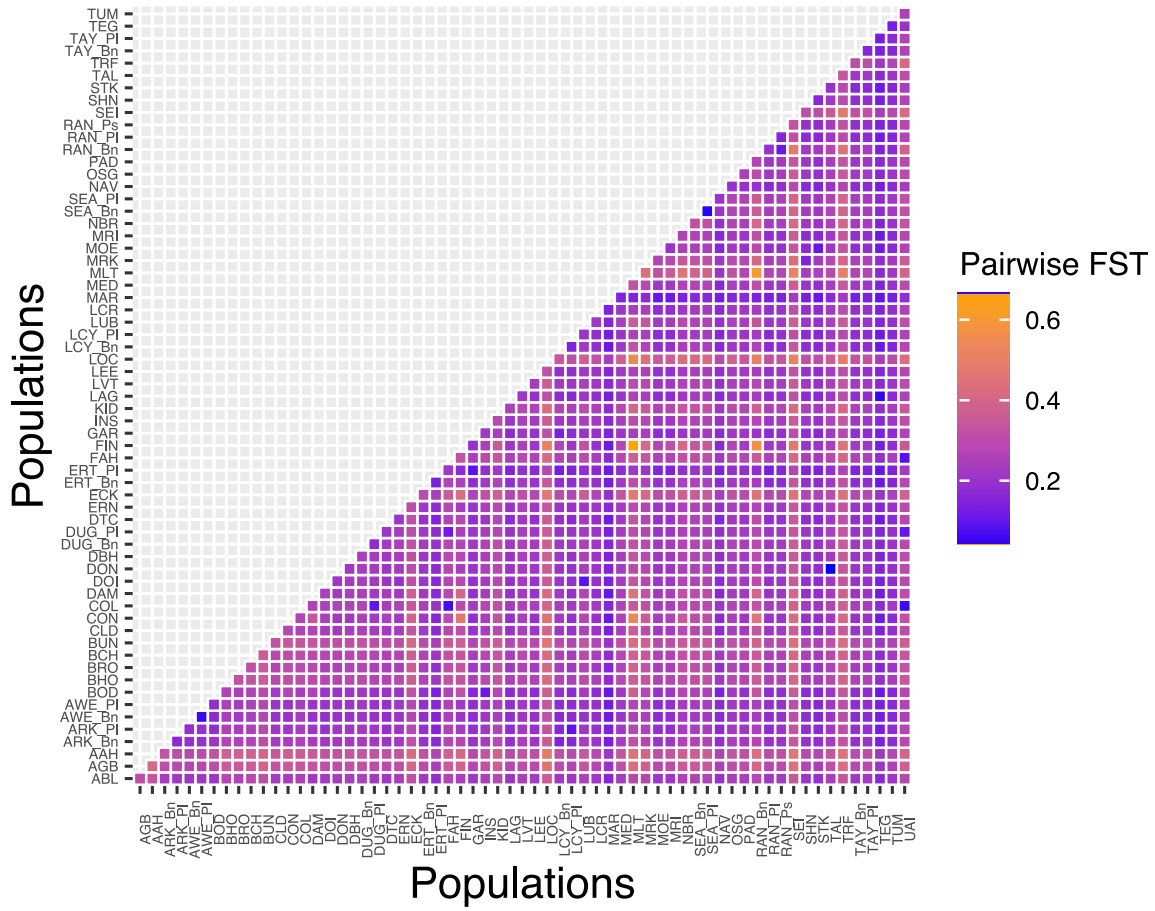


**Figure 2-4.** Plot of pairwise $F_{ST}$ comparisons. $F_{ST}$ scores are colours in a range of blue to orange going from low to high.

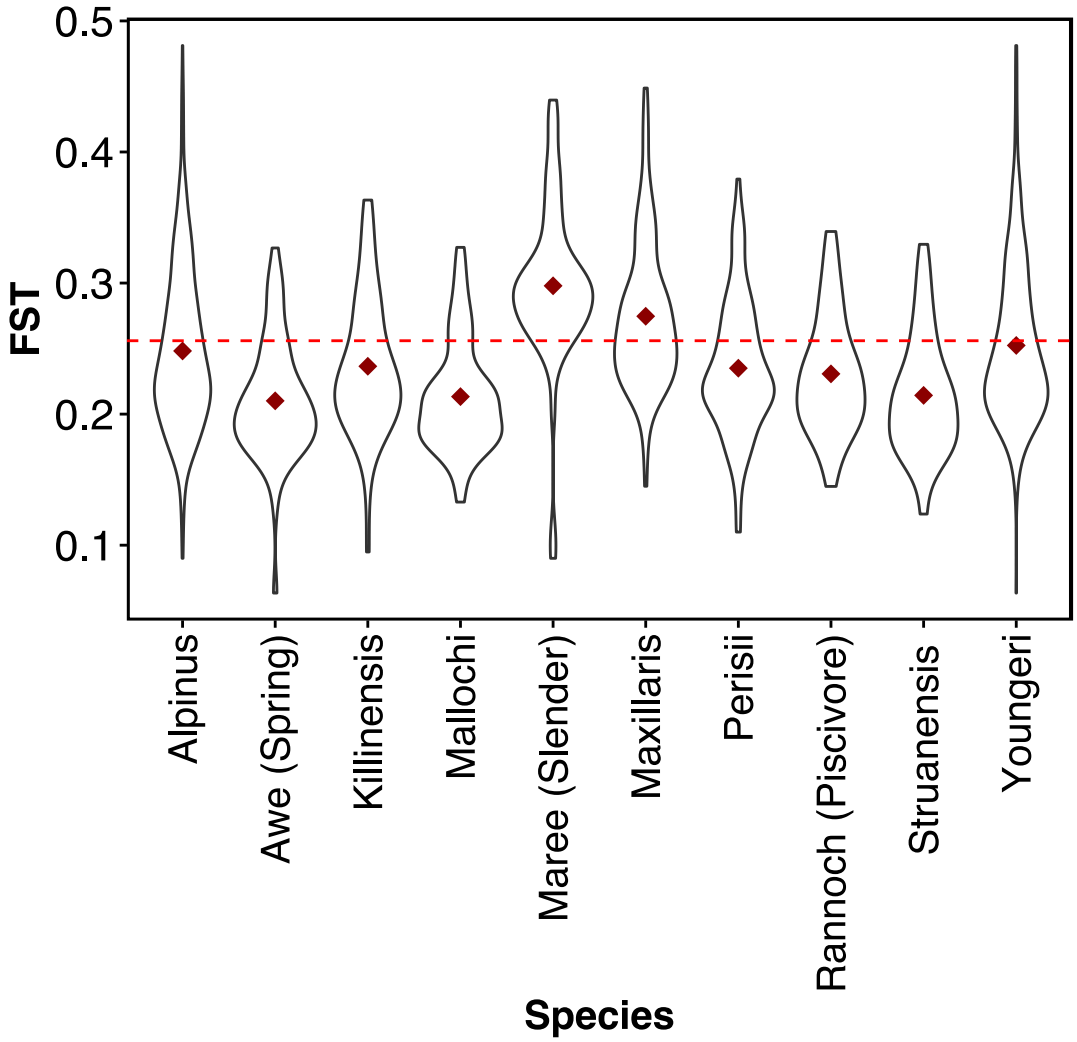


**Figure 2-5.** Violin plot of genetic differentiation (pairwise $F_{ST}$) between populations of the different endemic species as well as any populations still considered *Salvelinus alpinus*. The red diamond indicates the mean value with the red dotted line ($F_{ST}$=0.256) indicates the mean pairwise $F_{ST}$ for the whole dataset when all populations are considered to be *S.alpinus*. Comparisons are not included to the *Salvelinus colii*, *inframadus*, Windermere spring spawn, and Rannoch benthivore species due to low sample numbers for these populations. Comparisons involving the Loch Maree and Loch Coulin populations are also not included due to them containing multiple endemic species.

Admixture analysis was used to determine the number of genetic clusters (K) in the dataset to quantify the number of genetically differentiated groups in our dataset. 50 distinct clusters were identified, with most clusters representing a single population (Figure 2-6). While this was true for most populations, those in lochs Treig, Maree, Stack, Garry, and the Ericht planktivore each contained two different genetic clusters. At lower values of K, populations largely clustered by their Hydrometric Area before splitting off to form their own distinct clusters as the number of allowed genetic clusters increases.

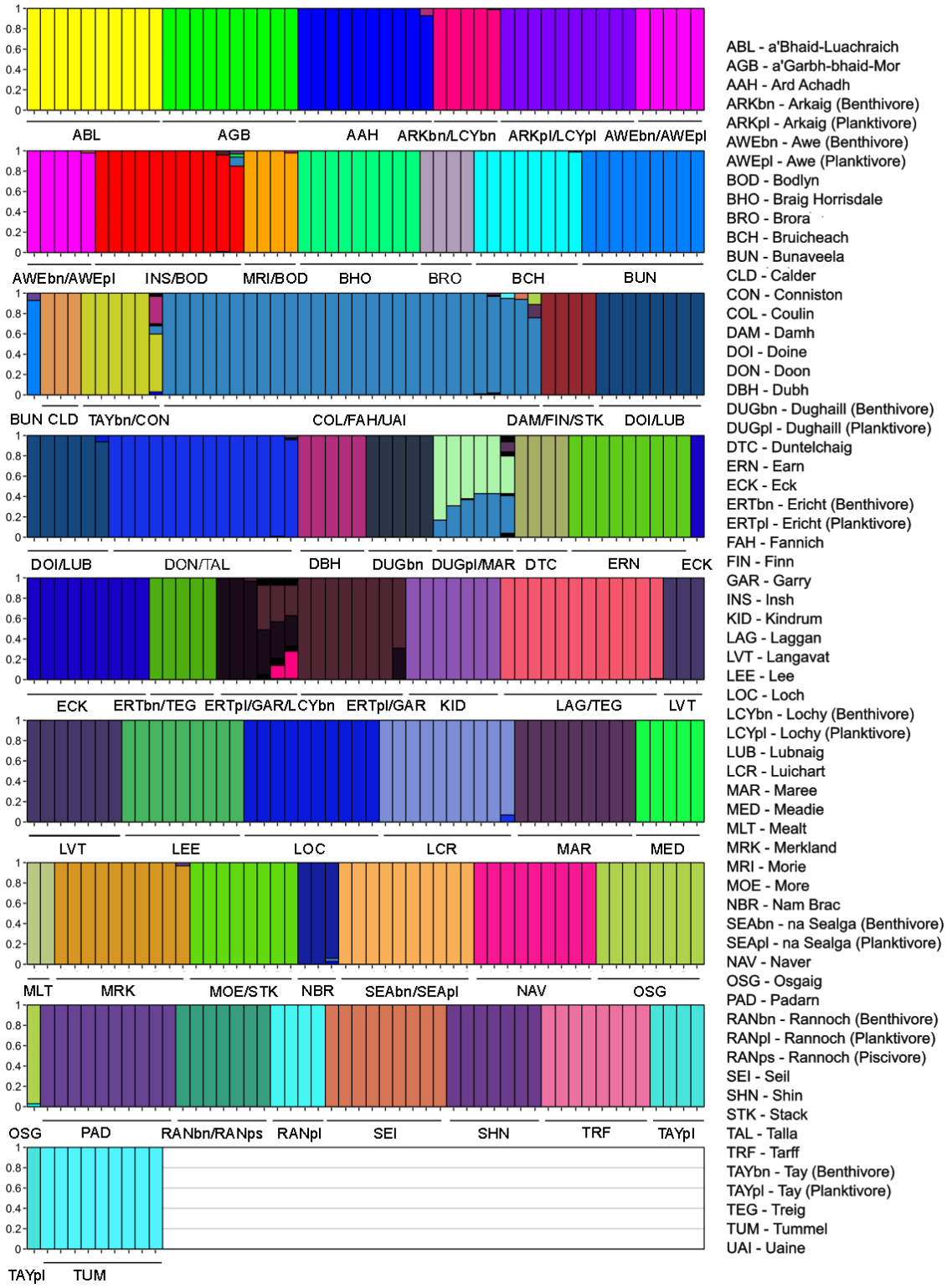


**Figure 2-6.** Admixture plot for most likely number of genetic clusters (K), K=50. Individuals are ordered by their genetic cluster. Most populations split into their own distinct genetic clusters with only some proximal populations sharing the same clusters. 3 letter codes for each population are provided. Ecotypes for relevant populations are indicated, bn for benthivore, pl for planktivore, ps for piscivore.

### *2.4.3 Mitochondrial haplotype network*

We observed a total of 25 ND1 mtDNA haplotypes from our mitochondrial data of 208 individuals covering 57 populations from across the British Isles (Figure 2-7) (Table 2-3). 39 populations contained the same shared haplotype (Brit_1) with 22 of these populations consisting solely of this haplotype. The Brit_1 haplotype was present across our entire British Isles range, with most other haplotypes a single mutational difference from this large haplotype (Figure A2-2). A second major shared haplotype (Scot_1) is found in eight different Scottish populations including lochs Coulin, Fannich, Uaine, the Dughaill planktivore, and Rannoch planktivore populations. Most other identified haplotypes were found in a single population and a total of 36 populations contained just one uniform haplotype. Given that most populations are uniform in their haplotypes, within population diversity and p distance were relatively low ranging from 0 to 0.0031. The majority of populations containing the same single haplotype resulted in relatively low p-distances between many populations with p-distances ranging from 0 to 0.53% across the dataset.

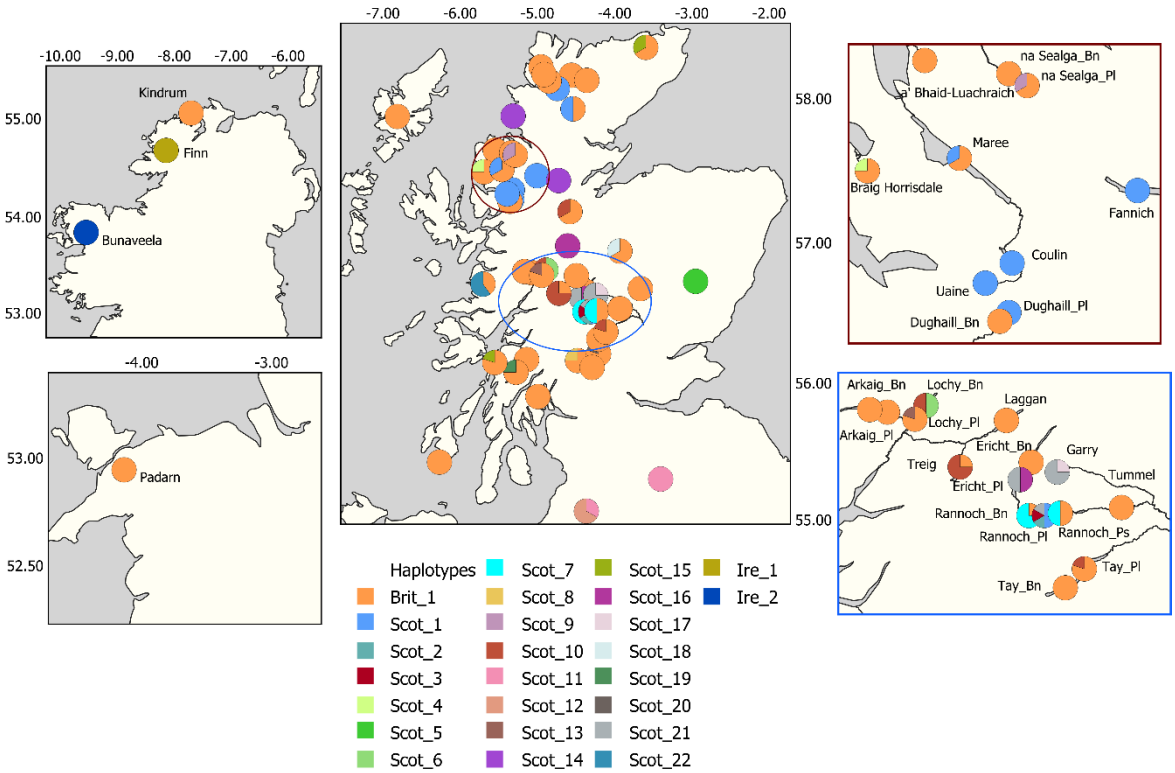


**Figure 2-7.** Map of haplotypes across our dataset. Colours indicate the haplotypes in the network. Full haplotype network can be found in Figure A2-2. No sequences were generated for our English population (Conniston water).

Populations of the same putative species were found across multiple haplotypes rather than forming a distinct one of their own (Table A2-5). The single populations we have that represent *Salvelinus killinensis*, *Salvelinus mallochi*, and *Salvelinus perisii* all belong to solely to the large Brit_1 haplotype. Only the population representing *Salvelinus colii*, Lough Finn (Ireland), formed its own distinct haplotype (Ire_1). Loch Coulin is said to contain two species, *S.maxillaris* and the Maree slender eye species, but showed only one uniform haplotype (Scot_1). Similarly, Loch Maree is said to contain three different species but only two different haplotypes (Brit_1 and Scot_1) were present.

### *2.4.4 Patterns of putatively adaptive variation*

We identified a set of 1,104 SNPs which showed strong association with a range of climatic and lake variables. EV1 from a PCA of using the adaptive SNPs explained 10.3% of total variance and clearly differentiated the Loch Seil population from all other populations which cluster together (Figure 2-3B). EV2 explained 8% of the total variance and most clearly separates the Irish Lough Bunaveela population from the other populations in one direction and a group including the lochs Coulin and Fannich, Lochan Uaine, and the Loch Dughaill planktivore populations in the other direction.

When comparing the positions of these SNPs to annotated genes in the *Salvelinus sp.* genome, we found that 586 known genes contained or were proximal to at least one putatively adaptive SNP. GO term analysis showed that amongst these genes, a number of processes relating to formation, development, and morphogenesis such as skeletal system development (GO:0001501), sagittal suture morphogenesis (GO:0060367), and lung lobe formation (GO:0060464) showed overrepresentation (Table A2-6). When we compared positions to a database of salmonid QTLs markers (Fenton et al., 2023a) (Chapter 6), we found that 35 QTLs contained or were found close to adaptive SNPs (Table A2-7). These QTLs were from a range of species including rainbow trout (*Oncorhynchus mykiss*), coho salmon (*Oncorhynchus kisutch*), lake whitefish (*Coregonus clupeaformis*), and lake trout (*Salvelinus namaycush*) with half of the QTLs related to body length, shape, weight or Fulton's condition.

### *2.4.5 Susceptibility metrics*

We used four measure to assess vulnerability of Arctic charr populations across Scotland to a range of threats. These were genetic offset, lake sensitivity, number of predatory fish species, and levels of observed heterozygosity. First, we used genetic offset, a measure of potential fitness maladaptation to changing climate. A higher genetic offset

indicates a higher level of fitness maladaptation to future climatic change and higher susceptibility to loss. Genetic offset ranged from 0.005 in the Loch Nam Brac population to 0.028 at Loch Insh under the RCP 4.5 emission scenario. There was a strong relationship between genetic offset scores and latitude with populations at the lower latitudes showing higher genetic offset scores ($P = 0.004$, $R^2 = 0.13$) (Figure 2-8).

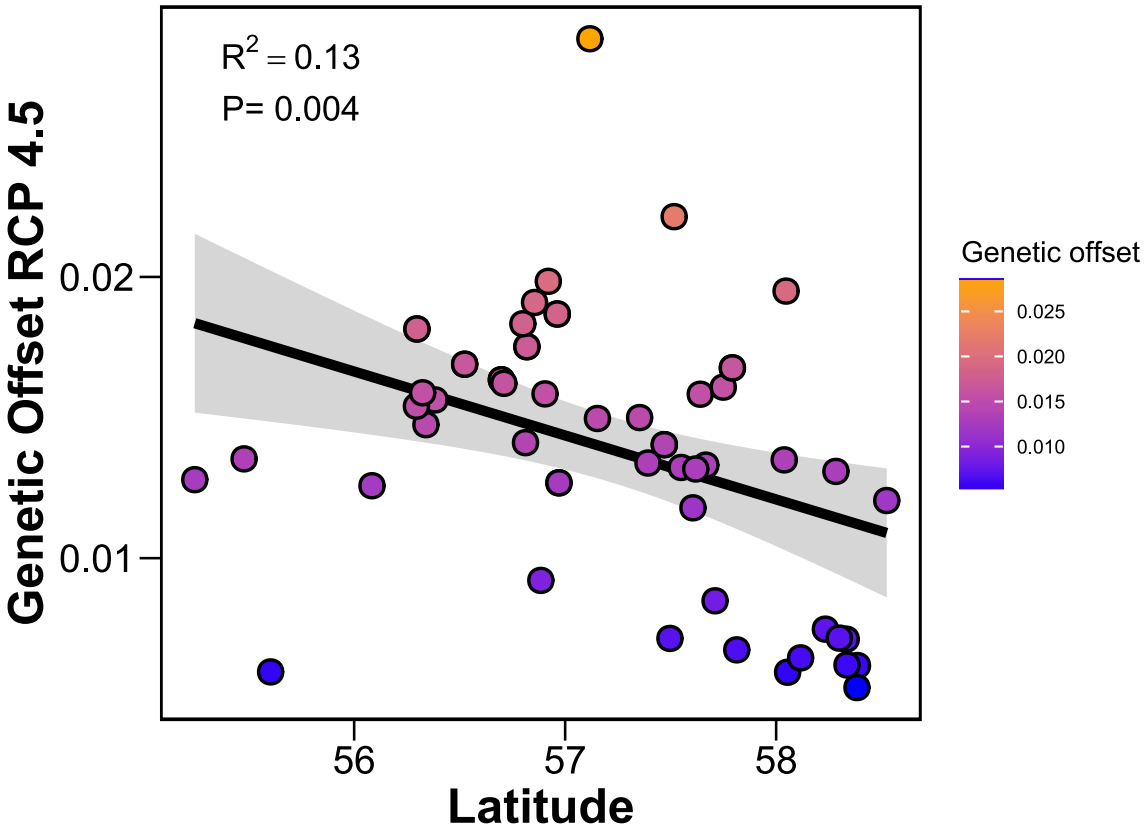


**Figure 2-7.** Genetic offset scores under RCP 4.5 scenario for each lake against latitude. Each point, population, is coloured by genetic offset score going from blue to orange as they go from low to high.

Next, we assessed population vulnerability based on lake bathymetry and altitude through the lake sensitivity score metric. Populations in lakes with higher scores were deemed more vulnerable to loss due to the small size and low depth and altitude of their lake environment. Three lakes, lochs Rannoch, Ericht, and Treig, had the lowest possible score of 3 while the highest scores seen was 8 in lochs Seil and Ard Achadh, out of possible 11, (Table 2-4) with an average score of 5. Populations in lakes with a higher number of predatory fish species were deemed more vulnerable to loss. The number of predatory fish species ranged from 0 to 4 with nearly all lakes having a score of at least 1 due to the near ubiquitous presence of brown trout in Arctic charr lakes across Scotland with an average score of 2. Our different metrics of susceptibility are largely uncorrelated with each other

(Table A2-8) allowing them each to tell us something different about population vulnerability. For example, there was no significant relationship between genetic offset scores and observed heterozygosity ($P = 0.249$, $R^2 = 0.01$) (Figure A2-3) while lake sensitivity scores significantly negatively correlated with number of predatory species ($P < 0.001$, $R^2 = 0.18$) (Figure A2-4). While no population performed poorly across all metrics, populations such as Loch Seil performed poorly on multiple metrics, showing low genetic diversity, a high sensitivity score, and a high genetic offset score (Table 2-4).

**Table 2-4.** Susceptibility metric scores for each population in Scotland. Due to how the metrics each ecotype in the same lake is given the same genetic offset, lake sensitivity scores and same number of predatory species. Observed heterozygosity based on all SNPs.

| Lake | Observed Heterozygosity | Genetic offset RCP4.5 | Lake sensitivity score | Number of predatory species |
|---|---|---|---|---|
| a'Bhadid-Luachraich | 0.09 | 0.007 | 5 | 2 |
| a'Garbh-bhaid Mor | 0.06 | 0.006 | 6 | 1 |
| Ard Achadh | 0.05 | 0.006 | 8 | 1 |
| Arkaig (Bn) | 0.12 | 0.013 | 5 | 3 |
| Arkaig (Pl) | 0.12 | | | |
| Awe (Bn) | 0.12 | 0.018 | 6 | 4 |
| Awe (Pl) | 0.12 | | | |
| Braig Horrisdale | 0.07 | 0.013 | 6 | 1 |
| Brora | 0.09 | 0.019 | 7 | 1 |
| Bruicheach | 0.07 | 0.013 | 5 | 2 |
| Calder | 0.11 | 0.012 | 6 | 1 |
| Coulin | 0.08 | 0.013 | 6 | 1 |
| Damh | 0.11 | 0.007 | 6 | 1 |
| Doine | 0.12 | 0.015 | 6 | 2 |
| Doon | 0.11 | 0.013 | 4 | 3 |
| Dubh | 0.12 | 0.009 | 6 | 1 |
| Dughaill (Bn) | 0.10 | 0.014 | 7 | 1 |
| Dughaill (Pl) | 0.10 | | | |
| Duntelchaig | 0.12 | 0.015 | 4 | 2 |
| Earn | 0.10 | 0.016 | 5 | 1 |
| Eck | 0.06 | 0.013 | 6 | 2 |
| Ericht (Bn) | 0.13 | 0.018 | 3 | 2 |
| Ericht (Pl) | 0.14 | | | |
| Fannich | 0.07 | 0.016 | 4 | 2 |
| Garry | 0.12 | 0.018 | 4 | 1 |
| Insh | 0.09 | 0.028 | 4 | 3 |
| Laggan | 0.12 | 0.020 | 4 | 3 |
| Langavat | 0.09 | 0.013 | 6 | 2 |
| Lee | 0.11 | 0.016 | 4 | 2 |
| Loch | 0.04 | 0.019 | 5 | 1 |
| Lochy (Bn) | 0.19 | 0.019 | 5 | 2 |

| | | | | |
|---|---|---|---|---|
| Lochy (Pl) | 0.12 | | | |
| Lubnaig | 0.10 | 0.015 | 5 | 3 |
| Luichart | 0.10 | 0.013 | 5 | 4 |
| Maree | 0.13 | 0.008 | 6 | 2 |
| Meadie | 0.10 | 0.007 | 6 | 1 |
| Mealt | 0.03 | 0.012 | 6 | 1 |
| Merkland | 0.07 | 0.007 | 6 | 1 |
| More | 0.12 | 0.013 | 6 | 2 |
| Morie | 0.12 | 0.016 | 5 | 1 |
| nam Brac | 0.09 | 0.005 | 6 | 1 |
| naSealga (Bn) | 0.09 | 0.017 | 5 | 1 |
| naSealga (Pl) | 0.09 | | | |
| Naver | 0.12 | 0.007 | 5 | 1 |
| Osgaig | 0.09 | 0.006 | 6 | 1 |
| Rannoch (Bn) | 0.10 | 0.016 | 3 | 3 |
| Rannoch (Pl) | 0.13 | | | |
| Rannoch (Ps) | 0.11 | | | |
| Seil | 0.05 | 0.016 | 8 | 1 |
| Shin | 0.11 | 0.006 | 5 | 2 |
| Stack | 0.11 | 0.006 | 6 | 2 |
| Talla | 0.10 | 0.014 | 5 | 1 |
| Tarff | 0.05 | 0.015 | 5 | 1 |
| Tay (Bn) | 0.13 | 0.017 | 4 | 4 |
| Tay (Pl) | 0.12 | | | |
| Treig | 0.13 | 0.014 | 3 | 2 |
| Tummel | 0.13 | 0.016 | 5 | 3 |
| Uaine | 0.07 | 0.022 | 7 | 1 |

## 2.5 Discussion

Our results highlight several different approaches to delineate important populations for conservation focus in a highly diverse species, the Arctic charr. Levels of genetic differentiation were high between populations indicating low levels of gene flow. As a result, many populations formed their own distinct genetic clusters and should be managed as such. We found little support for the existing delineation of endemic species in our genetic data with these proposed groups (Kottelat and Freyhof, 2007) showing levels of differentiation well within the range seen for Arctic charr and no evidence they represent significantly distinct groups in our genomic or mtDNA data. While no population could perform poorly across all our susceptibility metrics, we did identify several populations that performed poorly across multiple metrics suggesting that they may be of higher concern due to increased vulnerability to loss. We will now delineate ESUs and MUs using our patterns of genetic differentiation in tandem with past studies to indicate populations of evolutionary significance.

### *2.5.1 Patterns of genetic differentiation and structuring*

Non-anadromous fish species are generally expected to show higher levels of genetic differentiation than anadromous fish (DeWoody and Avise, 2000) and show more genetic clusters (Skovrind et al., 2016) due to the lack of gene flow between populations. High levels of genetic structuring between populations are often accompanied by low levels of genetic diversity particularly for the most isolated populations with small populations sizes due to the lack of gene flow (Furlan et al., 2012; Huey et al., 2017). Indeed, high levels of pairwise $F_{ST}$ have been seen previously in other non-anadromous populations of Arctic charr (Shikano et al., 2015).

Our results were in line with these general expectations, with many of the pairwise $F_{ST}$ comparisons between populations notably above the 0.15 value often used to determine significance of differentiation (Frankham et al., 2009). The patterns of high genetic differentiation are reflected in the admixture analysis with most populations forming their own distinct genetic clusters (Figure 2-5). Populations showing lower genetic differentiation, to the extent they exist in the same genetic cluster, were generally found between lakes in close proximity to each other or populations in the same lake. While it is currently unknown whether Arctic charr were anadromous when the species colonised Scotland *ca.* 11,000 years ago, it is likely that populations displayed some level of anadromy after colonisation (Fenton et al., 2023b) (Chapter 4). The N-J plot separated populations largely based on river system

flow, and therefore the probability of contact upon migration to the sea, is likely a reflection of this historical anadromy (Figure 2-2) (Chapter 5).

Of the ecotype pairs in our dataset, almost all ecotypes in the same lake show clear genetic differentiation from one another and represent different genetic clusters. Ecotypes are known to diverge under two broad scenarios: as a result of secondary contact between distinct colonising lineages when they come back into contact which other having diverged in allopatry, or as strict sympatric divergence (Butlin et al., 2008). The ecotypes at Tay and Dughaill are thought to have arisen from secondary contact scenarios (Jacobs et al., 2020) and the Rannoch planktivore is believed to be from a different colonising population to the other Rannoch ecotypes (Verspoor et al., 2010). The levels of genetic differentiation seen in our dataset reflect these histories with the ecotypes representing distinct genetic clusters. While less is known about the Ericht, Lochy, and Arkaig ecotype pair origins, the levels of genetic differentiation suggest secondary contact origins, however this needs to be investigated further. As has been suggested previously, the ecotypes at lochs Awe and na Sealga show low genetic differentiation and only one genetic cluster in seen in each lake reflecting the origins as recent sympatric divergences (Jacobs et al., 2020). While we did not explicitly sample the two known ecotypes at Loch Maree and Loch Stack, each lake has one individual assigned to a different admixture cluster to all others which likely represents the two ecotypes present. Loch Treig shows the same genetic pattern; however, Arctic charr from this lake have not been studied in detail and so it is currently unknown if multiple ecotypes are present there. The ecotypes at Maree and Stack were originally identified by genetics (Adams et al., 2008) so it is possible that further investigation into functional traits would confirm ecotype divergence at Loch Treig.

### *2.5.2 Adaptive differentiation*

Genomic datasets have often been used to help define ESUs through the identification of adaptive variation (Funk et al., 2012; Forester et al., 2022). As such we used our dataset to identify putatively adaptive SNPs correlated with change in climatic and lake variables identifying 1,104 associated SNPs. We then investigated the locations of these SNPs to understand the processes they were involved in.

Over half of the SNPs were found within or proximal to known genes in the *Salvelinus sp.* genome. Through analysis of the function of these genes and comparisons to a salmonid QTL database for a range of phenotypes, we identified a particularly important role for morphology in local adaptation and adaptive variance across populations of Arctic

charr. GO term analysis revealed a high number of terms relating to morphogenesis and development as overrepresented while half of the salmonid QTLs we identified to contain putatively adaptive SNPs were related directly to morphology (Table A2-6, A2-7). Of the genes found to contain putatively adaptive SNPs, 32 were shared with a recent study identifying key genes that underlie head and body shape morphology difference between different ecotype pairs of Arctic charr (Fenton et al., 2023a) (Chapter 6). Morphology is known to be highly variable across populations and our results confirm its importance in adaptation to a range of lake environments and climatic conditions (Bush and Adams, 2007). The identification of processes and regions related to head and bone morphology for SNPs associated with temperature variables supports previous studies which indicate that both phenotypes are influenced by temperature (Campbell et al., 2021; Hooker et al., 2023). The QTLs containing adaptive SNPs came from a number of different species thereby indicating to some extent that key regions underlying important phenotypes may be shared across species (Salisbury and Ruzzante, 2021; Fenton et al., 2023a) (Chapter 6). While morphology was the most prevalent in our dataset, several QTLs identified were related to other important phenotypes such as age of sexual maturation and spawning date. Spawning date is known to vary between these specific populations such as those observed in Loch Awe which has spring and autumn spawning ecotypes (Garduño-Paz et al., 2012). Overall, adaptive variation is facilitated through a number of different processes with morphology acting as a key aspect of local adaptation across populations in the British Isles.

Investigating patterns across these putatively adaptive SNPs saw that Loch Seil showed a notable distinct pattern separating from all other populations across EV1 in our PCA (Figure 2-3B). Loch Seil has never been studied before now and warrants further examination due to its potentially important adaptive variation. Overall, our PCA identifies populations which merit further exploration for their adaptive distinctiveness rather than confirmation of adaptive distinctiveness indicated from previous studies.

### *2.5.3 Population susceptibility*

A key part of conservation effort is identifying populations that may be more vulnerable to loss and are therefore of higher conservation priority. Populations can be more vulnerable for a variety of reasons, from low levels of standing diversity to higher risk of predation. As such, we assessed vulnerability across a number of distinct metrics.

Low levels of standing genetic variation suggests a population is less able to adapt to changes in their environment, for example due to climate change, and are at higher risk

of reduced population fitness due to inbreeding depression (Reed and Frankham, 2003). In our dataset many of the highly isolated populations, for example those that represent the only population in their HA like Loch Ard Achadh, were often found to have low heterozygosity with higher numbers of private alleles (Table 2-2). These populations are similar to isolated island species with no gene flow to other populations which have gained unique mutations and where genetic differentiation is most likely driven by genetic drift rather than selective pressures (Funk et al., 2016).

Arctic charr is adapted to cold, highly oxygenated environments and as such the species is highly susceptible to climate change, particularly increasing water temperatures, and eutrophication (Winfield et al., 2008; Kelly et al., 2020). Overall, we found that genetic offset was most strongly correlated with latitude (Figure 2-8), as previously reported in Arctic charr and Sockeye salmon (Layton et al., 2021; Tigano et al., 2023) with more southern populations suggested to be more vulnerable to loss. Genetic offset has been previously demonstrated to be a useful metric for fitness maladaptation to climatic change (Fitzpatrick et al., 2021). However, it is easy to over- or mis-interpret genetic offset and so conclusions drawn from it need to take a number of factors into consideration (Láruson et al., 2022). For example, changes in allele frequency, which genetic offset is based on, can simply reflect genetic drift rather than adaptive signals, which is why we focused on a subset of SNPs associated with climatic variables. Similarly, overall heterozygosity can drive genetic offset scores but we did not see a strong relationship between genetic offset and overall heterozygosity (Figure A2-3), suggesting that a change in environmental conditions, more so than levels of standing genetic variation, drives genetic offset scores (Tigano et al., 2023).

Lake sensitivity scores were inverse to number of predatory fish species as expected (Figure A2-4) as larger, deeper lakes, which are given low sensitivity scores, will naturally accommodate more complex environments and therefore a higher number of species. Our four metrics of susceptibility were largely independent of each other allowing us to investigate vulnerability across multiple axes (Table A2-8). While no single population can perform poorly across all the susceptibility metrics we provided although several performed poorly across several metrics. For example, Loch Seil has a high sensitivity score due to its small lake size, has low heterozygosity and a high genetic offset score. The small lake size seems to have limited the overall diversity of the population but also provides it less refuge from high water temperatures (Fischer and Lindenmayer, 2007; Campbell et al., 2021). Overall, these patterns highlight that populations like Loch Seil may be more vulnerable to

loss. The investigation of vulnerability across a range of metrics can help us identify vulnerable populations in ways that may be missed by the use of a single metric.

### *2.5.4 Endemic species classifications*

With the general patterns of genetic differentiation and structuring established we then determined the best criteria for identifying important populations for conservation action. We assessed whether the morphometric-based approach, used to classify endemic species of Arctic charr in the British Isles, represented an appropriate system for identifying important populations (Kottelat and Freyhof, 2007).

Overall, we found little support for the classification of these putative species in our dataset as was the case for a recent study of putative species of Arctic charr in Ireland (Barthelemy et al., 2023). Levels of genetic differentiation, and by proxy reproductive isolation, between different putative species was well within the range seen for the whole dataset (Figure 2-5) with some showing notably lower levels than dataset mean. Two species do show higher mean values, *S.maxillaris* and the Maree Slender eye species, however they had pairwise $F_{ST}$ to populations not belonging to their species much below the mean for the whole dataset. Instead of populations considered to be putative species, showing the highest levels of genetic differentiation the highest values were seen in lochs Loch, Tarff, and Seil. Similarly, none of the putative species clearly separated from the rest of the distribution in our PCA using all SNPs or just using putative adaptive SNPs with other populations or groups being those that separated clearly across the first few axes (Figure 2-3).

Our neighbour-joining tree split our dataset by difference in river system drainage, as previously discussed, instead of distinguishing the different species from the rest of the distribution (Figure 2-2). Populations of the same putative species, for *S.youngeri, S.maxillaris*, and *S.perisii*, were found sporadically throughout the tree with similar results seen when looking at the putative species found in Ireland (Barthelemy et al., 2023). Similarly, we found no clear divides for many putative species in the haplotype network, with many of the different putative species consisting solely, almost entirely, of the dominant haplotype (Brit_1) which likely represents a colonising sub-lineage (Chapter 5) (Table A2-5). Populations of the same putative species were found across multiple haplotypes rather than forming a distinct one of their own suggesting potentially different, rather than shared, ancestral origins (Figure 2-7). Only the population representing *S.colii*, Lough Finn (Ireland), formed its own distinct haplotype (Ire_1) (Figure A2-2). However, our dataset only

contains three populations from Ireland and so it is unclear whether this haplotype would appear as unique in a wider Ireland-focused dataset.

For the two populations that are said to contain two or more putative species, lochs Coulin and Maree, we found little evidence to support this. Loch Coulin showed no indication of multiple genetic groupings which could represent two species in either the ddRADseq or mtDNA data. Similarly, the two genetic groupings found in Loch Maree likely represent the two known ecotypes in this lake (Adams et al., 2008) and there is no indication of a third groups that would represent *S. maxillaris*. Neither of these lakes showed the highest levels of diversity in our dataset again suggesting that evidence for multiple species is limited.

Overall, we found little evidence to support the endemic species classifications based on genetics, as was the case for similar endemic species classifications for European whitefish (*Corgeonus lavearetus*) in the British Isles (Etheridge et al., 2012; Crotti et al., 2020). Recent research has also questioned the morphological distinctiveness used to describe the putative species, with several putative species in Ireland found to be indistinguishable in the larger range of Arctic charr based on morphological differentiation (Barthelemy et al., 2023). We found both that there is insufficient evidence that these groups should be considered distinct from *S.alpinus* based on patterns of genetic differentiation and structuring but there was also a lack of evidence supporting populations considered to be the same putative endemic species being classified as such with the phenotypic similarities between populations likely reflecting local adaptation and convergence (Præbel et al., 2013). We therefore argue that the classification of these endemic species has no genetic basis and urge the IUCN to reconsider their continued presence on the Red List. We recognise however that Arctic charr is clearly a phenotypically and genetically different fish in different locations which often differs in its functional role in the lake ecosystem (Adams and Maitland, 2007). The contribution of Arctic charr to the effective biodiversity of relatively depauperate northern freshwater systems cannot be overstated. However, our results find that a criterion based only on morphometrics is inappropriate for defining important populations in this species. The defining of intraspecific conservation units allows us a method to recognise the diversity of Arctic charr in a more pragmatic way than defining different species (Savvaitova, 1995; Reist et al., 2013).

### *2.5.5 Delineation of ESUs and MUs*

Using the information provided by our genetic dataset and existing phenotypic, environmental, and geographical information we argue for the following conservation unit delineations. Given the high levels of genetic differentiation, we argue that only populations in close proximity, conservation translocations, and populations with notable shared genetic origin should be considered within the same intraspecific conservation units (Table 2-5). While each may have high genetic differentiation to other groups, we need to be careful not to delineate every group as an ESU.

**Table 2-5.** Conservation unit classifications for populations in our dataset as evolutionarily significant units (ESUs) or management units (MUs) based on our analyses and previous findings.

| Population(s) | Conservation unit type | Delineation criteria for ESU |
|---|---|---|
| a'Bhaid-Luachraich | MU | |
| a'Garbh-bhaid Mor | MU | |
| Ard Achadh | MU | |
| Arkaig (Benthivore) | ESU | Diverged ecotype |
| Arkaig (Planktivore) | ESU | Diverged ecotype |
| Awe (Benthivore) | ESU | Diverged ecotype |
| Awe (Planktivore) | ESU | Diverged ecotype |
| Braig horrisdale | MU | |
| Brora | MU | |
| Bruicheach | ESU | Only extant HA pop. |
| Calder | MU | |
| Coulin, Dughaill (Planktivore), Fannich, Uaine | ESU | Diverged ecotype |
| Damh | MU | |
| Doine | ESU | Diverged ecotype |
| Doon, Talla | ESU | Only extant HA pop. |
| Dubh | MU | |
| Dughaill (Benthivore) | ESU | Diverged ecotype |
| Duntelchaig | MU | |
| Earn | ESU | Only extant HA pop. |
| Eck | ESU | Only extant HA pop. |
| Ericht (Benthivore) | ESU | Diverged ecotype |
| Ericht (Planktivore), Garry | ESU | Diverged ecotype |
| Insh | MU | |
| Laggan, Treig | MU | |
| Langavat | MU | |
| Lee | ESU | Only extant HA pop. |
| Loch | MU | |
| Lochy (Benthivore) | ESU | Diverged ecotype |
| Lochy (Planktivore) | ESU | Diverged ecotype |

| | | |
|---|---|---|
| Lubnaig | ESU | Diverged ecotype |
| Luichart | MU | |
| Maree | ESU | Diverged ecotype |
| Meadie | MU | |
| Mealt | MU | |
| Merkland | ESU | Diverged ecotype |
| More | ESU | Diverged ecotype |
| Morie | MU | |
| Nam Brac | MU | |
| naSealga (Benthivore) | ESU | Diverged ecotype |
| naSealga (Planktivore) | ESU | Diverged ecotype |
| Naver | MU | |
| Osgaig | MU | |
| Rannoch (Benthivore) | ESU | Diverged ecotype |
| Rannoch (Piscivore) | ESU | Diverged ecotype |
| Rannoch (Planktivore) | ESU | Diverged ecotype |
| Seil | MU | |
| Shin | ESU | Diverged ecotype |
| Stack | ESU | Diverged ecotype |
| Tarff | MU | |
| Tay (Benthivore) | ESU | Diverged ecotype |
| Tay (Planktivore) | ESU | Diverged ecotype |
| Tummel | MU | |

We argue that each instance of ecotype divergence, either in multimodal lakes or parapatric divergences, is evolutionary significant and should therefore be recognised as such due to their importance for the functional diversity of the species (Jonsson and Jonsson, 2001; Maitland and Adams, 2018) and considered separate ESUs. This includes the ecotype pairs at lochs Awe and na Sealga and the Loch Rannoch benthivore and piscivore ecotypes that do not show clear genetic differentiation and a represent single genetic cluster in our admixture analysis (Figure 2-6). While these divergences do not show clear genome-wide differentiation, there is sufficient evidence of differences in spawning habitat and timing and ecology distinctiveness to suggest these represent important incipient diverging groups (Adams et al., 1998; Alexander and Adams, 2000; Kettle-White, 2001; Jonsson and Jonsson, 2001; Adams and Huntingford, 2004; Garduño-Paz et al., 2010; Jacobs, 2018). Indeed, with their complete separation in spawning times, the ecotypes at Loch Awe arguably have the highest potential for divergence and speciation amongst ecotype pairs and should be protected as such. While only the Shin-Merkland parapatric divergence showed notable genetic differentiation, there is sufficient evidence of reproductive isolation in the Doine-Lubnaig and Stack-More pairs in previous works that suggests these should still be considered different ESUs (Adams et al., 2006; Adams et al., 2007b). Indeed, Arctic charr

in freshwater lakes seem to have very marked homing to their natal spawning sites, as seen in Lake Windermere (Frost, 1965), and so while these pairs are very geographically close, they represent distinct populations. The two Loch Maree and Stack ecotypes should be investigated more directly although these should still be considered separate ESUs. While most of these ESUs consist solely of one ecotype population, ecotypes such as the Loch Dughaill planktivore were placed in an ESU containing other populations, such as Loch Coulin and Fannich, due to patterns of genetic similarity seen in admixture and PCA analyses.

Finally, we considered populations which may have high divergence for evolutionary significance. For example, populations that are the only extant population in their respective HA have a higher potential to gain unique and important adaptive variation, due to their extreme isolation, and their loss would significantly affect the geographic range of the species. Lochs Eck, Doon, Bruicheach, Lee and Earn all represent the only populations in their respective Hydrometric Areas and therefore could be considered important groups with high divergence potential. The Loch of Girlsta, which is the sole extant population on the Shetland Islands, represents an ESU on the same criteria but needs further investigation as it was not included in our dataset (Maitland and Adams, 2018).

While populations like Loch Seil showed a notable adaptive pattern in our adaptive SNPs, the lack of previous work on this population means further investigation is needed before this can be considered an evolutionarily significant unit. On the basis of our current understanding, we defined all remaining groups the working designations of MUs. These designations should not be considered as definitive and more data on these populations as well as the many others not in our dataset should be used to further our understanding and help delineate accurate conservation units within the species.

A small number of Arctic charr populations are already included as notified features within protected sites (such as Sites of Special Scientific Interest) and are therefore protected in law. Whilst the formal designation of every ESU as a notified feature is unlikely, inclusion of some ESUs may be possible by adding them to the list of notified features in SSSIs protected for other purposes if their supporting habitat occurs within their boundaries. However, inclusion with the network of protected sites this is not the only way that Arctic charr populations can be protected. Protection of supporting habitats through Water Framework Directive (WFD) River Basin Management Plans, and the application of clear regulation to prevent the introduction of non-native species can all contribute to the health

and security of existing populations. This can also ensure the continuation of evolutionary processes.

## 2.6 Conclusions

Our results highlight the difficulty in defining important conservation units in a highly diverse species. The use of single criterion for defining conservation units, whether that be morphometrics or genetics, can lead to misleading assessment of conservation priority and we reject the endemic species classifications on this basis. Delineation of intraspecific conservation units should be based on data from a variety of sources but need for concordance between data types can limit what we define as evolutionary significant diverging groups or groups with high divergence potential that may not display concordance in data types but are important for the evolutionary potential of the species and deserve to be properly protected and managed as such. However, conflicting data should not be ignored in favour of preferred results and possible explanations as to why the data conflicts should be brought into consideration when defining important conservation units. We hope our study can act as a framework for identifying important and vulnerable sites in other highly diverse species.

# Chapter 3: Increased ecological opportunity leads to higher diversity and probability of trophic specialisation in Arctic charr

## 3.1 Abstract

Understanding the extrinsic factors that drive the emergence of diversity is critical to the protection and management of biodiversity. However, even superficially similar local environments can vary considerably, and so large-scale datasets are needed to understand the key drivers across a wider species distribution. The Arctic charr (*Salvelinus alpinus*) represents an excellent study system for these processes as the species shows substantial diversification to a range of freshwater environments and multiple instances of trophic specialisation in the form of ecotype populations. However, the processes behind this variation remain largely unknown due to a lack of large-scale genetic studies. To address this, we investigated the drivers of diversity using a national scale study of Arctic charr populations across Scotland (N = 64). Using a genome-wide dataset of SNPs (N = 24,878) and phenotypic data, we found that the extent of genetic and phenotypic diversity was strongly predicted by ecosystem size with larger, deeper lakes providing more ecological opportunity and diversification potential. Similarly, using environmental data from all lakes containing Arctic charr in Scotland (N = 187), we found that ecosystem size strongly predicted the potential for trophic specialisation. Our results show the importance of increased ecological opportunity in underlying adaptive radiations.

## 3.2 Introduction

A number of studies have investigated the drivers behind the emergence of diversity and the processes underlying local adaptation in a range of species (Primmer, 2011; Recknagel et al., 2014; Hoban et al., 2016; Franch-Gras et al., 2018; Manel et al., 2020). The genetic diversity of a population, and a species in general, is critical to its evolutionary and adaptive potential and therefore its viability (Reed and Frankham, 2003; Paz-Vinas et al., 2018; Ørsted et al., 2019). Thus, an understanding of the factors that facilitate or limit the emergence of diversity can help us better protect populations, species, and diversity itself.

Adaptive radiations act as excellent case studies for investigating the drivers of diversity as they often involve rapid diversification and adaptation to a range of similar yet distinct environments (Schluter, 2000; Grant and Grant, 2007; Losos, 2010; Elmer and Meyer, 2011). There are a number of different theories which try to explain the mechanics behind adaptive radiations, and place differing emphases on various spatial, temporal, environmental, and genomic factors (Schluter, 2000; Stroud and Losos, 2020; Richards and Martin, 2022). However, divergence in trophic niche, also known as trophic specialisation, is at the centre of numerous adaptive radiations, including those described in a number of fish species (Kocher, 2004; Litsios et al., 2012; Galvez et al., 2022). The divergence of populations to utilise unused resources allows for adaptation to novel environments leading to greater diversity. As such, a consistent factor across terrestrial and aquatic environments is that larger, often more complex, ecosystems show higher levels of genetic, phenotypic, and species diversity as they offer a larger number of niches and greater ecological opportunity (Tews et al., 2004; Nosil and Reimchen, 2005; Gavrilets and Losos, 2009; Wagner et al., 2012; Recknagel et al., 2014; Recknagel et al., 2017; Schmidt et al., 2022).

An example of an adaptive radiation can be seen in the salmonid species Arctic charr (*Salvelinus alpinus*) where populations diversified into a range of freshwater habitats across the Holarctic at the end of the last ice age (Maitland, 1995; Maitland and Adams, 2018). The species shows high levels of genetic diversity, as well as variation in a range of phenotypes including morphology, colouration, spawning time, and diet with the species described as ‘the most variable known vertebrate on Earth’ (Jonsson and Jonsson, 2001; Klemetsen, 2010; Klemetsen, 2013). Trophic specialisation of Arctic charr, has occurred in a number of lakes where multiple distinct populations, referred to as trophic ecotypes or ecomorphs, co-exist having diverged to specialise to different parts of their shared lake environment (Snorrason et al., 1994; Adams et al., 1998; Jonsson and Jonsson, 2001; Elmer, 2016). These divergences

have generally occurred along the benthic-limnetic axis with benthivorous and planktivorous ecotypes being expressed, although in some lakes a profundal benthivore ecotype or a large bodied piscivorous ecotype have been described (Malmquist et al., 1992; Adams et al., 1998; Alekseyev et al., 2002; Hooker et al., 2016). In Scotland, there are 10 lakes known to contain multiple ecotypes; however, there may well be more instances of trophic specialisation that are not yet discovered (Adams et al., 1998; Fraser et al., 1998; Adams et al., 2008; Garduño-Paz et al., 2010; Hooker et al., 2016; Jacobs et al., 2020). Studies have shown that these ecotype divergences have arisen through different ancestral histories across locations (Jacobs et al., 2020). The number of shared outlier and phenotype-associated loci is also known to be limited (Fenton et al., 2023a) (Chapter 6). These facts suggest that extrinsic factors such as local environment may play a particularly important role in underlying trophic specialisation.

Like in other adaptive radiations, ecological opportunity has been suggested to be key to the emergence of phenotypic and genetic diversity in Arctic charr (Hindar and Jonsson, 1982; Alekseyev et al., 2002). Larger lake size was recently shown to correlate with an increase in the number of ecotypes in polymorphic Arctic charr lakes, although this study focused on a localised part of Greenland (Doenz et al., 2019). Increasing ecosystem size has been shown to correlate with higher levels of phenotypic variability in Scotland (Recknagel et al., 2017) although the same study found no relationship with genetic diversity based on a limited number of microsatellites. Genetic studies on Arctic charr populations, particularly in Scotland, have generally either included a large number of populations but used a limited number of markers or used genome-wide datasets but focused on a small number of populations (Wilson et al., 2004; Jacobs et al., 2020). In Scotland there are approximately 187 lakes that contain known extant populations of Arctic charr (Maitland and Adams, 2018). Of these, only 35 have ever been included in a phenotypic study and only 22 in a genetic study (Table A3-1) thus much of the genetic and phenotypic variability in Scotland remains unexplored. While looking at a few select populations can allow for an in-depth understanding of the mechanics behind their variation, it can also lead to a distorted view on what drives diversity across a wider geographical distribution. By exploring a wide geographic range and a large number of lakes, we can more robustly test the factors that drive emergence of diversity and the processes behind local adaptation.

To address these questions, we investigated the key factors that underlie the emergence of diversity in one of the most highly diverse species in the world. To do this, we conducted a national-scale study of Arctic charr populations across Scotland in the British

Isles. We used a recently generated genome-wide dataset of SNPs and phenotypic data to characterise the extent of phenotypic and genetic diversity across populations and identify environmental drivers of that variation. We then used environmental data for all 187 Arctic charr containing lakes in Scotland to identify any consistent factors that underlie ecotype divergence to create a predictive model to identify potentially unknown ecotype divergences.

## 3.3 Material and Methods

### *3.3.1 Genomic data*

A ddRADseq dataset was generated, as part of Chapter 2 following the methodology of Jacobs et al. 2020 for 410 individuals covering 64 different populations across the British Isles (Figure 3-1). This included samples from 49 lakes in Scotland, three in Ireland, two in Wales, and one in England. Eight of the Scottish lakes contain multiple trophic specialist populations living in sympatry which for the purposes of our analyses were considered different populations. ddRADseq libraries of 90 samples plus five replicates were prepared using a modified version of Recknagel et al., 2015 ddRADseq protocol for Illumina sequencing platforms. Paired-end 75-bp sequencing was performed on the IlluminaNextSeq500 platform at Glasgow Polyomics. Raw ddRADseq data (ddRADseq NCBI short read bioproject: PRJNA607173) for eight of the Scottish lakes (Awe, Dughaill, na Sealga, Tay, Lubnaig, Merkland, Uaine, Eck) was generated as part of Jacobs et al. 2020. Reads were mapped to the annotated *Salvelinus* sp. genome (ASM291031v2) using *bwa mem v0.7.16* (Li, 2013). RAD loci building and SNP calling were performed in *Stacks v2.60* (Rochette et al., 2019). The *gstacks* module of *Stacks*, was run with the –gt-alpha set to 0.01 to ensure called genotypes were reliable when RAD-loci building. SNPs were retained in the populations module of *Stacks* if they met the following criteria: present in 66% of all individuals in each population and across all populations, a minimum minor allele frequency of 0.01, maximum observed heterozygosity of 0.5 with only the first SNP per locus retained. We used the script *filter_hwe_by_pop.pl* (https://github.com/jpuritz/dDocent) to filter any SNPs that were out of Hardy-Weinberg Equilibrium within populations, although no such SNPs were detected. *VCFtools v0.1.16* was used to identify and remove any SNPS with read depth <10x or >30x (Danecek et al., 2011). After filtering, we had a dataset of 24,878 SNPs. Measures of observed and expected heterozygosity were determined for all populations and loci using the –fstats parameter in the *populations* module of *Stacks* when calling SNPs.

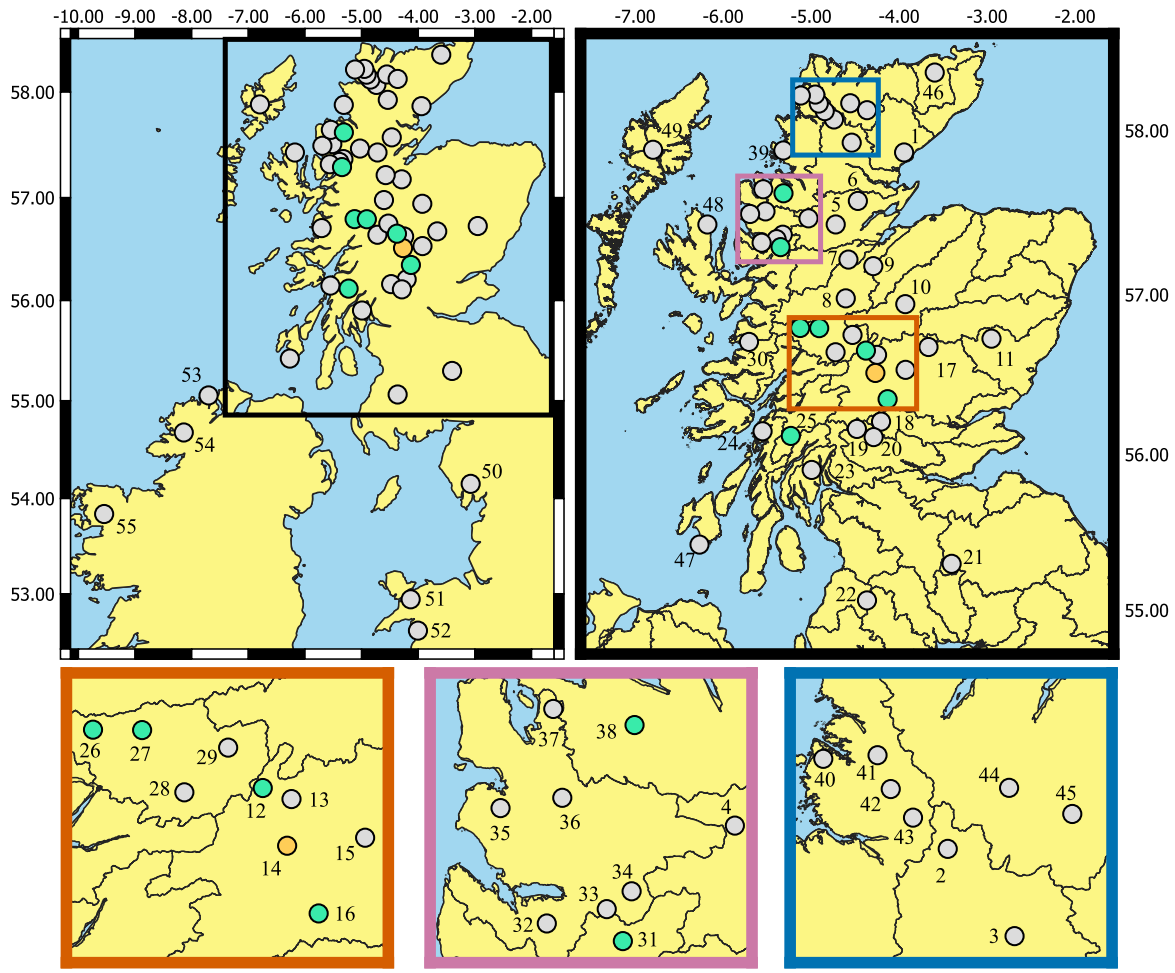


**Figure 3-1.** Map of populations included in our study, originally presented in Chapter 2. Grey points indicate unimodal populations, green indicate lakes with two ecotype populations (benthivore/planktivore) and orange indicates a lake with three ecotypes (benthivore/planktivore/piscivore). Lakes are numbered in ascending hydrometric area numbers for Scotland with outgroup populations thereafter.

### *3.3.2 Phenotypic variation*

We evaluated the level of variability in the expression of trophic morphology (or phenotype) amongst populations using a metric known as VarHD (Recknagel et al., 2017). This was determined from measurements of head depth, head length and fork length from full body photos using *ImageJ* (Table A3-2). To correct for differences in length, we regressed head depth and head length against fork length in linear regressions. We then calculated the residuals from these two regressions. The absolute difference between

residuals for each individual was then calculated to generate an index of head depth and head length (referred to as HD) free from body size effects, and which describes functional ecomorphology (Adams et al., 1998) following Recknagel et al., 2017. A high positive HD score indicates an individual with a bulky, deep head shape while a low negative score indicates a shallower, slenderer head. We then calculated four different morphological measures for each population: maximum (MaxHD), minimum (MinHD), mean (MeanHD) and variance in head depth (VarHD). VarHD was used as our measure of phenotypic diversity for each population. Head depth plots were made using the *ggpubr v0.6.0* R package (Kassambara A, 2023).

### *3.3.3 Environmental variables*

We investigated a range of factors from bathymetric, climatic, and geographical data as potential drivers of the extent of genetic and phenotypic diversity using linear regression models. Environmental data for 19 variables was collected from WorldClim using the *raster v3.5-15* R package with latitude and longitude co-ordinates for each lake (resolution 2.5 arcmin) (Hijmans and van Etten, 2012; Fick and Hijmans, 2017). Bathymetric data on lake surface area, mean and max depth, and altitude were collected for each lake from published work (Murray and Pullar, 1910; Maitland and Adams, 2018). An additional bathymetric variable littoral zone percentage (defined as the area of the lake shallower than 5 metres depth) was calculated using available bathymetric maps (Murray and Pullar, 1910) and *ImageJ v1.50i* (Schneider et al., 2012). We ran a principal components analysis (PCA) using these bathymetric variables to create a proxy for ecosystem size (Recknagel et al., 2017). PC1 from the PCA was used as the proxy, with littoral zone percentage loading in the opposite direction to the other bathymetric variables meaning PC1 acted as a gradient from smaller, shallower lakes with larger proportion of littoral zone to larger, deeper ones with smaller littoral zone percentage (Figure A3-1).

### *3.3.4 Predicting ecotype divergences*

Information relating to which lakes contained extant populations of Arctic charr in Scotland (N = 187) and bathymetric data for those lakes were taken from (Maitland and Adams, 2018) to identify unknown instances of ecotype divergence. We used maximum lake depth, mean lake depth, surface area, altitude, littoral zone percentage to create a metric for ecosystem size via a PCA. For lakes without bathymetric maps, littoral zone percentage was predicted using a linear regression for the relationship between littoral zone percentage and lake surface and maximum depth for lakes with bathymetric maps. We then created a

predictive one term model using ecosystem size, implemented with the predict function (Kuhn, 2008) to predict the possibility of currently unknown ecotype divergences in lesser studied lakes. For this model, the dataset of 187 lakes was split into two sets, training and test datasets. The training dataset (N = 24) contained all lakes with currently known sympatric ecotype divergences (N = 10 lakes) (Adams et al., 1998; Fraser et al., 1998; Adams et al., 2008; Garduño-Paz et al., 2010; Hooker et al., 2016; Jacobs et al., 2020) and any lakes that have been studied previously both in phenotypic and genetic studies but have shown no indication of ecotype divergence (N = 14 lakes) (Table A3-1). The test dataset contained all other lakes in Scotland known to hold Arctic charr (N = 163 lakes). A generalised linear model (GLM) was performed using the training dataset to create a predictive model. This model was then applied to the test dataset to give a probability of ecotype divergence for each of these lakes based on their ecosystem size.

## 3.4 Results

### *3.4.1 Patterns of genetic diversity*

Across our SNP dataset of 24,878 SNPs covering 64 populations (N = 410 individuals), observed heterozygosity ranged from 0.029 to 0.188, with a mean population heterozygosity of 0.099 (Table A3-1). There was no relationship between the level of heterozygosity and latitude (adjusted $R^2$ = -0.02, P = 0.86) (Figure A3-2) with the most diverse populations being found in the middle latitude range, specifically the Tay and Lochy catchments (lakes 12-17 and 26-29 on Figure 3-1). We explored a range of environmental and bathymetric variables for correlations with levels of diversity. While we found no significant relationships for many of our environmental variables related to temperature and precipitation with levels of genetic diversity, there were strong relationships seen with many of the bathymetric factors including maximum lake depth and lake surface area (Table A3-3). Amongst these, we found that ecosystem size showed a strong positive correlation with genetic diversity (adjusted $R^2$ = 0.32, P < 0.001) with larger, deeper lakes having more diverse populations (Figure 3-2A).

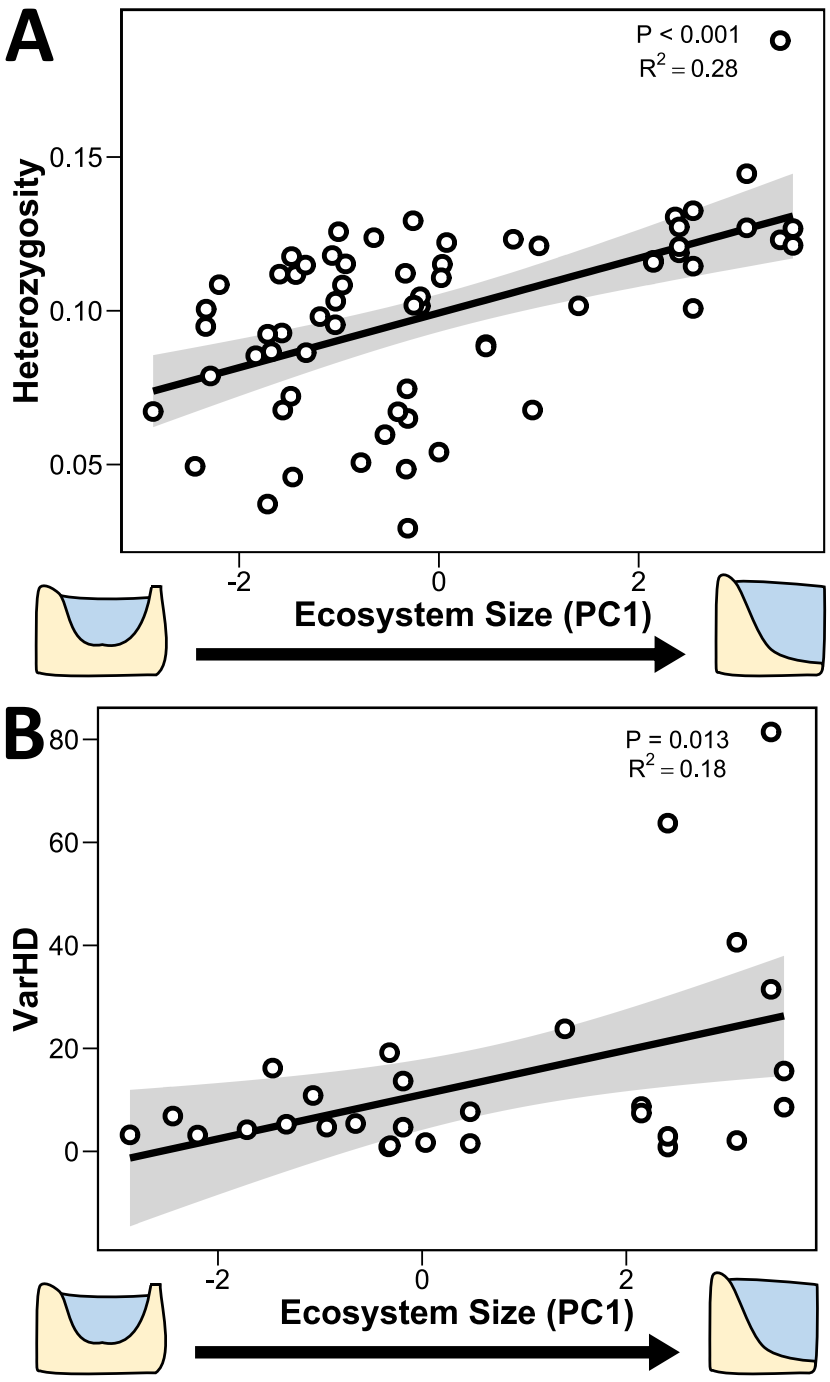


**Figure 3-2.** Ecosystem size plotted against observed heterozygosity in (A) and variation in HD (VarHD) in (B) for each population. PC1 from bathymetric PCA is used as proxy for ecosystem size (Figure A3-1). Ecosystem size increases with more positive PC1 scores going from small and shallow lakes to larger, deeper lakes with proportionally less littoral zone.

### *3.4.2 Patterns of phenotypic diversity*

There was notable variability in the extent of phenotypic variation across our populations (N= 29 populations and 195 individuals) (Figure 3-3A) (Table A3-2). The extent of head depth variability (VarHD) varied considerably across the dataset with the most variable population, the Lochy benthivore population (VarHD = 81.44) being considerably more variable than the least variable population, the Awe benthivore (VarHD = 0.85) with the mean VarHD across populations being 13.71. The individual with the slenderest head was found in the benthivore ecotype at Loch Ericht (MinHD = -13.69), while the individual with the deepest head was found at Loch Treig (MaxHD = 20.44). MeanHD varied from -5.36 in the Ericht benthivore population to 4.43 in the Loch Earn population. In lakes with multiple sympatric ecotypes, the benthivore populations on average had more negative HD scores than the planktivore populations however these differences were all non-significant (Table A3-4).

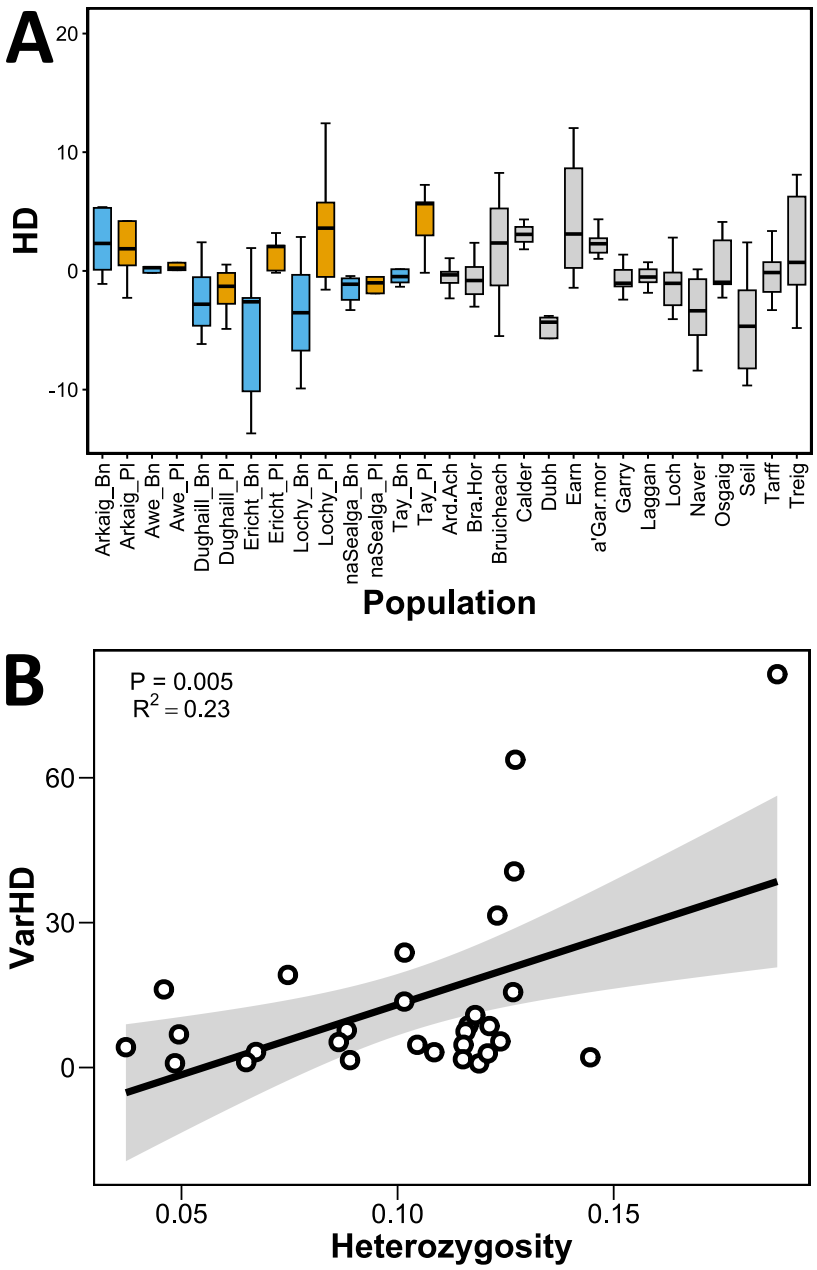


**Figure 3-3.** A) shows the extent of variation in HD (head depth index) across populations. Benthivore ecotype populations are coloured in blue and planktivore ecotype populations are coloured in orange. Distribution of HD is indicated by box and whiskers with median value indicate by black bars for each population. B) shows the relationship between VarHD and observed heterozygosity.

Phenotypic diversity (VarHD) correlated with ecosystem size, with larger, deeper lakes having more diverse populations (adjusted $R^2$ = 0.18, P = 0.013) (Figure 3-2B) consistent with previous findings (Recknagel et al., 2017). No significant relationship was seen for either MeanHD or MinHD with ecosystem size. MaxHD showed a positive relationship with ecosystem size suggesting that larger, deeper lakes generally support Arctic

charr individuals with deeper heads (adjusted $R^2 = 0.12$, $P = 0.037$) (Figure A3-3). We found a positive relationship between heterozygosity and VarHD with more genetically diverse populations showing higher levels of phenotypic variability in head depth (adjusted $R^2 = 0.23$, $P = 0.005$) (Figure 3-3B). Heterozygosity showed no relationship with any of other HD metrics (MeanHD, MaxHD, MinHD) (Figure A3-4).

### *3.4.3 Predicting ecotype divergence*

We investigated the importance of ecosystem size and its association with ecotype divergence and trophic specialisation using data from all 187 lakes known to contain extant Arctic charr populations in Scotland (Maitland and Adams, 2018). We explored ecosystem size as a predictive variable for potentially unknown ecotype divergences, as the lakes containing multiple ecotypes had a much higher average ecosystem size score (1.84 average PC1 score) than the rest of the lakes in our phenotypic dataset (-0.872 average PC1 score) ($P < 0.001$). Of the 161 lakes in the test dataset, eight showed notably high probability (Probability > 0.57) of ecotype divergence based on ecosystem size (Table 3-1) (Figure 3-4). These lakes all had larger ecosystem sizes than three of the known ecotype divergences (Stack = -0.20 PC1 score, Dughaill = 0.91 PC1 score, na Sealga = 1.70 PC1 score). The most likely lakes were Loch Ness (12.11 PC1 score), Loch Morar (6.67 PC1 score), and Loch Treig (4.43 PC1 score) with the former two having larger ecosystem sizes than Loch Tay which had the largest ecosystem size score of any of the known divergences (5.89 PC1 score).

**Table 3-1.** Lakes with high probability (Prob.) of ecotype divergence (>50%) based on ecosystem size (Eco.size). Data on latitude and longitude co-ordinates, altitude (m), max and mean depth (m), surface area ($km^2$), littoral zone percentage are also provided.

| Loch | Lat | Long | Alt | Max.dep | Mean.dep | Surf.area | Lit.zone | Eco.size | Prob. |
|---|---|---|---|---|---|---|---|---|---|
| Ness | 57.32 | -4.42 | 17 | 230 | 132 | 56.51 | 1.6 | 12.11 | 1.00 |
| Morar | 56.95 | -5.72 | 14 | 310 | 19 | 26.68 | 4.2 | 6.67 | 0.99 |
| Treig | 56.81 | -4.73 | 251 | 133 | 63 | 6.00 | 5.0 | 4.13 | 0.92 |
| Fannich | 57.64 | -4.96 | 256 | 86 | 33 | 9.93 | 9.0 | 2.43 | 0.65 |
| More | 58.29 | -4.85 | 43 | 96 | 38 | 3.78 | 8.5 | 2.38 | 0.64 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Assynt | 58.17 | -5.03 | 65 | 86 | 31 | 8.04 | 8.5 | 2.27 | 0.61 |
| Quoich | 57.07 | -5.29 | 203 | 86 | 32 | 7.14 | 10.7 | 2.15 | 0.58 |
| Morie | 57.75 | -4.47 | 193 | 82 | 38 | 2.39 | 7.0 | 2.08 | 0.57 |

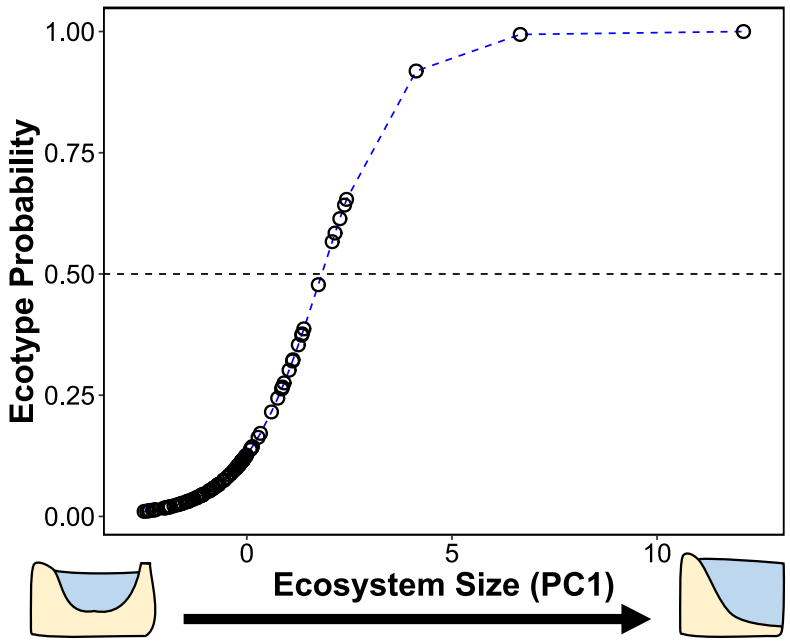


**Figure 3-4.** Probability of ecotype divergence based on ecosystem size. Dashed line at 50% ecotype probability. Individual points indicate lakes in the unknown dataset (N = 163 lakes). Ecosystem size increases from left to right along the X axis.

## 3.5 Discussion

Through our analyses, we identified increased ecological opportunity as a key driver of genetic and phenotypic diversity in Arctic charr. Larger, deeper lakes with proportionally less littoral zone displayed populations with higher levels of phenotypic and genetic diversity suggesting overall ecosystem size is key to levels of diversity. Similarly, we found ecosystem size was the strongest predictor of the probability of ecotype divergence within the lake suggesting that larger, deeper lakes provide the increased ecological opportunity needed for distinct trophic specialist populations to exist in the same environment.

### *3.5.1 Ecosystem size predicts levels of phenotypic and genetic diversity*

Studies across a range of terrestrial and aquatic environments have suggested that increasing ecosystem size provides higher ecological opportunity with more niches allowing adaptation to a more diverse range of resources and habitats and thus more phenotypically and genetically diverse populations (Nosil and Reimchen, 2005; Doenz et al., 2019; Schmidt et al., 2022). Lake surface area and maximum depth have both been demonstrated as important predictors of the degree of morphological variation and ecomorphological variation in a range of freshwater fish species including lake cichlids, European whitefish, and stickleback (Mcphail, 1993; Siwertsson et al., 2010; Recknagel et al., 2014)

We found that both phenotypic and genetic diversity in Arctic charr were strongly predicted by ecosystem size with larger, deeper lakes with proportionally less littoral zone housing phenotypically and genetically more diverse populations (Figure 3-2). These findings are consistent with previous work on Arctic charr which found a strong relationship between phenotypic diversity and ecosystem size although that study found no such relationship between ecosystem size and genetic diversity (Recknagel et al., 2017). This was due to the use of microsatellite data and the genome-wide approach adopted here shows that larger ecosystems allow for a greater range of viable genotypes.

While other studies have found that temperature and latitude can predict levels of genetic diversity (Martin and Mckay, 2004; Manel et al., 2020), we saw no such pattern in our dataset (Table A3-3). For the latter, this seems to be due to many of the highly diverse populations being found in the middle part of our range as seen in other species (Eckert et al., 2008; Willi et al., 2018), specifically the Tay and Lochy catchments (Lakes 12-17 and 26-29 on Figure 3-1). The Loch Tay and Rannoch ecotype populations, both found in the Tay catchment, have been suggested previously to be the result of secondary contact

scenarios (Verspoor et al., 2010; Jacobs et al., 2020) with secondary contact populations known to usually displaying higher levels of genetic diversity across species (Chiocchio et al., 2021; Salisbury et al., 2023; Fenton et al., 2023b) (Chapter 4). These populations being the result of secondary contact scenarios suggests a similar history for other populations in these catchments although these should be explored in detail. The lack of or weaker correlation between many of the climatic variables and extent of genetic diversity compared to the bathymetric variables is consistent with previous work that indicates that the diversity of freshwater species is driven more by biogeographic history and the shape of river basins than climate related variables such as surface temperature (Manel et al., 2020).

Across the lakes where multiple ecotypes have been described, a number of the planktivorous ecotypes showed more positive HD scores than their benthivorous counterparts suggesting they generally have bulkier, deeper heads (Figure 3-3A). This trend has been seen in past research on ecotype pairs (Garduño-Paz et al., 2012; Fenton et al., 2023a) (Chapter 6) however the opposing trend where planktivorous ecotypes show slimmer heads to aid with faster swimming and capturing evasive prey has also been seen (Adams et al., 1998; Adams and Huntingford, 2002; Recknagel et al., 2014; Hooker et al., 2016). The variability of this trend both within our dataset and across previous studies is expected as divergences between ecotypes are known to be quite variable across replicates likely due to a combination of differences in local environments and differences in ancestral histories (Fenton et al., 2023a; Blain et al., 2023) (Chapter 6). The positive correlation between MaxHD and ecosystem size across all lakes suggests that populations in larger, deeper lakes, have greater likelihood of deeper headed individual Arctic charr (Figure A3-3). Given that the planktivorous ecotypes generally have deeper heads, this indicates a shift towards increased use of the limnetic zone, which is proportionally greater in the larger ecosystems in our metric (Figure A3-1). This trend has previously been shown with populations of Arctic charr in larger ecosystems shifting their diet towards more limnetic sources and a more planktivorous diet (Eloranta et al., 2015).

### *3.5.2 Predicting trophic specialisation using ecosystem size*

An increasing number of studies have sought to examine the factors that drive the occurrence of co-existing ecotype populations particularly in salmonid fishes in freshwater systems (Seehausen and Wagner, 2014; Tiddy et al., 2023). Studies on *Coregonus* species have shown positive correlations between levels of phosphorous, nitrogen, and dissolved organic matter and the number of co-existing ecotypes (Landry et al., 2007; Siwertsson et al., 2010). Other factors such as predation, competition, and prey availability have been

implicated as being important contributors to ecotype divergence but detailed studies in this area have been limited (Vamosi, 2003; Landry and Bernatchez, 2010; Öhlund et al., 2020). Prey needs to be variable enough to facilitate the distinction in diet that underlies these ecotype divergences although the extent to which diet needs to vary can be relatively low (Adams et al., 2008; Galvez et al., 2022). Habitat heterogeneity is important for prey variability but also in creating enough variability in spawning locations to facilitate differentiation with several ecotype pairs known to show differences in spawning habitats (Jonsson and Jonsson, 2001). While shared outlier or phenotype-associated loci across ecotype pairs are known to be relatively limited (Jacobs et al., 2020; Fenton et al., 2023a) (Chapter 6), the importance of intrinsic factors, such as genetic variation, in facilitating ecotype divergence cannot be understated with key morphological differences between ecotypes known to be at least partly heritable (Adams and Huntingford, 2002). The most consistently important factors though are lake size and maximum lake depth which positively correlate with the number of ecotypes in numerous salmonid systems (Mcphail, 1993; Wagner et al., 2014; Gordeeva et al., 2015; Lucek et al., 2016).

In agreement with our other findings and previous research, we found that the probability of ecotype divergence was strongly predicted by overall ecosystem size with the majority of the lakes containing known co-existing ecotypes in Scotland being large, deep lakes. Overall, this suggests that these large ecosystems provide the increased ecological opportunity needed to allow the differentiation and co-existence of different trophic specialist populations. Notably, lochs Dughaill and Stack, two relatively small lakes (in terms of surface area), support sympatric polymorphisms of Arctic charr. Loch Dughaill displays a profundal ecotype rather than more typically seen benthivore ecotype (Hooker et al., 2016). The higher proportion of littoral zone and smaller surface area seen here may have led to the utilisation of the profundal zone to facilitate trophic specialisation. This could suggest that maximum lake depth may be more important than overall surface area in facilitating ecotype divergence in some situations. However, an important consideration is that we are applying predictive models using contemporary data to divergences that likely occurred thousands of years ago (Jacobs et al., 2020) (Chapter 5). It is entirely possible that these lakes have changed considerably over that timeframe and so lochs Dughaill and Stack may have been much larger and deeper than they currently are when these divergences occurred. While the relative importance of maximum lake depth and surface area may vary for each instance of trophic specialisation, our ecosystem size metric acts as an effective summarisation of both.

Several, largely unexplored, lakes had a high probability of supporting multiple ecotypes in sympatry based on our predictive model (Figure 3-4) (Table 3-2). These included the extremely deep lochs Ness and Morar whose maximum depths and surface areas far exceeded any lake with known ecotype divergences in Scotland. The populations in lochs Ness, Morar, Quioch, and Assynt need to be studied in detail to determine if there is any evidence for ecotype divergence with only the Ness population studied previously with no suggestion of multiple ecotypes at the time (Bean et al., 1996; Winfield et al., 2002). The Loch More, Morie, Fannich, and Treig populations all have individuals in our ddRADseq dataset and so we can evaluate the possibility of ecotype divergence further. Data from a previous study (Chapter 2) showed that the Loch Treig population comprised two distinct genetic clusters and show two mitochondrial haplotypes. This is seen in many known multiple ecotype lakes such as Loch Tay where Arctic charr display secondary contact ecotype divergence (Jacobs et al., 2020). The Loch More, Morie, and Fannich populations do not show these same pattens however, neither do the known sympatric ecotype divergences at lochs Awe and na Sealga (Jacobs et al., 2020) so the possibility of ecotype divergence remains. In this study, the Loch Treig population also shows high genetic and phenotypic diversity (Figures 3-2, 3-3), which further raises the possibility that this represents two distinct ecotype populations. However, all of these lakes should be studied in more detail before any conclusions are drawn. Our study highlights some populations of potential interest for research and further work into the factors underlying the emergence of diversity may yield further sites of interest.

## 3.6 Conclusions

Our results highlight the importance of using large-scale datasets containing numerous populations to investigate the factors and processes behind diversity. While a range of intrinsic and extrinsic factors will be involved, we found that ecosystem size was a critical predictor of the extent of phenotypic and genetic diversity and the probability of trophic specialisation. Our findings suggest that ecological opportunities available through larger ecosystems, in our case larger, deeper lakes, are the most significant component for a commonly seen diversification of freshwater fishes.

# Chapter 4: How glaciation impacted evolutionary history and contemporary genetic diversity of flora and fauna in the British Isles

Note: This chapter has been published as a literature review article in the *Scottish Geographical Journal*. (https://doi.org/10.1080/14702541.2023.2231407).

## 4.1 Abstract

Ice coverage not only affects the climate and landscape of geographical regions but has impacted species composition, the fragmentation and isolation of populations, and the colonisation routes of species post-glaciation. Major advancements in generating genetic data and applying sophisticated analyses to accurately model the demographic and colonisation history of animal and plant species has allowed evolutionary biologists novel insights. Meanwhile, physical geographers have made major advancements in reconstructing glacial history. However, the information flow between the geographic and evolutionary fields on this topic remains limited; consequently, evolutionary studies on contemporary biodiversity and species colonisation tend to be vague about the role of glacial history, while evolutionary and biodiversity information is rarely leveraged when discussing historical glaciation. In this review, we bring together current knowledge on the changing patterns of ice coverage in the British Isles from the maximum extent of the British-Irish ice sheet (BIIS) *ca.* 27 ka through to the end of large-scale ice coverage at the start of the Holocene period *ca.* 11.5 ka. We do so to then highlight how glaciation affected species composition during this time period and the subsequent colonisation of temperate flora and fauna and the patterns of genetic variation seen in them.

## 4.2 Introduction

In regions of the Northern Hemisphere that were previously glaciated, the patterns of glaciation and the timing of ice advance and retreat shaped the native flora and fauna and influenced biodiversity as we know it. Glaciers and ice sheets have drastically impacted the landscape of previously glaciated areas as well as nearby areas, affecting which species could colonise and the rate at which they could do so. The harsh environmental conditions associated with glaciated areas forced temperate species to survive in isolated populations, with gene flow between them interrupted and fostering conditions where populations diversified over time. This separation and distribution on the landscape, whether terrestrial or aquatic, impacted the genetic variation and distinctiveness of populations. Evidence of these historical patterns of genetic diversity and variation then remained when these populations colonised previously glaciated areas and at times mixed. In this paper, we review what is currently known about the complex patterns of glacial advance and retreat during the last glaciation of the British Isles in order to illustrate how glacial history has influenced patterns of biological diversity post-glaciation.

Throughout Earth's history, the cyclical process of glaciation has impacted the entire globe, most extremely in its dramatic local effects on the landscapes of temperate regions (Bickerdike et al., 2018). The global Last Glacial Maximum (gLGM) *ca.* 27,000 – 19,000 years ago (ka) represents the time in the last glacial period when ice sheets covered their maximum global extent. In the British Isles at the start of the gLGM, the entirety of Scotland, Ireland, the majority of Wales, and parts of England were covered by an ice sheet commonly referred to as the British-Irish Ice Sheet (BIIS). The impact of glacial history in the British Isles is particularly complex, due to parts of the British Isles not being covered by ice at any point during this time period. This means that ice sheets had major impacts on some areas and much less of an impact on others, although their influence can still be seen almost everywhere. Increasing levels of Northern Hemisphere solar insolation would lead to the BIIS retreating northwards over time. By the time of the Late Glacial Interstadial *ca.* 15 ka, a period of pronounced warming, ice was largely restricted to mainland Scotland (Ruddiman, 2008; Clark et al., 2012). However, this period of warming climate was short-lived and intense cold conditions returned between *ca.* 12.9–11.5 ka in period known as the Younger Dryas. The return of cold conditions was caused by the melting of ice into the Arctic Ocean from the Laurentide Ice Sheet which covered much of North America during the last ice age (Tarasov and Peltier, 2006). During this time, an ice cap formed over the western Highlands of Scotland known as the Loch Lomond Readvance (Bickerdike et al., 2018). Small cirques

and valley glaciers also formed in parts of England, Wales, and Ireland before the British Isles became ice-free *ca.* 11.5 ka, at the start of the Holocene period (Barr et al., 2017).

The environment of glaciers and ice sheets would have been too hostile for many of the plant and animal species that we now consider to be native to the British Isles to survive. Temperate flora and fauna species persisted during glacial periods in ice-free regions known as glacial refugia. These refugia acted as geographically isolated havens which protected various species from the harsh glacial conditions and were sources from which these species would colonise after glacial conditions ended. The isolated nature of these refugia means that different populations of the same species could evolve in different ways while in isolation, leading to genetic and phenotypic differentiation or even them becoming different species. Consequently, populations of the same species in different regions of the British Isles often have very different colonisation histories that are reflected in distinct evolutionary patterns that can be resolved today (McKeown et al., 2010; Crotti et al., 2020).

For its impact on landscapes and the separation of different refugial populations, glacial history has majorly affected biodiversity and the intraspecific (i.e., within-species) genetic variation we see today. However, the role of glacial history often remains underdiscussed or only loosely integrated into evolutionary studies looking at these patterns. With advances in genetic tools and analyses allowing us to infer patterns of demographic and colonisation histories with increasing accuracy, it is vital that we bring together the latest knowledge of glacial history, in high resolution detail in space and time, and relate it to evolutionary genetic diversity of extant flora and fauna.

In this review, we outline the current understanding of glacial history of the British Isles with sufficient depth to allow patterns of post-glacial invasion processes to be inferred. We first cover the general large-scale patterns of change in ice coverage from the maximum extent of the BIIS *ca.* 27 ka through to the final demise of ice coverage *ca.* 11.5 ka, at the end of the Younger Dryas. We emphasise which parts of the British Isles were covered by ice and when ice left these regions, specifically with a focus on how glacial history has affected biodiversity. We describe the current state of knowledge about how the species composition of the British Isles changed over this time period, from arctic assemblages to the temperate species of today. We show the current understanding of the colonisation routes that terrestrial and aquatic species would have taken into the British Isles and how these were influenced by glacial events across time and space. Finally, we provide three case studies which demonstrate the importance of tying glacial history into phylogeographic analysis by

illustrating the different ways glacial processes affected patterns of biodiversity in the British Isles.

## 4.3 Outlining changing ice coverage over time

Below, we overview how ice coverage in the British Isles changed from the start of gLGM *ca.* 27 ka to the start of the Holocene period *ca.* 11.5 ka. This is split into four sections: 1) detailing where ice was at the start of gLGM when the BIIS covered its maximum extent, covering more area than at any other point in time; 2) covering a small number of ice expansions into regions not covered by ice at the time of the gLGM; 3) covering patterns of retreating ice and detailing when major regions of the British Isles became ice-free; 4) covering the return of ice during the Younger Dryas period. Greater detail on how ice coverage changed over time can be found in Table 4-1, in the highlighted references, and a number of papers that summarise changes over these time periods (Clark et al., 2012; Clark et al., 2018; Bickerdike et al., 2018; Ballantyne and Small, 2019).

**Table 4-1.** A detailed chronology of the major changes in ice coverage in the British Isles from the gLGM through to the end of the Younger Dryas period. Approximate timings of these events are taken from the cited papers.

| Event | Time period | Reference |
|---|---|---|
| Local retreat of Irish Sea ice around the Preseli Mountains in South Wales | 26.7 ka | (Glasser et al., 2018) |
| A rise in relative sea levels led to ice retreating from the edge of the continental shelf west of Ireland, which had previously represented the westernmost extent of the BIIS | 26.3 ka | (Ó Cofaigh et al., 2019) |
| Extension of ice to the Isle of Scilly from the Irish Sea Basin | 25 ka | (Smedley et al., 2017) |
| Ice retreats northward from its southern extent in the Celtic Sea leaving much of southern Ireland deglaciated and an isolated ice cap over the Kerry Cork region | 23 ka | (Ó Cofaigh et al., 2012; Chiverrell et al., 2013; Barth et al., 2016) |
| Retreating BIIS begins to lose connection with the FIS in the northern part of the North Sea | 23 ka | (Bradwell et al., 2019) |
| Movement of ice from the North Sea into Yorkshire and Norfolk | 21 ka | (Evans et al., 2019) |
| Southward movement of ice through the Vale of York | 21 ka | (Bateman et al., 2015) |
| Ice from the Irish Sea moves into the Cheshire Basin | 21 ka | (Clark et al., 2012) |
| Retreat of ice in the North Sea seemed to trigger the initial and unstable re-growth of an independent ice cap centred around the Shetland Islands | 21 ka | (Bradwell et al., 2019) |
| Summits of the Wicklow Mountains in the east of Ireland become ice-free | 19-18 ka | (Ballantyne et al., 2006) |
| Cambrian, Rhinogydd, Moelwyns, and Arenigs mountains becoming exposed as nunataks in Wales | 19-18 ka | (Hughes et al., 2016) |
| Retreating ice leaves behind an isolated behind a Welsh Ice Cap | 19 ka | (Glasser et al., 2018) |
| Majority of the northern sector of the North Sea becomes ice-free as the Norwegian Channel Ice Stream undergoes a major retreat | 19 ka | (Clark et al., 2012) |
| Shetland Ice Cap underwent a significant reduction in size, from *ca.* 45,000 $km^2$ to 15,000 $km^2$, with substantial losses along its marine margin | 19 ka | (Bradwell et al., 2019) |
| Holyhead Mountain in Anglesey becomes exposed from ice coverage | 18.4 ka | (Hughes et al., 2016) |
| Northwest of England begins to emerge from ice coverage | 18-17 ka | (Bateman et al., 2015) |
| Loss of the Kerry-Cork ice cap | 18 ka | (Ballantyne et al., 2011; Clark et al., 2012) |
| Southern margin of the BIIS around the Isle of Man as ice retreats northward | 18 ka | (Clark et al., 2012) |

| | | |
|---|---|---|
| BIIS largely restricted to the northern half of Ireland, the Outer Hebrides, Orkney, and most of mainland Scotland | 17 ka | (Clark et al., 2012) |
| East coast of Scotland, from Peterhead down to the Tay estuary, becomes deglaciated | 17 ka | (Clark et al., 2012) |
| Any remnants of connection between the BIIS and FIS in the North Sea had now been severed | 17 ka | (Bradwell et al., 2019) |
| Irish and Scottish components of BIIS decouple as ice continues to retreat northwards | 16 ka | (Ballantyne et al., 2011) |
| Irish and North Sea are now largely deglaciated | 16 ka | (Ballantyne and Small, 2019) |
| Disappearance of the Welsh Ice Cap | 16 ka | (Glasser et al., 2018) |
| Shetland Ice Cap now wholly centred on the Shetland landmass | 16 ka | (Bradwell et al., 2019) |
| Irish and Scottish components of the BIIS decoupled | 16 ka | (Ballantyne et al., 2011) |
| Demise of the Shetland Ice Cap | 15 ka | (Bradwell et al., 2019) |
| Shrinking ice coverage leaves behind three small ice caps covering counties Dongeal, Sligo, Galway | 15 ka | (Clark et al., 2012) |
| No evidence of valley glaciers in the Lake District or the Vale of York suggests England was ice-free during the Late Glacial Interstadial | 15 ka | (Bateman et al., 2015) |
| BIIS is restricted to the NW Highlands with a separate ice cap over the Southern Uplands | 14.7 ka | (Clark et al., 2012) |
| Valley glaciers persisted in Wales as late as *ca.* 13.5 ka with Wales thereafter considered to be ice-free | 13.5 ka | (Glasser et al., 2018) |
| Ireland is ice free during the Late Glacial Interstadial | 13 ka | (Clark et al., 2012) |
| Return of Ice during the Younger Dryas period | 12.9 – 11.5 ka | (Bickerdike et al., 2018) |

Generally, there are two main approaches for reconstructing ice sheets: empirical reconstructions, which often use geomorphological and geochronological evidence to estimate ice sheet extent and dynamics, and numerical models which use ice physics to reconstruct ice sheet behaviour with the data discussed here largely coming from the former (Ely et al., 2019). Geomorphological information can detail ice-margin position and ice-flow direction using, for example, moraines which are ridges formed from glacial debris (Hughes et al., 2014). Geochronological data is used to determine the timing of ice-free conditions in different areas and can be derived, for example, from the time glacially eroded bedrock has been exposed (Small et al., 2017). As empirical evidence for the size and location of glaciers is not found continuously across the British Isles, there will remain some uncertainties over how the ice sheet margin changed over time and how this may have influenced the flora and fauna present today. Numerical models can be used to reflect known empirical evidence, and

the realistic implementation of ice physics can allow us to cast a light on regions where empirical evidence is lacking (Ely et al., 2019).

### 4.3.1 Ice coverage at the start of the global Last Glacial Maximum (gLGM) ca. 27 ka

At the start of the gLGM, the BIIS extended across the entirety of terrestrial regions of Ireland and Scotland while much of Wales, the north of England, and the area now occupied by the Irish Sea was also covered (Figure 4-1A) (Clark et al., 2012). In the North Sea, BIIS was conjoined with the larger Fennoscandian Ice Sheet (FIS), which covered large parts of northern Europe, leading to a fully glaciated northern sector of the North Sea (Clark et al., 2012). While a few expansions of ice occurred after this time, glacial retreats in other parts of its range mean that this was the maximum extent of the BIIS.

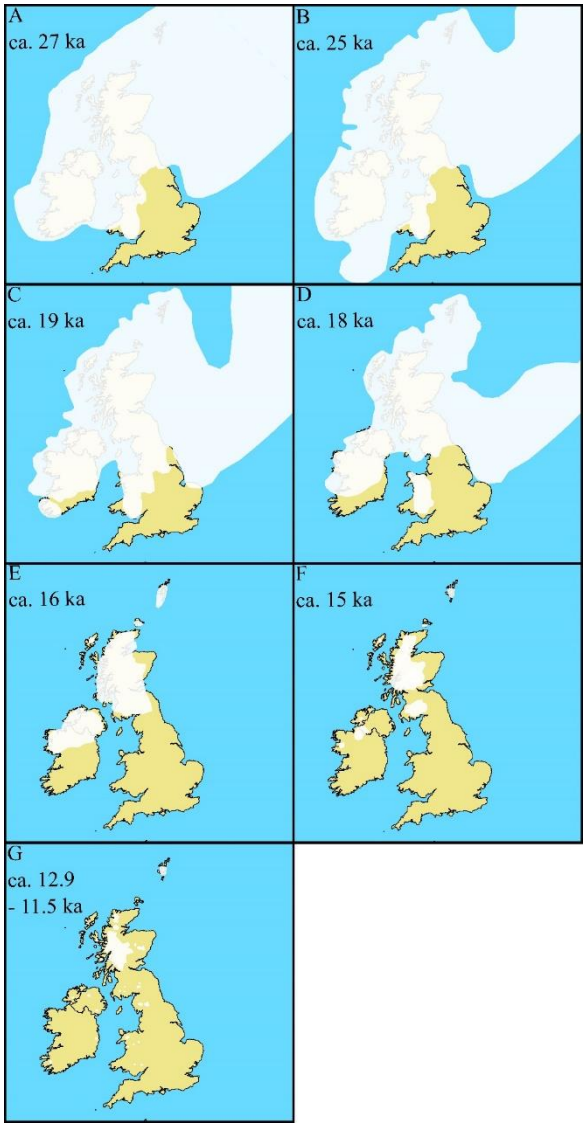


**Figure 4-1.** Map of the British Isles illustrating the extent of ice coverage at different time periods. (A) The extent of the British-Irish Ice Sheet (BIIS) at the time of the global Last Glacial Maximum (*ca*. 27 ka). The BIIS is confluent with ice from the Fennoscandian Ice Sheet (FIS) in the North Sea. (B) The advance of ice into the Celtic Sea as the BIIS reached its maximum southern extent (*ca.* 25 ka). (C) The rapid retreat of ice coverage northwards *ca.* 19 ka which resulted in the separation of the Kerry-Cork Ice Cap (KCIC) in southern Ireland. (D) The continued retreat of ice northwards which left behind a separate Welsh Ice Cap (*ca.* 18 ka). (E) The separation of the Irish and Scottish components of the BIIS *ca.* 16 ka. Connectivity between the BIIS and FIS has now been completely lost in the North Sea. (F) Ice coverage immediately prior to the Late Glacial Interstadial (*ca.* 15 ka). Large-scale ice coverage would be lost in all regions except the NW Highlands during the following period of warming. (G) Ice coverage during the Younger Dryas period (*ca.* 12.9 – 11.5 ka). All map backgrounds were generated using the Maps R package.

### *4.3.2 Southern advances*

The largest of the known southern advances was the expansion of ice through the Irish Sea Basin into the Celtic Sea, with the BIIS ultimately reaching approximately 150 km southwest of the Isles of Scilly *ca.* 25 ka (Figure 4-1B) (Smedley et al., 2017). Other advances were the movement of ice from the North Sea onto the Yorkshire coast and into the north of Norfolk (Evans et al., 2019), the southward advance of ice through the Vale of York (Bateman et al., 2015), and the penetration of ice from the Irish Sea into the Cheshire Basin all of which had begun by *ca.* 21 ka (Clark et al., 2012).

### *4.3.3 Rapid retreats*

By 23 ka, while small advances happened in other parts of its range, the BIIS had begun large-scale retreats. The first of these was from its short-lived advance into the Celtic Sea and the second from the northern most part of the North Sea losing connection with the FIS in this region (Chiverrell et al., 2013; Bradwell et al., 2019). The ice sheet rapidly retreated northward across southern Ireland, leaving much of the region deglaciated (Ó Cofaigh et al., 2012; Barth et al., 2016). The exception was the Kerry-Cork region which was covered by an isolated ice cap, the Kerry-Cork Ice Cap (Figure 4-1C) until its disappearance *ca.* 18 ka after which southern Ireland was ice-free (Ballantyne et al., 2011).

*Ca.* 18 ka, as the BIIS continued to retreat northward, ice from Ireland and Wales decoupled and left behind a Welsh Ice Cap, which covered much of central Wales (Figure 4-1D). The Welsh Ice Cap would decrease in size over time until it is believed to have disappeared *ca.* 16 ka (Glasser et al., 2018). Ice had retreated to such an extent by *ca.* 16 ka that most of the North Sea and Irish Sea were deglaciated and the Irish and Scottish components of the BIIS were by now decoupled from each other (Figure 4-1E) (Ballantyne and Small, 2019).

By the start of the Late Glacial Interstadial *ca.* 15 ka, the BIIS was mostly restricted to the northwest Highlands with a separate ice cap covering the Southern Uplands (Figure 4-1F) (Clark et al., 2012). The ice sheet which had remained in Ireland separated, leaving behind three smaller ice caps which covered the uplands of counties Donegal, Sligo, and Galway (Clark et al., 2012). Any remaining ice coverage or valley glaciers present in England, Wales, and Ireland would disappear during the Late Glacial Interstadial with ice only remaining in the northwest Highlands during this brief period of warming (Clark et al., 2012; Bateman et al., 2015; Glasser et al., 2018).

### *4.3.4 Return of ice coverage during the Younger Dryas*

The Younger Dryas, or the Loch Lomond Stadial (LLS) as it is known within the British Isles, was a period of abrupt cooling that occurred between *ca.* 12.9 – 11.5 ka (Carlson, 2013). During this period, ice returned to a number of regions in the British Isles (Figure 4-1G). This period is characterised by the formation of a large ice field in Scotland called the West Highland Glacier Complex or the Loch Lomond Readvance. This glacier complex stretched from the Slioch Mountain in the north to Loch Lomond in the south. Data suggests that the complex covered the entire width of the Highland Mountains, however patchy evidence along the west coast means it is possible ice coverage extended offshore in this region (Bickerdike et al., 2018). Other parts of Scotland show evidence of small ice fields, cirques, or valley glaciers forming during this time period (Bateman et al., 2015; Boston et al., 2015). A number of ice-dammed lakes also existed in Scotland during this time period. These lakes existed where competing ice masses met and were unable to drain properly. These include a series of ice-dammed lakes formed over Glen Spean and Glen Roy and a separate, more northern, lake situated near the village of Achnasheen (Sissons, 1977; Benn, 1989; Bickerdike et al., 2016; Bickerdike et al., 2018).

Ice coverage in England, Wales, and Ireland was restricted to small cirques and valley glaciers located in mountainous regions (Anderson et al., 1998; Barr et al., 2017; Wilson et al., 2018; Bickerdike et al., 2018). The return to glacial conditions was brief and by the beginning of the Holocene period, *ca.* 11.5 ka, all glacial features are thought to have disappeared.

## 4.4 Species composition during the last ice age

Due to the harsh cold conditions, the species composition of the British Isles would have looked very different during the last ice age than it does today. Rather than being barren, ice sheets and the surrounding tundra supported their own biomes of microorganisms and cold-adapted flora and fauna respectively (Ruddiman, 2008; Anesio and Laybourn-Parry, 2012; San-Miguel-Ayanz et al., 2016). The upper layer of ice sheets, like the BIIS, are dominated by glacial algae while water-filled depressions on the surface of glaciers, known as cryoconite holes, act as hotspots for insects, bacteria, and fungi (Zawierucha et al., 2015; Hoham and Remias, 2020). At ground level, the base of glaciers and sub-glacial lakes provide habitats for a wide range of prokaryotic microorganisms (Stibal et al., 2020).

The non-glaciated regions of the British Isles would have been tundra and mountain biomes which supported arctic and sub-arctic mammal communities including reindeer (*Rangifer tardanus*) and Arctic fox (*Alopex lagopus*), as well as sparse assemblages of arctic-alpine plants (Ruddiman, 2008; San-Miguel-Ayanz et al., 2016). Retreating ice sheets and increasing temperatures would have led to the demise of terrestrial cold-adapted species, such as the giant deer (*Megaloceros giganteus*) and the mammoth (*Mammuthus primigenius*), and eventually allowed temperate species to colonise (Woodman et al., 1997; Montgomery et al., 2014).

## 4.5 Glacial refugia

Until the end of large-scale ice coverage in the British Isles, temperate species would have persisted in ice-free havens known as glacial refugia. Whilst the traditional view is that these glacial refugia were found solely in southern Europe, particularly in the Balkan, Italian, and Iberian Peninsulas, there is increasing evidence for a number of glacial refugia in northern Europe as well as in the British Isles (Finnegan et al., 2013; Hewitt, 1999; Young et al., 2021). For example, pollen records from Bodmin Moor in the southwest of England suggest the area may have acted as a glacial refuge for spruce (*Picea*), pine (*Pinus*), and fir (*Abies*) trees during the gLGM while similar refugia existed for hazel (*Corylus*) and alder (*Alnus*) trees in other parts of the British Isles (Bush & Hall., 1987; Deacon, 1974; Kelly et al., 2010). The influence of these more northern glacial refugia can be seen in the colonisation patterns of a number of species. For example, elm (*Ulmus*) and lime (*Tilia*) trees show patterns suggesting expansion almost in parallel in the north and south of Europe rather than expanding south to north, following patterns of glacial retreat, due to the presence of both northern and southern European glacial refugia (Giesecke et al., 2017).

## 4.6 Colonisation routes

### *4.6.1 Terrestrial routes*

Many terrestrial species such as badgers (*Meles meles*), red squirrels (*Sciurus vulgaris*), and barn owls (*Tyto alba*) were able to colonise from glacial refugia on continental Europe due to lower sea levels allowing a number of low-lying landmasses to act as land bridges that connected Britain and Ireland to one another and to mainland Europe (Montgomery et al., 2014; Walker et al., 2020; Machado et al., 2022). The most notable of these land bridges was referred to as Doggerland and occupied much of the modern-day North Sea connecting Britain to mainland Europe (Coles, 2000; Ballin, 2017) while another

land bridge going from Land's End to Cork connected Ireland and Britain (Lambeck, 1995). The importance of these land bridges as colonisation routes for terrestrial species is illustrated in the difference in mammalian species diversity between Ireland and Britain. Around 15 ka, the release of massive volumes of meltwater from receding ice sheets resulted in a rapid rise of global sea levels (Edwards and Brooks, 2008). This would lead to the formation of the Irish Sea and flooded the land bridge that connected Ireland with Britain and mainland Europe. Britain meanwhile would remain physically connected to Europe through Doggerland until *ca.* 8 ka with the final connecting points between Holland and East Anglia (Weninger et al., 2008; Walker et al., 2020). The earlier cut-off of Ireland from mainland Europe led to fewer mammalian species colonising and establishing, resulting in lower mammalian species diversity in Ireland compared to Britain (Montgomery et al., 2014).

An important consideration for species colonisation is the rate at which different species arrive and establish and how this can be influenced by a number of factors. For example, in tree species, the species' optimum temperature, light requirement, and soil disturbance are all known to predict the post-glacial arrival time of species (Alsos et al., 2022). Differences in these factors created notable time lags in the appearances of different species in the post-glacial historical record (Bennett et al., 1991; Giesecke et al., 2017). The colonisation of flora can be greatly influenced by soil composition and drainage as these affect the rate at which soil goes from immature to mature to undergoing impoverishment and acidification with different species colonising better under different soil conditions. Areas nearer the sources of ice sheets are known to undergo soil maturation quicker, while areas of high precipitation, such as mountainous regions carved by glacial action, will subsequently undergo soil acidification quicker which hinders the establishment of species such as oak (Pennington, 1986). These factors would also affect fauna colonisation, because the flora composition of a region can directly affect fauna composition. As such it is important to consider the supporting habitat requirements of different species when discussing their patterns of colonisation.

### *4.6.2 Aquatic routes*

The colonisation of various freshwater fish species would have varied greatly based on their salinity tolerance (Cano-Barbacil et al., 2022). Anadromous species like Arctic charr (*Salvelinus alpinus*), which can freely migrate into wholly marine environments, would have been able to colonise rivers and lakes that were accessible by the sea once they became exposed from ice cover (Wheeler, 1977). As such, it is likely that salmonid species like Arctic

charr and brown trout (*Salmo trutta*) were the first to colonise British freshwaters after the last glacial period. This is supported by studies which have examined the postglacial colonisation of Alaskan freshwaters by anadromous salmonid species like Arctic charr (Milner and Bailey, 1989; Milner et al., 2000).

A number of glacial lakes may have acted as important connective routes for freshwater species unable to cross marine environments, known as stenohaline species. Between *ca.* 20 - 14 ka, a pro-glacial freshwater lake called Lake Hibernia is thought to have existed in the present-day Irish Sea, dammed to the north by the receding BIIS and to the south by a land bridge extending from Land's End to Cork (Lambeck, 1995). Similarly, a pro-glacial lake is believed to have existed in the southern part of the North Sea between *ca.* 20 - 18 ka dammed to the north by the still connected BIIS and FIS. A number of large rivers including the Thames, Rhine, and Scheldt are believed to have discharged into this lake (Quigley, 2014). Doggerland also contained a number of connected lakes and rivers which linked rivers such as the Humber, Thames, and the Rhine (Coles, 2000). These lakes and freshwater systems may have acted as connective routes to allow various species to colonise British freshwater environments at various points in time. The high number of populations of stenohaline species like roach (*Rutilus rutilus*) and barbel (*Barbus barbus*) in the southern parts of England likely reflects the connectivity river systems in these regions had with continental Europe (Wheeler, 1977).

While the creation of the Irish Sea resulted in the loss of the land bridge connecting Ireland and Britain, it may have been a key event in allowing the colonisation of some stenohaline freshwater species. Stenohaline species struggle to tolerate the wide fluctuations in the salinity of sea water and thus struggle to cross marine environments. The release of massive volumes of freshwater from receding ice sheets would have led to a major decrease in the salinity of the seas around the British Isles. This may have allowed the creation of temporary low salinity water corridors which would allow stenohaline species to cross marine environments, that they would otherwise be unable to do (Wheeler, 1977; Quigley, 2014).

## 4.7 How glaciation has affected intraspecific genetic variation

Glaciation has had major impacts on within-species variation through the isolation of populations in different glacial refugia, allowing for the occurrence of numerous differences between populations in different refugia. Therefore, differences seen in present

day populations of the same species across the British Isles may reflect patterns of colonisation from different glacial refugia populations.

Advancements in the field of genomics now mean that we can use methods like coalescence analysis and phylogenetic trees to investigate the demographic history of species using whole genome or genome-wide datasets in a level of detail simply not achievable even two decades ago (Excoffier et al., 2013; Excoffier et al., 2021). With the ability to investigate populations at high resolution, it is more vital than ever to place the findings of these analyses into the context of known glacial history to truly understand these patterns.

We present three example species – two aquatic and one terrestrial – where various genetic techniques were used to uncover the history of these species in the British Isles. We will then relate these findings to glacial history to demonstrate how glaciation affected their patterns of intraspecific variation in different ways.

### *4.7.1 Case 1: Different refugia populations colonising different parts of the British Isles*

For many fish species, populations across the British Isles have radically different ancestral histories owing to their colonisation from different glacial refugia (McKeown et al., 2010). If different parts of the British Isles were colonised by different refugia, then populations colonised by those different lineages will be genetically distinct from each other.

An example of this is seen in the European whitefish (*Coregonus lavaretus*) which is relatively common across northern continental Europe and has isolated populations in England, Wales, and Scotland. Previously, due to patchy sampling and applications of different genetic markers, little was known about how the populations across Britain related to each other, nor was much known about the colonisation history that gave rise to them (Winfield et al., 2013; Crotti et al., 2020). A genomic study was conducted to determine how populations across the British Isles related to one another. Using population diversity and differentiation analyses on a genome-wide dataset of more than 25,000 SNPs (single nucleotide polymorphisms) the study found the populations split into two distinct groups, one for the Scottish populations and one containing the English and Welsh populations (Crotti et al., 2020). This suggested a higher level of genetic similarity between populations in England and Wales to those in Scotland. This is, potentially, the result of England and Wales being colonised by the same ancestral lineage and one which was different to that which colonised Scotland. These, in turn, were both separate from the other European

populations, from Russia, Norway and the Baltic, which were also included in the study (Crotti et al., 2021).

As well as modelling the evolutionary relationships using the genome-wide data, a complementary approach using mitochondrial DNA was also used. This is based on fewer genetic data per individual but is a powerful way of looking at historical differences between populations because it has a clock-like mutation rate and a well understood evolutionary pattern (Avise, 1989). Genomic analyses show that the English and Welsh populations are more closely related to, and recently diverged from, continental European (alpine) lineages, whilst the Scottish populations are more distinct and have genetic contributions from an older Baltic ancestry (Crotti et al., 2021). Coalescence analysis, which estimates the time of divergence between populations, was then used to determine the timing of the splits between these different populations. The analysis showed that the populations found in Scotland diverged from a shared common ancestor with the English/Welsh populations *ca.* 65 ka and remained isolated from the other British populations until *ca*. 12 ka when low levels of contact (gene flow) began (Figure 4-2). The English and Welsh populations were from a more long-term shared lineage and indicated to have diverged from each other *ca.* 14 ka. These genetic signatures and timings show that the Scottish populations were colonised from a different ancestral lineage to the English/Welsh populations which were separated much earlier than the end of the ice age. The divergence of English and Welsh populations is estimated to have occurred after or around the period when England and Wales are assumed to have been ice free. As discussed earlier, this suggests that these populations came from the same colonising lineage with the populations then diverging in their respective freshwater habitats.

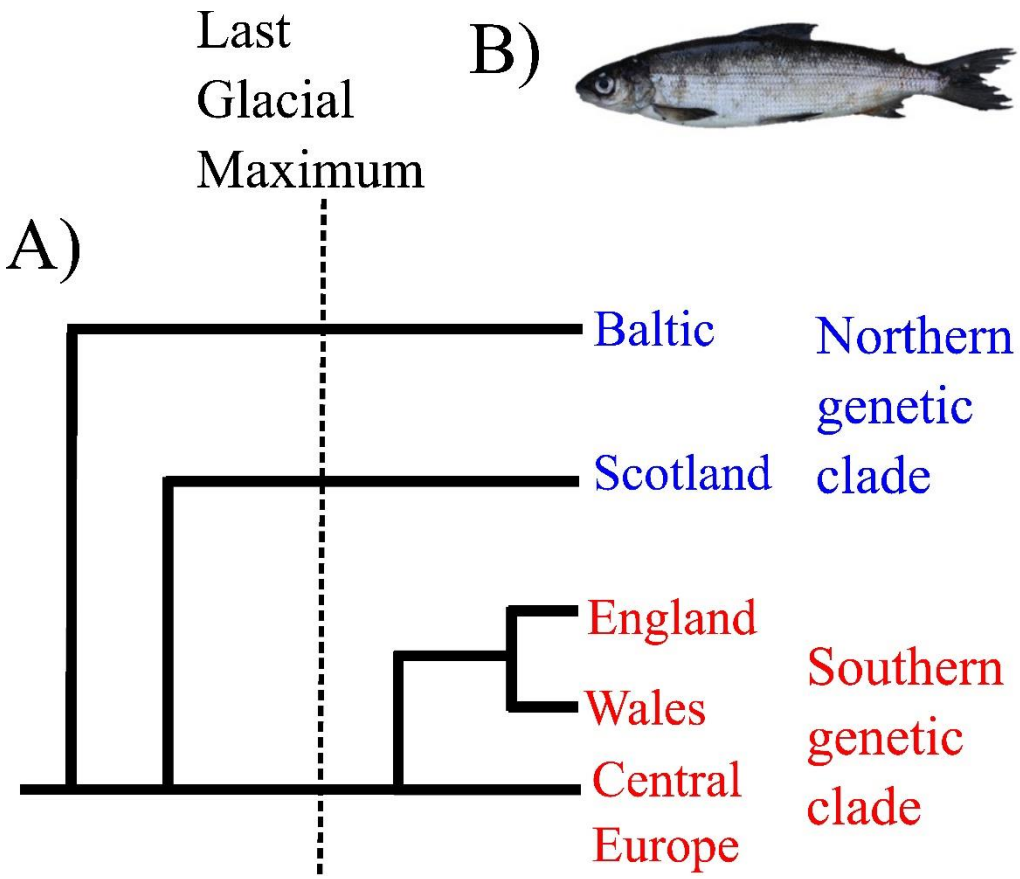


**Figure 4-2.** A) A tree showing the ancestral history of populations of European whitefish from different parts of the British Isles. Branches represent the extent of divergence between populations with the top of the branch being the most recent common ancestor of all populations. The shorter the branch, the more recently the populations diverged from each other. The dotted horizontal line reflected the timing of the global Last Glacial Maximum. Genetic analysis shows that populations from England and Wales were colonised by the same glacial refugia population which also covers central Europe while the Scottish populations were colonised from a more northern refugia which also colonised the Baltic after the end of the last ice age. Figure based on results from (Crotti et al., 2020; Crotti et al., 2021). B) A picture of European whitefish from Loch Eck in Scotland.

Examples like this, which show a distinct geographical separation of two different colonising lineages, even in relatively close contemporary geographic proximity, means there was little chance for genetic admixing between different ancestral populations/lineages. This can make inference of the ancestral history of different populations easier as historical patterns are not further complicated by recent admixing between populations (Gonçalves et al., 2007). These periods of isolation, and sometimes later contact, are pervasive signatures in the genetics of northern taxa and a source of

intraspecific diversity (Bernatchez and Wilson, 1998), and potentially even speciation (Hudson et al., 2011).

### *4.7.2 Case 2: Secondary contact between different colonising refugia populations*

Patterns of genetic variation across the British Isles are complex for many species due to different refugial lineages coming into contact with one another and then mixing genetically, such as, when a northern glacial refuge population expanded outwards and met with a population colonising from the south. This can have a number of different outcomes at the points of contact. One of the two lineages may fail to establish either due to maladaptation or from being outcompeted by the other lineage. Alternatively, the two lineages may mix and form a single gene pool, resulting in populations of higher genetic variation than would be found in other parts of the species' range (Finnegan et al., 2013; Chiocchio et al., 2021). Another potential outcome of such secondary contact is that the different lineages remain separate from one another and differentiate, usually after a period of some genetic mixing, to utilise different parts of their shared habitat, referred to as trophic specialisation and generating distinct ecotypes (Schluter, 2000).

A number of lakes in Scotland display multiple ecotypes of Arctic charr which are diverged along the benthic-limnetic axis, are reproductively isolated, and use different ecological niches (Adams et al., 1998; Verspoor et al., 2010; Jacobs et al., 2020). These repeated ecotypes, described as benthivorous and planktivorous due to their diets, show a number of phenotypic similarities across lakes (Klemetsen, 2010; Maitland and Adams, 2018). The repeated appearance of these ecotype pairs seen in contemporary populations may suggest that these pairs evolved from similar colonisation histories (Schluter, 2000). For example, that each lake with an ecotype pair represents a contact point between distinct colonising lineages, or that each pair is the result of sympatric divergence within the lake after colonisation by a single lineage (Elmer, 2016).

A recent study used high-resolution coalescence analysis and approximately 12,000 SNPs sampled across the genome to test different scenarios under which four different Arctic charr ecotype pairs, found in four separate lakes, diverged in Scotland (Jacobs et al., 2020). The study tested various models ranging from those representing a sympatric divergence within the lake under isolation with migration to those representing secondary contact, i.e., a period of isolation with no migration before coming back into contact, with the likelihoods of each model compared to one another.

The study found that populations found in lochs Tay and Dughaill, had ecotypes showing signatures of postglacial secondary contact followed by admixture or introgression, with ancestral populations that had most likely diverged prior to the gLGM. For example, the ancestral lineages of the ecotypes in Loch Tay diverged *ca.* 30 ka followed by more recent gene flow and differentiation. The ecotype pairs at lochs Awe and na Sealga showed much more recent divergences with the Loch Awe ecotypes for example diverging *ca.* 4 ka, a long time after that region in Scotland was ice-free. Given the timings of divergence and our knowledge of glacial history, the study concluded that the ecotype pairs at lochs Tay and Dughaill were the result of secondary contact between distinct colonising lineages, that diverged during the last ice age, and their current lakes as these were covered by ice around the time of their divergences (Ballantyne and Small, 2019). Meanwhile the pairs at Awe and na Sealga represent recent sympatric divergences within the lake, after colonisation of that lake by a single lineage (Figure 4-3). This highlights the importance and novel insights of genomic studies and estimations of historical divergence as solely examining the similar phenotypes of the ecotypes across lakes would have led to the incorrect conclusion of similar ancestral histories.

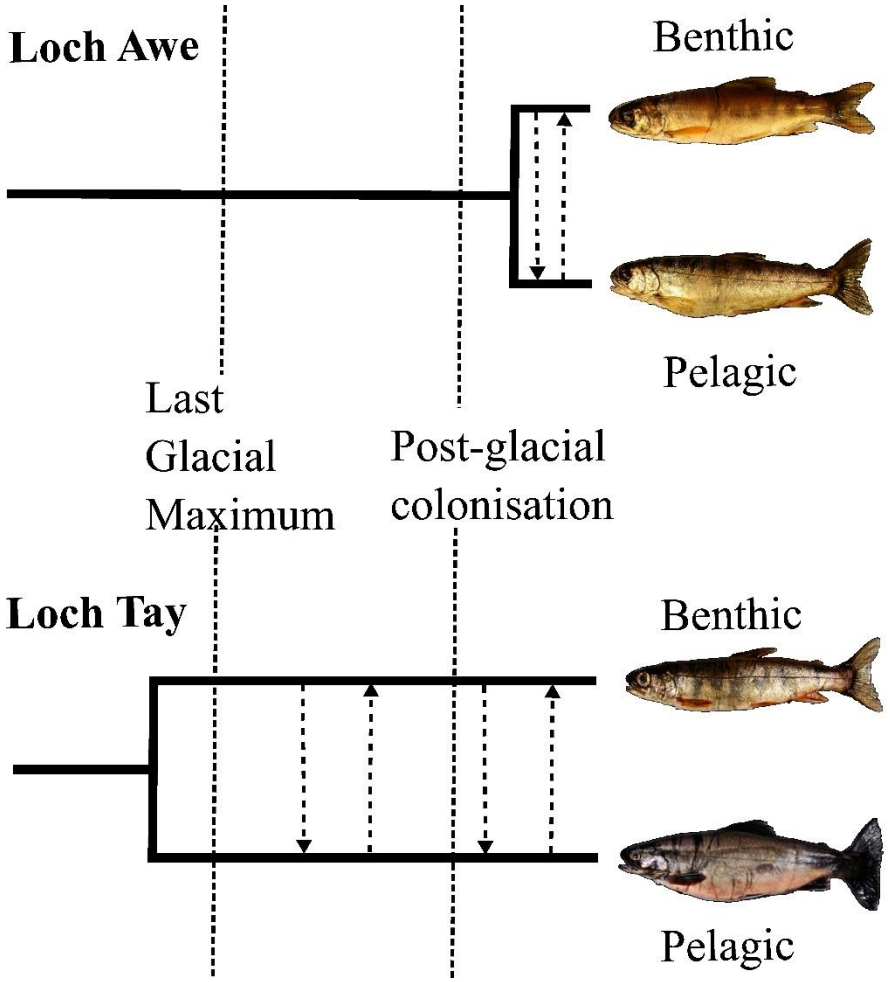


**Figure 4-3.** Trees demonstrating the differing divergences of benthic-pelagic ecotypes of Arctic charr in two Scottish lochs. The two dotted horizontal lines reflected the timing of the global Last Glacial Maximum and post-glacial colonisation of these lakes respectively. The branches represent the divergences of the ecotypes from one another. Dotted arrows reflected gene flow between the ecotypes after divergence. The ecotypes at Loch Awe are estimated to have diverged from one another *ca.* 4 ka making this a sympatric divergence from a single colonising population. The ecotype pair at Loch Tay are believed to have diverged *ca.* 30 ka meaning they diverged during the last ice age and represent a case of allopatric divergence followed by secondary contact after both lineages colonised the same loch. Figure based on results from (Jacobs et al., 2020). Pictures of benthic and pelagic individuals from each respective loch are shown.

### *4.7.3 Case 3: Population fragmentation during the Younger Dryas*

Another important factor, when considering the colonisation of species and patterns of genetic variation, is the impact of the Younger Dryas period. Whilst ice coverage certainly was not on the same scale as gLGM, cold conditions during the Younger Dryas would have created condition similar to the last ice age and thus the temperate species that had colonised the British Isles would have struggled to survive (Golledge et al., 2008). Records from Hawes Water in northern England for example, show the disappearance of birch (*Betula*) woodlands early in the Younger Dryas period after they had initially colonised during the Late Glacial Interstadial. The appearance of grassland and tall herb communities then followed which in turn gave way to Arctic taxa, such as purple saxifrage (*Saxifraga oppositifolia*), before the start of the Holocene when birch returned (Marshall et al., 2002). The return of ice and cold conditions during the Younger Dryas may have led to the fragmentation of newly established populations into temporary refuge populations. These different populations would have then expanded outwards at the start of the Holocene similar to what was seen at the end of the last ice age, potentially leading to distinct colonisation regions or secondary contact and admixture, as explored in the previous two case studies.

A species that highlights the importance of the Younger Dryas on intraspecific variation is the Eurasian field vole (*Microtus agrestis*). A study looking at populations across Europe found, through phylogenetic analysis using mitochondrial data, that there were six distinct genetic clades/groups across Europe with one clade containing only populations found in the north of Britain (Herman and Searle, 2011; Herman et al., 2014). This unique clade seems contradictory to what would be expected for colonisation of terrestrial species like field vole, which would have used the land bridges that connected Britain to Europe at the end of the last ice age. Some presence of this clade in the south of Britain or Europe would be expected as a contemporary reflection of the path the species took during colonisation.

The answer to how this clade could exist became clearer when coalescence simulations were run. Firstly, the study found that the most recent common ancestor between all six clades was dated to *ca.* 23 ka (Herman et al., 2014). Secondly, the unique clade to northern Britain was dated to have diverged from its closest neighbour, the clade containing the southern British populations and those found in Western Europe, *ca.* 12 ka. Similar timings are suggested for the divergence of several other clades from one another (Figure 4-4).

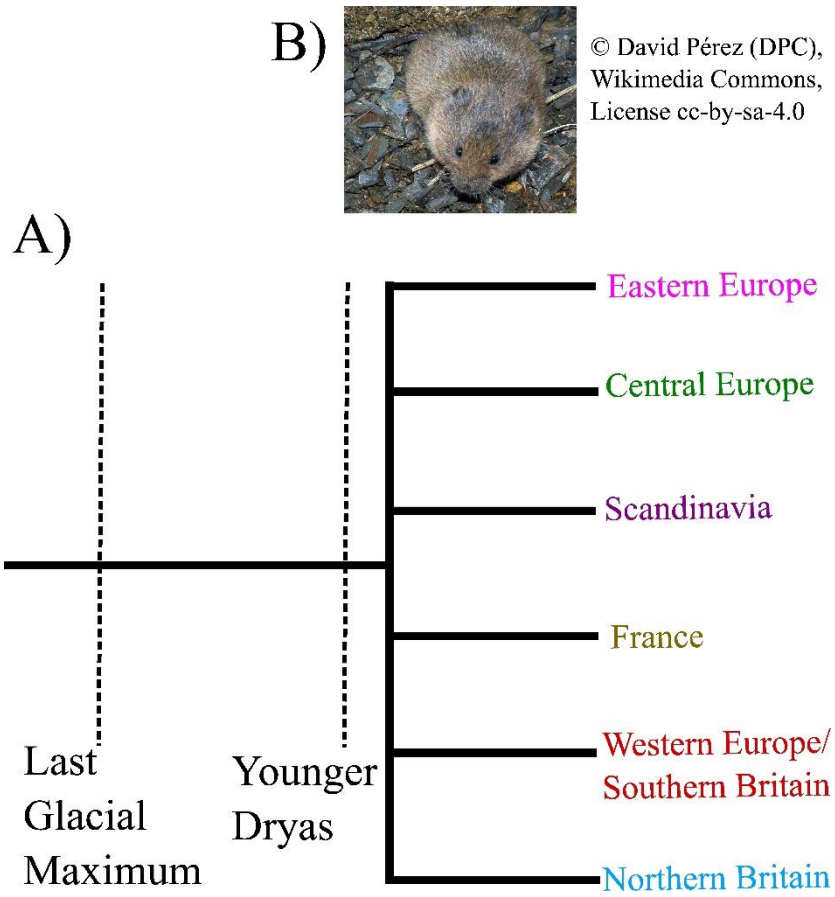


**Figure 4-4.** A) A tree showing the influence of ice readvances during the Younger Dryas on the colonisation history of field vole. Branches represent the extent of divergence between populations with the top of the branch being the most recent common ancestor. The dotted lines represent the timing of the global Last Glacial Maximum and the start of the Younger Dryas period respectively. Ancestral analysis shows that a single ancestral population existed during the last ice age which began to colonise much of Europe at the end of glacial conditions. Analyses suggest that several populations then became isolated from one another when glacial conditions returned during the Younger Dryas and diverged from one another. These populations then separately expanded at the start of the Holocene resulting in six distinct clades seen across Europe today. Figure based on results from (Herman and Searle, 2011). B) A picture of a Eurasian field vole.

These results when combined with knowledge of glacial history suggest a colonisation history along the lines of the following. A single ancestral population initially existed during the gLGM, which then expanded outwards at the end of the last ice age across Europe and into Britain. The return of ice and glacial conditions during the Younger Dryas period *ca.* 12.9 ka caused fragmentation of this new species range into six different groups/populations. These populations diverged from one another in isolation and then each

expanded outwards at the start of the Holocene, resulting in the six genetically distinct clades seen today. The divergence timing of the unique northern Britain clade indeed suggests a population became isolated in the north of Britain during the Younger Dryas, diverged, and subsequently colonised much of the surrounding area once warmer conditions returned. This species highlights the importance in considering the impact of the Younger Dryas when exploring colonisation history. Indeed, specific changes in ice masses will have had a major impact on contemporary genetic structuring. This can be applied to changes in the BIIS that we covered earlier. For example, a major difference in timing of ice-free conditions for a specific region or the short-lived advance of ice into previous unglaciated areas will likely have affected the genetic structuring for populations in and around those areas.

## 4.8 Conclusions

The species we have discussed, and many others, act as examples as to why an understanding of glacial history and the impact it has had on the landscape of the British Isles can provide vital information in understanding colonisation and demographic history (Kotlík et al., 2018; McDevitt et al., 2022). The examples we have presented here are simple in their design to discuss specific patterns affected by glacial processes. Many species will exhibit a number of the patterns we discussed across their species range in the British Isles. The extant flora and fauna of the British Isles reflect not only the genetic diversity and structuring promoted by contemporary ecology and environment but are also shaped by past isolation and connectivity during and after glaciation. The impact of glaciation is variable across species, as it is influenced by their natural history and dispersal potential across those changing and often depauperate landscapes.

However, we can link glacial history and genetics to a greater degree than what has been discussed in the examples we have provided. For example, the Arctic charr study used glacial history to demonstrate the differential colonisation histories of specific populations using knowledge of when the British Isles was covered in ice. Given that this species was likely to be one of the first colonisers of the British Isles post-glaciation and has a wide-spread contemporary distribution across Scotland, it is almost certain that the differential timing of when different regions were ice-free, and colonisable, would have affected contemporary patterns of genetic variation. Therefore, the application of the genetic techniques that we have discussed to the wider species distribution could allow us to start highlighting some of key changes in glacial coverage within the British Isles that have influenced contemporary patterns of genetic variation.

Of course, there remains challenges in doing this. While an extensive chronology of changes in ice coverage exists, there remains some uncertainty in the exact timings of different events due to the fragmented nature of empirical data. Similarly, the majority of genetic studies are conducted on contemporary samples meaning they are limited by the species current distribution and must disentangle patterns influenced by glacial history from those influenced by more recent selective pressures or those which are the result of over 10,000 years of history that occurred post-colonisation. Fossil or pollen records and DNA from ancient samples would provide valuable information to clear up some of these uncertainties.

The sharing of information between the evolutionary and glacial fields is vital for future studies and will lead to important advancements in each. Indeed, the combination of the two has allowed genomic data to highlight the importance of pre-glacial admixture and post-glacial gene flow on many patterns of contemporary variation. Knowledge of glacial history can provide valuable information in understanding the genetic patterns that are uncovered while knowledge of species distribution and timing of colonisation can provide key insight into how glacial coverage must have changed over time or affected certain regions. Glacial history is key to understanding contemporary patterns of biodiversity in the British Isles and we hope this review acts as a springboard to further discussions and integration of information from across these different fields.

# Chapter 5: The historical patterns that have shaped contemporary genetic differentiation across populations of Arctic charr in Scotland

## 5.1 Abstract

The impact of contemporary and historical processes can be seen in patterns of genetic differentiation across the geographic range of a species. Glaciation and glacial history through the fragmentation of populations during glacial conditions and their resulting diversification during isolation are important, but often undervalued, contributors to contemporary patterns. The Arctic charr is a non-anadromous species in the British Isles but likely exhibited anadromy when it first colonised around the end of the last ice age. The change of migratory behaviour, colonisation from multiple glacial refugia populations, and the asynchronous patterns of ice retreat at the last ice age likely all affected patterns of genetic differentiation but their influence has remained largely unexplored. To address this, we conducted a national-scale genetic study of Arctic charr populations across the British Isles. We found that several haplotypes shared across the British Isles were found elsewhere in the Holarctic suggesting the British Isles was colonised by multiple sub-lineages of the Atlantic lineage of this species. A neighbour-joining tree largely split populations by broad direction of river discharge (East vs West) and suggests a period of anadromy existed after colonisation where populations mixed. However, genetic differentiation was only correlated with geographic distance in comparisons within the same Hydrometric Area highlighting the strong effect of genetic drift on differentiation. Several groups of populations covering multiple Hydrometric Areas shows unique ancestral histories and evidence for genetic mixing likely affected by ice coverage during the Loch Lomond Stadial and the presence of ice-dammed lakes.

## 5.2 Introduction

Patterns of genetic differentiation between populations within the same species are a reflection of contemporary and historical processes. Differences in environments and subsequent local adaptations, and reduction of gene flow between populations, can lead to the accumulation of differences between populations (Orsini et al., 2013; Van Strien et al., 2015). Many species display patterns of genetic differentiation which correlate with the geographic distance between populations, known as Isolation-by-distance (IBD), where more geographically distant populations are more genetically differentiated from one another. These patterns can represent differential selection on population in different parts of the range of the species and an equilibrium between gene flow and genetic drift, often as a consequence of dispersal limitations (Hutchison and Templeton, 1999; Orsini et al., 2013). Other species show no relationship between geographic distance and genetic differentiation. This can be due to levels of gene flow between populations reflecting similarities in their environments independent of geographical distance, known as Isolation-by-environment (Orsini et al., 2013; Sexton et al., 2014). A lack of IBD can occur in patterns of panmixia where gene flow is high and genetic drift and exposure to differing selective pressures across the range of the species is low (Palm et al., 2009). Alternatively, they can reflect cases where gene flow between populations is low and genetic drift is high, as is often seen in non-migratory or poorly dispersing species (Waters et al., 2020).

However, patterns of genetic differentiation are not always a reflection of current processes and can be influenced by a number of historical processes (Rey and Turgeon, 2007; Vera-Escalona et al., 2015; Aguillon et al., 2017; Ruzzante et al., 2019). For example, lower levels of genetic differentiation between geographically distant populations could indicate colonisation by the same ancestral population (Vera-Escalona et al., 2015), the absence of contemporary natural or man-made geographic barriers to gene flow (Coleman et al., 2018), and/or a historical loss of migratory behaviour (Delgado and Ruzzante, 2020). An often-undervalued driver of these historical changes for species at higher latitudes is glacial history and glaciation related processes (Hewitt, 1996; Fenton et al., 2023b) (Chapter 4). The presence of ice sheets during glacial periods, in particular the last ice age (*ca.* 115,000–11,500 years ago (ka)), over large geographical scales confined many temperate species into isolated safe havens known as glacial refugia (Hewitt, 1999). Populations in different geographically isolated refugia had the potential to differentiate from other refuge populations over time, before separately colonising emerging new habitats in glaciated areas when they became ice-free (Finnegan et al., 2013; Roberts and Hamann, 2015). Patterns of

glacial retreat during the last ice age were asynchronous, which resulted in different regions becoming ice-free at different times (Tylmann et al., 2022). This would have affected the timing of colonisation of previously glaciated areas and potentially what lineages/refugia populations may have been able to colonise. Constrained drainage of meltwater from these ice sheets in some places resulted in glacial or ice-dammed lakes which may have provided important routes for eventual colonisation of freshwater species (Fenton et al., 2023b) (Chapter 4). Such historical lakes, that often no longer exist, may have allowed for genetic mixing across what are now separate river systems and geographic patterns of genetic connectivity that are difficult to interpret from simple contemporary freshwater catchments. This pattern has been shown in species like lake whitefish (*Coregonus clupeaformis*) where populations in unconnected river systems share localised ancestral histories due to colonisation through glacial lakes (Pielou, 1991; Ruzzante et al., 2019).

The salmonid species Arctic charr (*Salvelinus alpinus*) is an excellent species to study the effects of historical processes on contemporary differentiation between population, particularly in the British Isles (Maitland and Adams, 2018). The Arctic charr has a widespread distribution across a range of freshwater environments in the Holarctic and displays high levels of genetic differentiation and variation across populations (Adams et al., 2007b; Klobucar et al., 2021). Like in many other adaptive radiations in freshwater fish species (Taylor et al., 1996; Jones et al., 2012; Delgado et al., 2019), the loss of diadromy (migration between freshwater and marine environments during the life cycle) after colonisation of post-glacial environments is well documented in Arctic charr (Finstad and Hein, 2012; Jørgensen and Johnsen, 2014). While many populations in more northern latitudes still show a mixture of anadromous and non-anadromous individuals, in the southern parts of the distribution, including the British Isles, all populations are non-anadromous and no longer migrate to sea (Maitland & Adams, 2018; Salisbury et al., 2022). The species is believed to have colonised the British Isles *ca.* 11 ka towards the end of the last ice age with these ancestral colonising populations likely to have been anadromous (Maitland and Adams, 2018; Fenton et al., 2023b) (Chapter 4). As such, the contemporary, non-anadromous populations very likely arose from migratory ancestral populations, a pattern documented in other species such as alewife (*Alosa pseudoharengus*) (Palkovacs et al., 2008; Delgado et al., 2019). It is very likely therefore that there was a period of time when populations in the British Isles were still anadromous and populations in different Hydrometric Areas, integral river catchments, may have mixed for some time after colonisation before the anadromous phenotype was lost. The probability of straying into a non-natal river decreases with distance, and so populations in different river systems closer

to each other geographically are likely to be more genetically similar (Matala et al., 2012; Keefer and Caudill, 2014). As such, patterns of genetic differentiation across Hydrometric Areas may, to an extent reflect this period of historic anadromy, and show some correlation with geographic distance.

Another important factor that may have influenced patterns of genetic differentiation in Arctic charr, is the number of ancestral refugia populations that colonised the British Isles and the timing of colonisation. Across its Holarctic distribution, there are believed to be five major lineages of Arctic charr (Brunner et al., 2001; Moore et al., 2015) with all populations in the British Isles likely derived from the Atlantic lineage, which is prominent in Europe, although evidence for this has been limited to the examination of a single population in each instance. Multiple sub-lineages are known to exist within the Atlantic lineage (Jacobsen et al., 2022) diverging from each other *ca.* 60 ka during the last ice age and so the colonisation of the British Isles may have involved both of these known sub-lineages. Several lakes in Scotland contain multiple distinct ecotype populations, i.e., populations that are specialised to different ecological niches within the same lake (Jonsson and Jonsson, 2001; Maitland and Adams, 2018). While the origins of these ecotype divergences in each lake are not always known, they arise under two broad scenarios: sympatric divergence in the lake after a single colonisation event, or from secondary contact after multiple colonisation events from populations which diverged during the last ice age in allopatry. Several of these ecotype divergences in Scotland are suggested to have arisen from secondary contact between colonising populations (Verspoor et al., 2010; Jacobs et al., 2020). Given these findings, the British Isles was likely colonised by multiple diverged refugia populations at the end of the last ice age, influencing patterns of genetic differentiation. However, the colonisation history of populations across the British Isles remains unknown and therefore details of how colonisation by different refugia populations has impacted contemporary genetic differentiation are also unclear.

Arctic charr is believed to be one of the first freshwater colonisers of the British Isles around the end of the last ice age and so would have been greatly affected by the differential timing of ice-free conditions across the British Isles, particularly in Scotland where the majority of contemporary populations are found (Maitland and Adams, 2018). While much of the British Isles was ice-free by the start of the Late Glacial Interstadial *ca.* 15 ka (Ruddiman, 2008; Clark et al., 2012), the northwest Highlands of Scotland remained covered in ice through this brief period of warming and the subsequent Younger Dryas period (*ca.* 12.9 – 11.5 ka), also known as Loch Lomond Stadial (LLS), when glacial conditions in the

British Isles briefly returned before the start of the Holocene (Bickerdike et al., 2018). A number of contemporary populations of Arctic charr are found in regions that were still covered by ice during the LLS (Figure 5-1). Notably several of the lakes (lochs Arkaig, Ericht, Lochy) contain multiple ecotype populations for which the demographic histories have yet to have been explored (Fraser et al., 1998). Given that the rest of the British Isles was accessible for colonisation at this time, we anticipate that patterns of colonisation and genetic differentiation may be different for these ice-covered regions. For example, they may have been colonised by a different sub-lineage of Arctic charr to populations in nearby areas that were not still covered by ice. Additionally, a number of ice-dammed lakes are known to have existed in Scotland during the LLS (Sissons, 1977; Benn, 1989; Bickerdike et al., 2016; Bickerdike et al., 2018). As these ice-dammed lakes may have allowed colonisation with a common ancestor and historical mixing for a period of time between nearby populations that are not now connected which when examined now displays as a pattern of low genetic differences.

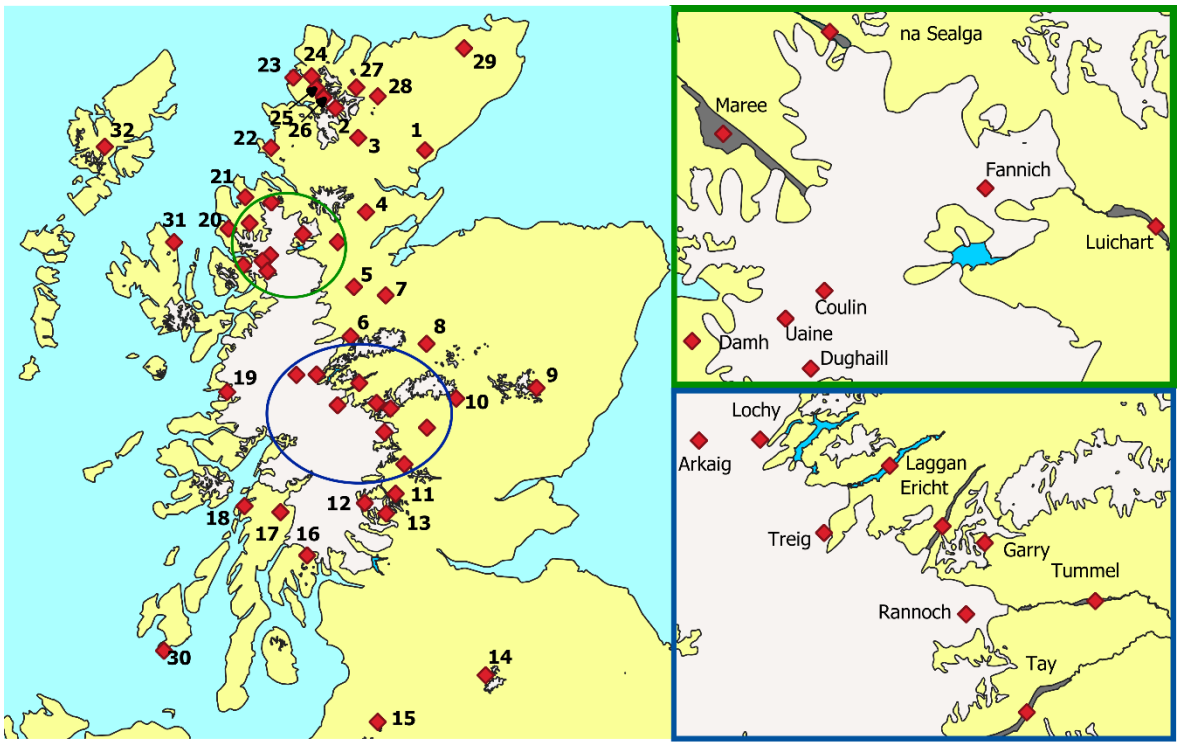


**Figure 5-1.** Map of ice coverage during the Loch Lomond Stadial (ca. 12.9-11.5 ka) in Scotland. Data on ice coverage and glacial lakes is reproduced from Bickerdike et al. 2016. Known ice cover is indicated in white and predicted ice-dammed lakes are in blue (Benn, 1989; MacLeod et al., 2011). Red dots indicate lakes in our dataset. Present-day lakes are shown in dark grey.

We investigated the impact of historical patterns on contemporary genetic differentiation between 64 populations of the highly diverse Arctic charr found in the British Isles. We hypothesised that populations in regions still covered by ice during the Loch Lomond Stadial would show distinct patterns of colonisation to the surrounding areas and evidence of genetic mixing across Hydrometric Areas through known ice-dammed lakes. We first investigated whether the British Isles was colonised by multiple refugia populations. To do this, we inferred the Holarctic lineages/sub-lineages that most likely colonised the British Isles using data from across the range of the species. We also reconstructed the timing and scenarios under which known ecotype pairs, particularly in these ice-covered regions, likely diverged to determine the prevalence of secondary contact divergences. We then explored contemporary patterns of genetic differentiation and investigated potential migration events to identify populations which showed evidence of genetic mixing. We then compiled our findings to show how patterns of ice coverage and the presence of ice-dammed lakes influenced contemporary patterns of genetic differentiation and structuring.

## 5.3 Methods and Materials

### *5.3.1 Mitochondrial ND1*

We used a recently generated dataset of mitochondrial NADH dehydrogenase 1 (ND1; 975bp in length) gene sequences for 56 populations across the British Isles (N=208 individuals) (Chapter 2). ND1 sequences from elsewhere in the Holarctic (Canada, Alaska, Russia, Iceland, Greenland, Norway and Sweden) were drawn from Jacobs et al., (2020) and extracted from whole mitochondrial genomes for *Salvelinus alpinus* available on GenBank (Oleinik et al., 2020; Schroeter et al., 2020; Jacobsen et al., 2022) (Table 5-1). For Greenland, as a large number of mtDNA sequences were available, we choose to use three representative populations for the notable lineages/sub-lineages mentioned in Jacobsen et al 2021 (Arctic lineage and Atlantic sub-lineages 1 and 2) with three sequences used for each population. We additionally generated sequences for individuals from Holar, Iceland and Lake Constance (Germany) using the methods outlined in (Chapter 2).

**Table 5-1.** Table of ND1 sequences added to the existing dataset of 208 sequences from across the British Isles (Chapter 2). Population names and localities are provided. NCBI accession numbers and paper references for whole mitochondrial genomes are provided where relevant. Sequences from Holar and Constance were generated as part of this project. No locality information was available for Alaskan samples. Samples from Holar are aquaculture samples.

| Population | Locality | Number | Accession number(s) | Paper |
|---|---|---|---|---|
| Jayko river | Canada | 2 | MW664921.1, MW664922.1 | Unpublished NCBI submission |
| Constance | Germany | 3 | Generated here | |
| Disko bay (DISK) | Greenland | 3 | MT880631.1, MT880632.1, MT880633.1 | (Jacobsen et al., 2022) |
| Scoresby Sound (SCOR) | Greenland | 3 | MT880708.1, MT880709.1, MT880710.1 | (Jacobsen et al., 2022) |
| Kapisilit river (NUUK) | Greenland | 3 | MT880640.1, MT880641.1, MT880642.1 | (Jacobsen et al., 2022) |
| Holar | Iceland | 2 | Generated here | |
| Vatnshlidarvatn | Iceland | 3 | MT880716.1, MT880717.1, MT880718.1 | (Jacobsen et al., 2022) |
| Luktvatn | Norway | 3 | MT880662.1, MT880660.1, MT880661.1 | (Jacobsen et al., 2022) |
| Lama | Russia | 1 | MZ571355.1 | Unpublished NCBI submission |
| Bol'shoe | Russia | 2 | | (Jacobs et al., 2020) |
| Leprindo | | | | |
| Sitasjaure | Sweden | 3 | MN957795.1, MN957796.1, MN957797.1 | (Oleinik et al., 2020) |
| Alaska | | 3 | MF621740.1, MF621741.1, MF621743.1 | (Schroeter et al., 2020) |

The Holarctic sequences were added to a multisequence alignment with the British Isles dataset using *MUSCLE* in *Uniprot UGENE v45* (Okonechnikov et al., 2012). In total, our ND1 dataset had 238 individuals. ND1 sequences from *Salvelinus fontinalis* (Accession number: MF621738.1) and *Salvelinus namaycush* (Accession number: OM736821.1) were added as outgroups. Templeton, Crandall, and Sing's (TCS) haplotype networks were formed using *PopArt v1.7* (Leigh and Bryant, 2015). We then ran a maximum likelihood tree using one sequence per haplotype from the haplotype network including the outgroup species (Moore et al., 2015). The best nucleotide substitution model for deriving a maximum likelihood tree was selected using *modeltest-NG v.0.1.7* (Darriba et al., 2020). *MEGA v11* was used to run the maximum likelihood tree using Tamura-Nei substitution model with 500 bootstraps (Tamura et al., 2021).

### *5.3.2 Genomic data*

We used a recently generated set of ddRADseq reads (Chapter 2) for 410 individuals covering 64 different populations across the British Isles to investigate demographic history and patterns of genetic differentiation and structuring. ddRADseq libraries of 90 samples plus five replicates were prepared using a modified version of (Recknagel et al., 2015) ddRADseq protocol for Illumina sequencing platforms. Paired-end 75-bp sequencing was performed on the IlluminaNextSeq500 platform at Glasgow Polyomics. Raw ddRADseq data (ddRADseq NCBI short read bioproject: PRJNA607173) for eight of the Scottish lakes (Awe, Dughaill, Eck, Lubnaig, Merkland, na Sealga, Tay, Uaine) was generated as part of Jacobs et al. 2020. Reads were mapped to the annotated *Salvelinus* sp. genome (ASM291031v2) using *bwa mem v0.7.16* (Li, 2013). RAD-loci building was done in Stacks *v2.60* (Rochette et al., 2019) with all SNP datasets used in this study generated from this catalogue of RAD loci. Subsequent, SNP filtering was performed using the *populations* module of *Stacks* and *VCFtools v0.1.16* (Danecek et al., 2011).

Genetic differentiation analysis was done using a dataset of 24,878 SNPs and 410 individuals from 64 populations across the British Isles. For this dataset, SNPs were retained if they met the following criteria: present in 66% of all individuals in each population and across all populations, a minimum minor allele frequency of 0.01, maximum observed heterozygosity of 0.5 with only the first SNP per locus retained and filtered to a minimum read depth of 10x and a maximum depth of 30x. For phylogenetic analyses, this dataset was further filtered to remove any SNPs with missing data along with those representing $F_{ST}$ outliers, 95$^{th}$ percentile for global $F_{ST}$ scores, or were previously identified as associated with

environmental and lake variables as described in (Chapter 2) leaving a dataset of 5,949 SNPs for our 64 populations.

SNP datasets for each ecotype pair comparison were made from the *gstacks* catalogue to analyse demographic history. We did this for eight ecotype divergences from across Scotland at lochs Arkaig, Awe, Dughaill, Ericht, Lochy, na Sealga, Rannoch, and Tay (Fraser et al., 1998; Adams et al., 1998; Garduño-Paz and Adams, 2010; Hooker et al., 2016; Maitland and Adams, 2018; Jacobs et al., 2020). Due to an insufficient number of samples, the Rannoch benthivore population was excluded and so the divergence at Rannoch was run with a 2-population model (planktivore-piscivore). SNPs were retained if they met the following criteria in each comparison: present in 90% of all individuals, showed a maximum observed heterozygosity of 0.5 and one SNP per locus with no minor allele frequency filter applied to ensure rare, informative sites were retained. All SNP datasets were filtered to remove SNPs with minimum read depth < 10x or a maximum depth > 30x.

### *5.3.3 Genetic differentiation and geographic distance*

We used pairwise $F_{ST}$ to investigate whether the extent of genetic differentiation between populations correlated with the geographic distance between them. Pairwise $F_{ST}$ was generated using the –fstats parameter in the *populations* module of *Stacks v2.60*. Geographic distance between the lakes was calculated using *QGIS v3.26.3* to calculate the distance through river systems and the *marmap v1.06* R package for distance through the marine environment (Pante and Simon-Bouhet, 2013; QGIS Development Team, 2022). An approximate point in the middle of the lake was used as the start point for each calculation. The same point was used for lakes that contain multiple ecotypes. To properly calculate the distance across marine environments, longitude and latitude co-ordinates for a point representing the discharge point of each river to the marine environment containing one of the study populations was determined using *QGIS.* Marine bathymetric data for the British Isles was gathered from the GEBCO database (General Bathymetric Chart of the Oceans) (https://www.gebco.net) and filtered from a minimum depth of 20 metres and a maximum depth of 200 metres to ensure the least-cost distance did not travel through landmasses. The least-cost distances between our marine outlet points was then calculated using the *lc.dist* function. Distance between each lake and their respective marine outlet point was calculated in *QGIS*. Marine and river distances were then combined to get the geographic distance between lakes in different river systems. For lakes in the same river system, the distance between the lakes along the river systems was just calculated in *QGIS*. Populations in the same lake, i.e., multiple ecotypes, were given a geographic distance of zero. Distance values

(d), genetic and geographic, were then log-transformed (D= log(d+1)) before use in Mantel tests implemented using the *vegan v2.6.2 R* package (François et al., 2021; Oksanen et al., 2022). Comparisons were split into three categories: between populations in the same lake (i.e., ecotype pairs), comparisons between lakes in the same Hydrometric Area (HA; integral river catchments or several small contiguous river catchments that have separate outlets to the sea (National River Flow Archive, 2014) and comparisons between lakes across Hydrometric Areas.

### *5.3.4 Phylogenetic trees*

A distance-based phylogenetic relationship between populations was constructed using an unrooted Neighbour-Joining (NJ) tree of the SNP data, constructed using the *poppr v2.9.3* R package (Kamvar et al., 2015). The tree was run with 1,000 bootstraps and no root. We used *Treemix v1.13* to investigate past admixture events between populations using the same subset of no missing data SNPs (Pickrell and Pritchard, 2012). Treemix analysis was run as described in Dahms et al., (2022) and original code can be found at (https://github.com/carolindahms/TreeMix). To summarise, 500 bootstraps were generated with a block size (-k) of 100 with no outgroup. A consensus tree was then formed using *PHYLIP consense v3.695* (Felsenstein, 2005). This consensus tree was then used to test for the optimal number of migration events. *Treemix* was run with the consensus tree as the starting tree (-tf) and 1-20 migration events were tested with 10 iterations each. Migration edges are added between populations in the tree to improve its likelihood. Any additional branch can be interpreted as a migration event that led to admixture. Model likelihoods for each number of migrations were compared with *OptM v0.1.*6 R package (Fitak, 2021). 30 independent runs were then performed for the optimal number of migrations using the consensus tree. The tree with the highest maximum likelihood was retained.

### *5.3.5 Demographic analyses*

Two-population SFS files were created using *dadi* in *easySFS* (https://github.com/isaacovercast/easySFS) for all eight ecotype pair SNP datasets. The preview option was to ensure that files were made with the maximum number of segregating sites possible. Demographic models were then run using *moments* in *GADMA2* v*2.0.0rc23* (Noskova et al., 2022). To account for the number of monomorphic sites, the total sequence length was determined using the number of segregating sites in the SFS, the number of variant sites used to make the SFS and the total number of genomic sites before the SFS file was made ((segregating sites/variant sites)*number of genomic sites)). As no known

mutation rate for Arctic charr is available, a mutation rate of $1x10^{-8}$ was used from lake whitefish (*Coregonus clupeaformis*) (Rougeux et al., 2017).

For each ecotype divergence, we tested seven different demographic models of different historical divergence scenarios. These models simulated were: *Strict Isolation* (population split followed by no migration), *Ancient Migration* (population split during gene flow, migration followed by a period of no migration), *Isolation with Migration* (population split with ongoing migration), *Secondary Contact* (population split with no migration followed by period of migration upon coming back into contact), *Isolation with change in Migration rate* (population split with ongoing migration but a notable change in migration rate after a certain time), *Isolation with Migration and Introgression* (population split with ongoing migration and a historical introgression event), and *Secondary Contact with Introgression* (population split with no migration followed by a introgression event upon secondary contact and continued migration thereafter). Each model was run 32 times with the most likely model for each divergence determined by the lowest AIC score (Excoffier et al., 2013).

## 5.4 Results

### *5.4.1 Mitochondrial ND1 haplotype network*

We used mtDNA sequences from across the Holarctic range to investigate the colonisation history of lake-dwelling populations across the British Isles. From our dataset of 238 sequences of mitochondrial ND1, we identified 32 haplotypes with 25 of these haplotypes found in the British Isles to some degree (Figure A5-1). Previous analysis of the British Isles data showed a dominant haplotype (Brit_1) shared across 121 of the 208 individuals analysed and was found in 39 of the 56 populations (Chapter 2) (Figure 5-2). With the inclusion of sequences from across the Holarctic, this haplotype is also found in Sweden, Germany, Iceland, Norway, and the Scoresby Sound (SCOR) Greenland populations with the sequences from Greenland, Iceland, and Norway previously identified to Atlantic sub-lineage 2 (Jacobsen et al., 2022). Three other previously identified haplotypes originally only found in Scotland (Scot_1, Scot_5, Scot_14) were found elsewhere in the Holarctic in Sweden, the Disko Bay (DISK) Greenland population, and Iceland respectively with the Disko Bay population previously identified as Atlantic sub-lineage 1 and the Icelandic population of Vatnshlidarvatn previously identified as Atlantic sub-lineage 2 (Jacobsen et al., 2022). In our maximum likelihood tree, the sequences from Canada, Alaska and the Kapisilit river (NUUK) in Greenland, previously identified as the Arctic lineage,

split off first with the Russian sequences representing the Siberian lineage, splitting off shortly afterwards (Figure 5-3). This suggests that all the haplotypes in the British Isles belong to the Atlantic lineage.

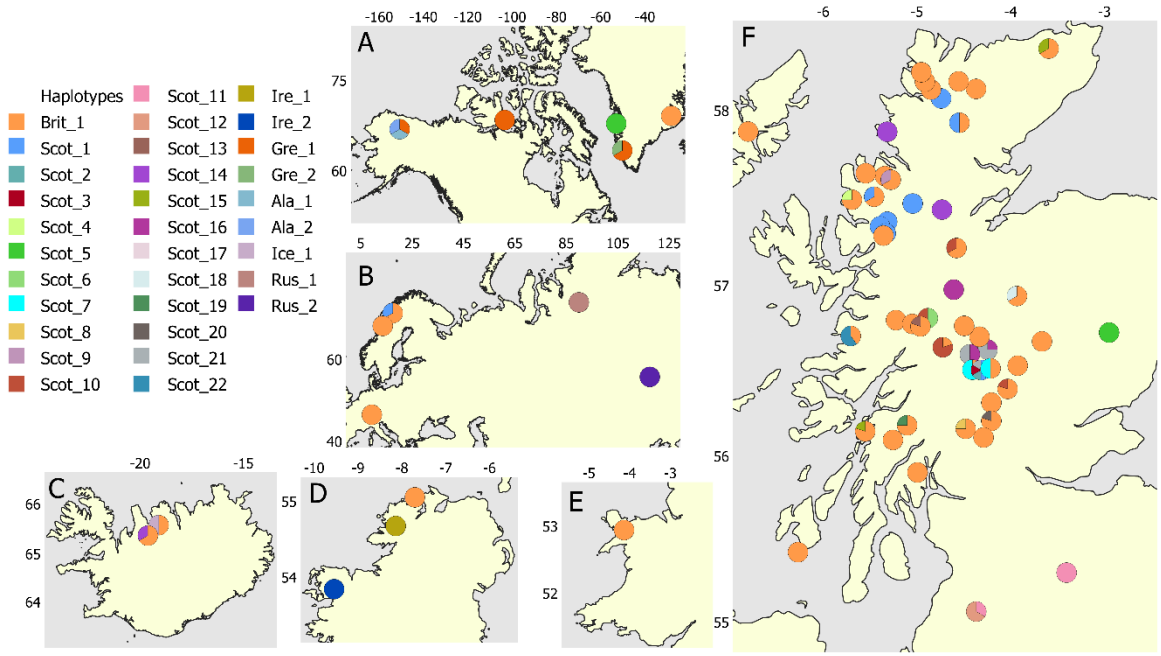


**Figure 5-2.** Haplotype map for our Holarctic dataset. Each point represents a population and is coloured based on haplotypes present as indicated in the legend. Panels A-F represents the different regions in our dataset: A shows North America and Greenland, B shows central Europe, C shows Iceland, D shows Ireland, E shows Wales, and F shows Scotland.

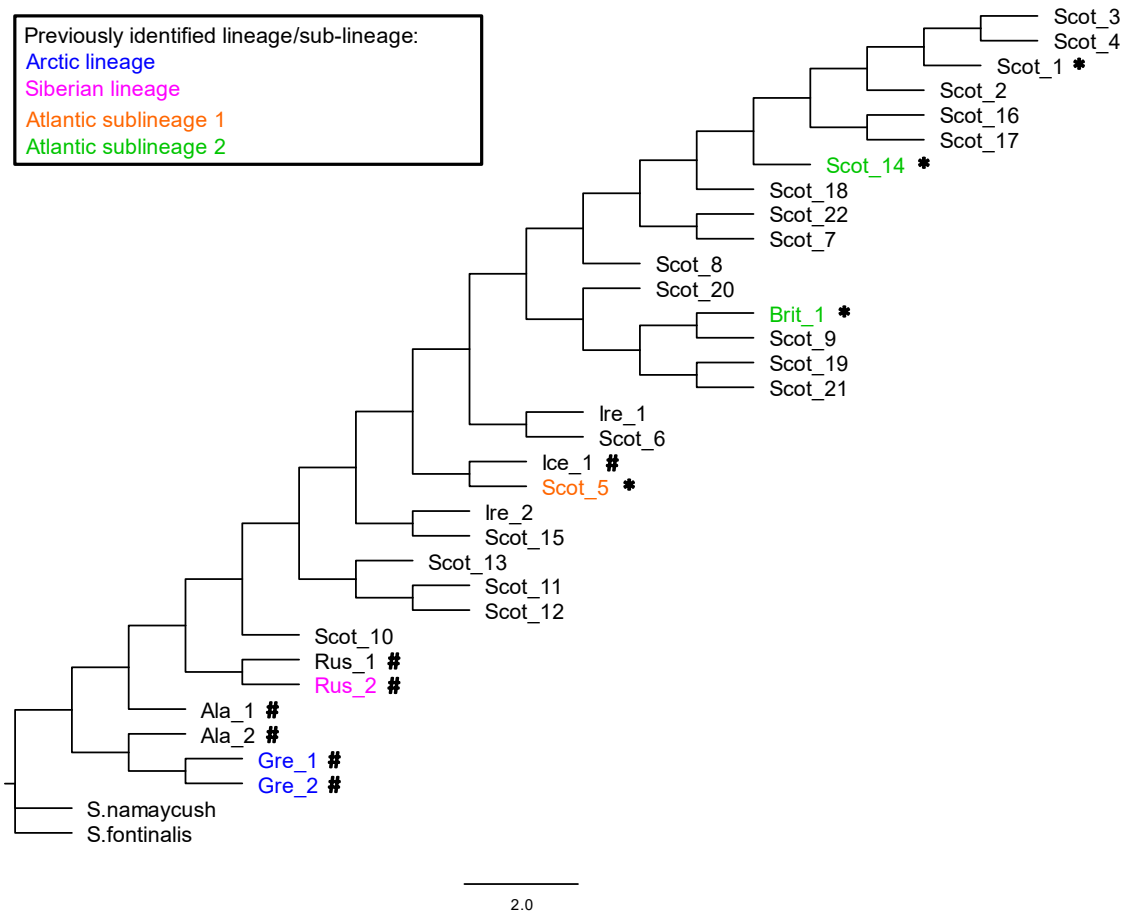


**Figure 5-3.** Maximum likelihood tree of haplotypes for the Holarctic dataset of mtDNA ND1. One sequence per haplotype was retained when making the tree. Haplotypes found in the British Isles are named with Brit, Scot, and Ire prefix based on where they are found. Other prefixes indicate haplotypes predominantly or exclusively found in Iceland, Greenland, Alaska, and Russia. Asterisk indicates haplotypes from the British Isles found elsewhere in the Holarctic. Hashtag indicates haplotypes not found in the British Isles. Haplotypes belonging to sequences previously identified as a known lineage or sub-lineage are coloured to indicate their origin.

### *5.4.2 Genetic differentiation*

We used a genome wide dataset of 24,878 SNPs covering 64 populations and 410 individuals across the British Isles to investigate comparisons between levels of genetic differentiation (pairwise $F_{ST}$) and geographic distance. In total, we had 2016 pairwise comparisons which can be split into three categories: comparisons between populations in the same lake (i.e., ecotype pairs), comparisons between lakes in the same Hydrometric Area and comparisons between lakes across Hydrometric Areas. Pairwise $F_{ST}$ between ecotypes occupying the same lake (N=10 comparisons) ranged from 0.053 (na Sealga benthivore – na

Sealga planktivore) to 0.200 (Rannoch benthivore-Rannoch planktivore) with an average pairwise $F_{ST}$ of 0.139. Pairwise $F_{ST}$ between lakes in the same Hydrometric Area (N=94 comparisons) ranged from 0.061 (Loch Laggan-Loch Treig) to 0.510 (Loch Loch-Loch Rannoch benthivore) with an average pairwise $F_{ST}$ of 0.217. Pairwise $F_{ST}$ between lakes in different Hydrometric Areas (N=1912 comparisons) ranged from 0.045 (Loch Doon-Talla Reservoir) to 0.663 (Lough Finn-Loch Mealt) with a mean pairwise $F_{ST}$ of 0.258. Geographic distance between lakes ranged from 7.42 km between lochs Stack and More to 1359.51 km between Llyn Bodlyn, in Wales, and the Talla Reservoir.

When looking across all comparisons, we saw a weak positive relationship between geographic distance and genetic differentiation (Mantel test P=0.03) with pairwise $F_{ST}$ increasing with greater geographic distance (Figure 5-4A). A much stronger relationship was evident in comparisons between populations in the same Hydrometric Area (Figure 5-4B), while no clear relationship was seen when looking only at comparisons across Hydrometric Areas (Figure 5-4C).

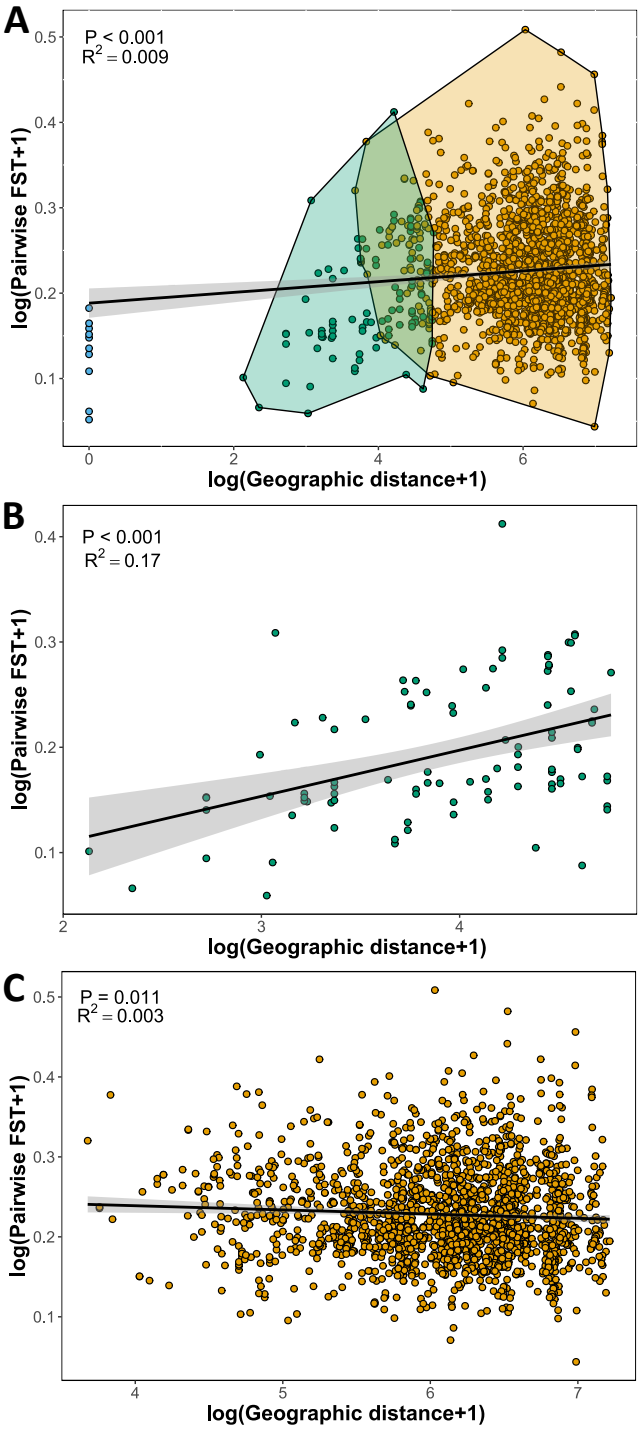


**Figure 5-4.** Correlations of genetic and geographic distance between populations within the British Isles (N=24,878 SNPs). Each point represents a comaprison between populations with the geographic distance between their lakes of origin on the X-axis and the genetic differentiation between the two populations, using pairwise $F_{ST}$, on the Y-axis.Three panels show all comparisons (A), comparisons for populations within the same Hydrometric Area (B), and comparison across Hydrometric Areas (C). Black lines represent linear models with R-squared and P values displayed. Comparisons between populations in the same lake are in blue, in the same Hydrometric Area in green, and across Hydrometric Areas in orange.

### *5.4.3 Population structuring*

A neighbour-joining (NJ) tree, originally presented in Chapter 2, was used to investigate the relationships between populations across the British Isles. The NJ tree predominantly separated populations by river system discharge (Figure 5-5) with populations in river systems that flow out to the east coast of Scotland forming a separate group to the rest of the dataset with all outgroup populations from the rest of the British Isles appearing alongside populations in west flowing river catchments. Alongside the Talla Reservoir population which is a transplant population of west-flowing Loch Doon (Maitland et al., 2007), two groups of populations show notable exception to this. First, all populations in the west-flowing Lochy catchment clustered with the east-flowing group among populations from the Tay catchment. The second are the populations from lochs Fannich and Luichart in the east flowing Conon catchment which are found clustered amongst the west-flowing populations.

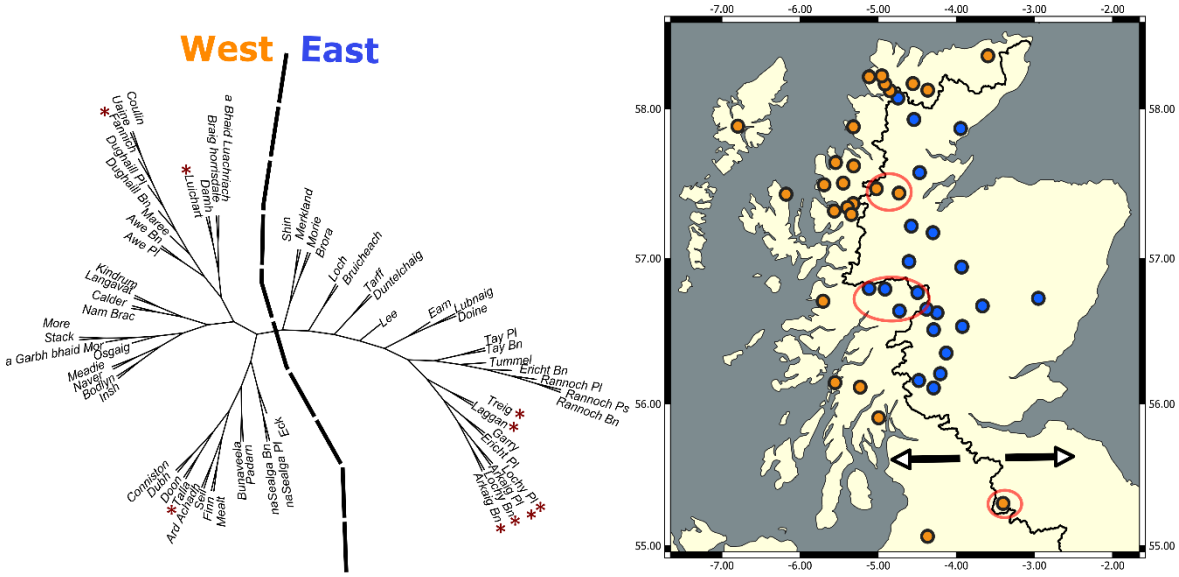


**Figure 5-5.** Neighbour-joining tree for populations found across the British Isles (N=5,949 SNPs) originally presented in Chapter 2. Populations largely spilt by river catchment discharge (East vs West). Map shows the locations of all Scottish lakes. Line indicates the actual boundary between east and west flowing river catchments. Populations are coloured by which part of the NJ plot they appear. Red circles indicate the three groups of populations that do not follow the trend. These are lochs Fannich and Luichart in the north, the Lochy HA populations in the middle and the Talla Reservoir in the south, which was translocated there from the west-flowing Loch Doon. These populations are indicated by red asterisks in the neighbour-joining tree.

### *5.4.4 Migration events between populations*

We used Treemix analysis to identify any notable migration events between populations in our dataset. When we tested a different number of migration events from 1-20, our linear models suggested no more than three notable migration events (Figure A5-2). These were: the migration of individuals from the Dughaill benthivore population into the Dughaill planktivore population and two migrations from populations in the Tay catchment into populations in the Lochy catchment (Loch Garry -> Loch Lochy benthivore and the Loch Ericht benthivore -> Loch Treig) (Figure 5-6).

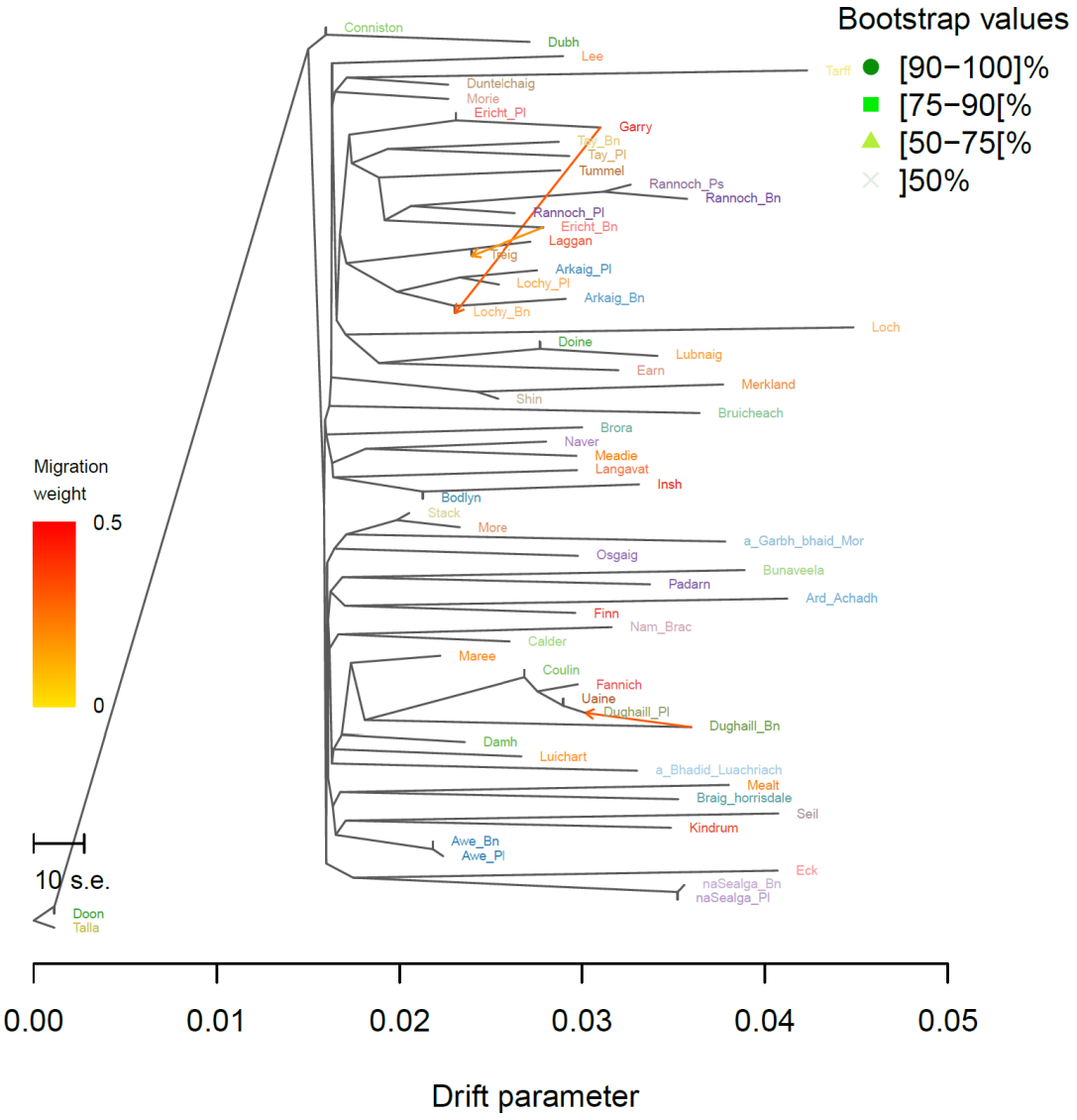


**Figure 5-6.** Treemix plot showing three notable migrations between populations in our dataset. Migration weight is coloured from yellow to red with increasing value.

### *5.4.5 Demographic history of ecotype pairs*

We tested the likely divergence scenarios of eight ecotype pairs across Scotland, using a range of models covering sympatric divergences to secondary contact scenarios between colonising lineages. We found that six of the eight ecotype pairs likely resulted from secondary contact by two or more ancestral colonizers (that had diverged in glacial refuges) (Table 5-2). The secondary contact model was the best fit for the loch Lochy ($DIV_{Loc}$ = 5927 generations since divergence, 17782 years since divergence), Ericht ($DIV_{Eri}$ = 3365 generations, 10095 years), and Rannoch (planktivore-piscivore) ($DIV_{Ran}$ = 7617 generations, 22850 years) divergences while isolation with migration and introgression models were the best fit for the Arkaig ($DIV_{Ark}$ = 5927 generations, 17781 years), Dughaill ($DIV_{Dug}$ = 3225 generations, 9675 years) and Tay ($DIV_{Tay}$ = 25483 generations, 76540 years) divergences. Sympatric divergence models, suggesting divergence under gene flow, were the best fit for the na Sealga ($DIV_{Sea}$ = 3319 generations, 9418 years) and Awe ($DIV_{Awe}$ = 1548 generations, 4644 years) divergences with the ancient migration and isolation with migration models the most likely respectively.

**Table 5.2.** Divergence model results for eight ecotype pairs found across Scotland. The number of segregating sites that the models were based on is listed, as well as the most likely model for each pair. AM- Ancient migration, SI – Strict isolation, IM – Isolation with migration, IMint – Isolation with migration and introgression, IMchange – Isolation with change in migration rate, SC – Secondary contract, SCint – Secondary contact with introgression. The second and third most likely models are also reported with difference in AIC shown.

| Lake (ecotypes) | Segregating Sites | Best Model | 2nd Best Model (ΔAIC) | 3rd Best Model (ΔAIC) |
|---|---|---|---|---|
| Awe (Bn-Pl) | 906 | IM | IMint (0.66) | AM (2.65) |
| Arkaig (Bn-Pl) | 1350 | IMint | SCint (2.05) | SC (2.67) |
| Dughaill (Bn-Pl) | 1967 | IMint | SC (1.46) | SCint (3.48) |
| Ericht (Bn-Pl) | 1002 | SC | IMint (1.29) | SCint (2.22) |
| Lochy (Bn-Pl) | 1452 | SC | IMchange (26.59) | SCint (28.29) |
| na Sealga (Bn-Pl) | 775 | AM | IM (4.65) | SC (4.78) |
| Rannoch (Pl-Ps) | 1452 | SC | IMint (1.99) | SCint (2.68) |
| Tay (Bn-Pl) | 1025 | IMint | IMchange (7.19) | SC (7.25) |

## 5.5 Discussion

Our results highlight how historical processes can provide insights that help explain contemporary patterns of genetic differentiation. We identified a number of mitochondrial haplotypes found in the British Isles and recorded elsewhere in the species distribution suggesting that the British Isles were colonised by a number of ancestral lineages which vary in their prominence. Patterns of genetic structuring indicate mixing between populations in nearby river systems before the anadromous phenotype was lost and evidence of some shared colonisation histories both within and across river catchments. Patterns of genetic differentiation were strongly correlated with geographic distance when looking at comparisons within the same Hydrometric Areas giving a strong pattern of Isolation-by-distance. Conversely, there was no relationship when looking across Hydrometric Areas only highlighting the lack of gene flow between populations in different river catchments.

We found that populations in regions still covered by ice during the likely time period for colonisation by Arctic charr, the Loch Lomond Stadial (*ca.* 12.7 ka - 11.5 ka), showed patterns not seen elsewhere in our dataset. This included shared mitochondrial haplotypes, evidence of genetic mixing, and low genetic differentiation across Hydrometric Areas with a number of known ice-dammed lakes that existed during this time period which likely facilitated the cross-Hydrometric Area patterns seen.

### *5.5.1 Colonisation history of populations in the British Isles*

Atlantic salmon (*Salmo salar*), European whitefish (*Coregonus lavaretus*), and brown trout (*Salmo trutta*) all show evidence that their range of the British Isles were colonised by multiple distinct refugia populations that diverged during the last ice age (McKeown et al., 2010; Finnegan et al., 2013; Crotti et al., 2020; Crotti et al., 2021). Previous work has suggested that British Isles were likely colonised by multiple ancestral refugia populations of Arctic charr based largely on the existence of ecotypes pairs that arose through secondary contact scenarios between refugia populations that diverged during the last ice age (Verspoor et al., 2010; Jacobs et al., 2020). These refugia populations likely belong to the Atlantic lineage of Arctic charr based on analysis of small number of populations suggesting the Atlantic lineage colonised the British Isles (Brunner et al., 2001; Moore et al., 2015).

Our results indeed confirm that the British Isles were colonised by multiple ancestral sub-lineages with these all belonging to the Atlantic lineage of Arctic charr. None of the haplotypes representing the Siberian or Arctic lineage in our mtDNA dataset were found in any of populations in the British Isles while all the haplotypes representing the two previously identified sub-lineages of the Atlantic lineage (sub-lineages 1 and 2) were found to some extent in the British Isles (Figure 5-2, Figure 5-3, A5-1) (Jacobsen et al., 2022). Of these two known sub-lineages, Atlantic sub-lineage 2 seems to be the most prevalent coloniser in the British Isles with 41 of the 56 populations analysed containing one of the two haplotypes (Brit_1 and Scot_14 haplotypes) shared with confirmed populations belonging to this sub-lineage. Many of the less common Scotland specific haplotypes also belonged to the same branch in our phylogenetic tree as these two haplotypes (Figure 5-3) suggesting these haplotypes may also belong to this sub-lineage, adding further evidence that this sub-lineage was a prominent coloniser of the British Isles. Given the prevalence of the sub-lineage 2 haplotypes which are also found in Greenland, Iceland, Norway, Sweden, and Germany, this was likely the sub-lineage that colonised much of the European range for the species. Atlantic sub-lineage 1 had previously only been identified in Greenland

(Jacobsen et al., 2022). In our dataset, we found that the haplotype representing sub-lineage 1 (Scot_5), was only found in one population in the British Isles, Loch Lee, suggesting that this sub-lineage is present in the British Isles but is far less prominent.

Of the other haplotypes found in the British Isles, only one was found elsewhere in Europe. The Scot_1 haplotype which is shared across eight populations in Scotland, most of which are found in the NW Highlands, was also found in the Swedish population Sitasjaure. Given that this is a haplotype shared across a number of populations, the second most common in the dataset, and present elsewhere in Europe this may represent another colonising sub-lineage. Whole mitochondrial genomes and a dataset covering more populations in Europe should be used before any concrete conclusions are drawn on the number of lineages present in the British Isles and the timing and route they took to colonise; however, our results do suggest that the British Isles was colonised by a number of sub-lineages which vary in their prevalence.

Further evidence for the colonisation of the British Isles involving multiple refugia populations can be found in our divergence timing analysis. Timings and scenarios for eight ecotype pair divergences seen in Scotland suggest that the majority of ecotype pairs diverged under secondary contact scenarios with six of the eight pairs tested having most likely scenarios involving two separate invasions of the lake by populations that diverged during the last ice age (Table 5-2). The finding that the majority of ecotype pairs diverged under secondary contact scenarios is generally in-line with expectations of speciation most commonly involving a period of isolation (Coyne and Allen Orr, 2004). Our findings are in agreement with previous findings that suggested that the ecotype pairs at lochs Dughaill and Tay and the Rannoch planktivore-piscivore ecotypes all likely arose under secondary contact scenarios and the lochs Awe and na Sealga represent sympatric divergence after a single colonisation event (Verspoor et al., 2010; Jacobs et al., 2020). Our timings for the divergence at lochs Dughaill and Tay were notably different to the original timings however which were originally timed to be *ca.* 50 ka and 34 ka. This is almost certainly due to the much smaller number of individuals used in our dataset and as such our timing of divergence for all pairs should be taken with some caution. We also provided the first estimates of divergence timing and scenario for the ecotype pairs found at lochs Arkaig, Ericht, and Lochy. Our findings suggest that each of these pairs represents secondary contact divergences which originally diverged during the last ice age. However, analysis using large numbers of individuals should be done before any concrete conclusions are drawn on the exact timings of these divergences. Overall, our results clearly indicate that the British Isles were colonised by multiple ancestral

refugia populations, some of which represent the two known sub-lineages of the Atlantic lineage.

### *5.5.2 Contemporary genetic differentiation within and across Hydrometric Areas*

Patterns of Isolation-by-distance have been seen in other freshwater fish in the British Isles, such as common roach (*Rutilus rutilus*) although the strength of the pattern varied by river catchment (Crookes and Shaw, 2016). Research in Brown trout has also shown strong Isolation-by-distance within river catchments although only when barriers to movement were removed (Griffiths et al., 2009). Patterns of Isolation-by-distance have been seen in non-anadromous Arctic charr populations in the Labrador region of Canada before although this was suggested to reflect colonisation history and high levels of admixture between colonising lineages rather than contemporary gene flow (Salisbury et al., 2023).

Across our dataset, we show differing relationships between the extent of genetic differentiation and geographical distance across populations (Figure 5-3). Looking at comparisons between populations in the same Hydrometric Area showed a strong pattern of Isolation-by-distance (Figure 5-3B). This pattern is in-line with the evidence that migration of Arctic charr between lakes is limited (Maitland and Adams, 2018) and previous work that suggests notable gene flow only occurs between proximal populations (Chapter 2). Isolation-by-distance was much weaker when looking across all comparisons in the dataset (Figure 5-3A). Given the lack of anadromy in Arctic charr in the British Isles, the need to migrate through the marine environment effectively acts as a barrier to gene flow. As such, we only see a strong pattern of IBD when we look at comparisons within the same Hydrometric Areas (Figure 5-3B) as was the same in brown trout (Griffiths et al., 2009). We saw no relationship between geographic distance and genetic differentiation when looking at populations in different Hydrometric Areas as populations are isolated from one another by the marine environment (Figure 5-3C). This reflects the lack of contemporary gene flow and highlights other processes such as genetic drift as more predominately driving genetic differentiation. Extent of genetic differentiation may also be driven by differences in local environment, i.e. Isolation-by-environment which should be explored going forward.

### *5.5.3 The effect of the loss of anadromy on genetic differentiation*

A number of contemporary freshwater fish species, show evidence of populations with no anadromous phenotype having arisen from anadromous ancestors (Palkovacs et al., 2008; Delgado et al., 2019; Delgado and Ruzzante, 2020). The loss of marine migration is

both well-documented in freshwater fish and believed to be a key instigator of their speciation (Delgado et al., 2019; Waters et al., 2020). Arctic charr were likely anadromous when they first colonised the British Isles and the broader patterns of genetic structuring in our dataset seemingly reflect that. Our Neighbour-joining tree primarily split populations by river system flow with populations in Scotland splitting into east and west catchment flowing groups (Figure 5-4). As previously suggested (Chapter 2), this may reflect a period of historic anadromy as the direction of river system flow creates a notable east-west divide in Scotland and likely limited contact between populations on opposing sides of Scotland when they migrated to sea. The outgroup populations, from Wales, Ireland, and the west of England, cluster amongst west-flowing populations, which makes sense geographically as there is a higher likelihood of mixing with these populations. The finding that populations in nearby river catchments show more evidence for historic contact, in line with the probability of straying into a non-natal river decreasing with distance has also been shown in studies of Chinook salmon (*Oncorhynchus tshawytscha*) and Steelhead (*Oncorhynchus mykiss*) (Candy & Beacham, 2000; Westley et al., 2013).

Given that these populations still have access to the sea and are often in close proximity to it, the loss of anadromy is likely driven either by a decrease in fitness for migrants, for example due to increase in predatory and competing species in the marine environment, or an increase in fitness in not migrating (Hogan et al., 2014). Amongst salmonid fishes, anadromy is primarily driven by the increased productivity of the marine environment, particularly for species like Arctic charr that typically inhabit unproductive post-glacial lakes (Finstad and Hein, 2012). However, the warmer conditions at the southern end of the species range have led to higher primary productivity within the freshwater environments which arguably may have driven the loss of anadromy. Research suggests a a future loss of anadromy may occur for current anadromous populations in Canada due to increasing temperatures as a result of climate change (Layton et al., 2021). Given that Arctic charr likely first colonised during the Loch Lomond Stadial, the brief return of glacial conditions (Bickerdike et al., 2018) may have favoured the retention of anadromy for some time until conditions became much warmer during the Holocene and at this point these conditions no longer favoured anadromy (Kaufman et al., 2020). However, it remains unclear how long this anadromous period was and whether the anadromous phenotype was lost for all populations around the same time or whether, for example the more northern populations, retained the phenotype for longer.

### *5.5.4 How glacial history shaped colonisation history and contemporary genetic patterns*

Glacial history is a key driver of patterns of genetic differentiation (Fenton et al., 2023b) (Chapter 4). Firstly, the presence of large-scale ice cover forced many temperate species to survive in isolated safe havens known as glacial refugia (Hewitt, 1996). These refugia populations resulted in differentiation in isolation and the resultant different lineages/sub-lineages that colonised their contemporary ranges which ultimately led to varying patterns of genetic differentiation and diversification based on colonisation history (Fenton et al., 2023b) (Chapter 4). Some species show clear separation in where each ancestral lineage colonised (Luquet et al., 2019; Crotti et al., 2021), in others clear points of secondary contact between the two lineages can be seen resulting in hotspots of genetic diversity (Roux et al., 2013; Chiocchio et al., 2021; Chiocchio et al., 2022). Secondly, the patterns of ice advance and retreat were asynchronous and as a result different parts of the contemporary ranges of many temperate species were first accessible to be colonised at very different times (Clark et al., 2012; Fenton et al., 2023b) (Chapter 4). This in turn can affect what lineages colonised where which can greatly influence patterns of genetic differentiation.

The evidence of multiple sub-lineages in our dataset indeed suggests colonisation of the British Isles involved populations from multiple refugia given the estimated divergence timing of these sub-lineages is during the last ice age (Jacobsen et al., 2022). There is also evidence in our dataset that the differential patterns of ice coverage influenced patterns of colonisation. The only population in Scotland that shared an mtDNA haplotype with a known Atlantic sub-lineage 1 population was Loch Lee (Figure 5-2). Loch Lee is located in an eastern region of Scotland (Figure 5-1) which was one of the first parts of Scotland be free from ice cover, *ca.* 17 ka (Clark et al., 2012), which was several thousand years before much of the rest of Scotland. It has previously been theorised that sub-lineage 1 may represent an earlier colonisation event in Greenland than sub-lineage 2 (Jacobsen et al., 2022). Our results suggest the same may be true in Scotland, with the sub-lineage potentially colonising before much of Scotland was ice-free and as such is restricted to a few areas that were accessible during its time of colonisation.

Given Arctic charr is estimated to have colonised during the Loch Lomond Stadial, we theorised that populations in regions still covered by ice during this time and proximal to known ice-dammed lakes would show distinct patterns of genetic differentiation and mixing. Colonisation through ice-dammed and glacial lakes has been seen in other freshwater species

including lake trout (*Salvelinus namaycush*), longnose sucker (*Catostomus catostomus*), round whitefish (*Prosopium cylindraceum*), and lake chub (*Couesius plumbeus*) (Ruzzante et al., 2019) and can result in patterns of genetic mixing and lower genetic differentiation between populations in different river catchments (Pielou, 1991).

For our dataset, there were two key areas of interest in Scotland both of which were covered by ice during the LLS (Figure 5-1) (Bickerdike et al., 2016; Bickerdike et al., 2018). The first central Scotland is four populations in the Lochy Hydrometric Area, lochs Treig, Laggan, Lochy and Arkaig, which are all near a series of lakes over Glen Spey and Glen Roy (MacLeod et al., 2011). The second in the Northwest of Scotland, is a group of populations across several different Hydrometric Areas, including lochs Fannich, Dughaill, Coulin and Lochan Uaine, which all surround an ice-dammed lake located near the village of Achnasheen (Benn, 1989).

Evidence that these populations were likely colonised through these ice-dammed lakes can be seen in our study. For example, the Loch Fannich, Loch Coulin, Lochan Uaine, and Loch Dughaill planktivore populations all consist solely of the same shared mitochondrial haplotype despite being located in three separate Hydrometric Areas (Figure 5-2). Notably this shared haplotype is not seen in Loch Luichart, into which Loch Fannich flows. A similar pattern has been seen in lake whitefish in North America, where populations in part of the Mackenzie drainage show more contemporary genetic similarity to populations in the Yukon drainage than to other populations found within their own catchment. These populations are notably close to where two proglacial lakes were found, suggesting that colonisation occurred through those lakes from the Yukon drainage resulting in the patterns of genetic differentiation seen (Pielou, 1991). Genetic differentiation between these populations is relatively low despite a relatively large geographic distance in some comparisons and appear in the same section of our neighbour-joining tree (Figure 5-5). Due to the lack of anadromy, these similarities cannot be the result of contemporary gene flow. Their relative proximity to one another, all found approximately within 25 km of each other, could suggest human movement of fish between lakes, however, there is little evidence of this occurring in Arctic charr in the British Isles (Maitland and Adams, 2018). As such, these similarities are more likely the result of a shared colonisation history or postglacial origin which occurred through an ice-dammed lake during the LLS.

The populations in the Lochy Hydrometric Area also show some unique patterns. When we looked at possible migration events across lakes, we found that two of them

involved movement between populations in the Tay Hydrometric Area into the Lochy Hydrometric Area (Figure 5-6). While geographically proximal, these two Hydrometric Areas flow out to the opposite sides of Scotland. As such, it is unlikely these potential migrations, and gene flow between populations, occurred due to mixing when populations were still anadromous. It is therefore more likely that this mixing occurred through the series of ice-dammed lakes in the region during the LLS. Populations in the two Hydrometric Areas show much lower genetic differentiation than would be expected given the geographical distance between them and all populations in the west-flowing Lochy Hydrometric Area appear in the east-flowing group in the neighbour-joining plot (Figure 5-5). Given the ice coverage over the Lochy Hydrometric Area during the LLS, it is entirely possible that the populations in the nearby Tay Hydrometric Area came into the ice-dammed lakes and then colonised these lakes once free from ice coverage, similar to how a refugium population colonises the immediate area around it once colonisable (Herman and Searle, 2011), and as such we see lower genetic differentiation that might be expected.

While we focused on the patterns for these populations, there are other populations proximal to these ice-dammed lakes in these two areas of interest not covered by our dataset and these warrant further investigation to understand the true extent of these cross Hydrometric Area patterns. Overall, our results highlight the value in using glacial history when investigating colonisation history and contemporary genetic differentiation.

## 5.6 Conclusions

Our results highlight the importance of glacial history in shaping contemporary patterns of genetic differentiation. Arctic charr populations from different Hydrometric Areas that showed low genetic differentiation and evidence of mixing despite a lack of anadromy were populations covered by ice during the Loch Lomond Stadial and had shared colonisation histories or population mixing facilitated by known ice-dammed lakes. Patterns of Isolation-by-distance were strong within Hydrometric Areas but a lack of anadromy prevents any pattern of Isolation-by-distance across Hydrometric Areas with genetic drift the more predominant factor. The British Isles were colonised by multiple refugia populations which includes the two known sub-lineages of the Atlantic lineage. The sub-lineages vary in their prevalence suggesting notably different timing for their colonisation events. These results highlight how we must look into the past in order to truly understand the patterns of the present.

# Chapter 6: Genomic underpinnings of head and body shape in Arctic charr ecomorph pairs

Note: At the time of submission this article was in revision as an article in Molecular Ecology. Supplementary tables will be available as part of the published article. A pre-print is available on *Authorea*. (*https://doi.org/10.22541/au.169650386.60612041/v1*).

## 6.1 Abstract

Across its Holarctic range, Arctic charr (*Salvelinus alpinus*) populations have diverged into distinct trophic specialists across independent replicate lakes. The major aspect of divergence between ecomorphs is in head shape and body shape, which are ecomorphological traits reflecting niche use. However, whether the genomic underpinnings of these parallel divergences are consistent across replicates was unknown but key for resolving the substrate of parallel evolution. We investigated the genomic basis of head shape and body shape morphology across four benthivore-planktivore ecomorph pairs of Arctic charr in Scotland. Through genome-wide association analyses, we found genomic regions associated with head shape (89 SNPs) or body shape (180 SNPs) separately and 50 of these SNPs were strongly associated with both body and head shape morphology. For each trait separately, only a small number of SNPs were shared across all ecomorph pairs (3 SNPs for head shape and 10 SNPs for body shape). Signs of selection on the associated genomic regions varied across pairs, consistent with evolutionary demography differing considerably across lakes. Using a comprehensive database of salmonid QTLs newly augmented and mapped to a charr genome, we found several of the head and body shape associated SNPs were within or near morphology QTLs from other salmonid species, reflecting a shared genetic basis for these phenotypes across species. Overall, our results demonstrate how parallel ecotype divergences can have both population-specific and deeply shared genomic underpinnings across replicates, influenced by differences in their environments and demographic histories.

## 6.2 Introduction

Parallel evolution is an evolutionary process and outcome by which similar phenotypes arise and establish in multiple independent populations in separate environments (Schluter, 1996; Elmer and Meyer, 2011; Bolnick et al., 2018). These replicate evolutionary outcomes suggest that similar environments impose similar selective pressures on organismal phenotypes, with only a small number of phenotypic solutions favoured in that context. However, replicate populations often show deviations from what could be described as strict parallel evolution (Bolnick et al., 2018). These instances of non-parallelism can be driven by a range of factors (Elmer and Meyer, 2011; Oke et al., 2017; Stuart et al., 2017). While replicates can exist in similar environments, particular spatial and temporal selection pressures can and do vary even in environments that seem equivalent (Oke et al., 2017). The extent of parallelism can be influenced by a range of ecological factors (Landry et al., 2007; Riesch et al., 2016). For example, ecomorphs of three-spine stickleback (*Gasterosteus aculeatus*) that show higher divergence in diet also show high divergence in ecologically relevant traits (Berner et al., 2008; Kaeuffer et al., 2012). Further, the extent of contemporary gene flow and differing demographic histories can directly influence the extent of parallelism seen (Langerhans and DeWitt, 2004; Hendry, 2017; Weber et al., 2021). Replicates may also find alternative solutions to the same functional problem resulting in deviations from strict parallelism (Alfaro et al., 2004).

While the existence of parallel phenotypes is well established in natural populations (Siwertsson et al., 2013; Oke et al., 2017), the extent to which those are associated with similarly shared genomic underpinnings is rarely examined (Ravinet et al., 2016; McGirr and Martin, 2018; Magalhaes et al., 2021). Indeed, the appearance of the same phenotypes through parallel evolution processes does not mean that the same genomic processes underpin those similar evolutionary outcomes across replicates (Conte et al., 2012; Elmer & Meyer, 2011). Similar phenotypic outcomes could result from alternative genetic pathways with multiple genes influencing a single phenotype, known as polygenicity (Láruson et al., 2020). This might arise because of differing demographic histories (Elmer et al., 2014), variable genetic backgrounds (Arendt and Reznick, 2008; Kowalko et al., 2013), while relative effect size and hard and soft selective sweeps can also influence the genomic pathways used (Pritchard et al., 2010; Láruson et al., 2020). Similarly, the involvement of alternative splicing, differential gene expression, or post-translational modifications can also result in phenotypic parallelism through a different molecular mechanism (Filteau et al., 2013; McGirr and Martin, 2018; Jacobs and Elmer, 2021).

One classic example of parallel evolution is the replicated divergence across northern freshwater lakes of distinct trophic specialists, also known as ecomorphs or ecotypes. These occur abundantly in salmonid fishes in recently glaciated lakes, such as lake whitefish (*Coregonus clupeaformis*) (Landry et al., 2007), lake trout or lake charr (*Salvelinus namaycush*) (Baillie et al., 2016), and Arctic charr (*Salvelinus alpinus*) (Jensen et al., 2017; Doenz et al., 2019). In Arctic charr, ecomorphs associated with divergence along the depth axis and ecological niche, typically forming pelagic and benthic foraging specialists (Klemetsen, 2010; Elmer, 2016). Two key traits involved in this divergence are head shape and body shape, both having functional significance; head shape being important for foraging and prey specialisation and body shape important for swimming behaviour and niche use (Adams & Huntingford, 2002; Skoglund et al., 2015). Additionally, ecomorphs often differ in other complex traits such as body size, colouration, and spawning time (Jonsson and Jonsson, 2001; Garduño-Paz et al., 2012).

Previous research on benthivorous-planktivorous ecomorph pairs of Arctic charr has shown that individual traits related to head shape, such as eye diameter, can show a strong degree of parallelism across replicates however the extent of parallelism is highly variable across individual traits (Adams et al., 2008; Jacobs et al., 2020). Research into benthic ecomorphs in Iceland has shown that while these ecomorphs are morphologically similar across locations, they are distinguishable in several characteristics, most prominently in head shape, suggesting some specialised adaptations to their local environment (Sigursteinsdóttir and Kristjánsson, 2005). Genomic analysis to-date suggested limited genetic parallelism across lakes, with many differences in demography and colonisation history across the breadth of the Arctic charr distribution (Jacobs et al., 2020; Salisbury et al., 2020; Brachmann et al., 2022). However, this previous research on parallelism between Arctic charr ecomorphs focused on the patterns of the genomic response to selection such as outlier loci, which are known to have high rates of false-positives and false-negatives (Kelley et al., 2006; Excoffier et al., 2009; Narum and Hess, 2011) and are indirect approaches to exploring adaptive divergence in morphology, strongly influenced by evolutionary genomic and demographic histories (Ravinet et al., 2017). Therefore, to properly understand the genomic underpinnings of key phenotypes we need to apply different approaches that identify loci associated with phenotypic differences through genome-wide association studies (GWAS), in Arctic charr and determine the extent to which these are shared across ecomorph pairs (Elmer, 2016).

Because of their vital role in foraging and swimming, the genetics of head and body morphology has been explored as QTL studies in many fish species, including salmonids (Boulding et al., 2008; Küttner et al., 2014; Laporte et al., 2015; Smith et al., 2020). For example, in lake whitefish, as many as 138 different quantitative trait loci (QTLs) related to body shape have been identified, implying that the trait is highly polygenic in this species (Laporte et al., 2015). In other species, a fewer number of significant QTLs have been found for body shape, with a small number of QTLs identified related to benthic-limnetic differences between ecomorphs of lake trout (Smith et al., 2020) and a single region that differentiates ecomorphs in Atlantic cod (*Gadus morhua)* (Hemmer-Hansen et al., 2013). However, the extent to which parallel phenotypes have the same genomic underpinnings across replicates seems to vary greatly between species and locality. Studies in three-spine stickleback, rainbow trout (*Oncorhynchus mykiss*), and sockeye salmon (*Oncorhynchus nerka*) all suggest that while some genetic variation that underlies morphological variation can be shared across replicates, this is rarely the case across the species' entire range and the degree of genetic parallelism that is shared is often low (Weinstein et al., 2019; Larson et al., 2019; Fang et al., 2020). Often only a few key genes or loci are shared across replicates, and in some cases, they underlie morph differentiation in multiple species (Jacobs et al., 2017; Salisbury and Ruzzante, 2021).

In this study, we investigated the genomic underpinnings of head and body shape morphology in replicate ecomorph pairs of Arctic charr across four independent lakes in Scotland. First, we determined the extent of phenotypic parallelism in head and body shape morphology across ecomorph pairs. We examined these two traits separately as they are known to involve different axes of ecological specialisation and can have different genetic bases (Boulding et al., 2008; Küttner et al., 2014; Smith et al., 2020). Second, we investigated the genome-wide underpinnings of these phenotypes by identifying SNPs strongly associated with head shape variation and body shape variation. We then evaluated the genomic organisation of these head shape- and body shape-associated SNPs, specifically to determine if they were distributed widely across the genome, co-localised in genes or genomic regions, or within known salmonid QTLs. To do this, we updated and augmented an existing QTL database (Jacobs et al., 2017) and mapped this to the orthologous location in the *Salvelinus* sp. genome. Third, we examined evolutionary genomic background by analysis of selection on those loci we found to be associated with head shape and with body shape. Shared signals of selection and elevated differentiation and divergence in independent ecomorph pairs would suggest similar processes across evolutionary replicates. Finally, we synthesised the findings to identify shared genomic underpinnings and similarities in

phenotypic divergence between head and body shape, as two physically interlinked components of ecomorphology.

## 6.3 Materials and Methods

### *6.3.1 Study populations*

Fish populations were examined from four lakes in Scotland: Loch Awe, Loch Dughaill, Loch na Sealga, and Loch Tay (Figure 6-1A). Each of these lakes are found in different river systems and have distinct demographic histories and contemporary differentiation, as revealed by previous genetic structuring analysis (Maitland and Adams, 2018; Jacobs et al., 2020) and thus represent independent replicates. Each lake contains two ecomorph populations, one benthivorous and the other planktivorous, with each lake representing an ecomorph pair (Garduño-Paz et al., 2012; Hooker et al., 2016; Jacobs et al., 2020). Explicitly, we investigated the similarities in divergences of head and body shape between the two different ecomorphs across pairs. Short-read genomic data (ddRADseq NCBI short read bioproject: PRJNA607173) and fish photographs were drawn from previous research (Jacobs et al., 2020) and reanalysed here.

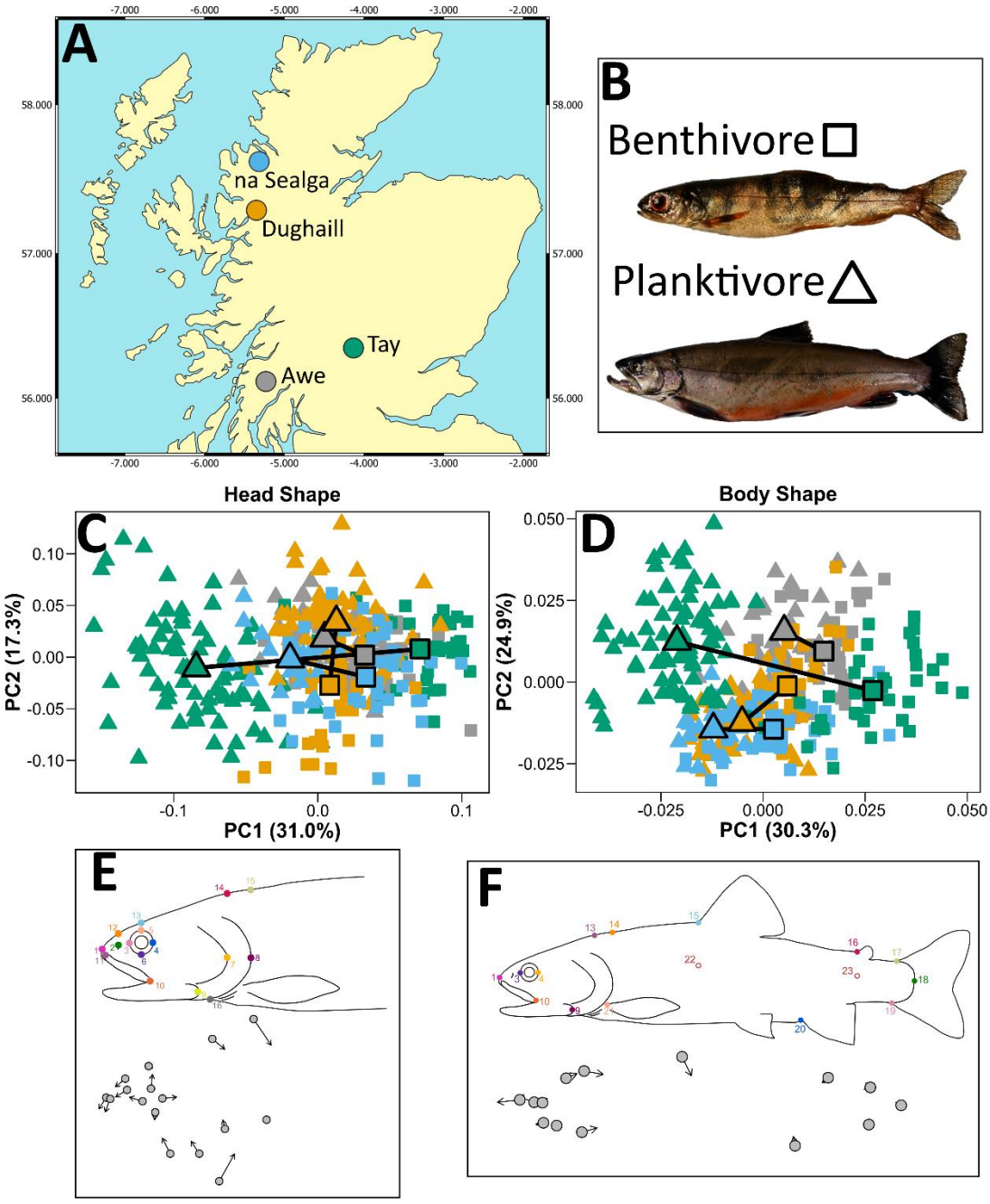


**Figure 6-1.** Sample locations and morphological analysis. (A) Map of Scotland showing the sampling locations of Arctic charr. (B) Pictures showing an example ecomorph pair, here from Loch Tay. Principal component analysis for landmark analysis for Head shape (C) and Body shape (D). Individuals are coloured by lake with larger points representing the mean values for each ecomorph with a line connecting means in each pair. Vector plots display how head shape (E) and body shape (F) morphology changes across PC1 with diagrams showing where each landmark was placed on the fish for each phenotype. Grey points show landmark positions at the minimum PC1 score and vectors show how landmark positions changes at the maximum PC1 score.

### *6.3.2 Morphological analyses*

To investigate patterns in head morphology and body shape separately, landmarks for head (16 landmarks) and body shape (14 landmarks) were placed using *ImageJ v1.50i* (Schneider et al., 2012) (Figure 6-1E-F). Raw linear measurements for each individual, including fork length can be found as part of Jacobs et al., 2020. To allow us to cover the whole-body shape of each individual, several markers related to head shape were retained in the body shape analysis. As such, the head and body shape analyses are not completely independent but are focused on each specific phenotype and use a distinct set of landmarks. In total, 341 individual fish were landmarked for head shape (Awe_Bn=37, Awe_Pl=31, Dughaill_Bn=43, Dughaill_Pl=54, naSealga_Bn=42, naSealga_Pl=20, Tay_Bn=41, Tay_Pl=73) and 335 were landmarked for body shape (Awe_Bn=34, Awe_Pl=30, Dughaill_Bn=43, Dughaill_Pl=54, naSealga_Bn=42, naSealga_Pl=19, Tay_Bn=40, Tay_Pl=73). Separate analyses were run for head shape and body shape to allow for comparisons between the two phenotypes.

General Procrustes Analysis (GPA) was performed to standardise the orientation and remove the isometric effect of size on shape for each individual using *geomorph v3.0.7* (R package) (Adams and Otárola-Castillo, 2013). Following this procedure, shape data were standardised for any residual size effects using the log of centroid size to correct for allometry after determining all lake ecotype populations had parallel slopes using the *procD.allometry()* function (Klingenberg, 2016). Size corrected data then were used in all subsequent analyses. Principal Component Analyses (PCA) using the *plotTangentSpace* and *arrayspecs* functions with plots made in *ggplot2 v3.3.5* (R package) (Wickham, 2016; R Core Team, 2021).

Analysis of variance (ANOVA) models were used to determine the relative contributions of parallel and non-parallel aspects of the morphological divergence across ecomorph pairs (Langerhans and DeWitt, 2004). Our ANOVA models (ANOVA model: PC ~ ecomorph + lake + ecomorph x lake), were run for PC1, as it explained the most variance in shape. The *EtaSq* function in the *BaylorEdPsych v0.5* (R package) (Beaujean, 2012) was used to estimate the effect size of each model term. In our model, the ecomorph term represents the parallel or shared term, the ecomorph*lake interaction is the non-parallel (non-shared) term, and the lake term represents unique evolutionary history.

Phenotypic Trajectory Analyses (PTA) were performed on the procrustes scores using the *trajectory.analysis* function in *geomorph* to look at the extent and direction of

phenotypic difference between ecomorphs. Magnitude of divergence is described by the length of trajectories (L) while the angle between trajectories (θ) describes their direction in phenotypic space. This approach allows us to determine how parallel the trajectories of each ecomorph pair are to one another by using the difference in phenotypic vector length ($\Delta L_P$) and the direction of phenotypic trajectories ($\theta_P$) where P is either head shape (Head) or body shape (Body) (See Stuart et al., 2017). The significance of differences in vector lengths and differences in vector direction were assessed using 1,000 permutations (Adams and Collyer, 2013; Stuart et al., 2017). Phenotypic divergence between ecomorphs in different lakes was considered to be parallel if the direction and/or magnitude of phenotype difference did not differ significantly between the pairs ($P > 0.05$).

### *6.3.3 Population genomics*

Filtered 75 bp reads for each individual, generated via ddRADseq from Jacobs *et al*. (2020) and accessed from (ddRADseq NCBI short read bioproject: PRJNA607173), were mapped using *bwa mem* and *SAMtools* using settings described in that paper to the *Salvelinus* sp. genome from NCBI (ASM291031v2). The number of reads per individual ranged from 1 to 3.5 million. RAD loci were built in the *gstacks* module of *Stacks v2.53* (Rochette et al., 2019) for 200 individuals (Awe_Bn= 26, Awe_Pl=29, Dughaill_Bn=28, Dughaill_Pl=27, naSealga_Bn=18, naSealga_Pl=20, Tay_Bn=21, Tay_Pl=31). Biallelic SNPs were retained in the *populations* module of *Stacks* if they met the following criteria: present in 66% of all individuals in each population and across all populations, a minimum minor allele frequency of 0.05, maximum observed heterozygosity of 0.5. Each ecomorph within a lake was considered to be a discrete population (Jacobs et al., 2020). The script filter_hwe_by_pop.pl to filter out sites outside Hardy-Weinberg equilibrium within populations using a p value of 0.001 (available at https://github.com/jpuritz/dDocent). *vcftools v0.1.13* (Danecek et al., 2011) was used to filter to a minimum coverage of 5x and a maximum of 50x. 13,071 SNPs were retained for analyses at a mean genotyping rate of 74.8%. A principal component analysis was performed to identify the major axes of genetic variation using *SNPRelate v1.22.0* (R package) (Zheng et al., 2012).

### *6.3.4 Genotype-phenotype association analyses*

To determine the association between genotypes and phenotypic variation in head or body shape, we ran Linear Mixed Models (LMM) in Gemma v0.98.1 (Zhou and Stephens, 2012). Univariate and Multivariate LMMs with Wald's test were run using PCs 1-5 for head shape and body shape as each explained more than 5% of the total variance (Figure A6-1),

the SNP dataset generated for the population genomics analyses, and lake of origin as a co-variate (Zhou and Stephens, 2014). Missing genotypes were imputed using *LinkImpute v1.1.4* (Money et al., 2015) and a relatedness matrix was generated using *Gemma* and included in the models to correct for population structure. Changes in allele frequency after imputation were checked with 85% of SNPs showing less than a 5% change in allele frequency. We determined significant associations using bonferroni-corrected P-values (0.05/7329 unlinked SNPs) from the Wald's tests. The number of unlinked SNPs was determined by LD-pruning the full SNP dataset using the snpgdsLDpruning function in *SNPRelate*. Bayesian Sparse Linear Mixed Models (BSLMM) (Zhou et al., 2013) were run using PC1-5 variables to determine how much of the phenotypic variation is explained by the SNPs in our dataset (PVE) and secondly how much of that variation is explained by large-effect loci (PGE), and finally how polygenic each phenotype is (ρ). Manhattan plots were made using *CMplot v4.0* (R package) (https://github.com/YinLiLin/CMplot).

We subsequently determined if SNPs showing significant associations with head shape or body shape morphology were found within annotated genes in the *Salvelinus* sp. reference genome using *BEDtools v2.27.1* (Quinlan and Hall, 2010). The functions of genes containing, or ± 1kbp of, associated SNPs were investigated using GO term overrepresentation analysis (ORA) and gene set enrichment analysis (GSEA). These analyses were run using *topGO v2.40.0* (R package) (Alexa and Rahnenfuhrer, 2020) with all genes containing any RAD loci as the full comparison dataset. Results were summarised using *REVIGO* (Supek et al., 2011) before visualisation in *Cytoscape v3.91* (Shannon et al., 2003).

### *6.3.5 Comparisons to known QTLs*

Using existing information on the genetics of important phenotypes from other salmonid species, we mapped a database of 1,338 QTL markers to the *Salvelinus* sp. genome. This was based on a previously published database of QTLs involved in traits related to morphology and life history, derived from a range of salmonid species and previously mapped to the *Salmo salar* genome (Jacobs et al., 2017). Additionally, a literature search was conducted up to April 2021 to augment the existing database with more recently published QTLs. This literature search was conducted in Web of Science and Google Scholar using the search terms "QTL", "quantitative trait loci", "salmonid", and the common and scientific names for rainbow trout, Atlantic salmon, Arctic charr, lake whitefish, Chinook salmon, coho salmon, brook trout, and lake trout. QTL marker sequences were gathered for 17 different phenotypes: body length, body shape, body weight, Fulton's condition factor,

directional change, disease resistance, embryonic development, gill rakers, growth rate, hatching time, head shape, parasite resistance, salinity tolerance, sexual maturation, smolting, spawning time, upper temperature tolerance (Table A6-1).

Following Jacobs et al. (2017), the strategy of mapping the QTL-linked markers to the *Salvelinus* sp. genome depended on the QTL marker type: RAD loci were mapped using *Bowtie2 v2.4.4* and the *very sensitive* pre-set; microsatellite primer sequences, which are shorter, were mapped using *Bowtie v1.3.1* allowing for 3 mismatches. QTLs for which the flanking markers mapped to different chromosomes were removed. Redundant QTLs, i.e., where two QTLs for the same trait from the same species mapped to the same location, were removed only keeping the QTL with the higher PVE or LOD score (following Jacobs et al., 2017). For QTLs where more than one marker was reported, we attempted to map all markers. Position values for the QTLs markers were then compared to positions of the phenotype associated SNPs using *BEDtools*, with a cut-off of ±100kbp. This value was used so that we could consider SNPs within the range to be proximal to a QTL peak while also accounting for the large size of many of the QTLs in the database. In total, we successfully mapped 669 QTL-linked sets of markers to the *Salvelinus* sp. genome after removing redundant QTLs (Table A6-2).

### *6.3.6 Genomic response to selection*

We investigated if the phenotype-associated SNPs identified in our analyses showed signals of genomic differentiation potentially caused by response to selection and if those signals were replicated across ecomorph pairs. To test this, for each ecomorph pair we compared $F_{ST}$ and $D_{XY}$ values for phenotype-associated SNPs to a random background subset of SNPs. This random subset was 100 SNPs randomly selected from the whole dataset and the mean $F_{ST}$ and $D_{XY}$ values for those SNPs were calculated. This was repeated 10,000 times and the means for $F_{ST}$ and $D_{XY}$ were taken across all permutations. These permuted values were then compared to the empirical mean $F_{ST}$ and $D_{XY}$ values for the phenotype-associated SNPs using the *t.test* function in R.

### *6.3.7 Analyses of recombination rate variation*

To test the effect of the recombination landscape on phenotype-genotype association, we first estimated recombination rates using the published Arctic charr linkage map (N=3,636) (Christensen et al., 2018) using *MareyMap v1.3.6* (Siberchicot et al., 2020). RAD loci from the linkage map were aligned to the *Salvelinus* sp. reference genome with Bowtie2 (Langmead and Salzberg, 2012) using the *-very-sensitive* setting. Loci were kept if they

uniquely mapped to one location, mapped to the same chromosome as all other loci on their linkage group, and followed the orientation of the linkage map (i.e., not reversed). The filtered dataset was used to estimate the recombination rate across each chromosome using a spline algorithm. Spar values were varied for each chromosome from 0.5 to 0.9, depending on chromosome size, to best fit the data (Berloff et al., 2002). Subsequently, *WindowScanR v0.1* (available at: https://github.com/tavareshugo/WindowScanR) was used to summarize recombination rate values in 1 MB windows along the genome. All SNPs were assigned to these windows using *BEDtools*. A random subset of 100 SNPs was then selected and the mean recombination rate for those SNPs was calculated based on their windows. This was repeated 10,000 times to generate a background mean recombination rate, which was then compared to the mean recombination rate of the phenotype-associated SNPs (based on their windows) using a t-test.

## 6.4 Results

### *6.4.1 Ecomorph divergence in head shape*

For head shape the benthivore and planktivore ecomorphs were separated across PC1 (31% variance explained), except from Loch Dughaill where the ecomorphs separate along PC2 (17.3%) (Figure 6-1C). Individuals with a positive PC1 score had shallower heads with larger eyes than those with negative PC1 scores (Figure 6-1E). For PC2, a more positive score suggested a longer head shape. The benthivore ecomorphs generally have a more negative PC2 score suggesting their heads are shorter than the planktivore ecomorphs (Figure A6-2).

We compared the magnitude and direction of phenotypic difference for head shape between the ecomorphs across pairs, to determine how similar the divergences were, through a phenotypic trajectory analysis (PTA). We found that for all pairwise comparisons, the angle of difference in phenotypic trajectories ($\theta$) was significantly different ($P < 0.05$) (Table 6-1). Similarly, almost all differences in vector lengths between ecomorphs pairs ($\Delta L$) were significantly different with the exception of the Awe vs na Sealga comparison. These results suggest that head shape morphology is variable across lakes. While the ecomorph term explained the most variance along PC1 in our ANOVA model (PC ~ ecomorph + lake + ecomorph x lake) ($\eta^2_{Eco} = 0.595$), the ecomorph x lake interaction term explained a similar amount of variance ($\eta^2_{Eco*Lake} = 0.568$) suggesting that the effect of lake environment and/or evolutionary history strongly impacts the direction and magnitude of head shape divergence

between ecomorphs, and that there are unique elements of divergence in head shape across lakes (Table A6-3).

**Table 6-1.** Phenotypic trajectory analysis comparisons across ecomorph pairs for head shape and body shape. The difference in vector length, the magnitude of difference, between ecomorph pair is indicated by ΔL. The angle between trajectories is indicated as θ. The significance values are provided for each comparison.

| Phenotype | Ecomorph pair comparison | ΔL (magnitude of difference) | p value | θ (direction of difference) | p value |
|---|---|---|---|---|---|
| Head shape | Awe-Dughaill | 0.0347 | 0.003 | 67.51° | 0.001 |
| | Awe-naSealga | 0.0032 | 0.777 | 65.74° | 0.002 |
| | Awe-Tay | 0.0901 | 0.001 | 57.29° | 0.007 |
| | Dughaill-naSealga | 0.0379 | 0.002 | 78.37° | 0.001 |
| | Dughaill-Tay | 0.0554 | 0.001 | 100.56° | 0.001 |
| | naSealga-Tay | 0.0933 | 0.001 | 47.88° | 0.027 |
| Body shape | Awe-Dughaill | 0.0097 | 0.004 | 92.48° | 0.001 |
| | Awe-naSealga | 0.0013 | 0.726 | 55.27° | 0.004 |
| | Awe-Tay | 0.028 | 0.001 | 39.64° | 0.038 |
| | Dughaill-naSealga | 0.011 | 0.001 | 70.65° | 0.001 |
| | Dughaill-Tay | 0.0183 | 0.001 | 79.24° | 0.001 |
| | naSealga-Tay | 0.0293 | 0.001 | 34.20° | 0.118 |

### *6.4.2 Genomic regions associated with head shape*

To determine the genomic variation underpinning head shape, we performed a genome-wide association analysis (GWAS) on a set of 13,071 SNPs (PCA of ecomorph and lake variation shown in Figure A6-3). Using the PC scores from the head shape analysis (PCs 1-5), a Bayesian Sparse Linear Mixed Model (BSLMM) showed that the proportion of phenotypic variance in head shape explained by genetic variation ($PVE_{Head}$) in the SNP dataset was 0.62 with the proportion of phenotypic variance explained by large ("sparse") effect loci ($PGE_{Head}$) was 0.82. This is supported by the ρ (rho) value for head shape that

suggests that the head shape phenotype is controlled largely by a few large-effect loci ($\rho_{Head}$ = 0.792).

Applying Linear Mixed Models to identify SNPs highly associated with head shape variation found a total of 82 SNPs (66 SNPs mapped on 27 of 39 chromosomes, 16 mapped to unanchored scaffolds) that showed a significant association with variation in head shape (Bonferroni corrected P-value < 0.05; Fig. 6-2A) with these SNPs broadly distributed across the genome.

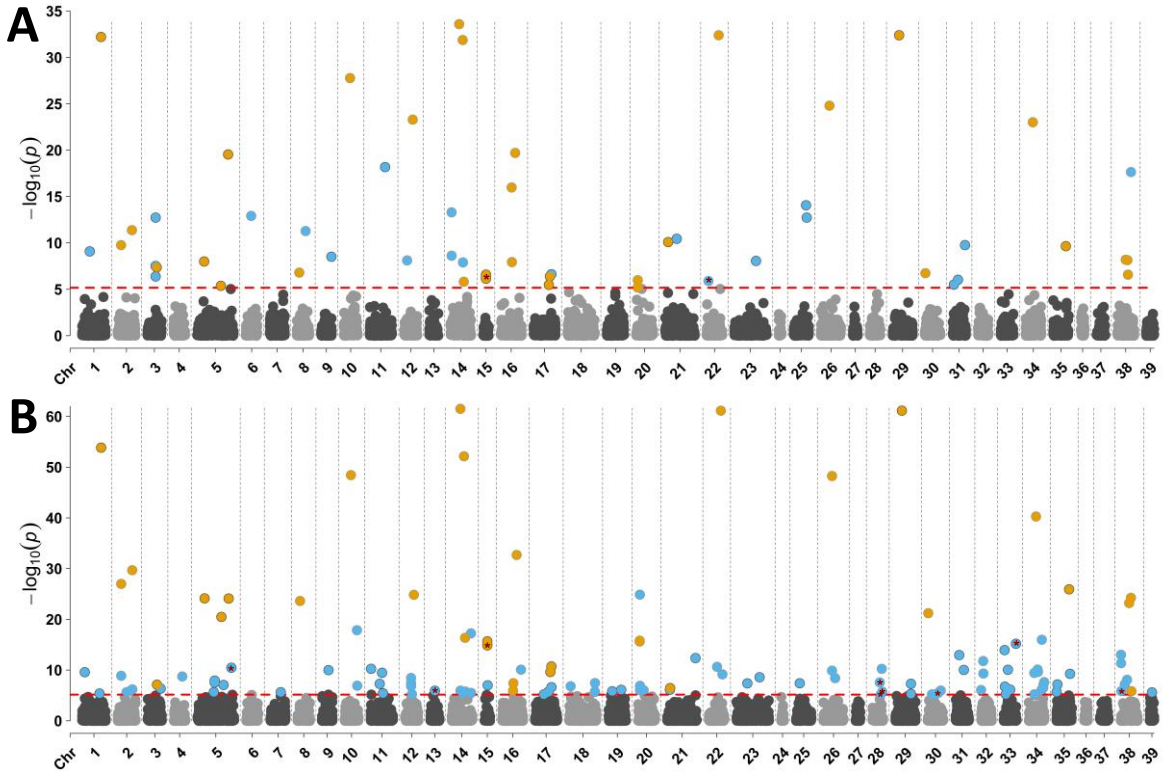


**Figure 6-2.** Manhattan plots for genomic location of SNPs highly associated with morphology. A) Shown are 67 SNPs associated with head shape across the four lakes of two ecomorphs of charr. B) Shown are 144 SNPs associated with body shape. SNPs associated with head or body shape are highlighted in blue; SNPs associated with both head shape and body shape are highlighted in orange (38 SNPs). Red asterisks indicate associated SNPs that show some frequency difference between ecomorphs in each of the four ecomorph pairs for head shape (N=2) and body shape (N=9) (referred to later as shared SNPs). SNPs on unanchored scaffolds are not pictured. Red dashed line shows cut-off for significance of association with the respective phenotype.

### *6.4.3 Genomic differentiation at SNPs associated with head shape*

We investigated whether these head shape-associated SNPs were highly diverged between the ecomorphs in all lakes, consistent with a shared genomic bases for these phenotypes, or whether they were specific to certain populations suggesting the deployment of different genetic pathways leading to the similar phenotypes across pairs. We found a total of three SNPs that were diverged in all four lakes (Figure 6-3A).

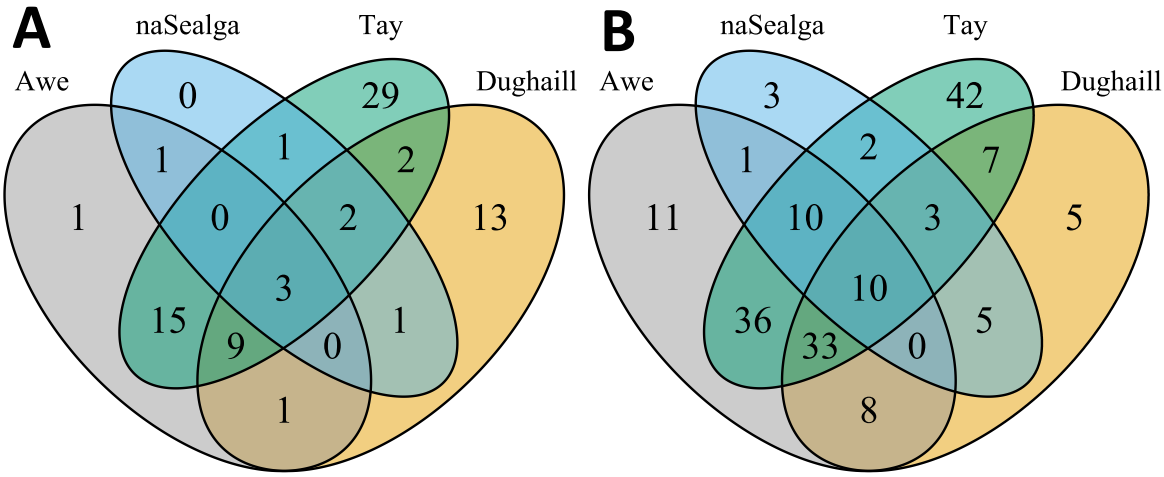


**Figure 6-3.** Venn diagram of SNPs associated with head shape (A) and body shape (B) and how they are shared across each combination of different lake pairs.

We aimed to identify if those SNPs associated with head shape showed signs of response to divergent or positive selection in all four lakes. Mean genetic differentiation ($F_{ST}$) and absolute divergence ($D_{XY}$) between ecomorphs in lochs Dughaill and Tay were elevated amongst associated SNPs when compared to the background (Figure 6-4, 6-5) (Table A6-4). $F_{ST}$ and $D_{XY}$ were significantly lower than background between ecomorphs in Loch Awe for the associated SNPs. There was no significant difference in $F_{ST}$ or $D_{XY}$ between associated SNPs and the background in na Sealga. These results suggest that the SNPs associated with head shape are not similarly responding to, or are under selection, across all lakes. These patterns were not influenced by linkage, because recombination rates for genomic regions around head shape-associated SNPs did not differ significantly from genomic background (Figure A6-4).

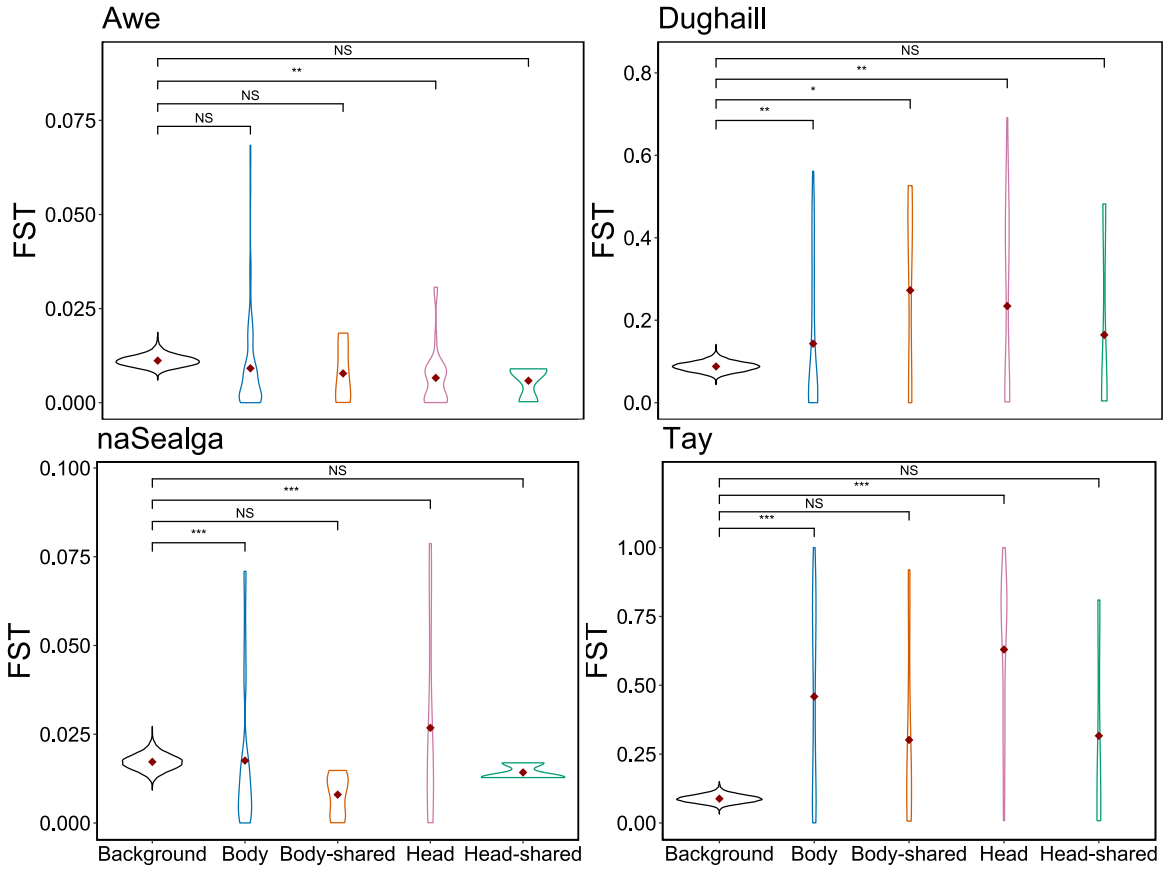


**Figure 6-4.** Genetic differentiation ($F_{ST}$) between ecomorphs at each lake for different associated SNPs datasets. Body refers to all SNPs associated for body shape and Head refers to all SNPs associated with head shape. Body-shared dataset is just the body shape SNPs found in all four ecomorph pairs and Head-shared is the equivalent for head shape SNPs. Background SNPs refers to a randomly selected background subset of SNPs used for comparisons. Red diamonds are the mean value for that dataset. Significance of difference in means is indicated by NS ($P > 0.05$), * ($P < 0.05$), ** ($P < 0.01$), *** ($P < 0.001$).

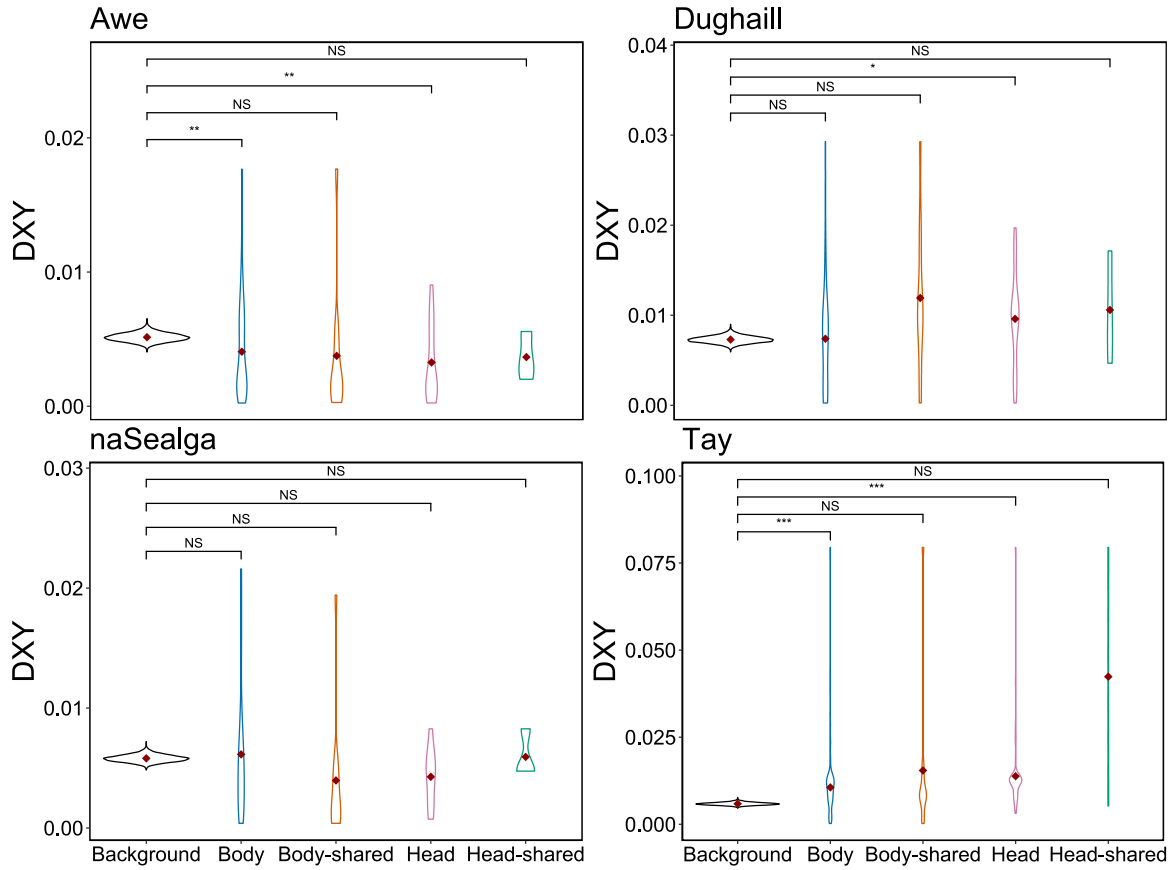


**Figure 6-5.** Absolute divergence ($D_{XY}$) between ecomorphs at each lake for different associated SNPs datasets. Body refers to all SNPs associated for body shape and Head refers to all SNPs associated with head shape. Body-shared dataset is just the body shape SNPs found in all four ecomorph pairs and Head-shared is the equivalent for head shape SNPs. Background SNPs refers to a randomly selected background subset of SNPs used for comparisons. Red diamonds are the mean value for that dataset. Significance of difference in means is indicated by NS ($P > 0.05$), * ($P < 0.05$), ** ($P < 0.01$), *** ($P < 0.001$).

### *6.4.4 Genes and QTLs associated with head shape variation*

Focusing on the location of the head shape-associated SNPs relative to known genes in the charr genome, we found that 38 of the 82 SNPs were located within annotated genes (Table A6-5). GO term analyses found that the genes containing head-shape associated SNPs showed over-enrichment for GO terms related to odontogenesis (GO:0042476), cranial skeleton system development (GO:1904888) among other processes and functions (Table A6-6) when compared to all genes containing SNPs in our dataset.

To examine if any of these genomic associations were shared across other species, we compared the positions of the head shape-associated SNPs to QTL markers from across salmonid species. We found that three of the head shape-associated SNPs were found within ±100,000 bp of the peak positions of two mapped QTLs (Table 6-2). These two QTLs were previously found to be associated with body shape morphology in lake trout and lake whitefish (Laporte et al., 2015; Smith et al., 2020).

**Table 6-2.** Table of associated SNPs found within ±100kbp of salmonid QTLs mapped to the Salvelinus sp. genome. Which phenotype the SNP is associated with, its position, and position of QTL marker in question are all indicated. The QTL type and species of origin are indicated. QTL name refers to the designated the QTL was given in the whole QTL database found in Table A6-1.

| SNP phenotype | QTL species | QTL marker | QTL type | Chromosome | SNP position | QTL position |
|---|---|---|---|---|---|---|
| Head and Body | *C.clupeaformis* | Cocl_BS_096 | Body shape | NC_036838.1 | 46764741 | 46685809 |
| Head and Body | *S.namaycush* | Sna_BS_069 | Body shape | NW_019942687.1 | 358354 | 361478 |
| Head and Body | *S.namaycush* | Sna_BS_069 | Body shape | NW_019942687.1 | 358355 | 361478 |
| Body | *S.alpinus* | Sal_BW_053 | Body weight | NC_036871.1 | 32030838 | 31993282 |
| Body | *S.namaycush* | Sna_BS_092 | Body shape | NC_036854.1 | 21330886 | 21293280 |

### 6.4.5 Ecomorph divergence in body shape

For body shape, all four ecomorph pairs showed separation along PC1 (30.3%) (Figure 1D) however, the pair from Loch Dughaill diverged in a different direction, along PC2 (24.9%). Individuals with a positive PC1 score (e.g., the benthivore morphs at Awe, na Sealga, and Tay) have shallower, more elongated body shapes (Figure 6-1F). The patterns across PC2 suggest that a more positive score is associated with a deeper body (Figure A6-5).

In the PTA, we again found that all differences in the magnitude of phenotypic difference between the pairs ($\Delta L$) was significant with the exception of the Awe-na Sealga comparison (Table 6-1). When comparing angle of difference in trajectories ($\theta$), we found all angles between pairs were significant with the exception of the na Sealga-Tay comparison (Table 6-1) suggesting body shape may show some parallelism across lakes. The ecomorph term in the ANOVA model ($\eta^2_{Eco} = 0.63$) for body shape (PC1) explained the most variation which indicates some shared elements of body shape across lakes (Table A6-3).

### 6.4.6 Genomic regions associated with body shape

A BSLMM found that the proportion of phenotypic variance in body shape explained by the SNP dataset ($PVE_{Body}$) was 0.82 and the proportion of phenotypic variance explained by large ("sparse") effect loci ($PGE_{Body}$) was 0.39 and the $\rho$ (rho) value ($\rho_{Body}$) was 0.425. By analysis with LMMs, we found 180 SNPs significantly associated with body shape variation (144 SNPs mapped to 34 chromosomes, 36 SNPs mapped to unplaced scaffolds) (Figure 6-2B). We found that 10 of these SNPs were present in all four ecomorph pairs (Figure 6-3B) and broadly distributed across the genome (on 7 chromosomes).

### 6.4.7 Genomic differentiation at SNPs associated with body shape

For $F_{ST}$, we found that the body shape-associated SNPs had a higher mean value than the background at Loch Dughaill and Tay (Figure 6-4, 6-5, Table A6-4) but no notable difference at Loch Awe or na Sealga. $D_{XY}$ was significantly higher at Tay for the associated SNPs while it was significantly lower at Loch Awe. There was no significant difference in $D_{XY}$ between associated SNPs and the background at Loch Dughaill and na Sealga. The mean recombination rate in regions containing body shape SNPs did not differ from the background (Figure A6-4).

### 6.4.8 Genes and QTLs associated with body shape

Relative to known genes in the *Salvelinus* sp. genome, 89 of the body shape-associated SNPs were located within annotated genes (Table A6-5). The body shape-associated genes showed overrepresentation for genes involved in skeletal system (GO:0001501), face (GO:0060324), eye (GO:0060041, GO:0001745), and mouth development (GO:0060021) among other processes and functions (Table A6-6). We found five SNPs in close proximity to four known QTLs (Table 6-2). These QTLs were previously found to be associated with body shape morphology in lake whitefish, body weight in Arctic charr and body shape morphology in lake trout (Norman et al., 2011; Laporte et al., 2015; Smith et al., 2020).

### 6.4.9 Comparisons between head shape and body shape

We found in our PTA that comparisons in head shape showed greater mean differences in the magnitude and direction of phenotypic difference ($\Delta L_{Head}$ = 0.053 ± 0.035 s.d., mean $\theta_{Head}$ = 69.87° ± 17.54 s.d. mean) compared to body shape (mean $\Delta L_{Body}$ = 0.016 ± 0.011 s.d., mean $\theta_{Body}$ = 61.62° ± 22.63 s.d.) (Table A6-7). Both our PTA and ANOVA suggest that body shape shows more elements of shared divergence than head shape, however, both phenotypes show substantial deviations from strict parallelism across lakes.

Our association analyses indicated a significant shared genetic basis behind body shape and head shape. 50 of the SNPs found in our study appeared associated both with head and body shape morphology (212 SNPs identified: 32 SNPs associated with head shape, 130 associated with body shape, and 50 associated with head and body shape), which exceeds random expectation (hypergeometric test; P = $6.633e^{-16}$). Of these head- and body-shape shared SNPs, 38 mapped to 20 chromosomes and 12 mapped to unplaced scaffolds (Figure 6-2). These SNPs show overrepresentation for terms related to brain (GO:0030900, GO:0021575) and heart development (GO:0003007), and regulation of cell shape (GO:0008360) among other processes (Table A6-6). With a number of SNPs shared between both head and body shape, two of the QTLs we identified as near associated SNPs were near SNPs shared for both phenotypes. These QTLs related to body shape in lake trout and lake whitefish respectively (Table 6-2) (Laporte et al., 2015; Smith et al., 2020).

This shared genetic basis is reflected in the PTA, which showed that there was a positive linear relationship when comparing vector lengths for head shape and body shape across lakes. Specifically, pairs in which the ecomorphs have diverged to a similar extent in head shape have also diverged to a similar extent in body shape (adjusted $R^2$ = 0.99, P <

0.001) (Figure 6-6). A similar positive relationship was seen when comparing differences in the direction of phenotypic difference, although this was non-significant (adjusted $R^2 = 0.29$, $P = 0.154$) (Figure A6-6).

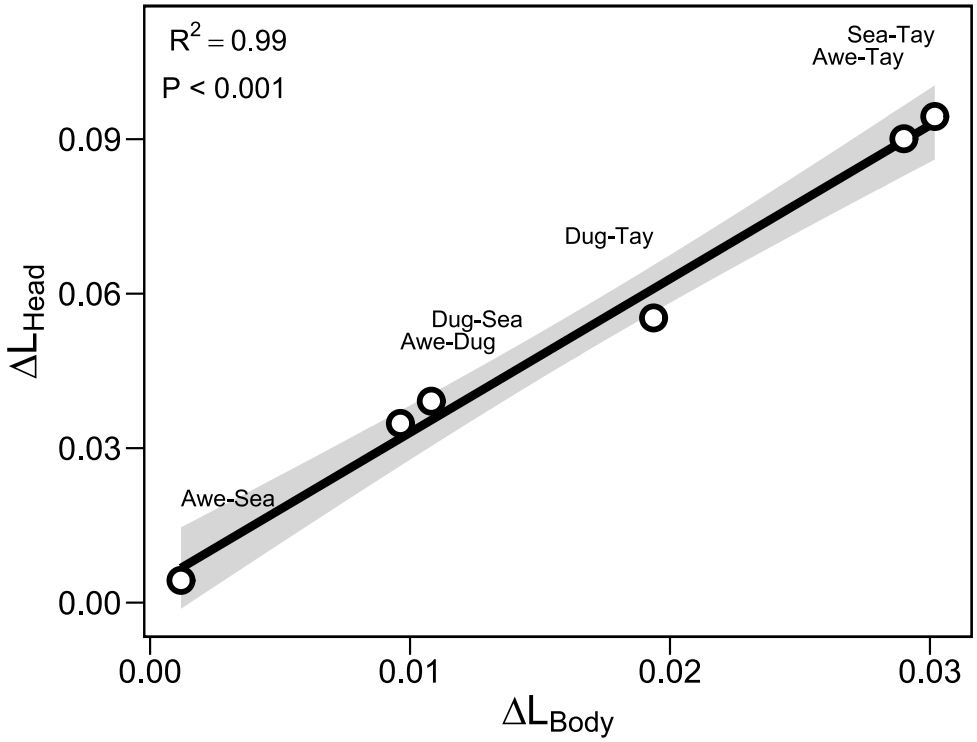


**Figure 6-6.** Correlation between the magnitude of difference in vector lengths between head shape (ΔLhead) and body shape ($\Delta L_{Body}$) in each sympatric ecomorph pair within each lake and compared pairwise between ecomorph pairs across all lakes (lakes Awe, Dughaill [Dug], Na Sealga [Sea], and Tay).

## 6.5 Discussion

Through our analyses, we identified genomic regions that underlie head shape and/or body shape morphology in ecomorph pairs of Arctic charr. Of these phenotype-associated SNPs, many (approximately one quarter: 50 of 212) were shared between head and body shape. The extent and direction of divergence in head and body shape morphology were positively correlated, suggesting a shared developmental basis for the two phenotypes. The SNPs we found associated to each phenotype were often found in genes related to morphology or anatomical development. Indeed, independently previously identified QTLs in the genomic region of body or head associated SNPs were often those related to morphology.

We found limited parallelism in shape morphology and in genomic underpinnings with many population-specific patterns in the divergence of head and body shape morphology between ecomorphs across pairs. Many of the phenotype-associated SNPs were not present in all four pairs likely due to polygenic genomic architectures and the incomplete representation of the genome in our approach. The phenotype-associated SNPs that were found to be shared were not highly diverged in all pairs and did not appear to be under the same selective pressures. Body shape in particular appears to be rather polygenic allowing for the genomic underpinnings of the phenotypes to vary across lakes.

### *6.5.1 Limited parallelism in head and body shape divergences across replicates*

While previous work on parallelism in these pairs suggested substantial phenotypic parallelism in some linear traits related to head shape (Jacobs et al., 2020) our more sensitive geomorphometric approach to describe shape suggests considerable phenotypic variation across replicated ecomorphs in multivariate space for both head shape and body shape.

The lack of phenotypic parallelism between Loch Dughaill and the other ecomorphs pairs may arise from its ecological distinctiveness compared to the others. While benthivore morphs typically occupy the shallow littoral zone of lakes as they do at the other three lakes in our study, the benthivore morph at Loch Dughaill is a 'profundal' benthivore and as such utilises a much deeper part of the lake environment, and the lake has a smaller area (Jonsson and Jonsson, 2001; Hooker et al., 2016). The divergence we see between the ecomorphs at Loch Dughaill, as indicated by the PCA, is in line with other 'profundal' charr morphs found across the Holarctic, with the benthivore ecomorph having a deeper head and body than the planktivore morph (Klemetsen, 2010; Skoglund et al., 2015). This is the inverse pattern of the more common benthivore-planktivore divergence where the benthivore morph has the shallower, longer head and body shape that is seen in the other three ecomorph pairs (Garduño-Paz et al., 2012).

Evidence for the parallel evolution of head and body shape morphology across ecomorphs has been shown in previous studies of Arctic charr (Adams et al., 2008; Kristjánsson et al., 2012; Saltykova et al., 2017). However, even disregarding the notably distinct Dughaill ecomorph pair, the other three ecomorph pairs show limited evidence of parallelism in shape morphology despite the parallel evolution of ecomorphs themselves. Evolutionary divergence and thus phenotypic trajectories are influenced by the interaction between environmental variation and adaptive genetic variation. As a result, repeated

ecomorph divergences often have very different phenotypic trajectories for key components of phenotypes, as seen in three-spine stickleback (Stuart et al., 2017). Thus, the lack of strict parallelism seen in our study is likely the result of known differences in the evolutionary histories of these pairs and differing selective pressures in their local environments, for example from differences in their ecosystems or diets (Jacobs et al., 2020). Parallelism can also be generated by other mechanisms, such as differences in gene expression, post-translational modifications, and/or alternate splicing (Filteau et al., 2013; McGirr and Martin, 2018). Indeed, differences in splicing and gene expression patterns showed parallel patterns across ecomorph pairs in a study on three of the ecomorph pairs we investigated (Jacobs and Elmer, 2021) and provide explanations of alternative adaptive paths.

Of the two traits we tested, head shape had longer phenotypic vector lengths and greater angles, suggesting that the ecomorphs are more differentiated in head shape from one another and that head shape has evolved in more distinct directions across lakes. Head shape is well recognised as an important phenotype for foraging and prey specialisation, while body shape is important for swimming behaviour and habitat complexity in these and other fishes (Webb, 1984; Adams et al., 1998; Skoglund et al., 2015). The considerable divergence in head shape between ecomorphs within lakes suggests different prey specialisations (Garduño-Paz et al., 2012; Hooker et al., 2016) while their body shapes and therefore perhaps swimming behaviours are more subtly different. The high vector angles for head shape suggest notable difference in foraging across lakes, whether that be due to the lake environments or the actual species available as prey (Garduño-Paz and Adams, 2010). For both traits, Loch Tay and Dughaill showed notably higher vector lengths than Awe and na Sealga and more evolutionary divergence, also seen in the ecomorphs' genomic divergence ($F_{ST}$ ~1% between ecomorphs in Awe and na Sealga vs 9% in Dughaill and Tay) (Table A6-8). This reflects likely what were previously inferred to be recent sympatric divergences of ecomorph pairs in Awe and na Sealga while Tay and Dughaill each have complex histories of divergence and secondary contact between colonising lineages (Jacobs et al., 2020; Fenton et al., 2023b) (Chapter 4).

### *6.5.2 Genomic underpinnings of head and body shape across lakes*

From our total of 212 SNPs that showed high associations with head and/or body shape (Figure 6-2), we found more SNPs associated with body shape than for head shape. Head shape was controlled by more large-effect loci relative to body shape and may suggest that head shape is controlled by fewer genes/pathways. In both cases these will be an underestimate of actual associations because we have reduced representation of the genome

captured. We found for that for both head and body shape only a small number of associated SNPs were diverged between ecomorphs in all four pairs (Figure 6-3). This is line with what has been suggested both in other Arctic charr studies and other salmonid species, in which genetic differentiation between ecomorphs is largely nonparallel across pairs bar at a few key genes (Salisbury et al., 2020; Salisbury and Ruzzante, 2021; Brachmann et al., 2022). Further to previous work, we found that the SNPs shared across pairs were not highly differentiated between ecomorphs in all pairs suggesting that while present, they are not critical to underlying the phenotypic differences in each pair (Figure 6-4). These results also suggest that the genomic underpinnings of each phenotype varies across the lakes, likely contributing to the phenotypic differences we see between pairs. The polygenic genomic underpinnings of both phenotypes, as indicated by the numbers of associated SNPs identified, indicate that there are multiple pathways that can achieve the same phenotypes hence the lack of high divergence for the same SNPs across all lakes (Salisbury and Ruzzante, 2021).

Loch Dughaill and in particular Loch Tay often showed notable high genomic divergence between ecomorphs for many of the associated SNPs for each trait. Additionally, the associated SNPs for both traits showed high $D_{XY}$ values compared to the background subsets at both of these lakes. While increased levels of $D_{XY}$ or $F_{ST}$ compared to genomic background can be indicative of positive selection (Bamshad and Wooding, 2003), they might also be expected for loci resisting introgression following secondary contact (Cruickshank and Hahn, 2014), as is likely the case in Loch Tay and Loch Dughaill (Jacobs et al., 2020). The associated SNPs we found are widespread across the genome (Figure 6-2) indicating these are not single linked regions of divergence as found in studies on Atlantic cod and rainbow trout (Hemmer-Hansen et al., 2013; Pearse et al., 2019) but instead are diffused and highly polygenic, similar to patterns for body shape in lake whitefish (Laporte et al., 2015).

### *6.5.3 Functional genomic regions for head and body shape*

Roughly half of the associated SNPs identified for each of head and body shape were found within or proximal to an annotated gene in the charr genome. A number of the GO terms that appeared as significantly overrepresented or enriched in our study have been identified in other studies investigating adaptive divergences or parallel evolution in various fish species. Odontogenesis (GO:0042476), sensory perception of sound (GO:0007605), blood vessel remodelling (GO:0001974), response to muscle activity (GO:0014850), ventricular trabecula myocardium morphogenesis (GO:0003222), common-partner SMAD

protein phosphorylation (GO:0007182), cellular response to ethanol (GO:0071361), and neuromuscular synaptic transmission (GO:0007274) have shown significance in other Arctic charr studies investigating ecomorph divergence (Guðbrandsson et al., 2019; Salisbury et al., 2020). The GO terms for associative learning (GO:0008306), regulation of cell shape (GO:0008360), and UDP-glucuronate biosynthetic process (GO:0006065) also appear in overrepresented groups in a study on the divergence of a sympatric lake whitefish species pair (*Corgeonus clupeaformis*) in the USA (Hebert et al., 2013). Finally, in pupfishes (*Cyprinodon* sp.), the divergent expression of a number of genes involved in cranial skeletal system development was seen between different trophic specialists with the GO term for this process (GO:1904888) significant in our study (McGirr and Martin, 2018). Differences in ossification rate have been related to adaptive morphological differentiation in other freshwater fish and the over enrichment or overexpression of genes related to formation of various bones in our study indicates a similarly important role in adaptive divergences between ecomorphs of Arctic charr (Esin et al., 2018). Indeed, previous work has noted the importance of differences in bone structure and sizes between different ecomorphs of Arctic charr (Kapralova et al., 2015; Jónsdóttir et al., 2023).

The QTL database that we have developed allows us to explore whether rapid replicated diversification of ecomorphs in different salmonid species is underlined by the use of the same functional regions as has been previously suggested for salinity tolerance (Norman et al., 2012; Jacobs et al., 2017). Our results suggest that this is true to some extent with QTLs related to body shape in lake trout and whitefish found in proximity to SNPs that we identified as being associated with phenotypic differences in Arctic charr ecomorphs. Whilst we only identified a small number of QTLs located near the associated SNPs, this is line with other work which suggests that shared basis for ecomorph divergence across species may be limited (Salisbury and Ruzzante, 2021). This QTL marker database will be a valuable resource for future salmonid research.

## 6.6 Conclusions

Our results indicate differences in head and body shape responses to ecological selection regimes across four replicate lakes. These differing responses are likely enabled through the use of largely different genetic bases across independent replicate ecomorph pairs. Specifically, we found that only a small number of SNPs were shared across all four pairs, suggesting limited genetic parallelism with these shared SNPs under varying selective pressures across lakes. We found that head and body shape morphology have a level of

shared genetic underpinnings in Arctic charr and that the genetics of these phenotypes is shared to an extent across different salmonid species. Our analyses highlight the complexity of the evolutionary genetics that underlie parallel phenotypes across replicates. Further it we demonstrate the power of using population replicates to resolve fundamental genetic and evolutionary patterns from the noise of local variation.

# Chapter 7: General Discussion

The protection of biodiversity requires conservation across a number of hierarchical levels (genes, populations, species, and ecosystems) and various temporal elements (past, present, and future biodiversity), and requires an understanding of the drivers behind the emergence of diversity. The aim of this thesis was to investigate the various elements of diversity and their drivers in a highly variable salmonid fish species the Arctic charr to help inform decisions on the protection and management of this species.

In Chapter 2, I focused on population genetic structuring and differentiation to identify important units of conservation. In Chapter 3, I investigated contemporary patterns of genetic and phenotypic diversity to understand the environmental drivers behind the emergence of diversity. Chapters 4 and 5 focused on the impacts of more historical processes on genetic diversity and differentiation by exploring the glacial history of the British Isles and the colonisation history of the species. Finally, in Chapter 6, I focused on the genes underlying key phenotypes to investigate whether similar phenotypes had similar genomic underpinnings across replicates. In this chapter, I will lay out the key findings of each chapter before synthesising across studies to draw general conclusions on the key questions of this thesis. I will also address the limitations of these studies and lay out future research to answer questions that arose as part of this work.

## 7.1 Key findings

In Chapter 2, we delineated important units of conservation for Arctic charr in the British Isles with a particular focus on Scotland. We generated a genome-wide dataset of SNPs through ddRADseq, and used mitochondrial DNA data, to analyse patterns of genetic structuring and differentiation. The ddRADseq data generated here formed the basis of the genetic analyses performed in Chapter 3 and was used alongside the mtDNA data in Chapter 5. We found that genetic differentiation between populations was high, suggesting there is minimal gene flow between many populations with most populations representing their own reproductively isolated group and forming their own distinct genetic clusters. We found little evidence in genetics to support the existing classification of previously described endemic species based on morphometrics (Kottelat and Freyhof, 2007). There were no distinct patterns seen in our phylogenetic tree or haplotype network for these putative species and patterns of genetic differentiation were within the normal range for Arctic charr suggesting that these populations should not be considered distinct species separate from Arctic charr.

In order to protect the diversity and geographical range of the species we highlighted ecotype populations, those living in sympatry or parapatry, and populations that represent an important part of the geographical range of the species to be prioritised for conservation.

In Chapter 3, we investigated the factors that underlie the emergence of genetic and phenotypic diversity and trophic specialisation across populations of Arctic charr in the British Isles. We found that increased ecological opportunity, in our case measured as greater ecosystem size, was the key driver of levels of phenotypic and genetic diversity with populations found in deeper, larger lakes being more diverse. Similarly, we found that overall ecosystem size was a key driver of trophic specialisation with larger, deeper lakes more likely to contain multiple ecotype populations.

In Chapter 4, we conducted a literature review on how glacial history has influenced patterns of genetic diversity and differentiation in the British Isles. Through our literature research, we showed how ice coverage across the British Isles underwent asynchronous patterns of advance and retreat after the global Last Glacial Maximum (*ca.* 27 ka). Ice eventually retreated northward with only the northwest highlands of Scotland remaining covered by ice through the Late Glacial Interstadial *ca.* 15 ka. Ice briefly returned during the Younger Dryas period (*ca.* 12-9 - 11.5 ka) before the start of the Holocene period when the British Isles were considered ice-free. During glacial conditions, temperate species were forced to survive in isolated safe havens known as glacial refugia where populations differentiated over time. These differentiated populations then acted as the sources of colonisation once ice coverage ended. Patterns of genetic differentiation and diversity across different species in the British Isles were influenced by the number of refugia populations involved in colonisation, the possible routes for colonisation, and the variable patterns of ice retreat. The understanding from this literature review acted as the basis for the investigation of how glacial history influenced the colonisation and subsequent genetic differentiation of populations of Arctic charr in Chapter 5.

In Chapter 5, we analysed large-scale patterns of genetic differentiation and colonisation history of Arctic charr populations in the British Isles to understand the influence of glacial history and differential ice coverage at the time of colonisation. We saw a strong pattern of Isolation-by-distance when comparing populations in the same Hydrometric Area suggesting notable gene flow only occurs between proximal populations. We found there was no relationship between genetic differentiation and geographic distance across Hydrometric Areas suggesting that genetic differentiation is predominantly driven by

genetic drift due to a lack of anadromy in the species and therefore a lack of contemporary gene flow between many populations. Results from our neighbour-joining tree suggest that the anadromous phenotype may have been maintained for some time post-colonisation with the tree splitting populations by river system flow in Scotland (East vs West). Through a mtDNA dataset covering much of the Holarctic distribution of the species we found that colonisation of the British Isles occurred through multiple sub-lineages of the Atlantic lineage. These sub-lineages varied considerably in their prevalence with many populations representing the previously identified Atlantic sub-lineage 2 (Jacobsen et al., 2022). We found that populations in regions still covered by ice during the time period of Arctic charr colonisation, the Younger Dryas/Loch Lomond Stadial (*ca.* 12.9 - 11.5 ka), and proximal to known ice-dammed lakes showed distinct patterns across our analyses. These included shared mtDNA haplotypes, evidence of genetic mixing, and low genetic differentiation across multiple Hydrometric Areas with these patterns the result of historical processes due to the lack of anadromy in the species.

In Chapter 6, we investigated the genomic underpinnings of parallel evolution of ecomorph pairs found across Scotland. We also augmented and updated an existing database of salmonid quantitative trait loci (QTLs) (Jacobs et al., 2017) and successfully mapped 669 QTL markers, covering eight different salmonid species and 17 different phenotypes, to the *Salvelinus sp.* genome. We found through our phenotypic analysis, that divergence in head and body shape morphology between benthivorous and planktivorous ecomorphs of Arctic charr were quite variable across replicates. These deviations from phenotypic parallelism were influenced by differences in local environments and the ancestral histories of the different ecomorph pairs. We found that only a small number of SNPs associated with head and/or body shape morphology were shared across all four ecomorphs pairs. Selection on these associated SNPs also varied considerably across lakes suggesting that the genomic underpinnings were variable, with different genomic pathways used to achieve similar phenotypes. However, we did identify a small number of associated SNPs that were found in QTLs related to body morphology in other salmonid species suggesting that some regions are key to these phenotypes across species.

## 7.2 Project limitations

While the studies and results of this thesis are robust, there were limitations to several analyses. The number of individuals we had per population for large-scale genetic analyses in the British Isles (Chapters 2, 3, 5) was ultimately small. This was done to allow us to cover

a large number of populations and thus examine wide-scale patterns but can limit our interpretations of each individual population. For example, the reliability of our demographic analysis of the divergence of ecotype pairs is somewhat constrained by the small sample sizes. This limited our ability to infer the exact scenario that they diverged under, although analysis gave sufficient information to infer the broad scenarios (sympatric vs secondary contact). While the vast majority of populations had sufficient numbers of individuals for our analyses, some only had one or two individuals, meaning they had to be excluded from some analyses (for example our genetic offset analysis). This in particular limited some of the work on the endemic species classifications where we only had a single population from some species, which due to small sample size, had to be excluded from certain analyses. However, the more thorough analysis of the other endemic species means we could still strongly infer the viability of the overall taxonomy. Some of the Irish putative species, e.g. *Salvelinus colii*, which we could not cover in detail were included as part of a recent study focused on Ireland, with the general conclusions largely in-line with our findings which strengthens the findings we had here (Barthelemy et al., 2023).

Our ddRADseq datasets, while cost-effective, are reduced representations of the genome. This results in a low density of SNPs and as such we can only investigate wider-scale patterns of differentiation rather than those on fine-scale regions. As such in our analyses identifying SNPs associated with phenotypes or environment (Chapter 2 and 6), we likely underestimated the number of associations and could have missed regions of association which would then not be included in downstream analyses.

Our QTL analysis (Chapter 6) and colonisation history analysis (Chapter 5) were both limited by the biases in pre-existing available data. We could only use QTLs for phenotypes that had been published with the sequence data needed to properly map. As such some key regions across species may have been missed. Many of the QTLs were from papers that focused only on one population and so it is unclear how consistent these regions would be underlying these phenotypes across populations. Thus, we had to compare available QTLs with these limitations in mind. Similarly, we could only investigate the colonisation history of Arctic charr in the British Isles (Chapter 5) using available mitochondrial ND1 sequences from the rest of the Holarctic. As a result, our coverage of the species distribution was limited to regions that had available sequences, meaning we could not cover all the major lineages of Arctic charr. Given the patterns in the haplotype network though, it is reasonable to assume that all sequences from the British Isles would still belong to the Atlantic lineage as we and others have previously suggested (Brunner et al., 2001; Moore et al., 2015). Given

the limited coverage of some regions in Europe, there may well have been more potential colonising sub-lineages than we suggested hence why we focused on the finding that were multiple sub-lineages involved rather than the specific number of them. The ND1 gene also only represents a small part of the mitochondrial genome and so we assumed that the haplotype patterns seen would be reflected across the wider mitochondrial genome. A wider more in-depth dataset examining more of the species distribution and the whole mitochondrial genome would eliminate some of the uncertainties present within this work.

## 7.3 Future work

### *7.3.1 Conservation of Arctic charr and highly diverse species*

The importance of protecting intraspecific diversity is being recognised through the increasing body of work focused on the delineation of intraspecific conservation units in numerous species (De Guia and Saitoh, 2007; Mee et al., 2015; Shaney et al., 2020; Forester et al., 2022). However, a number of different types of conservation units exist, often with multiple different definitions, which reflects the need for different appropriate units in different species but also the differing aims of different conservation efforts and that no one system will fit all cases (Moritz, 1994; Crandall et al., 2000; Coates et al., 2018; Minter et al., 2021). Determining an appropriate system for delineation of conservation units can therefore be challenging. These problems are particularly magnified in highly diverse species where determining what populations should be deemed to be important for conservation creates its own challenges.

While different types of intraspecific conservation units exist, they all fundamentally aim to protect species through the conservation of genetically distinct populations or groups of populations to ensure the maintenance of diversity (Allendorf et al., 2012). The preservation of genetic diversity is particularly important in ensuring the populations and species can adapt to changes in their environments (Hughes et al., 1997; Funk et al., 2012; Allendorf et al., 2012). Our analysis of the highly diverse Arctic charr (See Chapter 2) showed genetic differentiation between populations was particularly high suggesting that almost all populations should be independently managed in regard to their conservation. This creates a challenge for conservation as with so many distinct populations and limited resources we need a way to distinguish the populations of highest importance to prioritise for conservation.

With this in mind, we argued that conservation of Arctic charr should focus on populations whose conservation would preserve the ecological versatility and diversity of

the species while also protecting the potential for future diversity. We therefore primarily focused attention on the instances of trophic specialisation seen in Scotland and assigning them evolutionary significance (Fraser et al., 1998; Adams et al., 2008; Verspoor et al., 2010; Jacobs et al., 2020). Not only would protecting these populations conserve the ecological versatility of the species but these populations also often show high genetic and phenotypic diversity, as revealed by our analyses in Chapter 3. As such, protecting these populations would also maintain high diversity allowing the species to adapt to future changes in environment. Our analyses in Chapters 5 and 6 showed that the ancestral histories of these ecotypes are highly variable across lakes as are the genomic underpinnings of key phenotypes underlying trophic specialisation. This highlights the independent nature of these divergences and that while similar, they are different to one another and as such each instance of trophic specialisation should be protected. The replicated evolution of the same trophic specialists, with variable ancestral histories and extents of divergences between ecotypes across pairs, provides a vital system for studying the population genomics of ecological speciation and so should be protected as such.

To protect the adaptive potential of the species and its geographic range, we also argued that populations that represent the only population in their Hydrometric Area should be considered as high priority to conserve. These populations are highly isolated and represent the only population in their region, such as the Loch of Girlsta in Shetland. As such their loss would majorly impact the species distribution and the potential for unique adaptive divergence. Given the variable colonisation history of the species we revealed in Chapter 5, we would also argue that distinct and unique ancestral lineages/sub-lineages should be protected under the same basis of adaptive potential (Allendorf et al., 2012). From information from our dataset this most prominently applies to Loch Lee, which represents the sole Atlantic sub-lineage 1 population we identified in the British Isles. This population is already prioritised due to its geography, so this does not change our assessments. Future work should aim to further explore the colonisation history of the British Isles to find other populations worth prioritising. Future research should also focus on identifying other important aspects of the species that we should be prioritising for conservation. For example, a number of populations display evidence of stream spawning which may be an important phenotype to conserve (Walker, 2007).

Identification of these important units of conservation requires large-scale studies to contextualise the patterns and the use of data beyond genetics such as ecology and geographic data. We showed the limitations of not using such an approach in Chapter 2 with

the previously defined putative endemic species. These putative species were defined solely from the use of morphological data and our analyses revealed that patterns of genetic differentiation for these species was well within the wider range seen for Arctic charr suggesting these should not be considered distinct species. Work on a number of similar putative species defined in Ireland showed that the morphological differentiation used to define them as also well within the wider range seen (Barthelemy et al., 2023). This highlights how phenotypic and genetic patterns seen in a population need to be contextualized in a wider species distribution before making assigning conservation significance. Given the work of Kottelat and Freyhof also defined numerous other freshwater species across Europe, we argue for a review of these species (Kottelat and Freyhof, 2007). We believe many of these species will show similar findings to the work done here, and work done by others investigating similar defined species of European whitefish in the British Isles (Etheridge et al., 2012; Crotti et al., 2020; Crotti et al., 2021).

### *7.3.2 Mechanisms of ecotype divergence*

An increasing body of work has gone into understanding the mechanics of trophic specialisation and ecotype divergence across a range of species (Seehausen and Wagner, 2014; Tiddy et al., 2023). Our analyses in Chapter 3 highlighted the importance of ecological opportunity, represented by ecosystem size in our analyses, in providing the diversification of resources and environments needed for different ecotypes/trophic specialists to co-exist, in agreement with a range of previous studies (Mcphail, 1993; Wagner et al., 2014; Gordeeva et al., 2015; Lucek et al., 2016). While ecological opportunity was found to be the major driving force in our study, future work should go into investigating other factors that may be important that we did not cover. For example, predation, competition, nutrient availability, and prey availability have all been implicated as potentially important in underlying ecotype divergences but remain largely unexplored in many systems including Arctic charr (Vamosi, 2003; Landry and Bernatchez, 2010; Öhlund et al., 2020; Tiddy et al., 2023).

Despite the consistent importance of ecological opportunity across the divergences in our study, the histories of these ecotype pairs, as revealed by our work in Chapter 5 and previous studies (Verspoor et al., 2010; Jacobs et al., 2020), varied quite considerably. Our findings in Chapter 5 indicated that most ecotype pairs (six out of the eight tested) diverged under secondary contact scenarios involving multiple invasions of the same lake. The remaining two pairs in lochs Awe and na Sealga, in agreement with previous findings, represent sympatric divergence that occurred post-colonisation (Jacobs et al., 2020). Our findings that the majority of these divergences involve secondary contact is in line with the

general expectations of speciation which is that it most commonly occurs when a period of isolation between populations is involved (Coyne and Allen Orr, 2004). Future work should examine the other known ecotype pairs in lochs Maree and Stack (Adams et al., 2008) with our initial conclusions, in Chapter 2, suggesting these would also likely be secondary contact divergences. Given that the divergences in lochs Awe and na Sealga occurred in the face of gene flow the question remains, as is the case in many instances of sympatric divergences (Via, 2001; Niemiller et al., 2008), of how the ancestral population initially diversified to utilise different ecological niches and how this subsequently led to the build-up of genetic differences and the differentiation into two distinct populations.

Along with the varying ancestral histories of the ecotype pairs, our analyses in Chapter 6 revealed that genomic underpinnings behind the phenotypic divergence of ecotypes is highly variable across replicates. Only a small number of SNPs were shared across all pairs suggesting the use of different genomic pathways to achieve similar phenotypes across replicates. The presence of some regions being key to these phenotypes across species does suggest some level of parallelism in genomic underpinnings across populations. However, high density SNP datasets should be used to truly identify shared regions across populations or the lack thereof. Although we found that there were only a few shared associated SNPs across replicates, other types of potentially shared genomic architecture should be explored for the consistency across replicates in the future. While studies of the genomics of parallel evolution have often focused on shared outlier or phenotype-associated loci, differences in gene expression, epigenetic modifications, and alternative splicing have all been suggested as potentially important shared architecture (Filteau et al., 2013; McGirr and Martin, 2018; Jacobs and Elmer, 2021). Future studies should focus on trying to explore the genomics of parallel evolution across all these different mechanisms as should studies trying to understand patterns of local adaptation seen in a wider range of a species.

### *7.3.3 Impact of colonisation and glacial history*

While studies of patterns of genetic differentiation have often focused on geographic distance, environmental differences, and differences in colonisation history between populations as the driving forces (Orsini et al., 2013; Van Strien et al., 2015; Ruzzante et al., 2019), our analyses in Chapter 5 and our literature review in Chapter 4 demonstrated the importance of glacial history in shaping patterns of genetic diversity and differentiation with the differential timing of ice-free conditions majorly influencing the patterns we saw.

Datasets covering more of the range in the British Isles would allow us to investigate how the difference in timings from ice-free conditions across the British Isles affected patterns of colonisation on a wider scale. For example, in our dataset Loch Lee represented a different colonising sub-lineage to the rest of our populations potentially due to its earlier timing of ice-free conditions. A wider dataset may allow us to identify a more exact timing of colonisation based on where certain lineages are found and the timing of ice-free conditions for those areas. On this theme, while our analysis provides information on the broad patterns of colonisation, the exact routes of colonisation and the locations of the glacial refugia involved still remain to be found. Use of a dataset with more coverage of the species distribution, particularly in Europe, would give us more information on the number and locations of different colonising sub-lineages. We could then theorise where the glacial refugia may have been for each and determine the routes they colonised through our understanding of how ice coverage changed over time in Europe.

## 7.4 Conclusions

This thesis represents both an important advancement in our understanding of the diversity seen in Arctic charr and more broadly in the processes and factors the underlie the emergence and patterns of genetic diversity. Our results highlight how local environment, particularly ecosystem size, can influence the extent of phenotypic and genetic diversity and the diversification of populations into trophic specialists. We highlighted the ways glacial history has influenced colonisation history and subsequent patterns of genetic diversity and differentiation through the isolation of populations in glacial refugia and differential timing of ice-free conditions. Our analyses into ecotype pair demonstrated the variability of the genetics underlying similar phenotypes across replicates highlighting the significance of difference in local environment and demographic histories. Finally, we used patterns of genetic differentiation and structuring along with phenotypic and geographic data to highlight the importance of considering all available sources of information when defining important units of conservation.

# Appendices

# Appendix: Chapter 2

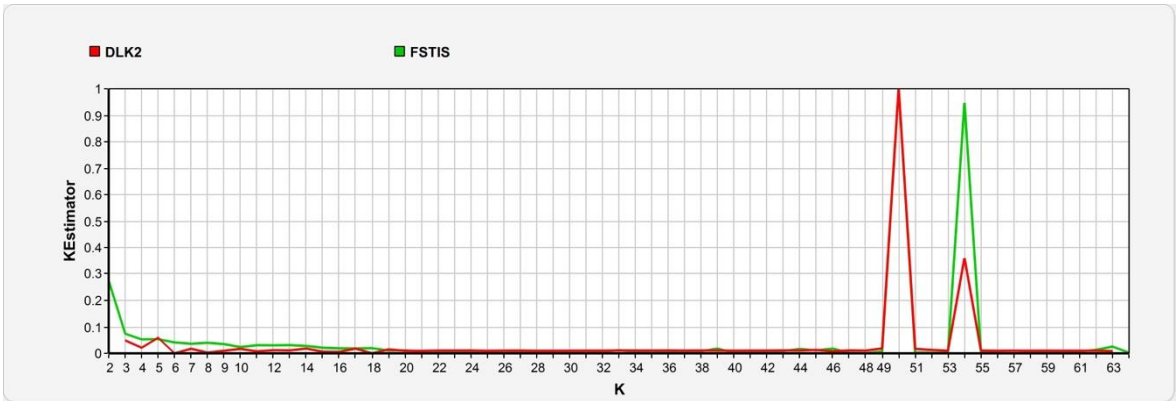


**Figure A2-1.** Admixture results showing the most likely number of genetic clusters in our dataset. Optimal number of K was determined with two statistics DLK2 (second order rate of change) and FSTIS (mean FST/FIS). A higher peak indicates a higher likelihood.

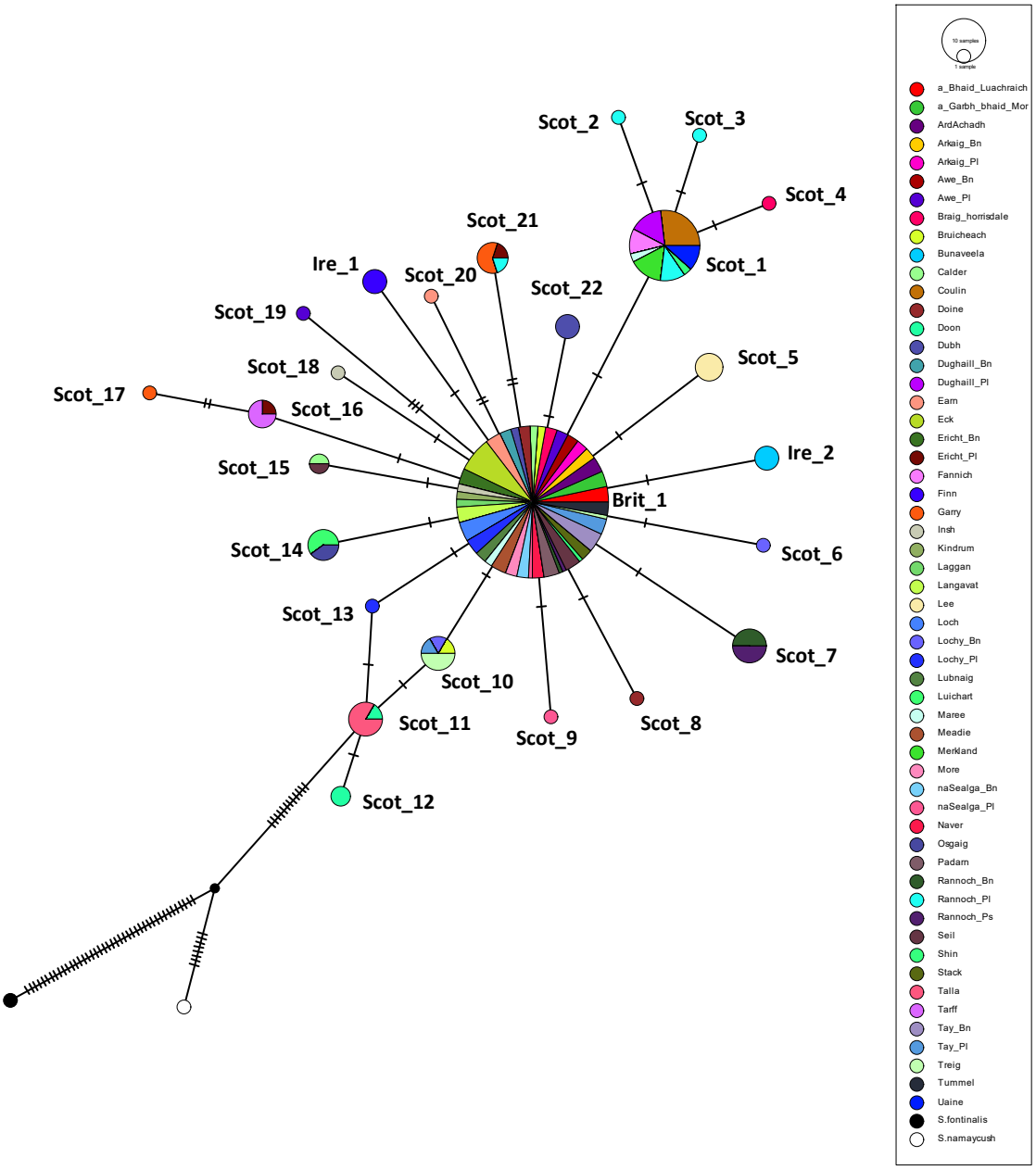


**Figure A2-2.** Haplotype network of mitochondrial ND1 gene for populations across the British Isles. Each haplotype is named by where it is found, e.g. Scot_ haplotypes are only found in Scotland. Dashes indicate the number of mutational changes between haplotypes. *Salvelinus fontinalis* and *namaycush* are used as outgroups.

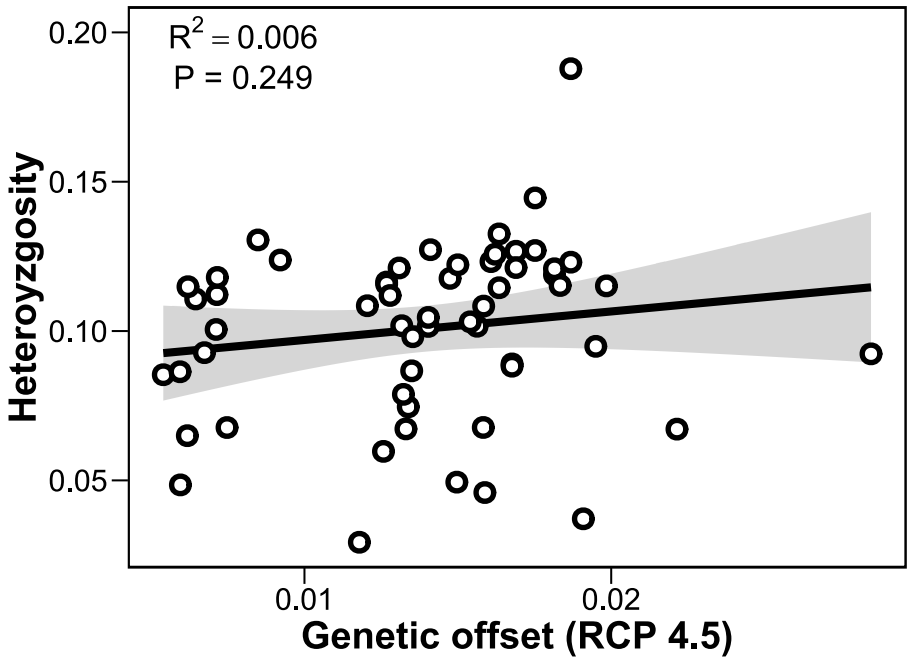


**Figure A2-3.** Correlation between genetic offset scores and observed heterozygosity for each population.

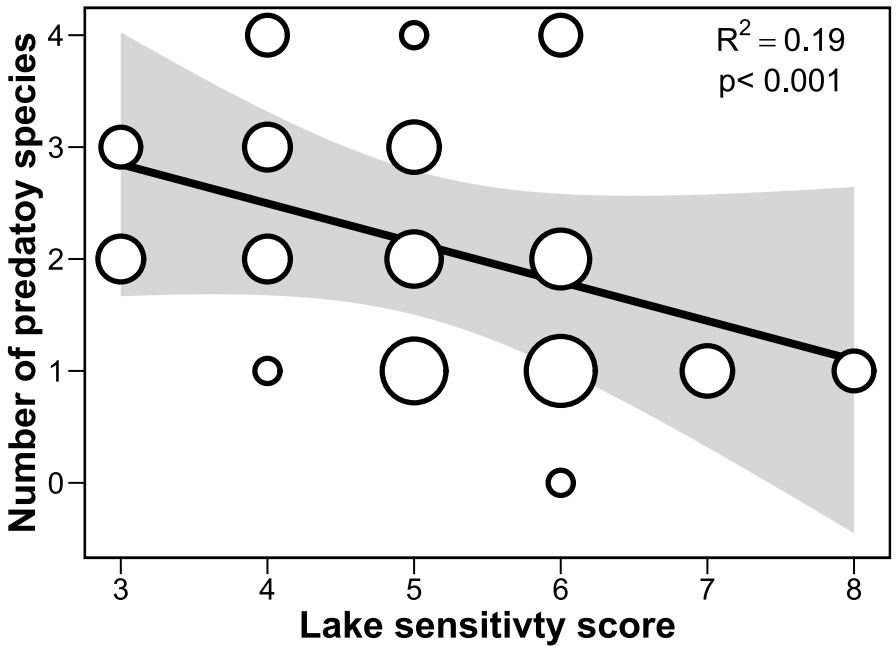


**Figure A2-4.** Correlation between lake sensitivity scores and number of predatory species. Circle size indicates the number of populations at that point.

**Table A2-1.** PCR master mix and thermal profile for generation of mitochondrial ND1.

| Component | Volume per reaction |
|---|---|
| 10x Buffer | 2.5ul |
| dNTPs (10 mM) | 0.5ul |
| $MgCl_2$ (50 mM) | 1.25ul |
| B1NDF (10mM) | 0.5ul |
| B1NDR (10mM) | 0.5ul |
| $ddH_20$ | 17.5ul |
| Taq | 0.25ul |
| DNA | 2ul |
| Total | 25ul |

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initialisation | 95ºC | 5 min | 1 |
| Amplification | 95ºC | 30 sec | 25 |
| | 55ºC | 1 min 30 sec | |
| | 72ºC | 30 sec | |
| Final extension | 60ºC | 30 min | 1 |

**Table A2-2.** List of environmental and bathymetric variables collected. Variables in bold were deemed to be uncorrelated and were the ones used for identification of adaptive SNPs. BIO2, BIO2, BIO5, BIO8, BIO9, BIO12 were the variables used for the genetic offset calculations.

| Code | Variable |
|---|---|
| Bio1 | Annual mean temperature |
| **Bio2** | **Mean diurnal range** |
| **Bio3** | **Isothermality (Bio2/Bio7)** |
| Bio4 | Temperature Seasonality |
| **Bio5** | **Maximum temperature of the warmest month** |
| Bio6 | Minimum temperature of the coldest month |
| Bio7 | Temperature annual range (Bio5-Bio6) |
| **Bio8** | **Mean temperature of wettest quarter** |
| **Bio9** | **Mean temperature of driest quarter** |
| Bio10 | Mean temperature of warmest quarter |
| Bio11 | Mean temperature of coldest quarter |
| **Bio12** | **Annual precipitation** |
| Bio13 | Precipitation of wettest month |
| Bio14 | Precipitation of driest month |
| Bio15 | Precipitation seasonality |
| Bio16 | Precipitation of wettest quarter |
| Bio17 | Precipitation of driest quarter |
| Bio18 | Precipitation of warmest quarter |
| Bio19 | Precipitation of coldest quarter |
| **Alt** | **Altitude** |
| Max.Depth | Maximum lake depth |
| **Mean.Depth** | **Mean lake depth** |
| **S.Area** | **Lake surface area** |
| **Lit.Zone** | **Percentage of the lake that is < 5 metres deep (Littoral zone)** |

**Table A2-3.** Criteria for lake sensitivity scores using measurements for maximum depth, surface area, and altitude.

| Max depth (m) | Depth sensitivity score | Surface Area (ha) | Surface Area sensitivity score | Altitude (m) | Altitude sensitivity score |
|---|---|---|---|---|---|
| <10 | 4 | | | <10 | 4 |
| 10 to 30 | 3 | <10 | 3 | 10 to 50 | 3 |
| 30 to 100 | 2 | 10 to 30 | 2 | 50 to 200 | 2 |
| >100 | 1 | >30 | 1 | >200 | 1 |

**Table A2-4.** Recorded species of relevance for each lake. Number of predatory (Brown trout (Salmo trutta), Northern pike (Esox lucius), European eel (Anguilla anguilla), European perch (Perca fluviatilis)) and competing species (Brown trout (Salmo trutta), Atlantic salmon (Salmo salar), Roach (Rutilus rutilus), Rainbow trout (Oncorhynchus mykiss), Powan (Coregonus lavaretus)) are indicated. Data was not available for outgroup populations.

| Pop ID | Species composition | Community Number | Number of competing species | Number of predatory species |
|---|---|---|---|---|
| a'Bhaid-Luachraich | Eel, Brown trout | 2 | 1 | 2 |
| a'Garbh-bhaid Mor | Brown trout | 1 | 1 | 1 |
| Ard Achadh | Brown trout | 1 | 1 | 1 |
| Arkaig | Eel, Pike, Brown trout | 3 | 1 | 3 |
| Awe | Eel, Minnow, Roach, Pike, Rainbow trout, Atlantic salmon, Brown trout, Three-spine stickleback, Perch | 9 | 4 | 4 |
| Bodlyn | NA | | | |
| Braig horrisdale | Brown trout | 1 | 1 | 1 |
| Brora | Atlantic salmon, Brown trout | 2 | 2 | 1 |
| Bruicheach | Eel, Brown trout | 2 | 1 | 2 |
| Bunaveela | NA | | | |
| Calder | Minnow, Brown trout, Three-spine stickleback | 3 | 1 | 1 |
| Conniston | NA | | | |
| Coulin | Brown trout | 1 | 1 | 1 |
| Damh | Brown trout | 1 | 1 | 1 |
| Doine | Brook Lamprey, Eel, Brown trout | 3 | 1 | 2 |
| Doon | Eel, Atlantic Salmon, Brown trout, Perch | 4 | 2 | 3 |
| Dubh | Brown trout | 1 | 1 | 1 |
| Dughaill | Minnow, Atlantic salmon, Brown trout, Three spine stickleback | 4 | 2 | 1 |
| Duntelchaig | Pike, Brown trout | 2 | 1 | 2 |
| Earn | Minnow, Rainbow trout, Atlantic salmon, Brown trout, Three-spine stickleback | 5 | 3 | 1 |
| Eck | Brook Lamprey, Eel, Minnow, Powan, Atlantic salmon, Brown trout, Three-spine stickleback | 7 | 3 | 2 |
| Ericht | Eel, Brown trout, Three-spine stickleback | 3 | 1 | 2 |
| Fannich | Pike, Brown trout | 2 | 1 | 2 |
| Finn | NA | | | |
| Garry | Brown trout | 1 | 1 | 1 |
| Insh | Eel, Pike, Atlantic salmon, Brown trout, Three-spine stickleback | 5 | 2 | 3 |

| | | | | |
|---|---|---|---|---|
| Kindrum | NA | | | |
| Laggan | Eel, Pike, Atlantic salmon, Brown trout, Three-spine stickleback | 5 | 1 | 3 |
| Langavat | Eel, Atlantic salmon, Brown trout | 3 | 2 | 2 |
| Lee | Eel, Brown trout | 2 | 1 | 2 |
| Loch | Atlantic salmon, Brown trout | 2 | 2 | 1 |
| Lochy | Pike, Brown trout | 2 | 1 | 2 |
| Lubnaig | Brook Lamprey, Eel, Atlantic salmon, Brown trout, Minnow, Perch | 6 | 2 | 3 |
| Luichart | Eel, Pike, Brown trout, Perch | 4 | 1 | 4 |
| Maree | Eel, Minnow, Atlantic salmon, Brown trout, Three-spine stickleback | 5 | 2 | 2 |
| Meadie | Atlantic salmon, Brown trout | 2 | 2 | 1 |
| Mealt | Three-spine stickleback | 1 | 1 | 1 |
| Merkland | Minnow, Atlantic salmon, Brown trout | 3 | 2 | 1 |
| More | Eel, Atlantic salmon, Brown trout | 3 | 2 | 2 |
| Morie | Atlantic salmon, Brown trout | 2 | 2 | 1 |
| Nam Brac | Brown trout | 1 | 1 | 1 |
| naSealga | Brown trout | 1 | 1 | 1 |
| Naver | Minnow, Atlantic salmon, Brown trout | 3 | 2 | 1 |
| Osgaig | Brown trout | 1 | 1 | 1 |
| Padarn | NA | | | |
| Rannoch | Pike, Atlantic salmon, Brown trout, Perch | 4 | 2 | 3 |
| Seil | Brown trout | 1 | 1 | 1 |
| Shin | Eel, Atlantic Salmon, Brown trout, Three-spine stickleback | 4 | 2 | 2 |
| Stack | Eel, Brown trout, Atlantic salmon, Three-spine stickleback, Flounder | 5 | 2 | 2 |
| Talla | Brook lamprey, Brown trout, Three-spine stickleback | 3 | 1 | 1 |
| Tarff | Minnow, Brown trout, Three-spine stickleback | 3 | 1 | 1 |
| Tay | Brook Lamprey, Eel, Minnow, Roach, Pike, Atlantic salmon, Brown trout, Perch | 8 | 3 | 4 |
| Treig | Pike, Brown trout | 2 | 1 | 2 |
| Tummel | Pike, Atlantic salmon, Brown trout, Perch | 4 | 2 | 3 |
| Uaine | Brown trout | 1 | 1 | 1 |

**Table A2-5.** Haplotypes for populations representing endemic species in our dataset.

| Locations in our dataset | Species | Haplotype(s) |
|---|---|---|
| Doine | Killinensis | Brit_1, Scot_15 |
| Shin | Mallochi | Brit_1, Scot_1 |
| Merkland | Maxillaris | Scot_1 |
| Coulin | Maxillaris, Maree (Slender) | Scot_1 |
| Maree | Maxillaris, Maree (Slender), Maree (Large eye) | Brit_1, Scot_1 |
| Lanagvat | Maxillaris | Brit_1 |
| Padarn | Perisii | Brit_1 |
| Rannoch (Pl) | Struanensis | Scot_1, Scot_2, Scot_3, Scot_21 |
| Rannoch (Bn) | Rannoch benthivore | Brit_1, Scot_7 |
| Rannoch (Ps) | Rannoch piscivore | Brit_1, Scot_7 |
| Eck | Youngeri | Brit_1 |
| Awe (Pl) | Youngeri | Brit_1, Scot_19 |
| Lee | Youngeri | Scot_5 |
| Earn | Youngeri | Brit_1, Scot_20 |
| Doon | Youngeri | Scot_11, Scot_12 |
| Talla | Youngeri | Scot_11 |
| Awe (Bn) | Awe spring spawn | Brit_1 |
| Uaine | Maree (slender) | Scot_1 |
| Finn | Colii | Ire_1 |

**Table A2-6.** Overrepresentation analysis results for GO terms from genes containing SNPs associated with climatic and/or lake environment. WeightFisher refers to the p value. GO terms are reported that have p value <0.05.

| GO ID | Term | Annotated | Significant | Expected | weight Fisher |
|---|---|---|---|---|---|
| GO:0003290 | atrial septum secundum morphogenesis | 5 | 2 | 0.09 | 0.0029 |
| GO:0006266 | DNA ligation | 6 | 2 | 0.1 | 0.0042 |
| GO:1904647 | response to rotenone | 6 | 2 | 0.1 | 0.0042 |
| GO:0048617 | embryonic foregut morphogenesis | 7 | 2 | 0.12 | 0.0059 |
| GO:0001702 | gastrulation with mouth forming second | 44 | 4 | 0.76 | 0.0068 |
| GO:0003229 | ventricular cardiac muscle tissue development | 94 | 4 | 1.62 | 0.0077 |
| GO:0035307 | positive regulation of protein dephosphorylation | 44 | 4 | 0.76 | 0.0077 |
| GO:0002028 | regulation of sodium ion transport | 97 | 4 | 1.68 | 0.0109 |
| GO:0006260 | DNA replication | 162 | 7 | 2.8 | 0.0116 |
| GO:2001045 | negative regulation of integrin-mediated signaling pathway | 10 | 2 | 0.17 | 0.0122 |
| GO:1903671 | negative regulation of sprouting angiogenesis | 14 | 3 | 0.24 | 0.0145 |
| GO:0001886 | endothelial cell morphogenesis | 11 | 2 | 0.19 | 0.0147 |
| GO:0003289 | atrial septum primum morphogenesis | 11 | 2 | 0.19 | 0.0147 |
| GO:0030514 | negative regulation of BMP signaling pathway | 57 | 4 | 0.98 | 0.0166 |
| GO:0033002 | muscle cell proliferation | 246 | 8 | 4.25 | 0.0169 |
| GO:0060840 | artery development | 1 | 1 | 0.02 | 0.0172 |
| GO:0003017 | lymph circulation | 1 | 1 | 0.02 | 0.0173 |
| GO:0006589 | octopamine biosynthetic process | 1 | 1 | 0.02 | 0.0173 |
| GO:0010215 | cellulose microfibril organization | 1 | 1 | 0.02 | 0.0173 |
| GO:0015714 | phosphoenolpyruvate transport | 1 | 1 | 0.02 | 0.0173 |
| GO:0015964 | diadenosine triphosphate catabolic process | 1 | 1 | 0.02 | 0.0173 |
| GO:0021775 | smoothened signaling pathway involved in ventral spinal cord interneuron specification | 1 | 1 | 0.02 | 0.0173 |
| GO:0021776 | smoothened signaling pathway involved in spinal cord motor neuron cell fate specification | 1 | 1 | 0.02 | 0.0173 |
| GO:0021919 | BMP signaling pathway involved in spinal cord dorsal/ventral patterning | 1 | 1 | 0.02 | 0.0173 |
| GO:0022018 | lateral ganglionic eminence cell proliferation | 1 | 1 | 0.02 | 0.0173 |
| GO:0030242 | autophagy of peroxisome | 1 | 1 | 0.02 | 0.0173 |
| GO:0033621 | nuclear-transcribed mRNA catabolic process, meiosis-specific transcripts | 1 | 1 | 0.02 | 0.0173 |
| GO:0040031 | snRNA modification | 1 | 1 | 0.02 | 0.0173 |
| GO:0043137 | DNA replication, removal of RNA primer | 1 | 1 | 0.02 | 0.0173 |
| GO:0060366 | lambdoid suture morphogenesis | 1 | 1 | 0.02 | 0.0173 |
| GO:0060367 | sagittal suture morphogenesis | 1 | 1 | 0.02 | 0.0173 |
| GO:0060464 | lung lobe formation | 1 | 1 | 0.02 | 0.0173 |

| | | | | | |
|---|---|---|---|---|---|
| GO:0060731 | positive regulation of intestinal epithelial structure maintenance | 139 | 5 | 2.4 | 0.0173 |
| GO:0060873 | anterior semicircular canal development | 1 | 1 | 0.02 | 0.0173 |
| GO:0070144 | mitochondrial arginyl-tRNA aminoacylation | 1 | 1 | 0.02 | 0.0173 |
| GO:0070413 | trehalose metabolism in response to stress | 1 | 1 | 0.02 | 0.0173 |
| GO:0070715 | sodium-dependent organic cation transport | 1 | 1 | 0.02 | 0.0173 |
| GO:1902603 | carnitine transmembrane transport | 1 | 1 | 0.02 | 0.0173 |
| GO:1902801 | regulation of siRNA-independent facultative heterochromatin formation | 1 | 1 | 0.02 | 0.0173 |
| GO:1903459 | mitotic DNA replication lagging strand elongation | 1 | 1 | 0.02 | 0.0173 |
| GO:1905204 | negative regulation of connective tissue replacement | 1 | 1 | 0.02 | 0.0173 |
| GO:1905905 | pharyngeal gland morphogenesis | 1 | 1 | 0.02 | 0.0173 |
| GO:0071374 | cellular response to parathyroid hormone stimulus | 12 | 2 | 0.21 | 0.0174 |
| GO:0045668 | negative regulation of osteoblast differentiation | 58 | 4 | 1 | 0.0176 |
| GO:0010666 | positive regulation of cardiac muscle cell apoptotic process | 13 | 2 | 0.22 | 0.0204 |
| GO:0021707 | cerebellar granule cell differentiation | 14 | 2 | 0.24 | 0.0235 |
| GO:0032930 | positive regulation of superoxide anion generation | 14 | 2 | 0.24 | 0.0235 |
| GO:0032940 | secretion by cell | 995 | 17 | 17.19 | 0.0242 |
| GO:0042149 | cellular response to glucose starvation | 38 | 3 | 0.66 | 0.0273 |
| GO:0014898 | cardiac muscle hypertrophy in response to stress | 26 | 3 | 0.45 | 0.03 |
| GO:0021819 | layer formation in cerebral cortex | 16 | 2 | 0.28 | 0.0303 |
| GO:0014066 | regulation of phosphatidylinositol 3-kinase signaling | 1942 | 46 | 33.56 | 0.0311 |
| GO:0051128 | regulation of cellular component organization | 89 | 4 | 1.54 | 0.0316 |
| GO:0003143 | embryonic heart tube morphogenesis | 106 | 4 | 1.83 | 0.0337 |
| GO:0006278 | RNA-templated DNA biosynthetic process | 45 | 3 | 0.78 | 0.0339 |
| GO:0006541 | glutamine metabolic process | 17 | 2 | 0.29 | 0.034 |
| GO:0023052 | signaling | 30 | 2 | 0.52 | 0.0341 |
| GO:0046639 | negative regulation of alpha-beta T cell differentiation | 21 | 2 | 0.36 | 0.0341 |
| GO:0061550 | cranial ganglion development | 2 | 1 | 0.03 | 0.0341 |
| GO:0002860 | positive regulation of natural killer cell mediated cytotoxicity directed against tumor cell target | 2 | 1 | 0.03 | 0.0342 |
| GO:0006526 | arginine biosynthetic process | 2 | 1 | 0.03 | 0.0342 |
| GO:0010216 | maintenance of DNA methylation | 2 | 1 | 0.03 | 0.0342 |
| GO:0010223 | secondary shoot formation | 2 | 1 | 0.03 | 0.0342 |
| GO:0018057 | peptidyl-lysine oxidation | 2 | 1 | 0.03 | 0.0342 |
| GO:0020028 | endocytic hemoglobin import into cell | 2 | 1 | 0.03 | 0.0342 |
| GO:0031508 | pericentric heterochromatin formation | 2 | 1 | 0.03 | 0.0342 |
| GO:0034983 | peptidyl-lysine deacetylation | 2 | 1 | 0.03 | 0.0342 |
| GO:0042271 | susceptibility to natural killer cell mediated cytotoxicity | 2 | 1 | 0.03 | 0.0342 |

| | | | | | |
|---|---|---|---|---|---|
| GO:0044205 | 'de novo' UMP biosynthetic process | 2 | 1 | 0.03 | 0.0342 |
| GO:0044210 | 'de novo' CTP biosynthetic process | 2 | 1 | 0.03 | 0.0342 |
| GO:0044691 | tooth eruption | 2 | 1 | 0.03 | 0.0342 |
| GO:0045004 | DNA replication proofreading | 2 | 1 | 0.03 | 0.0342 |
| GO:0045315 | positive regulation of compound eye photoreceptor development | 2 | 1 | 0.03 | 0.0342 |
| GO:0045870 | positive regulation of single stranded viral RNA replication via double stranded DNA intermediate | 2 | 1 | 0.03 | 0.0342 |
| GO:0045984 | negative regulation of pyrimidine nucleobase metabolic process | 2 | 1 | 0.03 | 0.0342 |
| GO:0048653 | anther development | 2 | 1 | 0.03 | 0.0342 |
| GO:0060370 | susceptibility to T cell mediated cytotoxicity | 2 | 1 | 0.03 | 0.0342 |
| GO:0070478 | nuclear-transcribed mRNA catabolic process, 3'-5' exonucleolytic nonsense-mediated decay | 2 | 1 | 0.03 | 0.0342 |
| GO:0070481 | nuclear-transcribed mRNA catabolic process, non-stop decay | 2 | 1 | 0.03 | 0.0342 |
| GO:0090673 | endothelial cell-matrix adhesion | 2 | 1 | 0.03 | 0.0342 |
| GO:1900163 | positive regulation of phospholipid scramblase activity | 2 | 1 | 0.03 | 0.0342 |
| GO:1902616 | acyl carnitine transmembrane transport | 2 | 1 | 0.03 | 0.0342 |
| GO:1904447 | folate import across plasma membrane | 2 | 1 | 0.03 | 0.0342 |
| GO:1905463 | negative regulation of DNA duplex unwinding | 2 | 1 | 0.03 | 0.0342 |
| GO:1990575 | mitochondrial L-ornithine transmembrane transport | 2 | 1 | 0.03 | 0.0342 |
| GO:2000570 | positive regulation of T-helper 2 cell activation | 2 | 1 | 0.03 | 0.0342 |
| GO:2000753 | positive regulation of glucosylceramide catabolic process | 2 | 1 | 0.03 | 0.0342 |
| GO:2000755 | positive regulation of sphingomyelin catabolic process | 3943 | 73 | 68.14 | 0.0342 |
| GO:0120036 | plasma membrane bounded cell projection organization | 1551 | 40 | 26.8 | 0.0375 |
| GO:0003281 | ventricular septum development | 102 | 4 | 1.76 | 0.0376 |
| GO:0048672 | positive regulation of collateral sprouting | 18 | 2 | 0.31 | 0.0378 |
| GO:0060045 | positive regulation of cardiac muscle cell proliferation | 18 | 2 | 0.31 | 0.0378 |
| GO:0060134 | prepulse inhibition | 18 | 2 | 0.31 | 0.0378 |
| GO:0070932 | histone H3 deacetylation | 18 | 2 | 0.31 | 0.0378 |
| GO:0071803 | positive regulation of podosome assembly | 18 | 2 | 0.31 | 0.0378 |
| GO:0071277 | cellular response to calcium ion | 74 | 4 | 1.28 | 0.0386 |
| GO:0045652 | regulation of megakaryocyte differentiation | 44 | 3 | 0.76 | 0.0399 |
| GO:0007635 | chemosensory behavior | 51 | 4 | 0.88 | 0.041 |
| GO:0001764 | neuron migration | 214 | 8 | 3.7 | 0.0416 |
| GO:0030324 | lung development | 222 | 7 | 3.84 | 0.0442 |
| GO:0051092 | positive regulation of NF-kappaB transcription factor activity | 78 | 4 | 1.35 | 0.0455 |
| GO:0097421 | liver regeneration | 47 | 3 | 0.81 | 0.047 |
| GO:0001501 | skeletal system development | 585 | 12 | 10.11 | 0.0499 |

**Table A2-7.** List of QTLs from our QTL database that contained adaptive SNPs or were found proximal to them. Species of origin and the phenotype are list as well as the QTL start and end in its respective chromosome in the *Salvelinus* sp. genome.

| QTL marker | Species | Trait | Chromosome | QTL start | QTL end |
|---|---|---|---|---|---|
| Cocl_BS_096 | *Coregonus clupeaformis* | Body Shape | NC_036838.1 | 46685809 | 46685809 |
| Omy_SM_031 | *Oncorhynchus mykiss* | Sexual maturation | NC_036838.1 | 50898901 | 50898901 |
| Cocl_GR_004 | *Coregonus clupeaformis* | Directional change | NC_036839.1 | 10561411 | 10561411 |
| Cocl_BS_021 | *Coregonus clupeaformis* | Body Shape | NC_036842.1 | 34066589 | 34066589 |
| Sal_CF_011 | *Salvelinus alpinus* | Condition Factor | NC_036842.1 | 5751355 | 88198010 |
| Cocl_DirCh_008 | *Coregonus clupeaformis* | Directional change | NC_036842.1 | 11649948 | 11649948 |
| Omy_DR_003 | *Oncorhynchus mykiss* | Disease Resistance | NC_036842.1 | 5127503 | 5127503 |
| Cocl_BS_053 | *Coregonus clupeaformis* | Body Shape | NC_036843.1 | 22695609 | 22695609 |
| Omy_SM_014 | *Oncorhynchus mykiss* | Sexual maturation | NC_036845.1 | 29414044 | 29414200 |
| Oki_BL_009 | *Oncorhynchus kisutch* | Body Length | NC_036847.1 | 24097769 | 32996863 |
| Omy_DR_012 | *Oncorhynchus mykiss* | Disease Resistance | NC_036848.1 | 13363254 | 13363254 |
| Sal_UTT_004 | *Salvelinus alpinus* | Upper temperature tolerance | NC_036848.1 | 10119484 | 21971402 |
| Omy_BW_016 | *Oncorhynchus mykiss* | Body Weight | NC_036853.1 | 10667932 | 17403275 |
| Omy_CF_009 | *Oncorhynchus mykiss* | Condition Factor | NC_036853.1 | 10667932 | 17403275 |
| Oki_SP_005 | *Oncorhynchus kisutch* | Spawning date | NC_036853.1 | 10667932 | 17403275 |
| Sna_BS_006 | *Salvelinus namaycush* | Body Length | NC_036854.1 | 3663455 | 3663455 |
| Sna_BL_133 | *Salvelinus namaycush* | Body Length | NC_036855.1 | 50063281 | 50063281 |
| Sna_BL_043 | *Salvelinus namaycush* | Body Length | NC_036855.1 | 32096909 | 32096909 |
| Omy_OSMO_005 | *Oncorhynchus mykiss* | Sodium/Chloride concentration 2 | NC_036855.1 | 17674273 | 17674527 |
| Omy_BW_031 | *Oncorhynchus mykiss* | Body Weight | NC_036856.1 | 18229605 | 18229605 |
| Omy_SP_018 | *Oncorhynchus mykiss* | Spawning date | NC_036856.1 | 42026512 | 42027014 |
| Sal_CF_095 | *Salvelinus alpinus* | Condition Factor | NC_036861.1 | 875380 | 875380 |
| Sal_HS_019 | *Salvelinus alpinus* | Head Shape | NC_036861.1 | 875380 | 875380 |
| Oki_SP_011 | *Oncorhynchus kisutch* | Spawning date | NC_036862.1 | 27196736 | 35780323 |
| Sna_BL_221 | *Salvelinus namaycush* | Body Length | NC_036863.1 | 42690182 | 42690182 |

| | | | | | |
|---|---|---|---|---|---|
| Sfo_GR_005 | *Salvelinus alpinus* | Growth Rate | NC_036865.1 | 12786571 | 12786571 |
| Sal_BW_072 | *Salvelinus alpinus* | Body Weight | NC_036866.1 | 14035337 | 44737836 |
| Cocl_DirCh_001 | *Coregonus clupeaformis* | Directional change | NC_036869.1 | 19718333 | 19718333 |
| Cocl_DirCh_005 | *Coregonus clupeaformis* | Directional change | NC_036870.1 | 5780669 | 5780669 |
| Oki_SM_002 | *Oncorhynchus kisutch* | Sexual maturation | NC_036871.1 | 12958985 | 12958985 |
| Sna_BS_058 | *Salvelinus namaycush* | Body Length | NW_019942665.1 | 366314 | 366314 |
| Cocl_BS_169 | *Coregonus clupeaformis* | Body Shape | NW_019942665.1 | 364528 | 364528 |
| Cocl_BS_167 | *Coregonus clupeaformis* | Body Shape | NW_019943451.1 | 91315 | 91315 |
| Cocl_BS_101 | *Coregonus clupeaformis* | Body Shape | NW_019943516.1 | 79438 | 79438 |
| Sal_BL_005 | *Salvelinus alpinus* | Body length | NW_019945303.1 | 61402 | 61582 |

**Table A2-8.** Correlation matrix between the vulnerability metrics.

| | **Genetic offset** | **Sensitivity score** | **Heterozygosity** | **Predatory species** |
|---|---|---|---|---|
| Genetic offset | 1 | -0.4019439 | 0.25560742 | 0.295591 |
| Sensitivity score | -0.4019439 | 1 | -0.4532981 | -0.5092968 |
| Heterozygosity | 0.25560742 | -0.4532981 | 1 | 0.52977594 |
| Predatory species | 0.295591 | -0.5092968 | 0.52977594 | 1 |

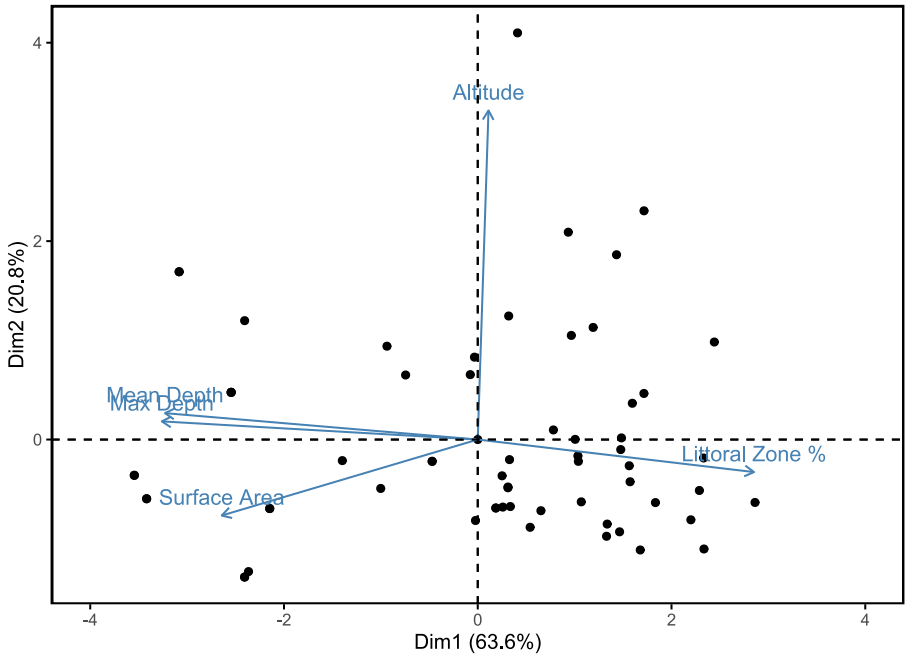


**Figure A3-1.** PCA loadings for bathymetric variables. PC1 scores were flipped for the use as our proxy for ecosystem size.

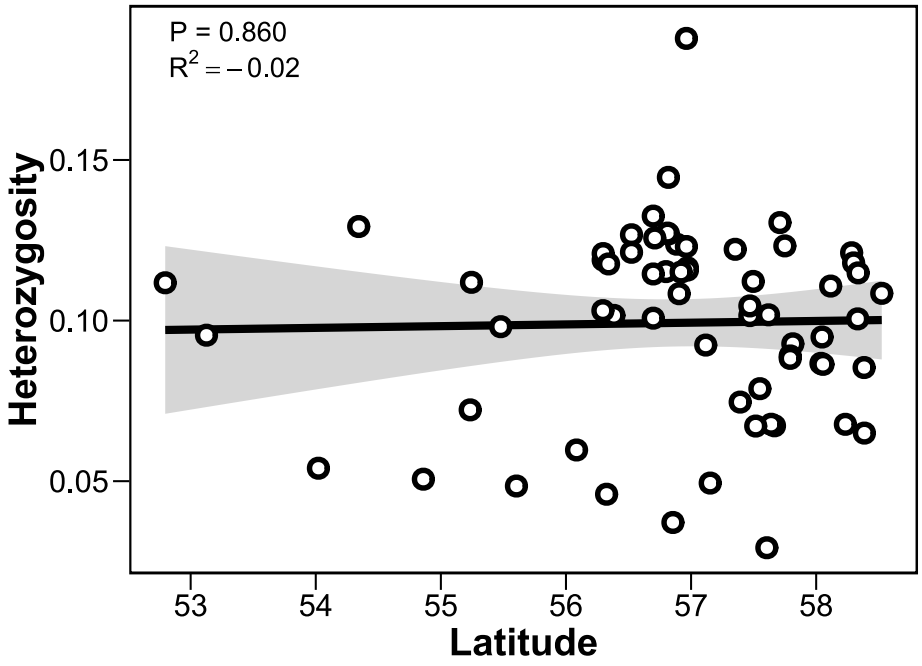


**Figure A3-2.** Correlation between observed heterozygosity and latitude.

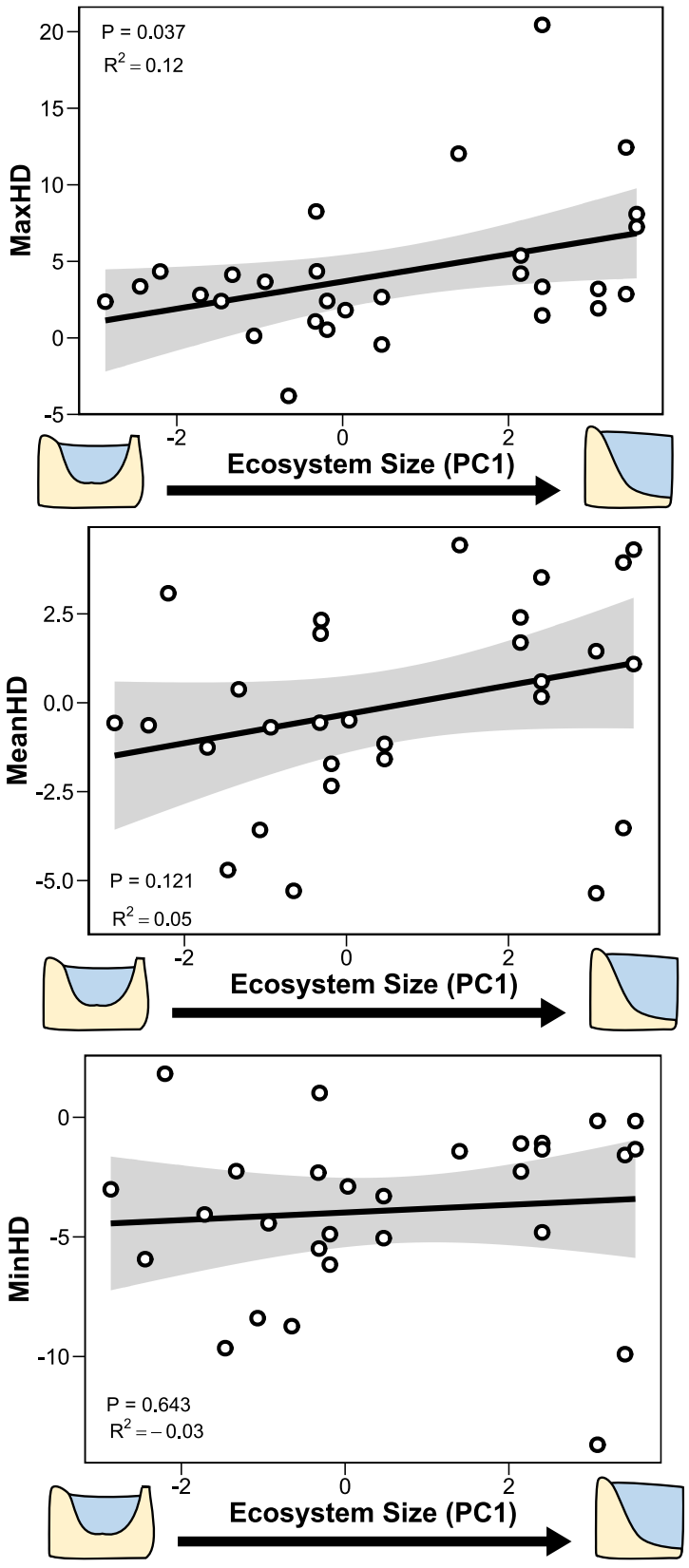


**Figure A3-3.** Relationship between ecosystem size and MaxHD, MeanHD, and MinHD respectively. Ecosystem size increases from left to right along the x-axis.

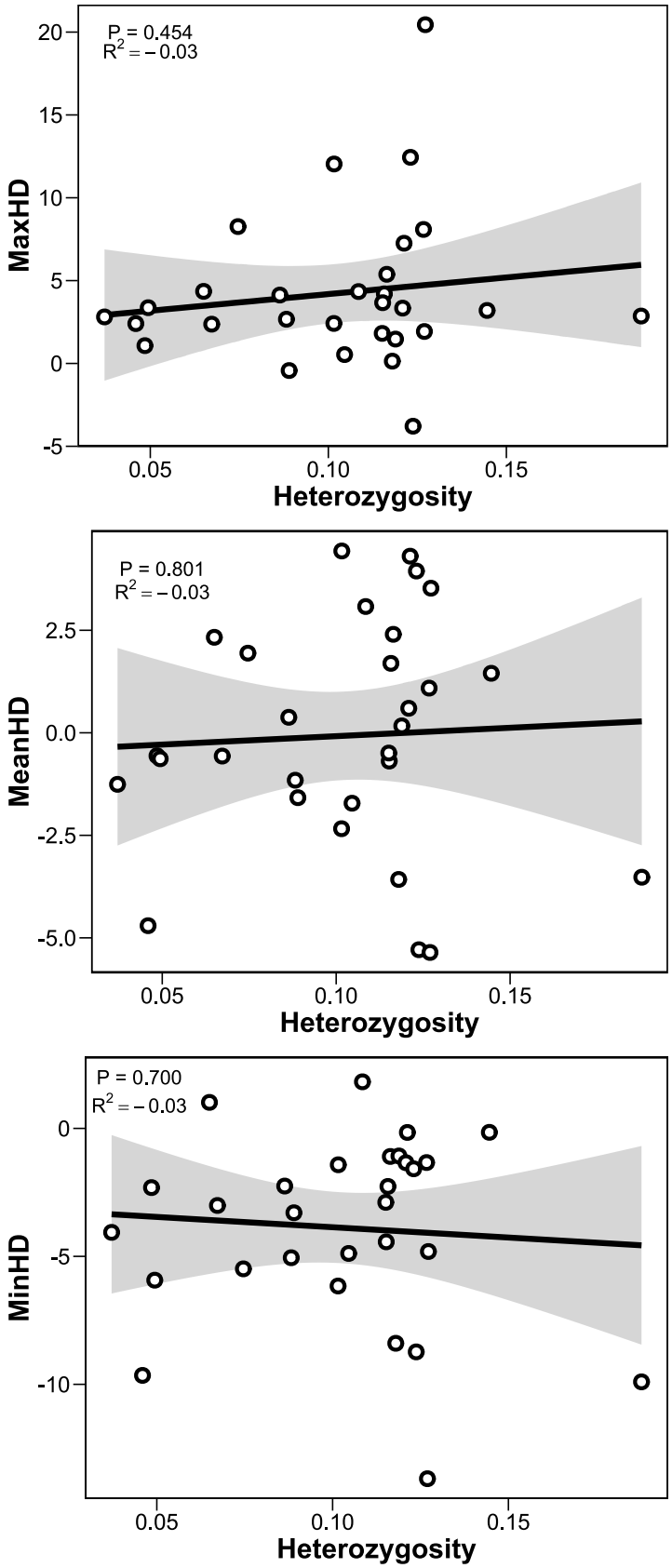


**Figure A3-4.** Relationship between observed heterozygosity and MaxHD, MeanHD, and MinHD.

**Table A3-1.** A list of lochs in Scotland that have been included in phenotypic and genetic studies including this dataset. One of the phenotypic and/or genetic studies including that population is provided as a reference. Asterisk indicate lakes studied in this work phenotypically and/or genetically but not in any previous study.

| Lake | Hydrometric Area | Phenotypic Paper | Genetic Paper |
|---|---|---|---|
| Brora | HA2 - Helmsdale | | * |
| Shin | HA3 - Shin | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| a' Ghriama | HA3 - Shin | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Merkland | HA3 - Shin | (Bush and Adams, 2007) | (Wilson et al., 2004) |
| Morie | HA4 - Conon | | * |
| Luichart | HA4 - Conon | | * |
| Fannich | HA4 - Conon | | * |
| Achilty | HA4 - Conon | (Alexander and Adams, 2004) | |
| Bruicheach | HA5 - Beauly | * | * |
| Ness | HA6 - Ness | (Winfield et al., 2002) | |
| Garry (Ness) | HA6 - Ness | (Walker, 2007) | |
| Tarff | HA6 - Ness | * | * |
| Duntelchaig | HA7 - Findhorn | | * |
| Moy | HA7 - Findhorn | (Alexander and Adams, 2004) | |
| Insh | HA8 - Spey | (Walker, 2007) | * |
| Builg | HA8 - Spey | (Alexander and Adams, 2000) | |
| Lee | HA13 - Esk | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Tay | HA15 - Tay | (Jacobs et al., 2020) | (Jacobs et al., 2020) |
| Tummel | HA15 - Tay | | * |
| Rannoch | HA15 - Tay | (Adams et al., 1998) | (Jacobs et al., 2020) |
| Loch | HA15 - Tay | * | * |
| Garry (Tay) | HA15 - Tay | * | * |
| Ericht | HA15 - Tay | (Fraser et al., 1998) | * |
| Earn | HA16 - Earn | (Bush and Adams, 2007) | (Wilson et al., 2004) |

| | | | |
|---|---|---|---|
| Lubnaig | HA18 - Forth | (Adams et al., 2006) | (Jacobs et al., 2020) |
| Voil | HA18 - Forth | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Doine | HA18 - Forth | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Talla | HA21 - Tweed | (Recknagel et al., 2017) | * |
| Doon | HA82 - Doon | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Eck | HA86 - Firth of Clyde | (Recknagel et al., 2017) | (Jacobs et al., 2020) |
| Seil | HA88 - Add | * | * |
| Awe | HA89 - Awe & Etive | (Jacobs et al., 2020) | (Jacobs et al., 2020) |
| Treig | HA91 - Lochy | * | * |
| Laggan | HA91 - Lochy | * | * |
| Lochy | HA91 - Lochy | * | * |
| Arkaig | HA91 - Lochy | * | * |
| Dubh | HA92 - Sheil | * | * |
| Dughaill | HA93 - Alsh | (Jacobs et al., 2020) | (Jacobs et al., 2020) |
| Damh | HA94 - Maree | | * |
| Braig Horrisdale | HA94 - Maree | * | * |
| Maree | HA94 - Maree | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Clair | HA94 - Maree | (Adams et al., 2006) | (Wilson et al., 2004) |
| Coulin | HA94 - Maree | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Uaine | HA94 - Maree | (Recknagel et al., 2017) | (Jacobs et al., 2020) |
| a' Bhaid-luachraich | HA94 - Maree | (Recknagel et al., 2017) | * |
| Bharranch | HA94 - Maree | (Walker, 2007) | |
| na Sealga | HA95 - Laxford | (Jacobs et al., 2020) | (Jacobs et al., 2020) |
| Osgaig | HA95 - Laxford | | * |
| a' Garbh-bhaid Mor | HA95 - Laxford | * | * |
| nam Brac | HA95 - Laxford | | * |
| Stack | HA95 - Laxford | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| More | HA95 - Laxford | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| Naver | HA96 - Naver | * | * |
| Meadie | HA96 - Naver | | * |

| | | | |
|---|---|---|---|
| Calder | HA97 - Thurso | * | * |
| Ard Achadh | HA104 - Kintyre | * | * |
| Mealt | HA105 - Inner Hebrides | (Recknagel et al., 2017) | (Wilson et al., 2004) |
| na Craobhaig | HA106 - Outer Hebrides | (Alexander and Adams, 2000) | |
| Langavat | HA106 - Outer Hebrides | (Recknagel et al., 2017) | * |
| Girlsta | HA108 - Shetland | (Recknagel et al., 2017) | |

**Table A3-2.** Dataset numbers of phenotypic (Nphenotypic) and genetic (NddRADseq) per population. Values of observed heterozygosity and variation in HD (VarHD) are provided.

| **Pop ID** | **NddRADseq** | **Heterozygosity** | **Nphenotypic** | **VarHD** |
|---|---|---|---|---|
| a'Bhaid-Luachraich | 10 | 0.093 | | |
| a'Garbh-bhaid Mor | 10 | 0.065 | 10 | 1.1 |
| Ard Achadh | 10 | 0.048 | 9 | 0.9 |
| Arkaig (Benthivore) | 4 | 0.116 | 5 | 8.7 |
| Arkaig (Planktivore) | 5 | 0.116 | 5 | 7.4 |
| Awe (Benthivore) | 5 | 0.119 | 5 | 0.8 |
| Awe (Planktivore) | 5 | 0.121 | 5 | 2.9 |
| Bodlyn | 2 | 0.112 | | |
| Braig horrisdale | 9 | 0.067 | 10 | 3.2 |
| Brora | 4 | 0.095 | | |
| Bruicheach | 8 | 0.075 | 10 | 19.2 |
| Bunaveela | 10 | 0.054 | | |
| Calder | 3 | 0.108 | 2 | 3.2 |
| Conniston | 1 | 0.129 | | |
| Coulin | 9 | 0.079 | | |
| Damh | 2 | 0.112 | | |
| Doine | 4 | 0.118 | | |
| Doon | 6 | 0.112 | | |
| Dubh | 5 | 0.124 | 4 | 5.4 |
| Dughaill (Benthivore) | 5 | 0.102 | 4 | 13.7 |
| Dughaill (Planktivore) | 5 | 0.105 | 5 | 4.7 |
| Duntelchaig | 4 | 0.122 | | |
| Earn | 9 | 0.102 | 9 | 23.8 |
| Eck | 10 | 0.060 | | |
| Ericht (Benthivore) | 4 | 0.127 | 5 | 40.6 |
| Ericht (Planktivore) | 5 | 0.145 | 5 | 2.1 |
| Fannich | 9 | 0.068 | | |
| Finn | 1 | 0.051 | | |
| Garry | 8 | 0.115 | 10 | 4.7 |

| | | | | |
|---|---|---|---|---|
| Insh | 10 | 0.092 | | |
| Kindrum | 7 | 0.072 | | |
| Laggan | 6 | 0.115 | 10 | 1.7 |
| Langavat | 10 | 0.087 | | |
| Lee | 9 | 0.108 | | |
| Loch | 10 | 0.037 | 10 | 4.2 |
| Lochy (Benthivore) | 2 | 0.188 | 2 | 81.4 |
| Lochy (Planktivore) | 5 | 0.123 | 5 | 31.5 |
| Lubnaig | 10 | 0.103 | | |
| Luichart | 10 | 0.102 | | |
| Maree | 10 | 0.131 | | |
| Meadie | 5 | 0.101 | | |
| Mealt | 2 | 0.029 | | |
| Merkland | 10 | 0.068 | | |
| More | 4 | 0.121 | | |
| Morie | 3 | 0.123 | | |
| Nam Brac | 3 | 0.085 | | |
| naSealga (Benthivore) | 5 | 0.089 | 5 | 1.5 |
| naSealga (Planktivore) | 5 | 0.088 | 5 | 7.7 |
| Naver | 9 | 0.118 | 8 | 10.9 |
| Osgaig | 9 | 0.086 | 10 | 5.3 |
| Padarn | 10 | 0.095 | | |
| Rannoch (Benthivore) | 1 | 0.101 | | |
| Rannoch (Piscivore) | 6 | 0.115 | | |
| Rannoch (Planktivore) | 4 | 0.133 | | |
| Seil | 9 | 0.046 | 9 | 16.2 |
| Shin | 7 | 0.111 | | |
| Stack | 5 | 0.115 | | |
| Talla | 8 | 0.098 | | |
| Tarff | 8 | 0.049 | 10 | 6.9 |
| Tay (Benthivore) | 5 | 0.127 | 5 | 15.6 |
| Tay (Planktivore) | 5 | 0.121 | 5 | 8.6 |
| Treig | 7 | 0.127 | 8 | 63.7 |
| Tummel | 9 | 0.126 | | |
| Uaine | 10 | 0.067 | | |

**Table A3-3.** Results of linear models for individuals environmental and lake variables as predictors for levels of genetic diversity.

| Predictor | p value | adjusted R squared |
|---|---|---|
| Latitude | 0.8602 | -0.01562 |
| Longitude | 0.000443 | 0.1686 |
| Max Depth | 6.03E-07 | 0.3714 |
| Ecosystem Size | 2.02E-06 | 0.3176 |
| Altitude | 0.4013 | -0.005021 |
| Mean Depth | 7.89E-07 | 0.3651 |
| Surface Area | 3.83E-05 | 0.2462 |
| Littoral Zone | 9.38E-05 | 0.2595 |
| Bio1 Annual Mean Temperature | 0.1497 | 0.01915 |
| Bio2 Mean Diurnal Range | 0.01115 | 0.09208 |
| Bio3 Isothermality (BIO2/BIO7) | 0.1007 | 0.02985 |
| Bio4 Temperature Seasonality | 0.007072 | 0.1051 |
| Bio5 Maximum Temperature of Warmest Month | 0.1034 | 0.02911 |
| Bio6 Minimum Temperature of Coldest Month | 0.02566 | 0.06828 |
| Bio7 Temperature Annual Range (BIO5-BIO6) | 0.006692 | 0.1066 |
| Bio8 Mean Temperature of Wettest Quarter | 0.01428 | 0.08502 |
| Bio9 Mean Temperature of Driest Quarter | 0.5765 | -0.01194 |
| Bio10 Mean Temperature of Warmest Quarter | 0.6615 | -0.0141 |
| Bio11 Mean Temperature of Coldest Quarter | 0.03031 | 0.06353 |
| Bio12 Annual Precipitation | 0.2008 | 0.0115 |
| Bio13 Precipitation of Wettest Month | 0.1168 | 0.02581 |
| Bio14 Precipitation of Driest Month | 0.2064 | 0.0108 |
| Bio15 Precipitation Seasonality | 0.2416 | 0.006852 |
| Bio16 Precipitation of Wettest Quarter | 0.1701 | 0.01579 |
| Bio17 Precipitation of Driest Quarter | 0.2188 | 0.009334 |
| Bio18 Precipitation of Warmest Quarter | 0.4517 | -0.007396 |
| Bio19 Precipitation of Coldest Quarter | 0.09936 | 0.03021 |

**Table A3-4.** Pairwise comparisons of meanHD between ecotypes in the same lake.

| Lake | HD mean (benthivore) | HD mean (planktivore) | T-test statistic | d.f | T-test p value |
|---|---|---|---|---|---|
| Arkaig | 2.40 | 1.69 | 0.39 | 7.95 | 0.7 |
| Awe | 0.17 | 0.60 | -0.49 | 6.14 | 0.64 |
| Dughaill | -2.34 | -1.72 | -0.30 | 4.62 | 0.78 |
| Ericht | -5.36 | 1.45 | -2.33 | 4.42 | 0.07 |
| Lochy | -3.52 | 3.94 | -1.09 | 1.33 | 0.44 |
| na Sealga | -1.58 | -1.16 | -0.31 | 5.53 | 0.77 |
| Tay | 1.09 | 4.31 | -1.46 | 7.38 | 0.18 |

# Appendix: Chapter 5

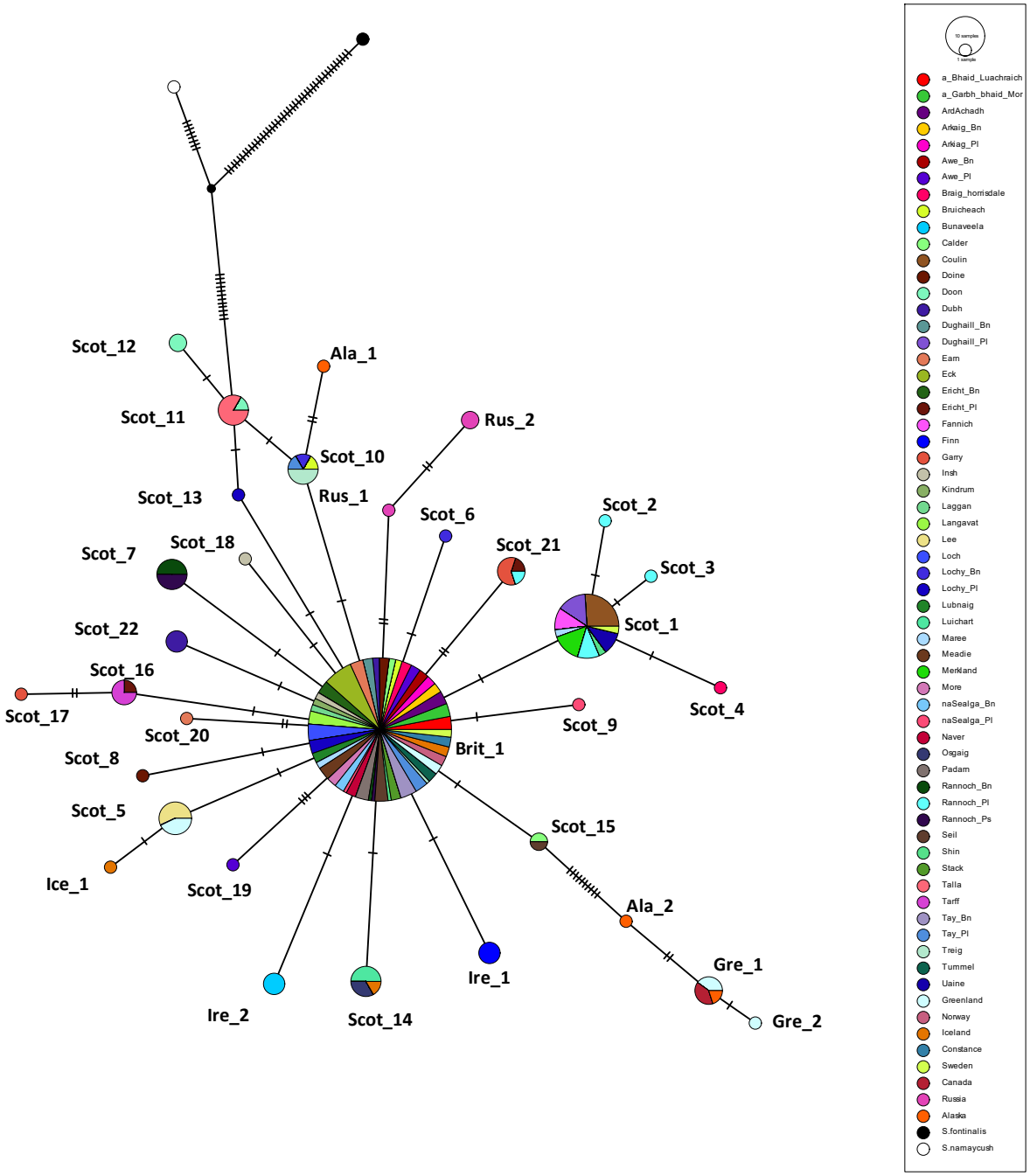


**Figure A5-1.** Haplotype network of ND1 gene for our Holarctic dataset. Haplotypes are named by where they are predominantly found (e.g. Scot_ haplotypes are found in Scotland). All British Isles haplotypes are named the same as in A2-2. *Salvelinus fontinalis* and *namaycush* were used as outgroups. Dashes indicate the number of mutational changes between haplotypes.

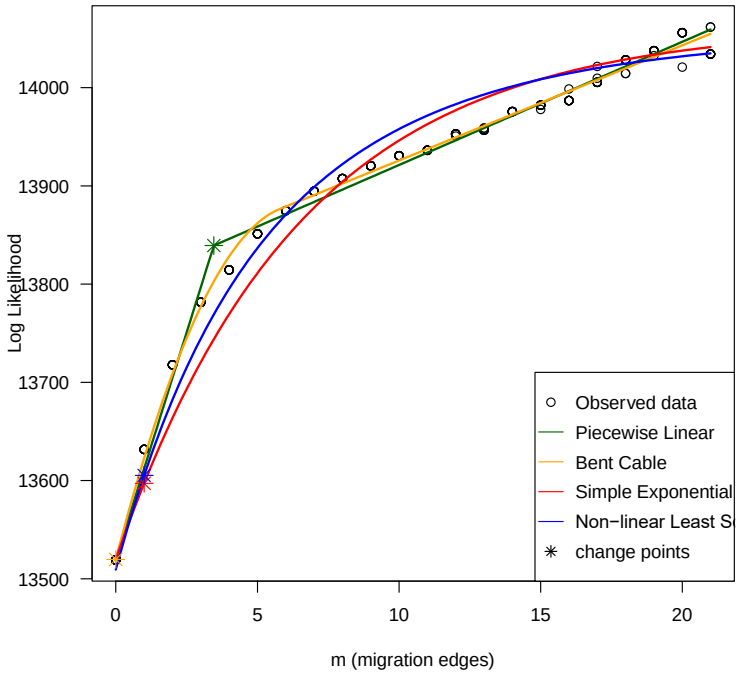


**Figure A5-2.** Plot of model results suggesting the optimal number of migration edges for *Treemix* analysis.

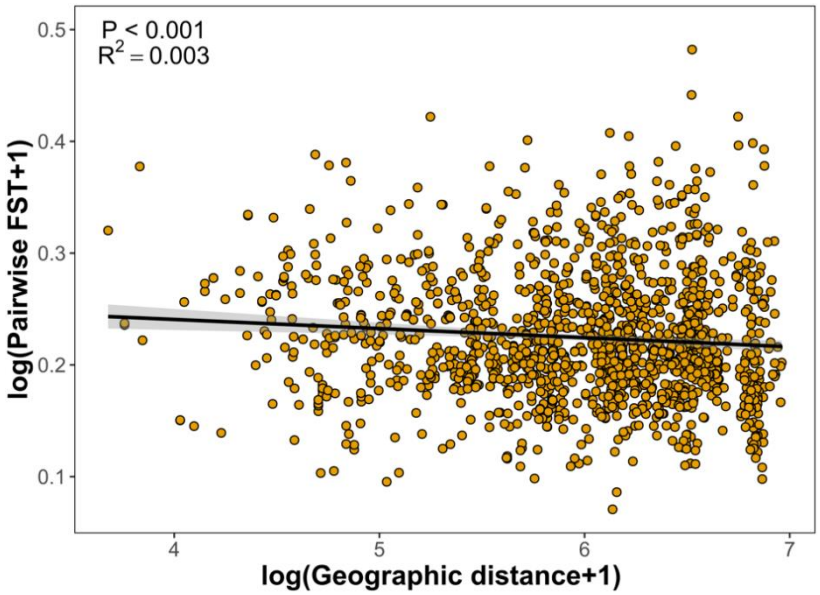


**Figure A5-3.** Isolation by distance plot for comparisons across Hydrometric Areas excluding outgroup populations and the Talla reservoir transplant population.

# Appendix: Chapter 6

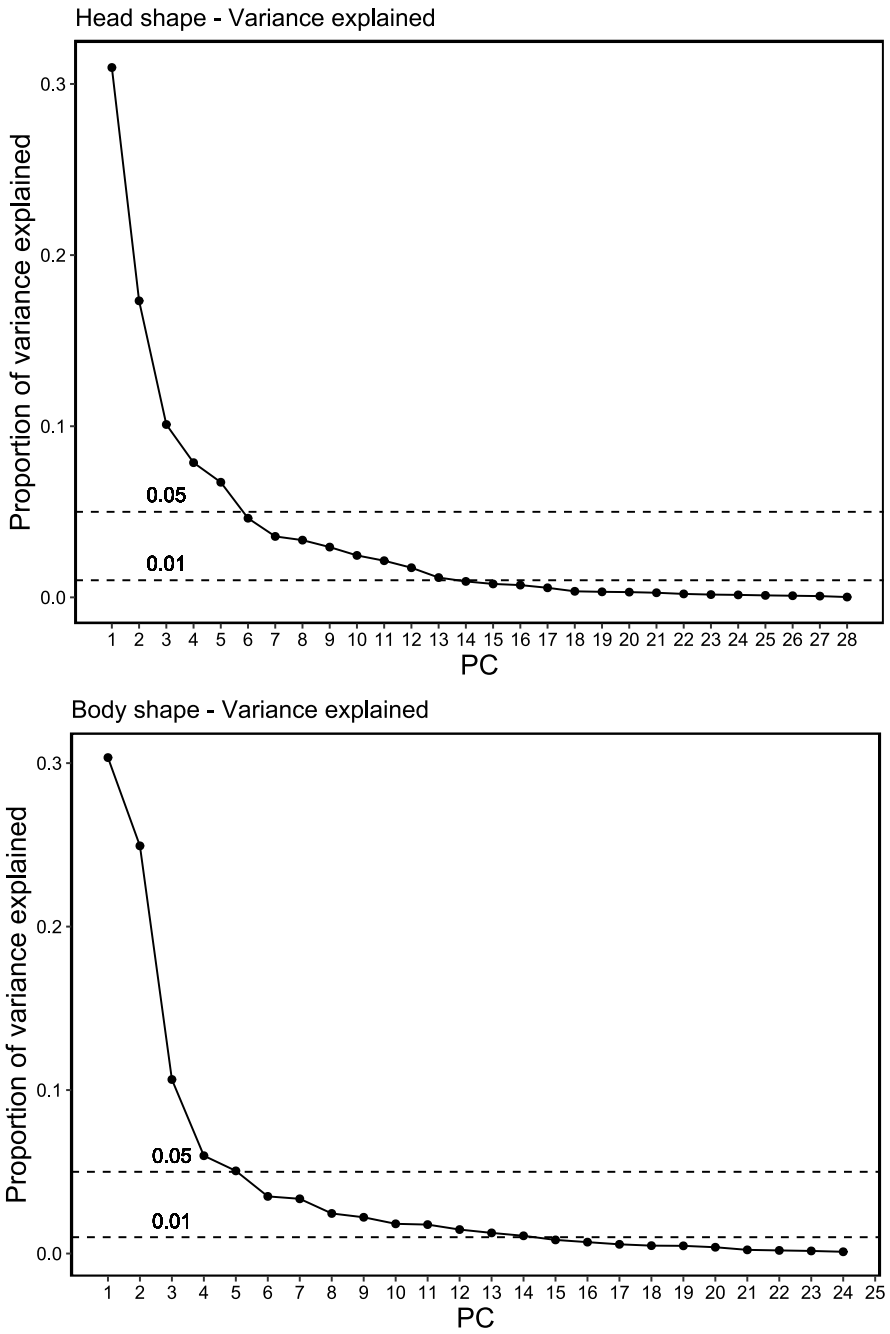


**Figure A6-1.** Proportion of variance explained by each PC axes for head and body shape respectively. Dotted lines indicate 5% and 1% of the total variance explained.

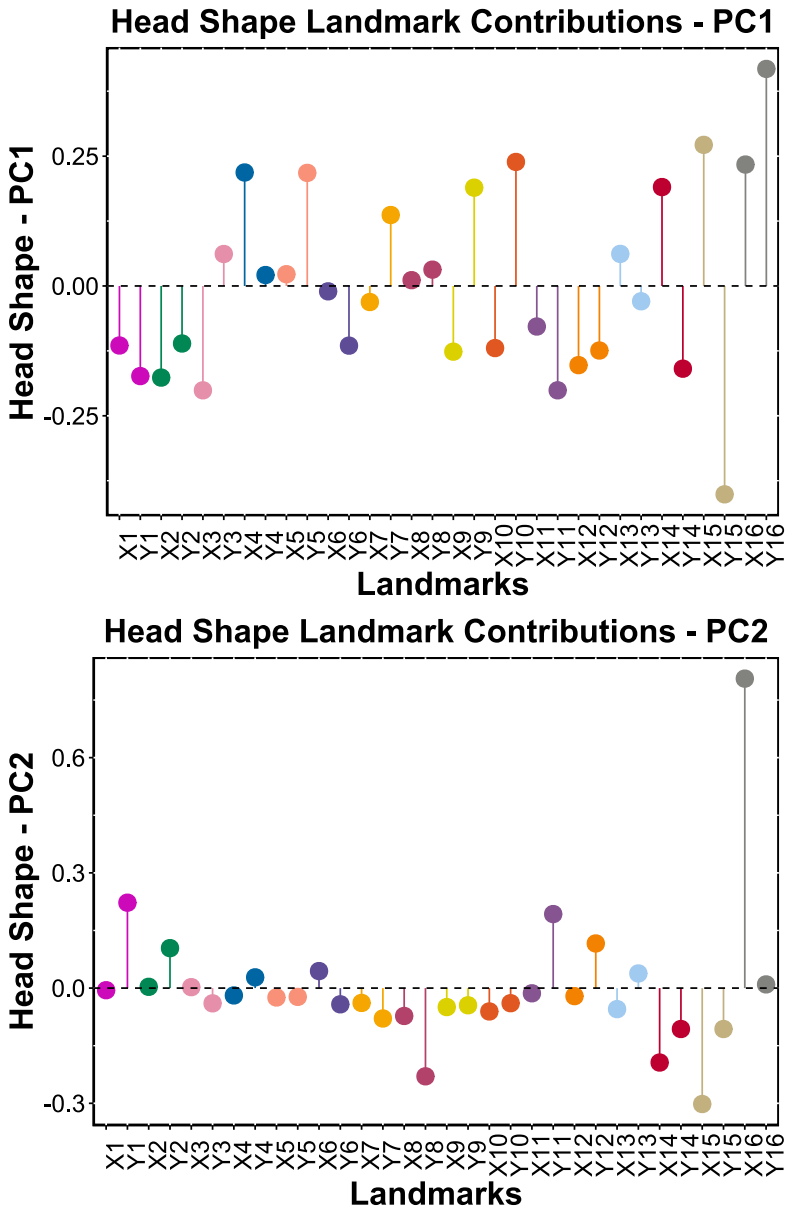


**Figure A6-2.** Lollipop plot showing contribution of each landmark to PC1 and PC2 scores for head shape. Each landmark is split into change in the X axis (e.g., X1) and Y axis (e.g., Y1). Colours correlate to landmark colours in Figure 6-1E.

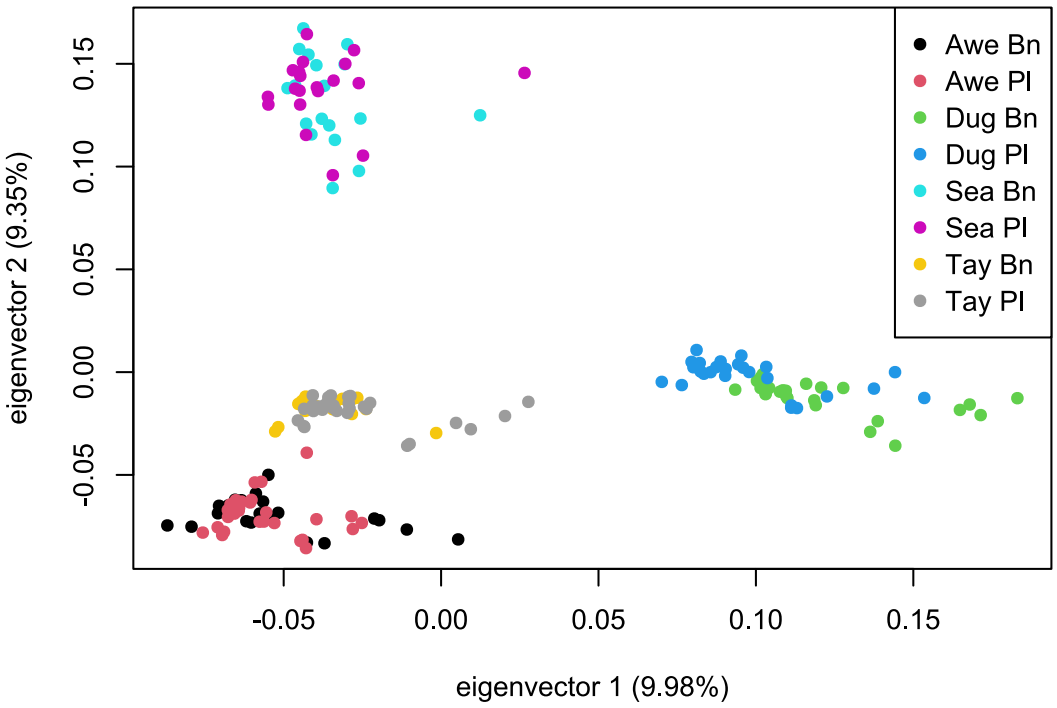


**Figure A6-3.** Principal component analysis for our SNP dataset of 13,071 SNPs. Points represent an individual and are coloured by ecotype population of origin as indicated in the legend. Bn refers to benthivore and Pl to planktivore.

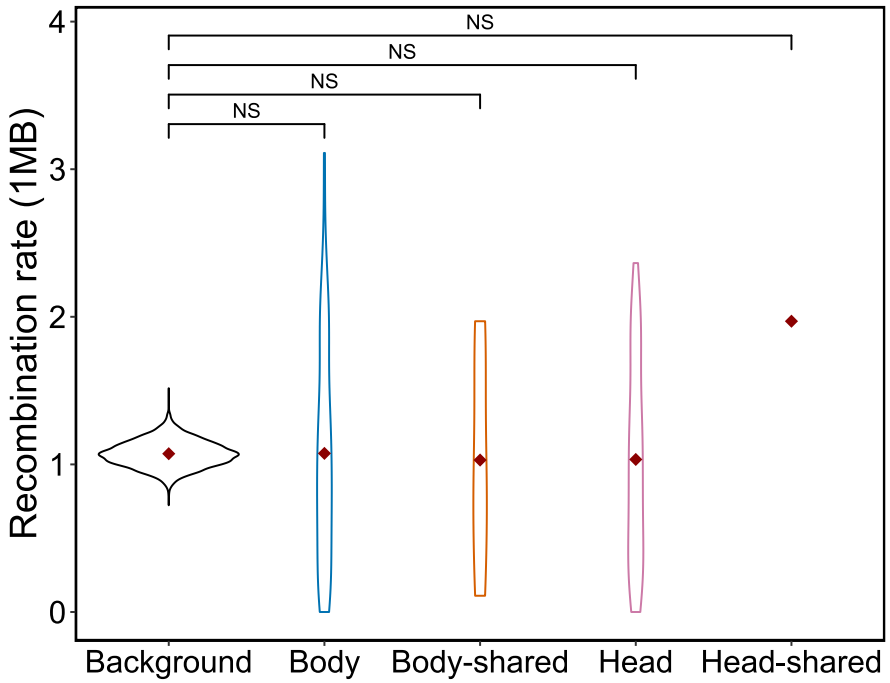


**Figure A6-4.** Recombination rate differences for different associated SNPs datasets. Body refers to all SNPs associated for body shape and Head refers to all SNPs associated with head shape. Body-shared dataset is just the body shape SNPs found in all four ecomorph pairs and Head-shared is the equivalent for head shape SNPs. Background SNPs refers to a randomly selected background subset of SNPs used for comparisons. Red diamonds are the mean value for that dataset. Significance of difference in means is indicated by NS ($P > 0.05$), * ($P < 0.05$), ** ($P < 0.01$), *** ($P < 0.001$).

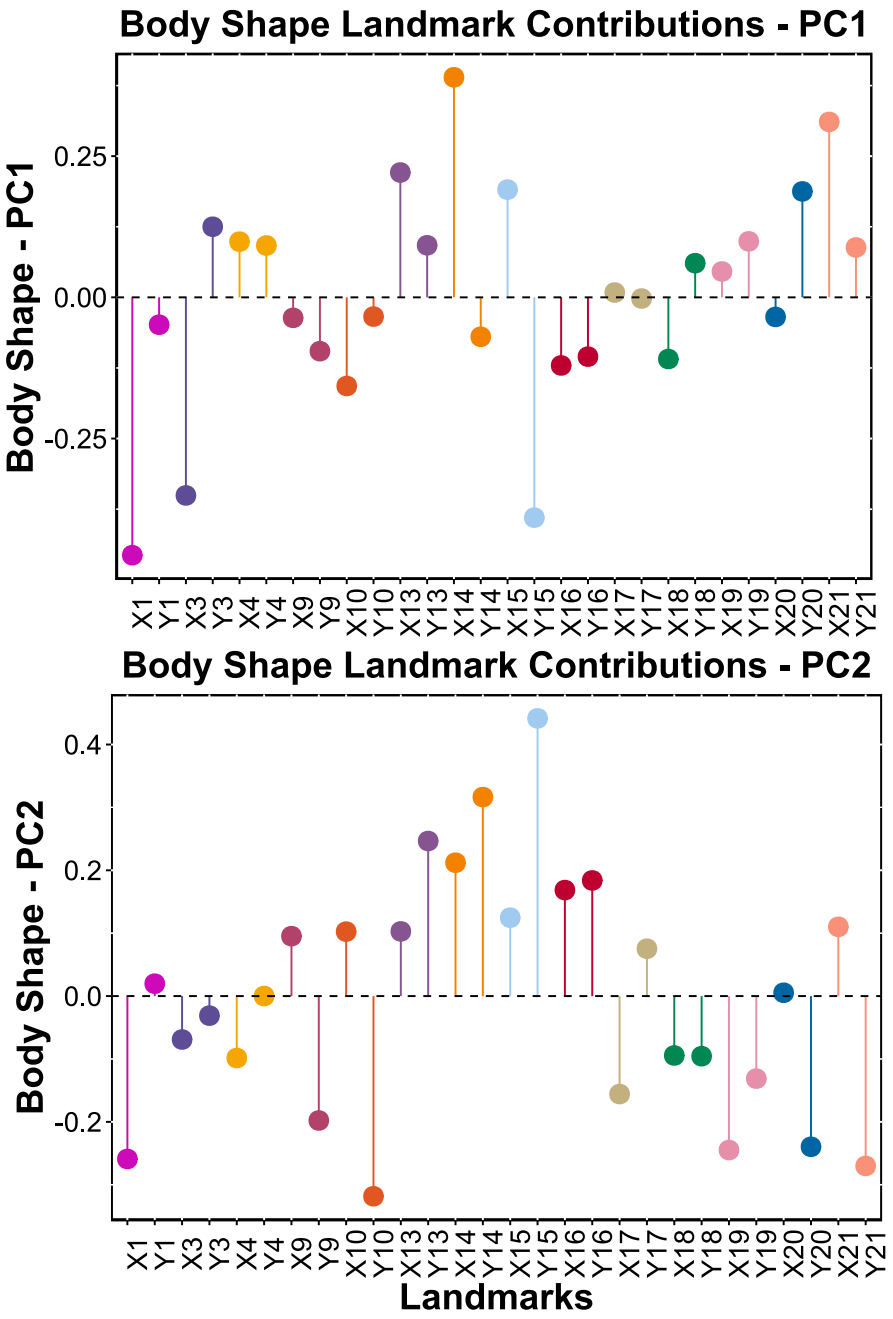


**Figure A6-5.** Lollipop plot showing contribution of each landmark to PC1 and PC2 scores for body shape. Each landmark is split into change in the X axis (e.g., X1) and Y axis (e.g., Y1). Colours correlate to landmark colours in Figure 6-1F.

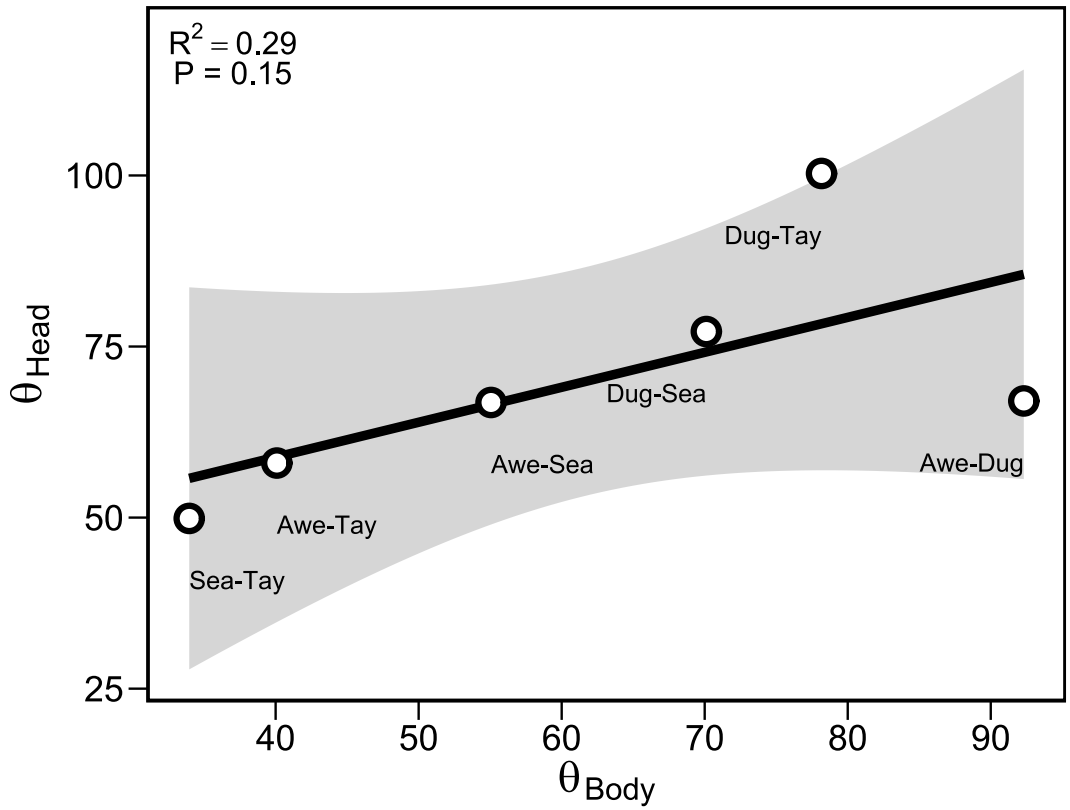


**Figure A6-6.** Comparing differences in the angle of phenotypic change between ecomorphs for head ($\theta_{Head}$) and body shape ($\theta_{Body}$) across ecomorph pairs. Each ecomorph pair comparison is indicated on the plot. Dug refers to Dughaill and Sea to na Sealga.

# References


Adams, C.E., Bean, C.W., Fraser, D. and Maitland, P.S. 2007a. Conservation and management of the Arctic charr: A forward view. *Ecology of Freshwater Fish*. **16**(1), pp.2–5.

Adams, C.E., Fraser, D., Huntingford, F.A., Greer, R.B., Askew, C.M. and Walker, A.F. 1998. Trophic polymorphism among Arctic charr from Loch Rannoch, Scotland. *Journal of Fish Biology*. **52**(6), pp.1259–1271.

Adams, C.E., Fraser, D., Wilson, A.J., Alexander, G., Ferguson, M.M. and Skulason, S. 2007b. Patterns of phenotypic and genetic variability show hidden diversity in Scottish Arctic charr. *Ecology of Freshwater Fish*. **16**(1), pp.78–86.

Adams, C.E., Hamilton, D.J., McCarthy, I., Wilson, A.J., Grant, A., Alexander, G., Waldron, S., Snorasson, S.S., Ferguson, M.M. and Skúlason, S. 2006. Does breeding site fidelity drive phenotypic and genetic sub-structuring of a population of Arctic charr? *Evolutionary Ecology*. **20**(1), pp.11–26.

Adams, C.E. and Huntingford, F.A. 2004. Incipient speciation driven by phenotypic plasticity? Evidence from sympatric populations of Arctic charr. *Biological Journal of the Linnean Society*. **81**(4), pp.611–618.

Adams, C.E. and Huntingford, F.A. 2002. Inherited differences in head allometry in polymorphic Arctic charr from Loch Rannoch, Scotland. *Journal of Fish Biology*. **60**(3), pp.515–520.

Adams, C.E. and Maitland, P.S. 2007. Arctic charr in Britain and Ireland - 15 species or one? *Ecology of Freshwater Fish*. **16**(1), pp.20–28.

Adams, C.E. and Maitland, P.S. 2018. *Arctic Charr in the Lochs of Scotland: An Assessment of Distribution and Status*. Fast-Print Publishing.

Adams, C.E., Wilson, A.J. and Ferguson, M.M. 2008. Parallel divergence of sympatric genetic and body size forms of Arctic charr, *Salvelinus alpinus*, from two Scottish lakes. *Biological Journal of the Linnean Society*. **95**(4), pp.748–757.

Adams, D.C. and Collyer, M.L. 2013. Phenotypic trajectory analysis: Comparison of shape change patterns in evolution and ecology. *Hystrix*. **24**(1), pp.75–83.

Adams, D.C. and Otárola-Castillo, E. 2013. Geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*. **4**(4), pp.393–399.

Aguillon, S.M., Fitzpatrick, J.W., Bowman, R., Schoech, S.J., Clark, A.G., Coop, G. and Chen, N. 2017. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genetics*. **13**(8).

Alekseyev, S.S., Samusenok, V.P., Matveev, A.N., Mikhail, & and Pichugin, Y. 2002. Diversification, sympatric speciation, and trophic polymorphism of Arctic charr, *Salvelinus alpinus* complex, in Transbaikalia. *Environmental Biology of Fishes*. **64**, pp.97–114.

Alexa, A. and Rahnenfuhrer, J. 2020. topGO: Enrichment Analysis for Gene Ontology.

Alexander, G.D. and Adams, C.E. 2004. Exposure to a common environment erodes inherited between-population trophic morphology differences in Arctic charr. *Journal of Fish Biology*. **64**(1), pp.253–257.

Alexander, G.D. and Adams, C.E. 2000. The phenotypic diversity of Arctic charr, *Salvelinus alpinus*, (Salmonidae) in Scotland and Ireland. *aqua, International Journal of Ichthyology*. **4**(2), pp.77–88.

Alfaro, M.E., Bolnick, D.I. and Wainwright, P.C. 2004. Evolutionary Dynamics of Complex Biomechanical Systems: An example using the four-bar mechanism. *Evolution*. **58**(3), pp.495–503.

Allendorf, F.W., Luikart, G. and Aitken, S.N. 2012. *Conservation and the Genetics of Populations* 2nd ed. Blackwell.

Alsos, I.G., Rijal, D.P., Ehrich, D., Karger, D.N., Yoccoz, N.G., Heintzman, P.D., Brown, A.G., Lammers, Y., Pellissier, L., Alm, T., Bråthen, K.A., Coissac, E., Merkel, M.K.F., Alberti, A., Denoeud, F. and Bakke, J. 2022. Postglacial species arrival and diversity buildup of northern ecosystems took millennia. *Science Advances*. **8**.

Anderson, E., Harrison, S., Passmore, D.G. and Mighall, T.M. 1998. Geomorphic evidence of Younger Dryas glaciation in the Macgillycuddy's Reeks, south west Ireland. *Quaternary Proceedings*. **13**(6), pp.75–90.

Anesio, A.M. and Laybourn-Parry, J. 2012. Glaciers and ice sheets as a biome. *Trends in Ecology and Evolution*. **27**(4), pp.219–225.

Angst, P., Ameline, C., Haag, C.R., Ben-Ami, F., Ebert, D. and Fields, P.D. 2022. Genetic Drift Shapes the Evolution of a Highly Dynamic Metapopulation. *Molecular Biology and Evolution*. **39**(12).

Apolônio Silva de Oliveira, D., Decraemer, W., Moens, T., dos Santos, G.A.P. and Derycke, S. 2017. Low genetic but high morphological variation over more than 1000 km coastline refutes omnipresence of cryptic diversity in marine nematodes. *BMC Evolutionary Biology*. **17**(1), p.71.

Arendt, J. and Reznick, D. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology and Evolution*. **23**(1), pp.26–32.

Arribas, L.P., Bagur, M., Klein, E., Penchaszadeh, P.E. and Palomo, M.G. 2013. Geographic distribution of two mussel species and associated assemblages along the northern Argentinean coast. *Aquatic Biology*. **18**(1), pp.91–103.

Auer, S.K., Dick, C.A., Metcalfe, N.B. and Reznick, D.N. 2018. Metabolic rate evolves rapidly and in parallel with the pace of life history. *Nature Communications*. **9**(1), p.14.

Avise, J.C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution*. **43**(6), pp.1192–1208.

Baillie, S.M., Muir, A.M., Hansen, M.J., Krueger, C.C. and Bentzen, P. 2016. Genetic and phenotypic variation along an ecological gradient in lake trout *Salvelinus namaycush*. *BMC Evolutionary Biology*. **16**(1).

Baines, J.F., Das, A., Mousset, S. and Stephan, W. 2004. The Role of Natural Selection in Genetic Differentiation of Worldwide Populations of Drosophila ananassae. *Genetics*. **168**(4), pp.1987–1998.

Ballantyne, C.K., McCarroll, D. and Stone, J.O. 2011. Periglacial trimlines and the extent of the Kerry-Cork Ice Cap, SW Ireland. *Quaternary Science Reviews*. **30**(27–28), pp.3834–3845.

Ballantyne, C.K., McCarroll, D. and Stone, J.O. 2006. Vertical dimensions and age of the Wicklow Mountains ice dome, Eastern Ireland, and implications for the extent of the last Irish Ice Sheet. *Quaternary Science Reviews*. **25**(17–18), pp.2048–2058.

Ballantyne, C.K. and Small, D. 2019. The last Scottish ice sheet. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*. **110**(1–2), pp.93–131.

Ballin, T.B. 2017. Rising waters and processes of diversification and unification in material culture: the flooding of Doggerland and its effect on north-west European prehistoric populations between ca. 13 000 and 1500 cal BC. *Journal of Quaternary Science*. **32**(2), pp.329–339.

Bamshad, M. and Wooding, S.P. 2003. Signatures of natural selection in the human genome. *Nature Reviews Genetics*. **4**(2), pp.99–111.

Barr, I.D., Roberson, S., Flood, R. and Dortch, J. 2017. Younger Dryas glaciers and climate in the Mourne Mountains, Northern Ireland. *Journal of Quaternary Science*. **32**(1), pp.104–115.

Barth, A.M., Clark, P.U., Clark, J., McCabe, A.M. and Caffee, M. 2016. Last Glacial Maximum cirque glaciation in Ireland and implications for reconstructions of the Irish Ice Sheet. *Quaternary Science Reviews*. **141**, pp.85–93.

Barthelemy, N., Hynes, R., Hutchinson, G. and Prodöhl, P.A. 2023. *Taxonomy and Phylogeography of the Irish Arctic Char (Salvelinus alpinus)* [Online]. Available from: www.epa.ie.

Barton, N.H. 2000. Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. **355**(1403), pp.1553–1562.

Bateman, M.D., Evans, D.J.A., Buckland, P.C., Connell, E.R., Friend, R.J., Hartmann, D., Moxon, H., Fairburn, W.A., Panagiotakopulu, E. and Ashurst, R.A. 2015. Last glacial dynamics of the Vale of York and North Sea lobes of the British and Irish ice sheet. *Proceedings of the Geologists' Association*. **126**(6), pp.712–730.

Bean, C.W., Mainstone, C.P., Hall, R.A., Hatton-Ellis, T.W., Lee, A.S. and Boon, P.J. 2018. Guidelines for the Selection of Biological SSSIs. Part 2: Detailed Guidelines for Habitats and Species Groups. Chapter 19 Freshwater Fish *In*: *Guidelines for the Selection of Biological SSSIs*. Joint Nature Conservation Committee.

Bean, C.W., Winfield, I.J. and Fletcher, J.W. 1996. Stock assessment of the Arctic charr (*Salvelinus alpinus*) population of Loch Ness, U.K. *In*: *Stock Assessment in Inland Fisheries*. Fishing News Books, pp.206–223.

Beaujean, A.A. 2012. BaylorEdPsych: R Package for Baylor University Educational Psychology Quantitative Courses.

Begun, D.J., Holloway, A.K., Stevens, K., Hillier, L.W., Poh, Y.-P., Hahn, M.W., Nista, P.M., Jones, C.D., Kern, A.D., Dewey, C.N., Pachter, L., Myers, E. and Langley, C.H. 2007. Population Genomics: Whole-Genome Analysis of Polymorphism and Divergence in Drosophila simulans. *PLoS Biology*. **5**(11), p.e310.

Benn, D.I. 1989. Controls on sedimentation in a late Devensian ice-dammed lake, Achnasheen, Scotland. *Boreas*. **18**(1), pp.31–42.

Bennett, A.K.D., Tzedakis, P.C. and Willis, K.J. 1991. Quaternary Refugia of North European Trees. *Journal of Biogeography*. **18**(1), pp.103–115.

Berloff, N., Perola, M. and Lange, K. 2002. Spline Methods for the Comparison of Physical and Genetic Maps. *Journal of Computational Biology*. **9**(3), pp.465–475.

Bernatchez, L. 2001. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*. **55**(2), pp.351–379.

Bernatchez, L. and Wilson, C.C. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*. **7**(4), pp.431–452.

Berner, D., Adams, D.C., Grandchamp, A. -C. and Hendry, A.P. 2008. Natural selection drives patterns of lake–stream divergence in stickleback foraging morphology. *Journal of Evolutionary Biology*. **21**(6), pp.1653–1665.

Bickerdike, H.L., Evans, D.J.A., Ó Cofaigh, C. and Stokes, C.R. 2016. The glacial geomorphology of the Loch Lomond Stadial in Britain: a map and geographic information system resource of published evidence. *Journal of Maps*. **12**(5), pp.1178–1186.

Bickerdike, H.L., Evans, D.J.A., Stokes, C.R. and Ó Cofaigh, C. 2018. The glacial geomorphology of the Loch Lomond (Younger Dryas) Stadial in Britain: a review. *Journal of Quaternary Science*. **33**(1), pp.1–54.

Blain, S.A., Schluter, D., Adams, C.E., Amundsen, P., Knudsen, R. and Chavarie, L. 2023. Patterns and repeatability of multi-ecotype assemblages of sympatric salmonids. *Global Ecology and Biogeography*. **32**(12), pp.2257–2270.

Bolger, A.M., Lohse, M. and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**(15), pp.2114–2120.

Bolnick, D.I., Barrett, R.D.H., Oke, K.B., Rennison, D.J. and Stuart, Y.E. 2018. (Non) Parallel Evolution. *Annual Review of Ecology, Evolution, and Systematics*. **49**, pp.303–330.

Bolnick, D.I. and Fitzpatrick, B.M. 2007. Sympatric Speciation: Models and Empirical Evidence. *Annual Review of Ecology, Evolution, and Systematics*. **38**(1), pp.459–487.

Bolnick, D.I. and Lau, O.L. 2008. Predictable Patterns of Disruptive Selection in Stickleback in Postglacial Lakes. *The American Naturalist*. **172**(1), pp.1–11.

Bonin, A., Nicole, F., Pompanon, F., Miaud, C. and Taberlet, P. 2007. Population Adaptive Index: a New Method to Help Measure Intraspecific Genetic Diversity and Prioritize Populations for Conservation. *Conservation Biology*. **21**(3), pp.697–708.

Boston, C.M., Lukas, S. and Carr, S.J. 2015. A Younger Dryas plateau icefield in the Monadhliath, Scotland, and implications for regional palaeoclimate. *Quaternary Science Reviews*. **108**, pp.139–162.

Boulding, E.G., Culling, M., Glebe, B., Berg, P.R., Lien, S. and Moen, T. 2008. Conservation genomics of Atlantic salmon: SNPs associated with QTLs for adaptive traits in parr from four trans-Atlantic backcrosses. *Heredity*. **101**(4), pp.381–391.

Bowen, B.W. 1999. Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Molecular Ecology*. **8**(s1).

Bowen, B.W. 1998. What is wrong with ESUs? The gap between evolutionary theory and conservation principles. *Journal of Shellfish Research*. **17**(5), pp.1355–1358.

Brachmann, M.K., Parsons, K., Skúlason, S., Gaggiotti, O. and Ferguson, M. 2022. Variation in the genomic basis of parallel phenotypic and ecological divergence in benthic and pelagic morphs of Icelandic Arctic charr (*Salvelinus alpinus*). *Molecular Ecology*. **31**(18), pp.4589–4599.

Bradwell, T., Small, D., Fabel, D., Clark, C.D., Chiverrell, R.C., Saher, M.H., Dove, D., Callard, S.L., Burke, M.J., Moreton, S.G., Medialdea, A., Bateman, M.D., Roberts, D.H., Golledge, N.R., Finlayson, A., Morgan, S. and Cofaigh, C. 2019. Pattern, style and timing of British–Irish Ice Sheet retreat: Shetland and northern North Sea sector. *Journal of Quaternary Science*., pp.1–42.

Brunner, P.C., Douglas, M.R., Osinov, A., Wilson, C.C. and Bernatchez, L. 2001. Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. *Evolution*. **55**(3), pp.573–586.

Bush, V. and Adams, C.E. 2007. Using phenotypic variation to determine conservation value: Application of a novel approach to Arctic charr. *Ecology of Freshwater Fish*. **16**(1), pp.29–33.

Butlin, R.K., Bridle, J.R. and Schluter, D. 2001. Speciation and patterns of biodiversity *In*: R. Butlin, J. Bridle and D. Schluter, eds. *Speciation and Patterns of Diversity* [Online]. Cambridge University Press, pp.1–14. Available from: https://www.cambridge.org/core/product/identifier/CBO9780511815683A010/type/book_part.

Butlin, R.K., Galindo, J. and Grahame, J.W. 2008. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*. **363**(1506), pp.2997–3007.

Campbell, C.S., Adams, C.E.., Bean, C.W. and Parsons, K.J. 2017. Conservation Evo-Devo: Preserving Biodiversity by Understanding Its Origins. *Trends in Ecology and Evolution*. **32**(10), pp.746–759.

Campbell, C.S., Adams, C.E., Bean, C.W., Pilakouta, N. and Parsons, K.J. 2021. Evolvability under climate change: Bone development and shape plasticity are heritable and correspond with performance in Arctic charr (*Salvelinus alpinus*). *Evolution and Development*. **23**(4), pp.333–350.

Candy, J.R. and Beacham, T.D. 2000. Patterns of homing and straying in southern British Columbia coded-wire tagged chinook salmon (*Oncorhynchus tshawytscha*) populations. *Fisheries Research*. **47**(1), pp.41–56.

Cano-Barbacil, C., Radinger, J. and García-Berthou, E. 2022. The importance of seawater tolerance and native status in mediating the distribution of inland fishes. *Journal of Biogeography*. **49**(11), pp.2037–2049.

Cardinale, B.J., Nelson, K. and Palmer, M.A. 2000. Linking species diversity to the functioning of ecosystems: on the importance of environmental context. *Oikos*. **91**(1), pp.175–183.

Carlson, A.E. 2013. The Younger Dryas Climate Event. *The Encyclopedia of Quaternary Science*. **29**(3), pp.126–134.

Casacci, L.P., Barbero, F. and Balletto, E. 2014. The 'Evolutionarily Significant Unit' concept and its applicability in biological conservation. *Italian Journal of Zoology*. **81**(2), pp.182–193.

CEN 2015. *Water quality - Sampling of fish with multi-mesh gillnets*. ICS.

Charlesworth, B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetical Research*. **63**(3), pp.213–227.

Chiocchio, A., Arntzen, Jan.W., Martínez-Solano, I., de Vries, W., Bisconti, R., Pezzarossa, A., Maiorano, L. and Canestrelli, D. 2021. Reconstructing hotspots of genetic diversity from glacial refugia and subsequent dispersal in Italian common toads (*Bufo bufo*). *Scientific Reports*. **11**(260).

Chiocchio, A., Zampiglia, M., Biaggini, M., Biello, R., Di Tizio, L., Leonetti, F.L., Olivieri, O., Sperone, E., Trabalza-Marinucci, M., Corti, C. and Canestrelli, D. 2022. Unveiling a hotspot of genetic diversity in southern Italy for the endangered Hermann's tortoise *Testudo hermanni*. *BMC Ecology and Evolution*. **22**(1).

Chiverrell, R.C., Thrasher, I.M., Thomas, G.S.P., Lang, A., Scourse, J.D., van Landeghem, K.J.J., McCarroll, D., Clark, C.D., Cofaigh, C.Ó., Evans, D.J.A. and Ballantyne, C.K. 2013. Bayesian modelling the retreat of the Irish Sea Ice Stream. *Journal of Quaternary Science*. **28**(2), pp.200–209.

Christensen, K.A., Rondeau, E.B., Minkley, D.R., Leong, J.S., Nugent, C.M., Danzmann, R.G., Ferguson, M.M., Stadnik, A., Devlin, R.H., Muzzerall, R., Edwards, M., Davidson, W.S. and Koop, B.F. 2018. The Arctic charr (*Salvelinus alpinus*) genome and transcriptome assembly. *PLoS ONE*. **13**(9), pp.1–30.

Clark, C.D., Ely, J.C., Greenwood, S.L., Hughes, A.L.C., Meehan, R., Barr, I.D., Bateman, M.D., Bradwell, T., Doole, J., Evans, D.J.A., Jordan, C.J., Monteys, X., Pellicer, X.M. and Sheehy, M. 2018. BRITICE Glacial Map, version 2: a map and GIS database of glacial landforms of the last British–Irish Ice Sheet. *Boreas*. **47**(1).

Clark, C.D., Hughes, A.L.C., Greenwood, S.L., Jordan, C. and Sejrup, H.P. 2012. Pattern and timing of retreat of the last British-Irish Ice Sheet. *Quaternary Science Reviews*. **44**, pp.112–146.

Coates, D.J., Byrne, M. and Moritz, C. 2018. Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*. **6**(OCT).

Le Coeur, C., Storkey, J. and Ramula, S. 2021. Population responses to observed climate variability across multiple organismal groups. *Oikos*. **130**(3), pp.476–487.

Coleman, R.A., Gauffre, B., Pavlova, A., Beheregaray, L.B., Kearns, J., Lyon, J., Sasaki, M., Leblois, R., Sgro, C. and Sunnucks, P. 2018. Artificial barriers prevent genetic recovery of small isolated populations of a low-mobility freshwater fish. *Heredity*. **120**(6), pp.515–532.

Coles, B.J. 2000. Doggerland: The cultural dynamics of a shifting coastline. *Geological Society Special Publication*. **175**, pp.393–401.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D. and Kingsley, D.M. 2005. Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science*. **307**(5717), pp.1928–1933.

Conte, G.L., Arnegard, M.E., Peichel, C.L. and Schluter, D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*. **279**(1749), pp.5039–5047.

Coyne, J.A. and Allen Orr, H. 2004. *Speciation*. Oxford University Press.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M. and Wayne, R.K. 2000. Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*. **15**(7), pp.290–295.

Crookes, S. and Shaw, P.W. 2016. Isolation by distance and non-identical patterns of gene flow within two river populations of the freshwater fish *Rutilus rutilus* (L. 1758). *Conservation Genetics*. **17**(4), pp.861–874.

Crotti, M., Adams, C.E., Etheridge, E.C., Bean, C.W., Gowans, A.R.D., Knudsen, R., Lyle, A.A., Maitland, P.S., Winfield, I.J., Elmer, K.R. and Præbel, K. 2020. Geographic hierarchical population genetic structuring in British European whitefish (*Coregonus lavaretus*) and its implications for conservation. *Conservation Genetics*. **21**(5), pp.927–939.

Crotti, M., Bean, C.W., Gowans, A.R.D., Winfield, I.J., Butowska, M., Wanzenböck, J., Bondarencko, G., Præbel, K., Adams, C.E. and Elmer, K.R. 2021. Complex and divergent histories gave rise to genome-wide divergence patterns amongst European whitefish (*Coregonus lavaretus*). *Journal of Evolutionary Biology*. **34**(12), pp.1954–1969.

Cruickshank, T.E. and Hahn, M.W. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*. **23**(13), pp.3133–3157.

Csapó, H., Krzywoźniak, P., Grabowski, M., Wattier, R., Bącela-Spychalska, K., Mamos, T., Jelić, M. and Rewicz, T. 2020. Successful post-glacial colonization of Europe by single lineage of freshwater amphipod from its Pannonian Plio-Pleistocene diversification hotspot. *Scientific Reports*. **10**(8695).

Dahms, C., Kemppainen, P., Zanella, L.N., Zanella, D., Carosi, A., Merilä, J. and Momigliano, P. 2022. Cast away in the Adriatic: Low degree of parallel genetic differentiation in three-spined sticklebacks. *Molecular Ecology*. **31**(4), pp.1234–1253.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G. and Durbin, R. 2011. The variant call format and VCFtools. *Bioinformatics*. **27**(15), pp.2156–2158.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A. and Davies, R.M. 2021. Twelve years of SAMtools and BCFtools. *GigaScience*. **10**(2).

Darriba, Di., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B. and Flouri, T. 2020. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Molecular Biology and Evolution*. **37**(1), pp.291–294.

Delgado, M.L., Górski, K., Habit, E. and Ruzzante, D.E. 2019. The effects of diadromy and its loss on genomic divergence: The case of amphidromous *Galaxias maculatus* populations. *Molecular Ecology*. **28**(24), pp.5217–5231.

Delgado, M.L. and Ruzzante, D.E. 2020. Investigating Diadromy in Fishes and Its Loss in an-Omics Era. *iScience*. **23**(12).

DeWoody, J.A. and Avise, J.C. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*. **56**(3), pp.461–473.

Doenz, C.J., Krähenbühl, A.K., Walker, J., Seehausen, O. and Brodersen, J. 2019. Ecological opportunity shapes a large Arctic charr species radiation. *Proceedings of the Royal Society B: Biological Sciences*. **286**(1913).

Doiron, S., Bernatchez, L. and Blier, P.U. 2002. A Comparative Mitogenomic Analysis of the Potential Adaptive Value of Arctic Charr mtDNA Introgression in Brook Charr Populations (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution*. **19**(11), pp.1902–1909.

Eckert, C.G., Samis, K.E. and Lougheed, S.C. 2008. Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology*. **17**(5), pp.1170–1188.

Edwards, R. and Brooks, A. 2008. The Island of Ireland: Drowning the Myth of an Irish Land-Bridge? *The Irish Naturalists' Journal*. **29**, pp.19–34.

Edwards, S. V., Kingan, S.B., Calkins, J.D., Balakrishnan, C.N., Jennings, W.B., Swanson, W.J. and Sorenson, M.D. 2005. Speciation in birds: Genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences*. **102**, pp.6550–6557.

Egan, S.P., Hood, G.R. and Ott, J.R. 2012. Testing the Role of Habitat Isolation among Ecologically Divergent Gall Wasp Populations. *International Journal of Ecology*. **2012**, pp.1–8.

Ellis, N., Smith, S.J. and Roland Pitcher, C. 2012. Gradient forests: Calculating importance gradients on physical predictors. *Ecology*. **93**(1), pp.156–168.

Elmer, K.R. 2016. Genomic tools for new insights to variation, adaptation, and evolution in the salmonid fishes: a perspective for charr. *Hydrobiologia*. **783**(1), pp.191–208.

Elmer, K.R., Fan, S., Kusche, H., Luise Spreitzer, M., Kautt, A.F., Franchini, P. and Meyer, A. 2014. Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communications*. **5**.

Elmer, K.R. and Meyer, A. 2011. Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution*. **26**, pp.298–306.

Eloranta, A.P., Kahilainen, K.K., Amundsen, P.A., Knudsen, R., Harrod, C. and Jones, R.I. 2015. Lake size and fish diversity determine resource use and trophic position of a top predator in high-latitude lakes. *Ecology and Evolution*. **5**(8), pp.1664–1675.

Ely, J.C., Clark, C.D., Hindmarsh, R.C.A., Hughes, A.L.C., Greenwood, S.L., Bradley, S.L., Gasson, E., Gregoire, L., Gandy, N., Stokes, C.R. and Small, D. 2019. Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last British–Irish Ice Sheet. *Journal of Quaternary Science*., pp.1–15.

Esin, E. V., Markevich, G.N. and Pichugin, M.Y. 2018. Juvenile divergence in adaptive traits among seven sympatric fish ecomorphs arises before moving to different lacustrine habitats. *Journal of Evolutionary Biology*. **31**(7), pp.1018–1034.

Etheridge, E.C., Adams, C.E., Bean, C.W., Durie, N.C., Gowans, A.R.D., Harrod, C., Lyle, A.A., Maitland, P.S. and Winfield, I.J. 2012. Are phenotypic traits useful for differentiating among a priori *Coregonus* taxa? *Journal of Fish Biology*. **80**(2), pp.387–407.

Evans, D.J.A., Roberts, D.H., Bateman, M.D., Ely, J., Medialdea, A., Burke, M.J., Chiverrell, R.C., Clark, C.D. and Fabel, D. 2019. A chronology for North Sea Lobe

advance and recession on the Lincolnshire and Norfolk coasts during MIS 2 and 6. *Proceedings of the Geologists' Association*. **130**(5), pp.523–540.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C. and Foll, M. 2013. Robust Demographic Inference from Genomic and SNP Data. *PLoS Genetics*. **9**(10).

Excoffier, L., Hofer, T. and Foll, M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity*. **103**(4), pp.285–298.

Excoffier, L., Marchi, N., Marques, D.A., Matthey-Doret, R., Gouy, A. and Sousa, V.C. 2021. Fastsimcoal2: Demographic inference under complex evolutionary scenarios. *Bioinformatics*. **37**(24), pp.4882–4885.

Fang, B., Kemppainen, P., Momigliano, P., Feng, X. and Merilä, J. 2020. On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nature Ecology and Evolution*. **4**, pp.1105–1115.

Feder, J.L., Opp, S.B., Wlazlo, B., Reynolds, K., Go, W. and Spisak, S. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Sciences*. **91**(17), pp.7990–7994.

Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package).

Fenton, S., Jacobs, A., Bean, C.W., Adams, C.E. and Elmer, K.R. 2023a. Genomic underpinnings of head and body shape in Arctic charr ecomorph pairs. *Authorea*.

Fenton, S., Elmer, K.R., Bean, C.W. and Adams, C.E. 2023b. How glaciation impacted evolutionary history and contemporary genetic diversity of flora and fauna in the British Isles. *Scottish Geographical Journal*. **139**(3–4), pp.445–465.

Fick, S.E. and Hijmans, R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. **37**(12), pp.4302–4315.

Filchak, K.E., Roethele, J.B. and Feder, J.L. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*. **407**(6805), pp.739–742.

Fillatre, E.K., Etherton, P. and Heath, D.D. 2003. Bimodal run distribution in a northern population of sockeye salmon (*Oncorhynchus nerka*): life history and genetic analysis on a temporal scale. *Molecular Ecology*. **12**(7), pp.1793–1805.

Filteau, M., Pavey, S.A., St-Cyr, J. and Bernatchez, L. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Molecular Biology and Evolution*. **30**(6), pp.1384–1396.

Finnegan, A.K., Griffiths, A.M., King, R.A., Machado-Schiaffino, G., Porcher, J.P., Garcia-Vazquez, E., Bright, D. and Stevens, J.R. 2013. Use of multiple markers demonstrates a cryptic western refugium and postglacial colonisation routes of Atlantic salmon (*Salmo salar* L.) in northwest Europe. *Heredity*. **111**(1), pp.34–43.

Finstad, A.G. and Hein, C.L. 2012. Migrate or stay: Terrestrial primary productivity and climate drive anadromy in Arctic char. *Global Change Biology*. **18**(8), pp.2487–2497.

Fischer, J. and Lindenmayer, D.B. 2007. Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography*. **16**, pp.265–280.

Fitak, R.R. 2021. OptM: Estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols*. **6**(1).

Fitzpatrick, M.C., Chhatre, V.E., Soolanayakanahally, R.Y. and Keller, S.R. 2021. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Molecular Ecology Resources*. **21**(8), pp.2749–2765.

Forester, B.R., Murphy, M., Mellison, C., Petersen, J., Pilliod, D.S., Van Horne, R., Harvey, J. and Funk, W.C. 2022. Genomics-informed delineation of conservation units in a desert amphibian. *Molecular Ecology*. **31**(20), pp.5249–5269.

Franch-Gras, L., Hahn, C., García-Roger, E.M., Carmona, M.J., Serra, M. and Gómez, A. 2018. Genomic signatures of local adaptation to the degree of environmental predictability in rotifers. *Scientific Reports*. **8**(1).

François, D., Ursenbacher, S., Boissinot, A., Ysnel, F. and Lourdais, O. 2021. Isolation-by-distance and male-biased dispersal at a fine spatial scale: a study of the common European adder (*Vipera berus*) in a rural landscape. *Conservation Genetics*. **22**(5), pp.823–837.

Frankham, R., Ballou, J.D. and Briscoe, D.A. 2009. *Introduction to Conservation Genetics* 2nd ed. Cambridge University Press.

Fraser, D., Adams, C.E. and Huntingford, F.A. 1998. Trophic polymorphism among Arctic charr *Salvelinus alpinus* L., from Loch Ericht, Scotland. *Ecology of Freshwater Fish*. **7**(4), pp.184–191.

Fraser, D.J. and Bernatchez, L. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*. **10**(12), pp.2741–2752.

Friend, G.F. 1956. A new sub-species of charr from Loch Eck. *The Glasgow Naturalist*. **17**, pp.219–220.

Frost, W.E. 1965. Breeding habits of Windermere charr, *Salvelinus willughbii* (Günther), and their bearing on speciation of these fish. *Proceedings of the Royal Society of London. Series B. Biological Sciences*. **163**(991), pp.232–284.

Fruciano, C., Franchini, P., Kovacova, V., Elmer, K.R., Henning, F. and Meyer, A. 2016. Genetic linkage of distinct adaptive traits in sympatrically speciating crater lake cichlid fish. *Nature Communications*. **7**(12736).

Funk, W.C., Lovich, R.E., Hohenlohe, P.A., Hofman, C.A., Morrison, S.A., Sillett, T.S., Ghalambor, C.K., Maldonado, J.E., Rick, T.C., Day, M.D., Polato, N.R., Fitzpatrick, S.W., Coonan, T.J., Crooks, K.R., Dillon, A., Garcelon, D.K., King, J.L., Boser, C.L., Gould, N. and Andelt, W.F. 2016. Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*. **25**(10), pp.2176–2194.

Funk, W.C., McKay, J.K., Hohenlohe, P.A. and Allendorf, F.W. 2012. Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*. **27**(9), pp.489–496.

Furlan, E., Stoklosa, J., Griffiths, J., Gust, N., Ellis, R., Huggins, R.M. and Weeks, A.R. 2012. Small population size and extremely low levels of genetic diversity in island populations of the platypus, Ornithorhynchus anatinus. *Ecology and Evolution*. **2**(4), pp.844–857.

Galvez, J.R., St. John, M.E., McLean, K., Dening Touokong, C., Gonwouo, L.N. and Martin, C.H. 2022. Trophic Specialization on unique resources despite limited niche divergence in a celebrated example of sympatric speciation. *Ecology of Freshwater Fish*. **31**(4), pp.675–692.

Garduño-Paz, M. V. and Adams, C.E. 2010. Discrete prey availability promotes foraging segregation and early divergence in Arctic charr, *Salvelinus alpinus*. *Hydrobiologia*. **650**(1), pp.15–26.

Garduño-Paz, M. V., Adams, C.E., Verspoor, E., Knox, D. and Harrod, C. 2012. Convergent evolutionary processes driven by foraging opportunity in two sympatric morph pairs of Arctic charr with contrasting post-glacial origins. *Biological Journal of the Linnean Society*. **106**(4), pp.794–806.

Garduño-Paz, M. V., Demetriou, M. and Adams, C.E. 2010. Variation in scale shape among alternative sympatric phenotypes of Arctic charr *Salvelinus alpinus* from two lakes in Scotland. *Journal of Fish Biology*. **76**(6), pp.1491–1497.

Gavrilets, S. and Losos, J.B. 2009. Adaptive Radiation: Contrasting Theory with Data. *Science*. **323**(5915), pp.728–732.

Giesecke, T., Brewer, S., Finsinger, W., Leydet, M. and Bradshaw, R.H.W. 2017. Patterns and dynamics of European vegetation change over the last 15,000 years. *Journal of Biogeography*. **44**(7), pp.1441–1456.

Glasser, N.F., Davies, J.R., Hambrey, M.J., Davies, B.J., Gheorghiu, D.M., Balfour, J., Smedley, R.K. and Duller, G.A.T. 2018. Late Devensian deglaciation of south-west Wales from luminescence and cosmogenic isotope dating. *Journal of Quaternary Science*. **33**(7), pp.804–818.

Golledge, N.R., Hubbard, A. and Sugden, D.E. 2008. High-resolution numerical simulation of Younger Dryas glaciation in Scotland. *Quaternary Science Reviews*. **27**(9–10), pp.888–904.

Gonçalves, H., Martínez-Solano, I., Ferrand, N. and García-París, M. 2007. Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (*Anura, Discoglossidae, Alytes*): Deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*. **44**(1), pp.494–500.

Gordeeva, N. V., Alekseyev, S.S., Matveev, A.N. and Samusenok, V.P. 2015. Parallel evolutionary divergence in Arctic char *Salvelinus alpinus* complex from Transbaikalia: variation in differentiation degree and segregation of genetic diversity among sympatric forms. *Canadian Journal of Fisheries and Aquatic Sciences*. **72**(1), pp.96–115.

Goudet, J. 2005. hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*. **5**, pp.184–186.

Gow, J.L., Peichel, C.L. and Taylor, E.B. 2007. Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. *Journal of Evolutionary Biology*. **20**(6), pp.2173–2180.

Grant, P.R. and Grant, B.R. 2007. *How and Why Species Multiply: The Radiation of Darwin's Finches*. Princeton University Press.

Griffiths, A.M., Koizumi, I., Bright, D. and Stevens, J.R. 2009. A case of isolation by distance and short-term temporal stability of population structure in brown trout (*Salmo trutta*) within the River Dart, southwest England. *Evolutionary Applications*. **2**(4), pp.537–554.

Guðbrandsson, J., Kapralova, K.H., Franzdóttir, S.R., Bergsveinsdóttir, Þ.M., Hafstað, V., Jónsson, Z.O., Snorrason, S.S. and Pálsson, A. 2019. Extensive genetic differentiation between recently evolved sympatric Arctic charr morphs. *Ecology and Evolution*. **9**(19), pp.10964–10983.

De Guia, A.P.O. and Saitoh, T. 2007. The gap between the concept and definitions in the Evolutionarily Significant Unit: The need to integrate neutral genetic variation and adaptive variation. *Ecological Research*. **22**(4), pp.604–612.

Gunther, A. 1862. Contribution to the knowledge of the British charrs. Part I. *Proceedings of the Zoological Society of London*., pp.37–54.

Gunther, A. 1865. Contributions to the knowledge of the British charrs. Part III. *Proceedings of the Zoological Society of London*., pp.698–699.

Hahn, C., Genner, M.J., Turner, G.F. and Joyce, D.A. 2017. The genomic basis of cichlid fish adaptation within the deepwater "twilight zone" of Lake Malawi. *Evolution Letters*. **1**(4), pp.184–198.

Harmon, L.J., Matthews, B., Des Roches, S., Chase, J.M., Shurin, J.B. and Schluter, D. 2009. Evolutionary diversification in stickleback affects ecosystem functioning. *Nature*. **458**(7242), pp.1167–1170.

Hastings, A., Byers, J.E., Crooks, J.A., Cuddington, K., Jones, C.G., Lambrinos, J.G., Talley, T.S. and Wilson, W.G. 2007. Ecosystem engineering in space and time. *Ecology Letters*. **10**(2), pp.153–164.

Hebert, F.O., Renaut, S. and Bernatchez, L. 2013. Targeted sequence capture and resequencing implies a predominant role of regulatory regions in the divergence of a sympatric lake whitefish species pair (*Coregonus clupeaformis*). *Molecular Ecology*. **22**(19), pp.4896–4914.

Hejase, H.A., Salman-Minkov, A., Campagna, L., Hubisz, M.J., Lovette, I.J., Gronau, I. and Siepel, A. 2020. Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps. *Proceedings of the National Academy of Sciences*. **117**(48), pp.30554–30565.

Hemmer-Hansen, J., Nielsen, E.E., Therkildsen, N.O., Taylor, M.I., Ogden, R., Geffen, A.J., Bekkevold, D., Helyar, S., Pampoulie, C., Johansen, T. and Carvalho, G.R. 2013. A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*. **22**(10), pp.2653–2667.

Hendry, A.P. 2017. *Eco-evolutionary Dynamics*. Princeton: Princeton University Press.

Hendry, A.P. 2009. Ecological speciation! Or the lack thereof? *Canadian Journal of Fisheries and Aquatic Sciences*. **66**(8), pp.1383–1398.

Hendry, A.P., Nosil, P. and Rieseberg, L.H. 2007. The speed of ecological speciation. *Functional Ecology*. **21**(3), pp.455–464.

Herman, J.S., McDevitt, A.D., Kawałko, A., Jaarola, M., Wójcik, J.M. and Searle, J.B. 2014. Land-Bridge calibration of molecular clocks and the post-glacial colonization of Scandinavia by the Eurasian field vole *Microtus agrestis*. *PLoS ONE*. **9**(8).

Herman, J.S. and Searle, J.B. 2011. Post-glacial partitioning of mitochondrial genetic variation in the field vole. *Proceedings of the Royal Society B: Biological Sciences*. **278**(1724), pp.3601–3607.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean society*. **68**, pp.87–112.

Hewitt, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*. **58**(3), pp.247–276.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature*. **405**(6789), pp.907–913.

Hijmans, R.J. and van Etten, J. 2012. raster: Geographic analysis and modeling with raster data.

Hindar, K. and Jonsson, B. 1982. Habitat and Food Segregation of Dwarf and Normal Arctic Charr (*Salvelinus alpinus*) from Vangsvatnet Lake, Western Norway. *Canadian Journal of Fisheries and Aquatic Sciences*. **39**, pp.1030–1045.

Hoarau, G., Coyer, J.A., Veldsink, J.H., Stam, W.T. and Olsen, J.L. 2007. Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology*. **16**(17), pp.3606–3616.

Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A. and Whitlock, M.C. 2016. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*. **188**(4), pp.379–397.

Hogan, J.D., Blum, M.J., Gilliam, J.F., Bickford, N. and McIntyre, P.B. 2014. Consequences of alternative dispersal strategies in a putatively amphidromous fish. *Ecology*. **95**(9), pp.2397–2408.

Hoham, R.W. and Remias, D. 2020. Snow and Glacial Algae: A Review1. *Journal of Phycology*. **56**(2), pp.264–282.

Holliday, J.A., Zhou, L., Bawa, R., Zhang, M. and Oubida, R.W. 2016. Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in *Populus trichocarpa*. *New Phytologist*. **209**(3), pp.1240–1251.

Hollingsworth, P.M., O'Brien, D., Ennos, R.A., Yahr, R., Neaves, L., Ahrends, A., Ballingall, K.T., Brooker, R.W., Burke, T., Cavers, S., Dawson, I.K., Elston, D.A., Kerr, J., Marshall, D.F., Pakeman, R.J., Trivedi, C., Wall, E., Wright, F. and Ogden, R. 2019. *Conserving Genetic Diversity: Development of a national approach for addressing Aichi Biodiversity Target 13 that includes wild species* [Online]. Available from:

https://www.nature.scot/doc/scotlands-biodiversity-progress-2020-aichi-targets-aichi-target-13-genetic-diversity-maintained.

Hooker, O.E., Adams, C.E. and Chavarie, L. 2023. Arctic charr phenotypic responses to abrupt generational scale temperature change: an insight into how cold-water fish could respond to extreme climatic events. *Environmental Biology of Fishes*. **106**(5), pp.909–922.

Hooker, O.E., Barry, J., Van Leeuwen, T.E., Lyle, A., Newton, J., Cunningham, P. and Adams, C.E. 2016. Morphological, ecological and behavioural differentiation of sympatric profundal and pelagic Arctic charr (*Salvelinus alpinus*) in Loch Dughaill Scotland. *Hydrobiologia*. **783**(1), pp.209–221.

Hudson, A.G., Vonlanthen, P. and Seehausen, O. 2011. Rapid parallel adaptive radiations from a single hybridogenic ancestral population. *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, **278**(1702), pp.58–66.

Huey, J.A., Balcombe, S.R., Real, K.M., Sternberg, D. and Hughes, J.M. 2017. Genetic structure and effective population size of the most northern population of the Australian River Blackfish, *Gadopsis marmoratus* (Richardson 1848): Implications for long-term population viability. *Freshwater Science*. **36**(1), pp.113–123.

Hughes, A.L.C., Clark, C.D. and Jordan, C.J. 2014. Flow-pattern evolution of the last British Ice Sheet. *Quaternary Science Reviews*. **89**, pp.148–168.

Hughes, J.B., Daily, G.C. and Ehrlich, P.R. 1997. Population Diversity: Its Extent and Extinction. *Science*. **278**(5338), pp.689–692.

Hughes, P.D., Glasser, N.F. and Fink, D. 2016. Rapid thinning of the Welsh Ice Cap at 20-19 ka based on 10Be ages. *Quaternary Research*. **85**(1), pp.107–117.

Hutchison, D.W. and Templeton, A.R. 1999. Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*. **53**(5), pp.1898–1914.

Hwang, H.-Y. and Wang, J. 2023. Effect of recombination on genetic diversity of Caenorhabditis elegans. *Scientific Reports*. **13**(1), p.16425.

IUCN 2023. The IUCN Red List of Threatened Species. Available from: https://www.iucnredlist.org.

Jacobs, A. 2018. *The population genomic origins of ecological specialisation in salmonid fishes*. Doctoral dissertation, University of Glasgow.

Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N. V., Alekseyev, S.S., Hooker, O., Leong, J.S., Minkley, D.R., Rondeau, E.B., Koop, B.F., Adams, C.E. and Elmer, K.R. 2020. Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLOS Genetics*. **16**(4), p.e1008658.

Jacobs, A. and Elmer, K.R. 2021. Alternative splicing and gene expression play contrasting roles in the parallel phenotypic evolution of a salmonid fish. *Molecular Ecology*. **30**(20), pp.4955–4969.

Jacobs, A., Womack, R., Chen, M., Gharbi, K. and Elmer, K.R. 2017. Significant synteny and colocalization of ecologically relevant quantitative trait loci within and across species of salmonid fishes. *Genetics*. **207**(2), pp.741–754.

Jacobsen, M.W., Jensen, N.W., Nygaard, R., Præbel, K., Jónsson, B., Nielsen, N.H., Pujolar, J.M., Fraser, D.J., Bernatchez, L. and Hansen, M.M. 2022. A melting pot in the Arctic: Analysis of mitogenome variation in Arctic char (*Salvelinus alpinus*) reveals a 1000-km contact zone between highly divergent lineages. *Ecology of Freshwater Fish*. **31**(2), pp.330–346.

Jensen, H., Kiljunen, M., Knudsen, R. and Amundsen, P.A. 2017. Resource partitioning in food, space and time between Arctic charr (*Salvelinus alpinus*), Brown trout (*Salmo trutta*) and European whitefish (*Coregonus lavaretus*) at the southern edge of their continuous coexistence. *PLoS ONE*. **12**(1), pp.1–18.

Jiggins, C.D. 2008. Ecological Speciation in Mimetic Butterflies. *BioScience*. **58**(6), pp.541–548.

Jombart, T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*. **24**(11), pp.1403–1405.

Jones, C.G., Lawton, J.H. and Shachak, M. 1994. Organisms as Ecosystem Engineers. *Oikos*. **69**(3), pp. 373-386.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M.C., Myers, R.M., Miller, C.T., Summers, B.R., Knecht, A.K., Brady, S.D., Zhang, H., Pollen, A.A., Howes, T., Amemiya, C., Baldwin, J., Bloom, T., Jaffe, D.B., Nicol, R., Wilkinson, J., Lander, E.S., Di Palma, F., Lindblad-Toh, K. and Kingsley, D.M. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*. **484**(7392), pp.55–61.

Jónsdóttir, G.Ó., von Elm, L.-M., Ingimarsson, F., Tersigni, S., Snoarrason, S.S., Pálsson, A. and Steele, S.E. 2023. Diversity in the internal functional feeding elements of sympatric morphs of Arctic charr (*Salvelinus alpinus*). *bioRxiv*.

Jonsson, B. and Jonsson, N. 2001. Polymorphism and speciation in Arctic charr. *Journal of Fish Biology*. **58**(3), pp.605–638.

Jørgensen, E.H. and Johnsen, H.K. 2014. Rhythmic life of the Arctic charr: Adaptations to life at the edge. *Marine Genomics*. **14**, pp.71–81.

Kaeuffer, R., Peichel, C.L., Bolnick, D.I. and Hendry, A.P. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution*. **66**(2), pp.402–418.

Kamvar, Z.N., Brooks, J.C. and Grünwald, N.J. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*. **6**(208).

Kapralova, K.H., Ias, Z., Onsson, O.J., Palsson, A., Idur, S., Le Deuff, S., Kristjánsson, B.K. and Snorrason, S.S. 2015. Bones in Motion: Ontogeny of Craniofacial Development in Sympatric Arctic Charr Morphs. *Developmental Dynamics*. **244**, pp.1168–1178.

Kapralova, K.H., Morrissey, M.B., Kristjánsson, B.K., Ólafsdóttir, G.Á., Snorrason, S.S. and Ferguson, M.M. 2011. Evolution of adaptive diversity and genetic connectivity in Arctic charr (*Salvelinus alpinus*) in Iceland. *Heredity*. **106**(3), pp.472–487.

Kassambara A 2023. ggpubr: 'ggplot2' Based Publication Ready Plots. Available from: https://rpkgs.datanovia.com/ggpubr/

Kaufman, D., McKay, N., Routson, C., Erb, M., Dätwyler, C., Sommer, P.S., Heiri, O. and Davis, B. 2020. Holocene global mean surface temperature, a multi-method reconstruction approach. *Scientific Data*. **7**(1).

Keefer, M.L. and Caudill, C.C. 2014. Homing and straying by anadromous salmonids: a review of mechanisms and rates. *Reviews in Fish Biology and Fisheries*. **24**(1), pp.333–368.

Kelley, J.L., Madeoy, J., Calhoun, J.C., Swanson, W. and Akey, J.M. 2006. Genomic signatures of positive selection in humans and the limits of outlier approaches. *Genome Research*. **16**(8), pp.980–989.

Kelly, S., Moore, T.N., de Eyto, E., Dillane, M., Goulon, C., Guillard, J., Lasne, E., McGinnity, P., Poole, R., Winfield, I.J., Woolway, R.I. and Jennings, E. 2020. Warming winters threaten peripheral Arctic charr populations of Europe. *Climatic Change*. **163**(1), pp.599–618.

Kettle-White, A. 2001. New fish records: Stoneloach, roach and a spring spawning Arctic charr population, Loch Awe, Argyll. *The Glasgow Naturalist*. **23**(6), pp.120–120.

Klemetsen, A. 2010. The Charr Problem Revisited: Exceptional Phenotypic Plasticity Promotes Ecological Speciation in Postglacial Lakes. *Freshwater Reviews*. **3**(1), pp.49–74.

Klemetsen, A. 2013. The most variable vertebrate on Earth. *Journal of Ichthyology*. **53**(10), pp.781–791.

Klingenberg, C.P. 2016. Size, shape, and form: concepts of allometry in geometric morphometrics. *Development Genes and Evolution*. **226**(3), pp.113–137.

Klobucar, S.L., Rick, J.A., Mandeville, E.G., Wagner, C.E. and Budy, P. 2021. Investigating the morphological and genetic divergence of arctic char (*Salvelinus alpinus*) populations in lakes of arctic Alaska. *Ecology and Evolution*. **11**(7), pp.3040–3057.

Kocher, T.D. 2004. Adaptive evolution and explosive speciation: The cichlid fish model. *Nature Reviews Genetics*. **5**(4), pp.288–298.

Kotlík, P., Marková, S., Konczal, M., Babik, W. and Searle, J.B. 2018. Genomics of end-pleistocene population replacement in a small mammal. *Proceedings of the Royal Society B: Biological Sciences*. **285**(1872).

Kottelat, M. and Freyhof, J. 2007. *Handbook of European Freshwater Fishes*. Self-published.

Kowalko, J.E., Rohner, N., Linden, T.A., Rompani, S.B., Warren, W.C., Borowsky, R., Tabin, C.J., Jeffery, W.R. and Yoshizawa, M. 2013. Convergence in feeding posture occurs through different genetic loci in independently evolved cave populations of *Astyanax mexicanus*. *Proceedings of the National Academy of Sciences*. **110**(42), pp.16933–16938.

Kristjánsson, B.K., Skúlason, S., Snorrason, S.S. and Noakes, D.L.G. 2012. Fine-scale parallel patterns in diversity of small benthic Arctic charr (*Salvelinus alpinus*) in relation to the ecology of lava/groundwater habitats. *Ecology and Evolution*. **2**(6), pp.1099–1112.

Kuhn, M. 2008. Building Predictive Models in *R* using the caret Package. *Journal of Statistical Software*. **28**(5).

Küttner, E., Parsons, K.J., Easton, A.A., Skúlason, S., Danzmann, R.G. and Ferguson, M.M. 2014. Hidden genetic variation evolves with ecological specialization: The genetic basis of phenotypic plasticity in Arctic charr ecomorphs. *Evolution and Development*. **16**(4), pp.247–257.

Lambeck, K. 1995. Late Devensian and Holocene shorelines of the British Isles and North Sea from models of glacio-hydro-isostatic rebound. *Journal of the Geological Society*. **152**(3), pp.437–448.

Landry, L. and Bernatchez, L. 2010. Role of epibenthic resource opportunities in the parallel evolution of lake whitefish species pairs (*Coregonus* sp.). *Journal of Evolutionary Biology*. **23**(12), pp.2602–2613.

Landry, L., Vincent, W.F. and Bernatchez, L. 2007. Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. *Journal of Evolutionary Biology*. **20**(3), pp.971–984.

Langerhans, R.B. and DeWitt, T.J. 2004. Shared and unique features of evolutionary diversification. *American Naturalist*. **164**(3), pp.335–349.

Langmead, B. and Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. **9**(4), pp.357–359.

Laporte, M., Rogers, S.M., Dion-Côté, A.M., Normandeau, E., Gagnaire, P.A., Dalziel, A.C., Chebib, J. and Bernatchez, L. 2015. RAD-QTL mapping reveals both genome-level

parallelism and different genetic architecture underlying the evolution of body shape in Lake whitefish (*Coregonus clupeaformis*) species pairs. *G3: Genes, Genomes, Genetics*. **5**(7), pp.1481–1491.

Larson, W.A., Dann, T.H., Limborg, M.T., McKinney, G.J., Seeb, J.E. and Seeb, L.W. 2019. Parallel signatures of selection at genomic islands of divergence and the major histocompatibility complex in ecotypes of sockeye salmon across Alaska. *Molecular Ecology*. **28**(9), pp.2254–2271.

Láruson, Á.J., Fitzpatrick, M.C., Keller, S.R., Haller, B.C. and Lotterhos, K.E. 2022. Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. *Evolutionary Applications*. **15**(3), pp.403–416.

Láruson, Á.J., Yeaman, S. and Lotterhos, K.E. 2020. The Importance of Genetic Redundancy in Evolution. *Trends in Ecology & Evolution*. **35**(9), pp.809–822.

Layton, K.K.S., Snelgrove, P.V.R., Dempson, J.B., Kess, T., Lehnert, S.J., Bentzen, P., Duffy, S.J., Messmer, A.M., Stanley, R.R.E., DiBacco, C., Salisbury, S.J., Ruzzante, D.E., Nugent, C.M., Ferguson, M.M., Leong, J.S., Koop, B.F. and Bradbury, I.R. 2021. Genomic evidence of past and future climate-linked loss in a migratory Arctic fish. *Nature Climate Change*. **11**(2), pp.158–165.

Lehnherr, I., St. Louis, V.L., Sharp, M., Gardner, A.S., Smol, J.P., Schiff, S.L., Muir, D.C.G., Mortimer, C.A., Michelutti, N., Tarnocai, C., St. Pierre, K.A., Emmerton, C.A., Wiklund, J.A., Köck, G., Lamoureux, S.F. and Talbot, C.H. 2018. The world's largest High Arctic lake responds rapidly to climate warming. *Nature Communications*. **9**(1), p.1290.

Leigh, J.W. and Bryant, D. 2015. POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*. **6**(9), pp.1110–1116.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li, H. and Roossinck, M.J. 2004. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. *Journal of Virology*. **78**(19), pp.10582–10587.

Litsios, G., Pellissier, L., Forest, F., Lexer, C., Pearman, P.B., Zimmermann, N.E. and Salamin, N. 2012. Trophic specialization influences the rate of environmental niche

evolution in damselfishes (*Pomacentridae*). *Proceedings of the Royal Society B: Biological Sciences*. **279**(1743), pp.3662–3669.

Loretán, G., Rueda, E.C., Cabrera, J.M., Pérez-Losada, M., Collins, P.A. and Giri, F. 2020. Geographical isolation and restricted gene flow drive speciation of *Aegla singularis* (*Decapoda: Anomura: Aeglidae*) in southern South America. *Biological Journal of the Linnean Society*. **129**(1), pp.177–189.

Losos, J.B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *American Naturalist*. **175**(6), pp.623–639.

Lucek, K., Kristjánsson, B.K., Skúlason, S. and Seehausen, O. 2016. Ecosystem size matters: the dimensionality of intralacustrine diversification in Icelandic stickleback is predicted by lake size. *Ecology and Evolution*. **6**(15), pp.5256–5272.

Luck, G.W., Daily, G.C. and Ehrlich, P.R. 2003. Population diversity and ecosystem services. *Trends in Ecology & Evolution*. **18**(7), pp.331–336.

Luquet, E., Rödin Mörch, P., Cortázar-Chinarro, M., Meyer-Lucht, Y., Höglund, J. and Laurila, A. 2019. Post-glacial colonization routes coincide with a life-history breakpoint along a latitudinal gradient. *Journal of Evolutionary Biology*. **32**(4), pp.356–368.

Lynch, M., Conery, J. and Burger, R. 1995. Mutation Accumulation and the Extinction of Small Populations. *The American Naturalist*. **146**(4), pp.489–518.

Machado, A.P., Cumer, T., Iseli, C., Beaudoing, E., Ducrest, A.L., Dupasquier, M., Guex, N., Dichmann, K., Lourenço, R., Lusby, J., Martens, H.D., Prévost, L., Ramsden, D., Roulin, A. and Goudet, J. 2022. Unexpected post-glacial colonisation route explains the white colour of barn owls (*Tyto alba*) from the British Isles. *Molecular Ecology*. **31**(2), pp.482–497.

MacLeod, A., Palmer, A., Lowe, J., Rose, J., Bryant, C. and Merritt, J. 2011. Timing of glacier response to Younger Dryas climatic cooling in Scotland. *Global and Planetary Change*. **79**(3–4), pp.264–274.

Magalhaes, I.S., Whiting, J.R., D'Agostino, D., Hohenlohe, P.A., Mahmud, M., Bell, M.A., Skúlason, S. and MacColl, A.D.C. 2021. Intercontinental genomic parallelism in

multiple three-spined stickleback adaptive radiations. *Nature Ecology and Evolution*. **5**(2), pp.251–261.

Maitland, J.R. 1881. A new char. *The Field*. **8**, p.516.

Maitland, P.S. 1995. World status and conservation of the Arctic charr *Salvelinus alpinus* (L.). *Nordic Journal of Freshwater Research*. **71**, pp.113–127.

Maitland, P.S. and Adams, C.E. 2018. *Arctic Charr in the Lochs of Scotland: An Assessment of Distribution and Status*. Fast-Print Publishing.

Maitland, P.S., Winfield, I.J., McCarthy, I.D. and Igoe, F. 2007. The status of Arctic charr *Salvelinus alpinus* in Britain and Ireland. *Ecology of Freshwater Fish*. **16**(1), pp.6–19.

Malmquist, H.J., Snorrason, S.S., Skúlason, S., Jonsson, B., Sandlund, O.T. and Jonasson, P.M. 1992. Diet Differentiation in Polymorphic Arctic Charr in Thingvallavatn, Iceland. *Journal of Animal Ecology*. **61**(1), pp.21–35.

Manel, S., Guerin, P.E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C. and Pellissier, L. 2020. Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*. **11**(1).

Marshall, J.D., Jones, R.T., Crowley, S.F., Oldfield, F., Nash, S. and Bedford, A. 2002. A high resolution Late-Glacial isotopic record from Hawes Water, Northwest England. Climatic oscillations: Calibration and comparison of palaeotemperature proxies. *Palaeogeography, Palaeoclimatology, Palaeoecology*. **185**(1–2), pp.25–40.

Martin, P.R. and Mckay, J.K. 2004. Latitudinal Variation in Genetic Divergence of Populations and the Potential for Future. *Evolution*. **58**(5), pp.938–945.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D. and Emerson, B.C. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*. **15**(1), pp.28–41.

Matala, A.P., Narum, S.R., Young, W. and Vogel, J.L. 2012. Influences of Hatchery Supplementation, Spawner Distribution, and Habitat on Genetic Structure of Chinook Salmon in the South Fork Salmon River, Idaho. *North American Journal of Fisheries Management*. **32**(2), pp.346–359.

McDevitt, A.D., Coscia, I., Browett, S.S., Ruiz-González, A., Statham, M.J., Ruczyńska, I., Roberts, L., Stojak, J., Frantz, A.C., Norén, K., Ågren, E.O., Learmount, J., Basto, M., Fernandes, C., Stuart, P., Tosh, D.G., Sindicic, M., Andreanszky, T., Isomursu, M., Panek, M., Korolev, A., Okhlopkov, I.M., Saveljev, A.P., Pokorny, B., Flajšman, K., Harrison, S.W.R., Lobkov, V., Ćirović, D., Mullins, J., Pertoldi, C., Randi, E., Sacks, B.N., Kowalczyk, R. and Wójcik, J.M. 2022. Next-generation phylogeography resolves post-glacial colonization patterns in a widespread carnivore, the red fox (*Vulpes vulpes*), in Europe. *Molecular Ecology*. **31**(3), pp.993–1006.

McGirr, J.A. and Martin, C.H. 2018. Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. *Evolution Letters*. **2**(2), pp.62–75.

McKeown, N.J., Hynes, R.A., Duguid, R.A., Ferguson, A. and Prodöhl, P.A. 2010. Phylogeographic structure of brown trout *Salmo trutta* in Britain and Ireland: Glacial refugia, postglacial colonization and origins of sympatric populations. *Journal of Fish Biology*. **76**(2), pp.319–347.

Mcphail, J.D. 1993. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): origin of the species pairs. *Canadian Journal of Zoology*. **75**, pp.515–523.

Mee, J.A., Bernatchez, L., Reist, J.D., Rogers, S.M. and Taylor, E.B. 2015. Identifying designatable units for intraspecific conservation prioritization: A hierarchical approach applied to the lake whitefish species complex (*Coregonus* spp.). *Evolutionary Applications*. **8**(5), pp.423–441.

Milner, A.M. and Bailey, R.G. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. *Aquaculture Research*. **20**(2), pp.179–192.

Milner, A.M., Knudsen, E.E., Soiseth, C., Robertson, A.L., Schell, D., Phillips, I.T. and Magnusson, K. 2000. Colonization and development of stream communities across a 200-year gradient in Glacier Bay National Park, Alaska, U.S.A. *Canadian Journal of Fisheries and Aquatic Sciences*. **57**(11), pp.2319–2335.

Minter, M., O'Brien, D., Cottrell, J., Ennos, R., Hill, J.K. and Hall, J. 2021. Exploring the potential for 'Gene Conservation Units' to conserve genetic diversity in wild populations. *Ecological Solutions and Evidence*. **2**(2).

Moccetti, P., Siwertsson, A., Kjær, R., Amundsen, P.A., Præbel, K., Tamayo, A.M.P., Power, M. and Knudsen, R. 2019. Contrasting patterns in trophic niche evolution of polymorphic Arctic charr populations in two subarctic Norwegian lakes. *Hydrobiologia*. **840**(1), pp.281–299.

Money, D., Gardner, K., Migicovsky, Z., Schwaninger, H., Zhong, G.Y. and Myles, S. 2015. LinkImpute: Fast and accurate genotype imputation for nonmodel organisms. *G3: Genes, Genomes, Genetics*. **5**(11), pp.2383–2390.

Montgomery, W.I., Provan, J., McCabe, A.M. and Yalden, D.W. 2014. Origin of British and Irish mammals: Disparate post-glacial colonisation and species introductions. *Quaternary Science Reviews*. **98**, pp.144–165.

Moore, J.S., Bajno, R., Reist, J.D. and Taylor, E.B. 2015. Post-glacial recolonization of the North American Arctic by Arctic char (*Salvelinus alpinus*): Genetic evidence of multiple northern refugia and hybridization between glacial lineages. *Journal of Biogeography*. **42**(11), pp.2089–2100.

Moritz, C. 1994. Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution*. **9**(10), pp.373–375.

Murray, J. and Pullar, L. 1910. *Bathymetrical Survey of the Scottish Fresh Water Lochs*. Edinburgh: Challenger Office.

Mussmann, S.M., Douglas, M.R., Oakey, D.D. and Douglas, M.E. 2020. Defining relictual biodiversity: Conservation units in speckled dace (Leuciscidae: *Rhinichthys osculus*) of the Greater Death Valley ecosystem. *Ecology and Evolution*. **10**(19), pp.10798–10817.

Nadeau, N.J., Whibley, A., Jones, R.T., Davey, J.W., Dasmahapatra, K.K., Baxter, S.W., Quail, M.A., Joron, M., ffrench-Constant, R.H., Blaxter, M.L., Mallet, J. and Jiggins, C.D. 2012. Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **367**(1587), pp.343–353.

Narum, S.R. and Hess, J.E. 2011. Comparison of $F_{ST}$ outlier tests for SNP loci under selection. *Molecular Ecology Resources*. **11**(s1), pp.184–194.

National Research Council 1999. *Perspectives on Biodiversity: Valuing Its Role in an Everchanging World*. Washington, D.C.: National Academies Press.

National River Flow Archive 2014. Integrated Hydrological Units of the United Kingdom: Hydrometric Areas with Coastline. *NERC Environmental Information Data Centre*.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press.

Nicastro, K.R., McQuaid, C.D., Dievart, A. and Zardi, G.I. 2020. Intraspecific diversity in an ecological engineer functionally trumps interspecific diversity in shaping community structure. *Science of The Total Environment*. **743**, p.140723.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M.J., Clark, A.G. and Bustamante, C. 2005. Genomic scans for selective sweeps using SNP data. *Genome Research*. **15**(11), pp.1566–1575.

Niemiller, M.L., Fitzpatrick, B.M. and Miller, B.T. 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Molecular Ecology*. **17**(9), pp.2258–2275.

Norman, J.D., Danzmann, R.G., Glebe, B. and Ferguson, M.M. 2011. The genetic basis of salinity tolerance traits in Arctic charr (*Salvelinus alpinus*). *BMC Genetics*. **12**, pp.1–12.

Norman, J.D., Robinson, M., Glebe, B., Ferguson, M.M. and Danzmann, R.G. 2012. Genomic arrangement of salinity tolerance QTLs in salmonids: A comparative analysis of Atlantic salmon (*Salmo salar*) with Arctic charr (*Salvelinus alpinus*) and rainbow trout (*Oncorhynchus mykiss*). *BMC Genomics*. **13**(1).

Nosil, P. 2012. *Ecological Speciation*. Oxford University Press.

Nosil, P., Harmon, L.J. and Seehausen, O. 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution*. **24**(3), pp.145–156.

Nosil, P. and Reimchen, T.E. 2005. Ecological opportunity and levels of morphological variance within freshwater stickleback populations. *Biological Journal of the Linnean Society*. **86**(3), pp.297–308.

Nosil, P., SandovalL, C.P. and Crespi, B.J. 2006. The evolution of host preference in allopatric vs. parapatric populations of *Timema cristinae* walking-sticks. *Journal of Evolutionary Biology*. **19**(3), pp.929–942.

Noskova, E., Abramov, N., Iliutkin, S., Sidorin, A., Dobrynin, P. and Ulyantsev, V.I. 2022. GADMA2: more efficient and flexible demographic inference from genetic data. *GigaScience*. **12**.

Ó Cofaigh, C., Telfer, M.W., Bailey, R.M. and Evans, D.J.A. 2012. Late Pleistocene chronostratigraphy and ice sheet limits, southern Ireland. *Quaternary Science Reviews*. **44**, pp.160–179.

Ó Cofaigh, C., Weilbach, K., Lloyd, J.M., Benetti, S., Callard, S.L., Purcell, C., Chiverrell, R.C., Dunlop, P., Saher, M., Livingstone, S.J., Van Landeghem, K.J.J., Moreton, S.G., Clark, C.D. and Fabel, D. 2019. Early deglaciation of the British-Irish Ice Sheet on the Atlantic shelf northwest of Ireland driven by glacioisostatic depression and high relative sea level. *Quaternary Science Reviews*. **208**, pp.76–96.

Öhlund, G., Bodin, M., Nilsson, K.A., Öhlund, S.-O., Mobley, K.B., Hudson, A.G., Peedu, M., Brännström, Å., Bartels, P., Præbel, K., Hein, C.L., Johansson, P. and Englund, G. 2020. Ecological speciation in European whitefish is driven by a large-gaped predator. *Evolution Letters*. **4**(3), pp.243–256.

Oke, K.B., Rolshausen, G., LeBlond, C. and Hendry, A.P. 2017. How parallel is parallel evolution? A comparative analysis in fishes. *American Naturalist*. **190**(1).

Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., German Grehov, O.G., Kandrov, D., Rasputin, K., Syabro, M. and Tleukenov, T. 2012. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics*. **28**(8), pp.1166–1167.

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C and Weedon J 2022. vegan: Community Ecology Package.

Oleinik, A.G., Skurikhina, L.A., Kukhlevsky, A.D. and Semenchenko, A.A. 2020. Complete mitochondrial genomes of the Arctic charr *Salvelinus alpinus alpinus Linnaeus* (*Salmoniformes: Salmonidae*). *Mitochondrial DNA Part B: Resources*. **5**(3), pp.2913–2915.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. and De Meester, L. 2013. Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*. **22**(24), pp.5983–5999.

Ørsted, M., Hoffmann, A.A., Sverrisdóttir, E., Nielsen, K.L. and Kristensen, T.N. 2019. Genomic variation predicts adaptive evolutionary responses better than population bottleneck history. *PLoS Genetics*. **15**(6).

Paetkau, D. 1999. Using Genetics to Identify Intraspecific Conservation Units: a Critique of Current Methods. *Conservation Biology*. **13**(6), pp.1507–1509.

Palkovacs, E.P., Dion, K.B., Post, D.M. and Caccone, A. 2008. Independent evolutionary origins of landlocked alewife populations and rapid parallel evolution of phenotypic traits. *Molecular Ecology*. **17**(2), pp.582–597.

Palm, S., Dannewitz, J., Prestegaard, T. and Wickström, H. 2009. Panmixia in European eel revisited: No genetic difference between maturing adults from southern and northern Europe. *Heredity*. **103**(1), pp.82–89.

Palsbøll, P., Berube, M. and Allendorf, F. 2007. Identification of management units using population genetic data. *Trends in Ecology & Evolution*. **22**(1), pp.11–16.

Pante, E. and Simon-Bouhet, B. 2013. marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. *PLoS ONE*. **8**(9).

Paz-Vinas, I., Loot, G., Hermoso, V., Veyssière, C., Poulet, N., Grenouillet, G. and Blanchet, S. 2018. Systematic conservation planning for intraspecific genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*. **285**(1877).

Pearse, D.E., Barson, N.J., Nome, T., Gao, G., Campbell, M.A., Abadía-Cardoso, A., Anderson, E.C., Rundio, D.E., Williams, T.H., Naish, K.A., Moen, T., Liu, S., Kent, M., Moser, M., Minkley, D.R., Rondeau, E.B., Brieuc, M.S.O., Sandve, S.R., Miller, M.R., Cedillo, L., Baruch, K., Hernandez, A.G., Ben-Zvi, G., Shem-Tov, D., Barad, O., Kuzishchin, K., Garza, J.C., Lindley, S.T., Koop, B.F., Thorgaard, G.H., Palti, Y. and Lien, S. 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology and Evolution*. **3**(12), pp.1731–1742.

Pennington, W. 1986. Lags in adjustment of vegetation to climate caused by the pace of soil development. Evidence from Britain. *Vegetatio*. **67**(2), pp.105–118.

Pickrell, J.K. and Pritchard, J.K. 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genetics*. **8**(11).

Pielou, E.C. 1991. *After the Ice Age: The Return of Life to Glaciated North America*. University of Chicago Press.

Poelstra, J.W., Richards, E.J. and Martin, C.H. 2018. Speciation in sympatry with ongoing secondary gene flow and a potential olfactory trigger in a radiation of Cameroon cichlids. *Molecular Ecology*. **27**(21), pp.4270–4288.

Præbel, K., Knudsen, R., Siwertsson, A., Karhunen, M., Kahilainen, K.K., Ovaskainen, O., Østbye, K., Peruzzi, S., Fevolden, S.E. and Amundsen, P.A. 2013. Ecological speciation in postglacial European whitefish: Rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecology and Evolution*. **3**(15), pp.4970–4986.

Presgraves, D.C. 2010. The molecular evolutionary basis of species formation. *Nature Reviews Genetics*. **11**(3), pp.175–180.

Primmer, C.R. 2011. Genetics of local adaptation in salmonid fishes. *Heredity*. **106**(3), pp.401–403.

Pritchard, J.K., Pickrell, J.K. and Coop, G. 2010. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*. **20**(4), pp.R208–R215.

QGIS Development Team 2022. QGIS Geographic Information System.

Quigley, D.T.G. 2014. Did melting ice sheets create temporary low- salinity corridors that facilitated the natural colonization of Ireland by stenohaline fishes? *The Irish Naturalists' Journal*. **33**, pp.124-137.

Quinlan, A.R. and Hall, I.M. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**(6), pp.841–842.

R Core Team 2021. R: A language and environment for statistical computing.

Ratner, B. 2009. The correlation coefficient: Its values range between 1/1, or do they. *Journal of Targeting, Measurement and Analysis for Marketing*. **17**(2), pp.139–142.

Ravinet, M., Faria, R., Butlin, R.K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M.A.F., Mehlig, B. and Westram, A.M. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology*. **30**(8), pp.1450–1477.

Ravinet, M., Westram, A., Johannesson, K., Butlin, R., André, C. and Panova, M. 2016. Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*. **25**(1), pp.287–305.

Recknagel, H., Elmer, K.R. and Meyer, A. 2014. Crater lake habitat predicts morphological diversity in adaptive radiations of cichlid fishes. *Evolution*. **68**(7), pp.2145–2155.

Recknagel, H., Hooker, O.E., Adams, C.E. and Elmer, K.R. 2017. Ecosystem size predicts eco-morphological variability in a postglacial diversification. *Ecology and Evolution*. **7**(15), pp.5560–5570.

Recknagel, H., Jacobs, A., Herzyk, P. and Elmer, K.R. 2015. Double-digest RAD sequencing using Ion Proton semiconductor platform (ddRADseq-ion) with nonmodel organisms. *Molecular Ecology Resources*. **15**(6), pp.1316–1329.

Reed, D.H. and Frankham, R. 2003. Correlation between Fitness and Genetic Diversity. *Conservation Biology*. **17**(1), pp.230–237.

Regan, C.T. 1909. The char of Ireland. *Irish Naturalist*. **18**, pp.3–6.

Reist, J.D., Power, M. and Dempson, J.B. 2013. Arctic charr (*Salvelinus alpinus*): a case study of the importance of understanding biodiversity and taxonomic issues in northern fishes. *Biodiversity*. **14**(1), pp.45–56.

Reusch, T.B.H., Ehlers, A., Hä, A. and Worm, B. 2005. Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *PNAS*. **102**, pp.2826–2831.

Rey, O. and Turgeon, J. 2007. Influence of historical events and contemporary estuarine circulation on the genetic structure of the banded killifish (*Fundulus diaphanus*) in the St. Lawrence River (Quebec, Canada). *Canadian Journal of Zoology*. **85**(8), pp.891–901.

Richards, E.J. and Martin, C.H. 2022. We get by with a little help from our friends: shared adaptive variation provides a bridge to novel ecological specialists during adaptive radiation. *Proceedings of the Royal Society B: Biological Sciences*. **289**(1975).

Riesch, R., Tobler, M., Lerp, H., Jourdan, J., Doumas, T., Nosil, P., Langerhans, R.B. and Plath, M. 2016. Extremophile Poeciliidae: multivariate insights into the complexity of speciation along replicated ecological gradients. *BMC Evolutionary Biology*. **16**(1), p.136.

Rivas, M.J., Saura, M., Pérez-Figueroa, A., Panova, M., Johansson, T., André, C., Caballero, A., Rolán-Alvarez, E., Johannesson, K. and Quesada, H. 2018. Population genomics of parallel evolution in gene expression and gene sequence during ecological adaptation. *Scientific Reports*. **8**(1), p.16147.

Roberts, D.R. and Hamann, A. 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. *Proceedings of the Royal Society B: Biological Sciences*. **282**(1804), p.20142903.

Robertson, J.M., Langin, K.M., Sillett, T.S., Morrison, S.A., Ghalambor, C.K. and Funk, W.C. 2014. Identifying Evolutionarily Significant Units and Prioritizing Populations for Management on Islands. *Monographs of the Western North American Naturalist*. **7**(1), pp.397–411.

Robinson, J., Kyriazis, C.C., Yuan, S.C. and Lohmueller, K.E. 2023. Deleterious Variation in Natural Populations and Implications for Conservation Genetics. *Annual Review of Animal Biosciences*. **11**(1), pp.93–114.

Des Roches, S., Post, D.M., Turley, N.E., Bailey, J.K., Hendry, A.P., Kinnison, M.T., Schweitzer, J.A. and Palkovacs, E.P. 2018. The ecological importance of intraspecific variation. *Nature Ecology and Evolution*. **2**(1), pp.57–64.

Rochette, N.C., Rivera-Colón, A.G. and Catchen, J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*. **28**(21), pp.4737–4754.

Rogers, S. and Bernatchez, L. 2007. The Genetic Architecture of Ecological Speciation and the Association with Signatures of Selection in Natural Lake Whitefish (*Coregonus* sp. Salmonidae) Species Pairs. *Molecular Biology and Evolution*. **24**(6), pp.1423–1438.

Rougeux, C., Bernatchez, L. and Gagnaire, P.A. 2017. Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of lake whitefish species pairs (*Coregonus clupeaformis*). *Genome Biology and Evolution*. **9**(8), pp.2057–2074.

Roux, C., Tsagkogeorga, G., Bierne, N. and Galtier, N. 2013. Crossing the species barrier: Genomic hotspots of introgression between two highly divergent *Ciona intestinalis* species. *Molecular Biology and Evolution*. **30**(7), pp.1574–1587.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. and Sanchez-Gracia, A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*. **34**(12), pp.3299–3302.

Ruddiman, W.F. 2008. *Earth's Climate: Past and Future* 2nd ed. W. H. Freeman, 2007.

Rundell, R.J. and Price, T.D. 2009. Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*. **24**(7), pp.394–399.

Ruzzante, D.E., McCracken, G.R., Salisbury, S.J., Brewis, H.T., Keefe, D., Gaggiotti, O.E. and Perry, R. 2019. Landscape, colonization, and life history: Their effects on genetic diversity in four sympatric species inhabiting a dendritic system. *Canadian Journal of Fisheries and Aquatic Sciences*. **76**(12), pp.2288–2302.

Ryder, O.A. 1986. Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology & Evolution*. **1**(1), pp.9–10.

Salisbury, S.J., McCracken, G.R., Perry, R., Keefe, D., Layton, K.K.S., Kess, T., Nugent, C.M., Leong, J.S., Bradbury, I.R., Koop, B.F., Ferguson, M.M. and Ruzzante, D.E. 2020. Limited genetic parallelism underlies recent, repeated incipient speciation in geographically proximate populations of an Arctic fish (*Salvelinus alpinus*). *Molecular Ecology*. **29**(22), pp.4280–4294.

Salisbury, S.J., Perry, R., Keefe, D., McCracken, G.R., Layton, K.K.S., Kess, T., Bradbury, I.R. and Ruzzante, D.E. 2023. Geography, environment, and colonization history interact with morph type to shape genomic variation in an Arctic fish. *Molecular Ecology*. **32**(12), pp.3025–3043.

Salisbury, S.J. and Ruzzante, D.E. 2021. Genetics causes and Consequences of Sympatric Morph Divergence in Salmonidae: A Search for Mechanisms. *Annual Review of Animal Biosciences*.

Saltykova, E., Siwertsson, A. and Knudsen, R. 2017. Parallel phenotypic evolution of skull-bone structures and head measurements of Arctic charr morphs in two subarctic lakes. *Environmental Biology of Fishes*. **100**(2), pp.137–148.

San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A., Tinner, W., Ballian, D., Beck, P., Birks, H.J.B., Eaton, E., Enescu, C.M., Pasta, S., Popescu, I., Ravazzi, C., Welk, E., Abad Viñas, R., Azevedo, J.C., Barbati, A., Barredo, J.I., Benham, S.E., Boca, R., Bosco, C., Caldeira, M.C., Cerasoli, S., Chirici, G., Cierjacks, A., Conedera, M., Da Ronch, F., Di Leo, M., García-Viñas, J.I., Gastón González, A., Giannetti, F., Guerrero Hue, N., Guerrero Maldonado, N., López, M.J., Jonsson, R., Krebs, P., Magni, D., Mubareka, S., Mulhern, G., Nieto Quintano, P., Oliveira, S., Pereira, J.S., Pividori, M., Räty, M., Rinaldi, F., Saura, S., Sikkema, R., Sitzia, T., Strona, G., Vidal, C., Vilar, L. and Zecchin, B. 2016. *European Atlas of Forest Tree Species* European Commission. Available from: https://forest.jrc.ec.europa.eu/en/european-atlas/.

Savvaitova, K. 1995. Patterns of diversity and process of speciation in Arctic char. *Nordic Journal of Freshwater Research*. **71**, pp.81–91.

Schenekar, T., Lerceteau-Köhler, E. and Weiss, S. 2014. Fine-scale phylogeographic contact zone in Austrian brown trout *Salmo trutta* reveals multiple waves of post-glacial colonization and a pre-dominance of natural versus anthropogenic admixture. *Conservation Genetics*. **15**(3), pp.561–572.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A. and Webster, M.S. 2010. Population diversity and the portfolio effect in an exploited species. *Nature*. **465**(7298), pp.609–612.

Schluter, D. 1996. Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **351**(1341), pp.807–814.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press.

Schmidt, C., Dray, S. and Garroway, C.J. 2022. Genetic and species-level biodiversity patterns are linked by demography and ecological opportunity. *Evolution*. **76**(1), pp.86–100.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. **9**(7), pp.671–675.

Schroeter, J.C., Maloy, A.P., Rees, C.B. and Bartron, M.L. 2020. Fish mitochondrial genome sequencing: expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA. *Conservation Genetics Resources*. **12**(3), pp.433–446.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C.S., Eroukhmanoff, F., Feder, J.L., Fischer, M.C., Foote, A.D., Franchini, P., Jiggins, C.D., Jones, F.C., Lindholm, A.K., Lucek, K., Maan, M.E., Marques, D.A., Martin, S.H., Matthews, B., Meier, J.I., Möst, M., Nachman, M.W., Nonaka, E., Rennison, D.J., Schwarzer, J., Watson, E.T., Westram, A.M. and Widmer, A. 2014. Genomics and the origin of species. *Nature Reviews Genetics*. **15**(3), pp.176–192.

Seehausen, O. and Wagner, C.E. 2014. Speciation in Freshwater Fishes. *Annual Review of Ecology, Evolution, and Systematics*. **45**(1), pp.621–651.

Sepkoski, J.J. 1997. Biodiversity: Past, Present, and Future. *Journal of Paleontology*. **71**(4), pp.533–539.

Sexton, J.P., Hangartner, S.B. and Hoffmann, A.A. 2014. Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*. **68**(1), pp.1–15.

Shaney, K.J., Diaz-Ramirez, L.G., Espindola, S., Castañeda-Rico, S., Berovides-Álvarez, V. and Vázquez-Domínguez, E. 2020. Defining intraspecific conservation units in the endemic Cuban Rock Iguanas (*Cyclura nubila nubila*). *Scientific Reports*. **10**(1).

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research*. **13**(11), pp.2498–2504.

Shikano, T., Järvinen, A., Marjamäki, P., Kahilainen, K.K. and Merilä, J. 2015. Genetic variability and structuring of Arctic Charr (*Salvelinus alpinus*) populations in northern Fennoscandia. *PLoS ONE*. **10**(10).

Siberchicot, A., Rezvoy, C., Charif, D., Gueguen, L. and Marais, G. 2020. MareyMap:Estimation of Meiotic Recombination Rates Using Marey Maps.

Sigursteinsdóttir, R.J. and Kristjánsson, B.K. 2005. Parallel Evolution, not Always so Parallel: Comparison of Small Benthic Charr, *Salvelinus alpinus*, from Grímsnes and Thingvallavatn, Iceland. *Environmental Biology of Fishes*. **74**(2), pp.239–244.

Sissons, J.B. 1977. Former Ice-Dammed Lakes in Glen Moriston, Inverness-Shire, and Their Significance in Upland Britain. *Transactions of the Institute of British Geographers*. **2**(2), p.224.

Siwertsson, A., Knudsen, R., Adams, C.E., Præbel, K. and Amundsen, P.A. 2013. Parallel and non-parallel morphological divergence among foraging specialists in European whitefish (*Coregonus lavaretus*). *Ecology and Evolution*. **3**(6), pp.1590–1602.

Siwertsson, A., Knudsen, R., Kahilainen, K.K., Praebel, K., Primicerio, R. and Amundsen, P.-A. 2010. Sympatric diversification as influenced by ecological opportunity and historical contingency in a young species lineage of whitefish. *Evolutionary Ecology Research*. **12**, pp.929–947.

Skoglund, S., Siwertsson, A., Amundsen, P.A. and Knudsen, R. 2015. Morphological divergence between three Arctic charr morphs - the significance of the deep-water environment. *Ecology and Evolution*. **5**(15), pp.3114–3129.

Skovrind, M., Olsen, M.T., Vieira, F.G., Pacheco, G., Carl, H., Gilbert, M.T.P. and Møller, P.R. 2016. Genomic population structure of freshwater-resident and anadromous ide (*Leuciscus idus*) in north-western Europe. *Ecology and Evolution*. **6**(4), pp.1064–1074.

van der Sluijs, I., Van Dooren, T.J.M., Hofker, K.D., van Alphen, J.J.M., Stelkens, R.B. and Seehausen, O. 2008. Female mating preference functions predict sexual selection against hybrids between sibling species of cichlid fish. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **363**(1505), pp.2871–2877.

Small, D., Clark, C.D., Chiverrell, R.C., Smedley, R.K., Bateman, M.D., Duller, G.A.T., Ely, J.C., Fabel, D., Medialdea, A. and Moreton, S.G. 2017. Devising quality assurance

procedures for assessment of legacy geochronological data relating to deglaciation of the last British-Irish Ice Sheet. *Earth-Science Reviews*. **164**, pp.232–250.

Smedley, R.K., Scourse, J.D., Small, D., Hiemstra, J.F., Duller, G.A.T., Bateman, M.D., Burke, M.J., Chiverrell, R.C., Clark, C.D., Davies, S.M., Fabel, D., Gheorghiu, D.M., McCarroll, D., Medialdea, A. and Xu, S. 2017. New age constraints for the limit of the British–Irish Ice Sheet on the Isles of Scilly. *Journal of Quaternary Science*. **32**(1), pp.48–62.

Smith, S.R., Amish, S.J., Bernatchez, L., Le Luyer, J., Wilson, C., Boeberitz, O., Luikart, G. and Scribner, K.T. 2020. Mapping of Adaptive Traits Enabled by a High-Density Linkage Map for Lake Trout. *G3: Genes, Genomes, Genetics*. **10**(June), g3.401184.2020.

Snorrason, S.S., Skúlason, S., Jonsson, B., Malmquist, H.J., Jónasson, P.M., Sandlund, O.T. and Lindem, T. 1994. Trophic specialization in Arctic charr *Salvelinus alpinus* (Pisces; Salmonidae): morphological divergence and ontogenetic niche shifts. *Biological Journal of the Linnean Society*. **52**(1), pp.1–18.

Soria-Carrasco, V., Gompert, Z., Comeault, A.A., Farkas, T.E., Parchman, T.L., Johnston, J.S., Buerkle, C.A., Feder, J.L., Bast, J., Schwander, T., Egan, S.P., Crespi, B.J. and Nosil, P. 2014. Stick Insect Genomes Reveal Natural Selection's Role in Parallel Speciation. *Science*. **344**(6185), pp.738–742.

Spencer, C.C.A., Deloukas, P., Hunt, S., Mullikin, J., Myers, S., Silverman, B., Donnelly, P., Bentley, D. and McVean, G. 2006. The Influence of Recombination on Human Genetic Diversity. *PLoS Genetics*. **2**(9), p.e148.

Stibal, M., Bradley, J.A., Edwards, A., Hotaling, S., Zawierucha, K., Rosvold, J., Lutz, S., Cameron, K.A., Mikucki, J.A., Kohler, T.J., Šabacká, M. and Anesio, A.M. 2020. Glacial ecosystems are essential to understanding biodiversity responses to glacier retreat. *Nature Ecology and Evolution*. **4**(May), pp.686–687.

Van Strien, M.J., Holderegger, R. and Van Heck, H.J. 2015. Isolation-by-distance in landscapes: Considerations for landscape genetics. *Heredity*. **114**(1), pp.27–37.

Stroud, J.T. and Losos, J.B. 2020. Bridging the Process-Pattern Divide to Understand the Origins and Early Stages of Adaptive Radiation: A Review of Approaches with Insights from Studies of Anolis Lizards *Journal of Heredity*. **111**(1), pp.33–42.

Stuart, Y.E., Veen, T., Weber, J.N., Hanson, D., Ravinet, M., Lohman, B.K., Thompson, C.J., Tasneem, T., Doggett, A., Izen, R., Ahmed, N., Barrett, R.D.H., Hendry, A.P., Peichel, C.L. and Bolnick, D.I. 2017. Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nature Ecology and Evolution*. **1**(6), pp.1–7.

Supek, F., Bošnjak, M., Škunca, N. and Šmuc, T. 2011. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*. **6**(7).

Taberlet, P., Fumagalli, L., Wust-Saucy, A. and Cosson, J. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*. **7**(4), pp.453–464.

Tamura, K., Stecher, G. and Kumar, S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*. **38**(7), pp.3022–3027.

Tarasov, L. and Peltier, W.R. 2006. A calibrated deglacial drainage chronology for the North American continent: evidence of an Arctic trigger for the Younger Dryas. *Quaternary Science Reviews*. **25**(7–8), pp.659–688.

Taylor, E.B., Foote, C.J. and Wood, C.C. 1996. Molecular genetic evidence for parallel life-history evolution within a pacific salmon (Sockeye Salmon and Kokanee, *Oncorhynchus Nerka*). *Evolution*. **50**(1), pp.401–416.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M. and Jeltsch, F. 2004. Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures. *Journal of Biogeography*. **31**(1), pp.79–92.

Tiddy, I.C., Schneider, K. and Elmer, K.R. 2023. Environmental correlates of adaptive diversification in post-glacial freshwater fishes. *Journal of Fish Biology*. **104**, pp.517-535.

Tigano, A., Weir, T., Ward, H.G.M., Gale, M.K., Wong, C.M., Eliason, E.J., Miller, K.M., Hinch, S.G. and Russello, M.A. 2023. Genomic vulnerability of a freshwater salmonid under climate change. *Evolutionary Applications*. **17**.

Turner, T.L., Hahn, M.W. and Nuzhdin, S. V 2005. Genomic Islands of Speciation in Anopheles gambiae. *PLoS Biology*. **3**(9), p.e285.

Tylmann, K., Rinterknecht, V.R., Woźniak, P.P., Guillou, V. and Team, A. 2022. Asynchronous dynamics of the last Scandinavian Ice Sheet along the Pomeranian ice-marginal belt: A new scenario inferred from surface exposure 10Be dating. *Quaternary Science Reviews*. **294**.

Vamosi, S.M. 2003. The presence of other fish species affects speciation in threespine sticklebacks. *Evolutionary Ecology Research*. **5**, pp.717–730.

Vera-Escalona, I., Habit, E. and Ruzzante, D.E. 2015. Echoes of a distant time: Effects of historical processes on contemporary genetic patterns in *Galaxias platei* in Patagonia. *Molecular Ecology*. **24**(16), pp.4112–4128.

Verspoor, E., Knox, D., Greer, R. and Hammar, J. 2010. Mitochondrial DNA variation in Arctic charr (*Salvelinus alpinus* (L.)) morphs from Loch Rannoch, Scotland: Evidence for allopatric and peripatric divergence. *Hydrobiologia*. **650**(1), pp.117–131.

Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology & Evolution*. **16**(7), pp.381–390.

Wagner, C.E., Harmon, L.J. and Seehausen, O. 2014. Cichlid species-area relationships are shaped by adaptive radiations that scale with area. *Ecology Letters*. **17**(5), pp.583–592.

Wagner, C.E., Harmon, L.J. and Seehausen, O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*. **487**(7407), pp.366–369.

Walker, A.F. 2007. Stream spawning of Arctic charr in Scotland. *Ecology of Freshwater Fish*. **16**(1), pp.47–53.

Walker, J., Gaffney, V., Fitch, S., Muru, M., Fraser, A., Bates, M. and Bates, R. 2020. A great wave: the Storegga tsunami and the end of Doggerland? *Antiquity*. **94**, pp.1409–1425.

Wang, J. 2022. Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs. *Heredity*. **129**(2), pp.79–92.

Waples, R. 1991. Pacific Salmon, Oncorhynchus spp., and the Definition of 'Species' Under the Endangered Species Act. *Marine Fisheries Review*. **53**(3), pp.11–22.

Waples, R. and Lindley, S.T. 2018. Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids, *Evolutionary Applications.* **11**, pp.1518-1526.

Waples, R., Ford, M.J., Nichols, K., Kardos, M., Myer, J., Thompson, T.Q., Anderson, E.C., Koch, I.J., McKinney, G., Miller, M.R., Naish, K., Narum, S.R., O'Malley, K.G., Pearse, D.E., Pess, G.R., Quinn, T.P., Seamons, T.R., Spidle, A., Warheit, K.I. and Willis, S.C. 2022. Implications of Large-Effect Loci for Conservation: A Review and Case Study with Pacific Salmon. *Journal of Heredity*. **113**(2), pp.121-144.

Waters, J.M., Emerson, B.C., Arribas, P. and McCulloch, G.A. 2020. Dispersal Reduction: Causes, Genomic Mechanisms, and Evolutionary Consequences. *Trends in Ecology and Evolution*. **35**(6), pp.512–522.

Webb, P.W. 1984. Body Form, Locomotion and Foraging in Aquatic Vertebrates. *American Zoologist*. **24**, pp.107–120.

Weber, A.A.-T., Rajkov, J., Smailus, K., Egger, B. and Salzburger, W. 2021. Speciation dynamics and extent of parallel evolution along a lake-stream environmental contrast in African cichlid fishes. *Science Advances*. **7**(45).

Weeks, A.R., Stoklosa, J. and Hoffmann, A.A. 2016. Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: The case of Australian mammals. *Frontiers in Zoology*. **13**(1).

Weinstein, S.Y., Thrower, F.P., Nichols, K.M. and Hale, M.C. 2019. A large-scale chromosomal inversion is not associated with life history development in rainbow trout from Southeast Alaska. *PLoS ONE*. **14**(9).

Weninger, B., Schulting, R., Bradtmöller, M., Clare, L., Collard, M., Edinborough, K., Hilpert, J., Jöris, O., Niekus, M., Rohling, E.J. and Wagner, B. 2008. The catastrophic final flooding of Doggerland by the Storegga Slide tsunami. *Documenta Praehistorica*. **35**, pp.1–24.

Westley, P.A.H., Quinn, T.P. and Dittman, A.H. 2013. Rates of straying by hatchery-produced Pacific salmon (*Oncorhynchus* spp.) and steelhead (*Oncorhynchus mykiss*) differ among species, life history types, and populations. *Canadian Journal of Fisheries and Aquatic Sciences*. **70**(5), pp.735–746.

Wheeler, A. 1977. The Origin and Distribution of the Freshwater Fishes of the British Isles. *Journal of Biogeography*. **4**(1), pp.1–24.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. [Online]. [Accessed 18 August 2023]. Available from: https://ggplot2.tidyverse.org.

Willi, Y., Fracassetti, M., Zoller, S. and Van Buskirk, J. 2018. Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution*. **35**(4), pp.781–791.

Wilson, A.J., Gíslason, D., Skúlason, S., Snorrason, S.S., Adams, C.E., Alexander, G., Danzmann, R.G. and Ferguson, M.M. 2004. Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. *Molecular Ecology*. **13**(5), pp.1129–1142.

Wilson, P., Ballantyne, C.K., Benetti, S., Small, D., Fabel, D. and Clark, C.D. 2018. Deglaciation chronology of the Donegal Ice Centre, north-west Ireland. *Journal of Quaternary Science*. **34**(1), pp.16–28.

Winfield, I.J., Bean, C.W., Gorst, J., Gowans, A.R.D., Robinson, M. and Thomas, R. 2013. Assessment and conservation of whitefish (*Coregonus lavaretus (L.)*) in the U.K. *Advances in Limnology*. **64**, pp.305–321.

Winfield, I.J., Bean, C.W. and Hewitt, D.P. 2002. The relationship between spatial distribution and diet of Arctic charr, *Salvelinus alpinus*, in Loch Ness, U.K. *Environmental Biology of Fishes*. **64**, pp.63–73.

Winfield, I.J., Fletcher, J.M. and James, J.B. 2008. The Arctic charr (*Salvelinus alpinus*) populations of Windermere, UK: Population trends associated with eutrophication, climate change and increased abundance of roach (*Rutilus rutilus*). *Environmental Biology of Fishes*. **83**(1), pp.25–35.

Winfield, I.J., Hateley, J., Fletcher, J.M., James, J.B., Bean, C.W. and Clabburn, P. 2010. Population trends of Arctic charr (*Salvelinus alpinus*) in the UK: Assessing the evidence for a widespread decline in response to climate change. *Hydrobiologia*. **650**(1), pp.55–65.

Wolf, J.B.W. and Ellegren, H. 2017. Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*. **18**(2), pp.87–100.

Woodman, P., McCarthy, M. and Monaghan, N. 1997. The Irish Quaternary Fauna Project. *Quaternary Science Reviews*. **16**(2), pp.129–159.

Woods, R., Barnes, I., Brace, S. and Turvey, S.T. 2021. Ancient DNA Suggests Single Colonization and Within-Archipelago Diversification of Caribbean Caviomorph Rodents. *Molecular Biology and Evolution*. **38**(1), pp.84–95.

Wright, S. 1931. Evolution in Mendelian Populations. *Genetics*. **16**(2), pp.97–159.

Yeaman, S. 2013. Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences*. **110**(19), pp.1743–1751.

Yeaman, S., Aeschbacher, S. and Bürger, R. 2016. The evolution of genomic islands by increased establishment probability of linked alleles. *Molecular Ecology*. **25**(11), pp.2542–2558.

Yeaman, S. and Whitlock, M.C. 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution*. **65**(7), pp.1897–1911.

Zawierucha, K., Kolicka, M., Takeuchi, N. and Kaczmarek, L. 2015. What animals can live in cryoconite holes? A faunal review. *Journal of Zoology*. **295**(3), pp.159–169.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C. and Weir, B.S. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. **28**(24), pp.3326–3328.

Zhou, X., Carbonetto, P. and Stephens, M. 2013. Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *PLoS Genetics*. **9**(2).

Zhou, X. and Stephens, M. 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*. **11**(4), pp.407–409.

Zhou, X. and Stephens, M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*. **44**(7), pp.821–824.