Environmental Adaptation in Three-spined Stickleback

Mahmuda Begum

Thesis submitted to the University of Nottingham for the fulfilment of the degree of Doctor of Philosophy

June 2021

Abstract

Environmental adaptation of a species driven by different agents of natural selection is a major focus of eco-evolutionary research. In this thesis, I aimed to understand the role of selection in shaping phenotypic and genotypic adaptation of a model fish, the three-spined stickleback (*Gasterosteus aculeatus*) in relation to different environmental factors. Firstly, I found significant associations between the pH of loch water and patterns of morphological diversity when comparing adaptive divergence on two neighbouring islands, North and South Uist. This suggests the influence of an abiotic agent on natural selection which might be responsible for the adaptive divergence of closely-related lineages. Secondly, I observed substantial variation in phenotypic and genotypic traits among hybrids of ancient marine anadromous and newly-adapted freshwater stickleback populations in a natural hybrid zone in North Uist. Phenotypic traits such as standard length of fish, total plate count and parasite abundance showedmosaic patterns, whereas body shape, all spines and pelvic structure exhibited clinal patterns across the hybrid zone due to adaptive divergence between parental populations. Genetic diversity was also observed using adaptive nuclear and mitochondrial markers of specific candidate loci (Eda, PPARA, WNT7B, NLRC5 and Cytochrome b gene) under selection. There was strong genotypic differentiation between marine and freshwater stickleback populations due to strong direct selection across the geographical range from sea to loch. The phenotype-genotype association of selected adaptive traits indicates the evolutionary consequences of adaptation, by identifying signatures of selection on genomic regions underlying phenotypic traits. Third, I found strong genetic links to parasite resistance underlying adaptation to infection, by examining parasitic abundance in stickleback in both natural and experimental environments. The genotypes of selected adaptive loci (Eda, PPARA, WNT7B and NLRC5) varied in relation to the common ectoparasite (*Gyrodactylus* sp.) and freshwater endoparasite (*Diplostomum pseudospathecum*), suggesting a role for parasite-mediated selection in the ecological adaptation of stickleback. Finally, I found both environmental and genetic influences on the relative abundance and diversity of bacterial species in the stickleback skin microbiome. In conclusion, the research presented in this thesis provides a significant insight into divergent phenotypic and genotypic adaptation driven by environmental selection to understand the mechanism of ecological speciation in nature.

Acknowledgments

All praise to Allah, the Almighty, who has granted me the blessings, the strength and the chance to accomplish my PhD study.

I would like to express my sincere gratitude to my supervisor, Professor Andrew MacColl, for his time, patient guidance and encouragement to complete my PhD project. I have been extremely lucky to have a supervisor who supported me at every stage of my project.

I also thank my second supervisor, Sharon Egan for her support and help for carrying out microbiological lab work in the first year of my PhD. I am very grateful to all other members of the MacColl group: Ann Lowe, Laura Dean, Jim Whiting, Talib Chitheer, Dan D'Agostino, Iain Hill and Anthony Ducker who provided continuous assistance both in the lab, and in the field. Many thanks to Dr. Angus Davison, Prof. Jan Bradley, Prof. John Brookfield and all the members of evolutionary and ecology research group for their cordial helps and critical suggestions at different levels of studies.

I am immensely grateful to my friend and research fellow, Victoria Nolan for her crucial supports and relentless encouragement at various stages of my research project. I would like to thank all of my office mates specially Lisa, Imad, Aram, Talib, Meshal, Chidinma, Mehrab and Abdul Aziz for their amiable cooperation at my office.

I could not have completed my whole journey without the love and support of my whole family, particularly two great ladies, my mother and my mother-in-law. I would like to dedicate my thesis to the memories of my beloved father and father-in-law, who let me dream for accomplishing PhD but left me before completion of the degree. I always express my gratitude to my siblings who believed in my ability to be successful in every sphere of life. Above all, I owe hugely to a very special person, my husband, Mizan for his enduring love, wholehearted support and understanding to complete this mission and to fulfil the dream of my life.

Finally, I am particularly appreciating University of Nottingham and Bangabandhu Science and Technology Fellowship Trust Bangladesh who awarded me the prestigious Vice-Chancellor's Scholarship for Research Excellence (International) for research grant and the Bangabandhu international fellowship for living cost and all other allowances.

Contributions

The research work presented in this thesis is my own, however, several people made important contributions to individual chapters as follows:

Chapter 2: Andrew MacColl and Laura Dean assisted me in sample collection from South Uist and Iain Hill and Anthony Ducker helped me on North Uist.

Chapter 3: Andrew MacColl, Laura Dean, Iain Hill and Anthony Ducker assisted me to collect and process samples from the Hosta hybrid zone on North Uist. It was Andrew MacColl's idea to apply the SNP-based assays for genetic analysis and Angus Davison guided me to develop the PCR-RFLP method targeting single gene. Laura Dean helped me in using Unix to select genes from the whole genome sequencing data provided by Andrew MacColl.

Chapter 4: Ann Lowe and Anthony Ducker helped me to conduct all parasite experiments, creating F2 hybrids, snail collection for *Diplostomum* and overall laboratory management.

Chapter 5: Andrew MacColl, Jim Whiting and Laura Dean assisted me in microbiome sample collection and processing on North Uist. Sharon Egan helped me to learn molecular techniques for individual bacteria identification and DGGE method.

For chapter 2-5, Andrew MacColl and Victoria Nolan advised me in conducting statistical analysis and made critical comments on overall writing of my thesis.

Table of contents

Chapter 1: General Introduction

1.1 Environmental adaptation

Adaptation (derived from the Latin word "adaptare": ad + aptus = "to" + "fitting") is simply the adjustment of an organism to fit into its environment. In evolutionary biology, adaptation or environmental adaptation, synonymous with adaptive evolution, is the biological process by which a species or population becomes accustomed to their environment, or changes to improve their chances of survival and reproduction under natural selection in a particular environment (Cheplick, 2020). According to Futuyma (2005): "an adaptation is a characteristic that changes to enhance the survival and reproduction of organisms that bear it, relative to alternative character states". An adaptive trait of an organism can be physical, physiological or behavioural, underlying a genetic basis of an inherited system that enables them to fit in their ecological niches (Dobzhansky, 1951). It may include changes in size or shape (e.g. long necks of giraffes, *Giraffa camelopardalis*), searching for food, mating (e.g., songbirds or attractive feathers in birds) or protective colouration or features (e.g., leafy sea dragon fish, *Phycodurus eques*) (Schaffner and Sabeti, 2008). The phenomenon of adaptation is the central topic of evolutionary biology because of its causal influence on evolutionary change through the process of natural selection (Amundson, 1996).

1.2 Natural selection and its role in environmental adaptation

After Charles Darwin and Alfred Russel Wallace first proposed the theory of evolution in 1858, natural selection is now considered the process that drives successful adaptation. This tends to increase an organisms' advantageous characteristics through reproduction in the face of environmental challenges (Darwin and Wallace, 1858; Darwin, 1859). Natural selection or "survival of the fittest" is one of the mechanisms of evolutionary change which drives the evolution of adaptation (Endler, 1986; Gregory, 2009). The contemporary evolutionary theory in the 20th century integrated Mendelian genetics with natural selection: genotypes with higher fitness increase in frequency in a population through differential reproduction (Barton et al., 2007). Positive natural selection occurs when the adaptive trait is beneficial and heritable to increase the chance of survival of a population (e.g., hinged jaw in snakes or grinding teeth in horses), while negative (or purifying) selection reduces the chance of survival, mating and reproduction by eliminating the harmful traits of an individual (e.g. sickle cell anaemia and haemophilia in humans reduces the survival chances after adulthood) (Schaffner and Sabeti, 2008; Zhang, 2008). The process of natural selection generally leads to the differential reproduction of genetically-distinct individuals by altering allele frequencies or the overall

genetic variation (Li, 1997). However, the environmental conditions experienced by a population influences the genetic variations (allele frequency) produced by mutations (a random change to genetic material) and contribute to selecting the fit or removing the unfit traits over a series of generations (Loewe, 2008).

Natural selection on adaptive traits can be studied by examining its effect on the distribution of phenotypes or genotypes of individuals in a population (Kingsolver et al., 2001; Gregory, 2009). Selection can influence the level of phenotypic variation and the variability of underlying genotypes within a population in several ways: 1. Directional selection, 2. Stabilizing selection and 3. Diversifying selection (Endler, 1986; Rolhauser and Pucheta, 2017). When different aspects of the environment changes, directional (or progressive) selection occurs by favouring a single phenotype to shift over time in one direction of the spectrum of existing variation (Endler, 1986). For example, smaller size of Atlantic salmon (*Salmo salar*) has been favoured over time due to several environmental factors including overfishing of large size fish (Karlsson and Karlström, 1994; Lobón-Cerviá, 2005; Jutila et al.,2006). This type of selection generally reduces genetic diversity by fixing a selected allele and monomorphic locus (Freeland, 2005). Stabilizing (or balancing) selection occurs when selection acts against two extremes of a trait and reduces diversity by favouring the intermediate phenotypes along with the associated genotypes (Bell, 1997). Several phenotypic traits in animals including height, weight, number of progeny and size of infants of human show a greater number or a higher chance of surviving when at the intermediate values (Sanjak et al., 2018). Selection acting against an intermediate trait favouring two or more diversifying phenotypes in an environment is known as disruptive or diversifying selection (e.g. morphologies and protective colouration or mimicry against predators) (Loewe, 2008). This type of selection is uncommon but is the most likely to increase variation and may cause speciation (Anon, 2020).

In addition, frequency-dependent-selection (a dynamic form of balancing selection that occurs when the fitness of a phenotype or genotype depends on its frequency in a population) and sexual selection (differential selection on males and females) are often demonstrated as ways of natural selection in local adaptation of a lineage or species (Pfennig et al., 2001; Figueredo et al., 2005; Noonan and Comeault, 2009). However, various selective pressures (or external agents that favour or disfavour traits) due to changes regarding different physical and biological aspects of environment may work together to influence the phenotypic and genetic variances within a population (Wadgymar et al., 2017). Therefore, natural selection as one of the key mechanisms of evolution (other than genetic drift and gene flow) can constantly make populations

adapted to their particular ecological circumstances by transferring heritable variation over generations.

1.3 Agents of selection in adaptive evolution

Ecological divergence through changes of a phenotype supported by the genotype of a species is a compelling concept to help understand the cause of evolution in nature. Following Darwin's theory of evolution as a result of natural selection, explicit research on evolutionary biology has been conducted to understand how adaptation, divergence and speciation occur in natural populations, but very little is known about the agent of selection (MacColl, 2011). It is challenging to address the interaction between the agents and the underlying traits in the process of natural selection leading to the evolution of a new species (Endler, 1986; Kingsolver et al., 2011; MacColl, 2011). Phenotypic variation in response to environmental factors favours natural selection and adaptation to a divergent habitat, but the changes in the genetic level along with such traitsare difficult to identify for a particular selective agent (Mills et al., 2014). Many traits of humans, animals and plants that vary among different populations along with genetic variation, are influenced by natural selection. A recent study on modern humans provided evidence of the action of natural selection acting on phenotypic and genotypic characteristics among populations. This study demosntrated the association of genetic variants with several phenotypic traits in women (15 traits) and men (10 traits). For instance, men with a higher body mass index had more children, and extreme traits such as height reduced their lifetime reproductive success by increasing or decreasing the fitness in the population indicating the action of directional and stabilizing selection respectively, yet did not identify any agent of selection (Sanjak et al., 2018). Therefore, it is necessary to study the causes of natural selection that exert variation in traits of a natural population, or adaptive evolution of a species in response to environmental changes.

Any environmental factor (biotic and abiotic factors) can act as an agent of selection resulting in covariance between traits and fitness of a population in local adaptation (Wade and Kalisz, 1990; MacColl, 2011; Wadgymar et al., 2017). Many studies on the physical aspects of environment (e.g. water, temperature, wind, light, soil, salinity, pH, nutrients) have documented spatiotemporal variation causing variation in the frequency of key traits (phenotypic and genotypic) and differential fitness in many plant and animal species (McCluskey et al., 1993; Little and Fabacher, 1994; Williamson et al., 1997; Garrido et al., 2012; Liancourt, et al., 2013; Maes et al., 2014). For example, the wind load of a cyclone can act as a selective pressure on the directional selection of new world *Arecaceae* (palm): trees from cyclone-prone areas exhibited a much lower range of tolerance to wind and significantly lower percentage of mortality than the cyclone-free

provinance (Griffith et al., 2008). Solar ultraviolet radiation (UVR) levels also have been shown to be an important abiotic agent of natural selection for the spawning depth, hatching success and recruitment success of a freshwater fish, yellow perch (*Perca flavescens*) (Williamson et al., 1997). In addition, genetic diversity within and between bacterial populations of two model marine microbes, *Prochlorococcus* and *Pelagibacter* in the Atlantic and Pacific Oceans has been linked to the availability of phosphorous as the dominant selective agent (Coleman and Chisholm, 2010).

Biotic agents including predators, competitors, prey, natural enemies, mating partners, availability of food sources, vegetation, parasites and disease also exert selection pressures resulting in local adaptation of organisms (Reznick and Endler, 1982; Grant and Grant, 2006; Calsbeek and Smith, 2007; Pfennig et al., 2007; Calsbeek, and Cox, 2010; Wadgymar et al., 2017). The most classic example is the adaptation of Darwin's finch beak size in relation to the availability of resources and intraspecific competition influencing the optimal fitness of finch species (*Geospiza fortis* compete with *G. magnirostris*) (Gibbs and Grant, 1987; Grant and Grant, 2006). Likewise, Pfennig et al., (2007) demonstrated divergence in the trophic morphology (body size) of two species of spadefoot toads (*Spea bombifrons* and *S. multiplicate*) in relation to natural selection driven by intraspecific and interspecific competition for food. This study found that when a single species was present in a pond, disruptive selection favoured extreme trophic phenotypes in both species due to resource polymorphism, but in the presence of two species and resource competition, hetero-specific competitors were selectively favoured with different phenotypes. Several studies on intra-speicific food competition in different species of fish such as humbug damselfish (*Dascyllus aruanus*), sunfish (*Lepomis* spp.), arctic charr (*Salvelinus alpinus*) and rainbow trout (*Oncorhynchus mykiss*) also found strong association with variation of body size (Forrester, 1991; Walton et al., 1994; Post et al., 1999), trophic morphology (Malmquist, 1992) and physiology (metabolic or growth rate) (McCarthy et al., 1994; McCarthy, 2001) in response to competitive environmental condition.

Predation especially can act as a strong causative agent to drive selection, and has shown significant impacts on the adaptive evolution of mammals, reptiles, amphibians, fish and insects (Brodie III, 1993; King and Lawson, 1997; Pfennig et al., 2001; Langerhans et al., 2004; Kuchta, 2005; Shepard, 2007; Noonan and Comeault, 2009; Vignieri et al., 2010). Adaptive radiation of plant-feeding stick insects (*Timema cristinae*) exhibited strong divergent selection in body size, shape and colour pattern, associated with the presence of visual predators (Nosil and Crespi, 2006). In response to the directional selection driven by the abundance of predators, natural populations of

guppies (*Poecilia reticulata*) experienced high mortality rates along with earlier age maturity and smaller size in high-predation communities than the low-predation communities (Reznick and Endler, 1982; Reznick et al., 1996; Reznick et al., 1997). Noonan and Comeault (2009) demonstrated how the bright colour phenotypic variation of Neotropical poison frogs, *Dendrobates tinctorius*, which assists with avoiding visual (avian) predators, is influenced by sexual selection and mate choice. However, parasites also play a significant role in the appearance, behaviour, condition and defence systems of host populations, but there has been little attention in parasite-mediated selection in comparison to other selective agents (Price et al., 1986; MacColl, 2011; Karvonen and Seehausen, 2012).

The extent of local adaptation can be influenced by the interaction of multiple agents of selection which can vary across the life history of an organism at diverse geographical ranges (McPeek, 1996; Wadgymar et al., 2017). Both biotic and abiotic can shape the distribution of a species across the spatial scales, but it is still challenging to distinguish interactions that drive environmental adaptation of local populations (Vamosi, 2005; Fraterrigo et al., 2014; Lewis et al., 2017). The ecological theory of adaptive evolution suggests that divergent natural selection emerging from the interaction of extrinsic (biotic and abiotic aspects of the environment) and intrinsic factors (lineage specific traits including genetic architecture) may be responsible for the bursts of diversification of many evolutionary lineages (Schluter, 2000; Wagner et al., 2012; Magalhaes et al., 2016). For example, Wagner et al., (2012) found an interaction between extrinsic environmental factors (e.g., lake depth, solar radiation, lake age) and intrinsic factors related to sexual selection (e.g. sexual dichromatism), that strongly influenced the adaptive diversification of cichlids lineages in African lakes. In addition, symbiotic microbes (such as endophyte fungi) have been detected as an internal agent of natural selection within the bodies of plants and animals, by exhibiting interactive effects on host genotype and phenotype to maximize their evolutionary fitness (Cheplick, 2015). Therefore, experimental and field studies on the identification of specific agents of selection need to be conducted to reveal the mechanism of environmental adaptation and speciation in nature.

1.4 Phenotypic and genotypic basis of adaptive evolution

Phenotypic diversity as a consequence of either genetic differentiation or phenotypic plasticity can represent the adaptive responses to divergent natural selection depending on the biotic and abiotic environment of a species (Endler, 1977; Schluter, 2000; Agrawal, 2001; Albertson et al., 2003; Langerhans et al., 2004). A number of studies on different fish species including the Trinidadian guppy (*P. reticulata*), Eastern mosquito

fish (*Gambusia affinis*), Swordtails (*Xiphophorus nigrensis*), Costa Rican live-bearing fish (*Brachyrhaphis rhabdophora*) and Jewelled Split fin (*Xenotoca variata*) have documented the interaction of predators with abiotic aspects of the environment that cause phenotypic divergence and differentiation (Reznick et al., 1997; Johnson and Belk, 2001; Magurran and Phillip, 2001; Langerhans et al., 2004; Horth, 2004; Moyaho et al., 2004). By studying the role of environmental agents in promoting phenotypic diversity between distinct habitat types such as saltwater and freshwater, the repeated evolution of closely-related lineages can be explored in eco-evolutionary studies of other species of fish (Berner et al., 2008).

About 80% of studies on selection have focused on morphological traits rather than other quantitative traits such as behaviour, physiology or life-history traits (Kingsolver et al., 2001). This might be because of the correlation of phenotypes with fitness due to the direct impact on reproduction and survival or indirect effect on other traits (Lande and Arnold, 1983; Arnold and Wade, 1984; Linnen and Hoekstra, 2009). A long-term observation on the survival, reproduction and phenotypic traits of Darwin's finches (*Geospiza fortis*) demonstrated that the strength of direct and total (direct and indirect) selection favoured changes in body and beak size, along with no reproduction and high mortality of the finches when faced with the challenges of drought and food availability (Grant and Grant, 1995, 2002, 2011). When only large and hard seeds were available, selection favoured large body and beak size, whereas an availability of small and soft seeds resulted in selection favouring smaller traits (Boag and Grant, 1981; Price et al., 1984). However, adaptive diversification observed in some closely-related lineages of animals, but not other lineages, have created a fundamental challenge to understand the process of evolution (Losos, 2010; Wagner et al., 2012; Meier et al., 2017). To understand the adaptive significance of natural selection between alternative environments, it is firstly important to identify the environmental agents driving selection in association with phenotypic and genotypic variation of closely-related populations or divergent lineages.

Understanding the genetic and molecular basis of adaptation driven by natural selection is the central goal of modern eco-evolutionary research (Jones et al., 2012a). Strong selection in association with species interactions may create inherited differences in fitness in terms of fecundity, survival rates or some other trait to increases the chances of survival in a particular set of conditions (Endler, 1986; Linnen and Hoekstra, 2009). Genetic changes occur as a consequence of phenotypic changes, because selection increases the frequency of the favoured allele (positive selection) or decreases it to eliminate the harmful allele (negative selection) as genetic fitness (Loewe, 2008). The

selective advantage of favoured alleles means the organism can better adapt to survive, reproduce and transfer these alleles to the next generation under challenging environmental conditions. The main causes of genetic variability of a species mainly include mutation (e.g., deletion, insertion or changes in the nucleotide in a specific position), random mating, migration and recombination (Griffiths et al., 2005).

The biotic and abiotic agents of selection can create evolutionary changes in a population gene pool by favouring particular combinations of genes that optimize the growth, development, morphology and reproductive ability of the individual (Pigliucci and Preston, 2004; Gowaty, 2005; Cheplick, 2015). For instance, variation in climate (seasons based on temperature, humidity, precipitation, UV radiation) exerts selective pressures on candidate genes of human metabolic disorders related to energy metabolism of heat or cold tolerance (Hancock et al., 2008). In another example, mutations in the *Agouti* locus were identified as responsible for the adaptive coat colouration in deermice (*Peromyscus maniculatus*), leading to a novel wider hair band and lighter camouflaging colour under environmental selection pressure (avian predator) in the light colour soil of the Nebraska Sand Hills, USA (Linnen et al., 2009, 2013). However, the adaptive loci or genes under selection involved in parallel divergence into two different habitats are still largely unknown (Jones et al., 2012b), and need to be identified as a signature of selection along with phenotypic divergence among populations of a species.

The role of natural selection in driving adaptive evolutionary change can also be observed from the spatial distribution of genotypes and phenotypes across a geographical gradation (Haldane, 1948; Endler, 1977; Barton, 1979). The variation in spatial patterns of allele frequencies or trait means is an indication of the strength of selection in determining the genetic structure of natural populations (Slatkin, 1973; Hoekstra et al., 2004). If gene flow is large enough due to migration, the populations will be homogenised in the absence of selection, whereas a step change in the environment creates selective pressure in gene flow which ultimately results in reproductive isolation and potential speciation events (Slarkin, 1985; Lenormand, 2002). Therefore, the structure of a hybrid zone can provide insight into the genetic basis of such traits, by maintaining distinct clines (gradation from one extreme form to the other along a geographical transition) of phenotypic and genotypic traits through the action of natural selection (Huxley, 1939; Barton and Hewitt, 1985).

Different types of selection combined with gene flow across populations may determine the shape and width of clines in the hybrid zone, where selection acts to generate

barriers to gene flow between two populations or species (Alexandrino et al., 2005). Disruptive selection occurs when selection on traits or alleles acts against hybrids whereas divergent selection favours different traits of two species located at opposite ends of the cline, depending on environmental selection (Campbell et al., 2018). Strong disruptive selection along with low dispersal rates produces a narrow cline resembling a step function in extreme cases. Conversely, weak selection produces wide, smooth sigmoid clines due to high dispersal and environmental selection against trait values of parental forms maintaining a stable equilibrium (Barton and Hewitt, 1985; Pedersen et al., 2017). There are several examples of clines only on phenotypic traits (Littlejohn and Watson, 1973; Endler, 1977; Hairston et al., 1992; Nürnberger et al., 1995; Gay et al., 2008; Campbell et al., 2018) or allele frequencies (Mallet et al., 1990; Ross and Harrison, 2002; Sotka et al., 2004; Hedrick, 2006; De La Torre et al., 2015), but clinal variation of both phenotypes and genotypes of specific genes are rarely examined in a natural hybrid zone (Mullen and Hoekstra, 2008; Linnen and Hoekstra, 2009; Vines et al., 2016; Pedersen et al., 2017). Therefore, to understand the association between environmental characteristics and the selection gradient, the analysis of a natural hybrid zone should be conducted, targeting both phenotypic and genotypic traits among populations of diversifying taxa.
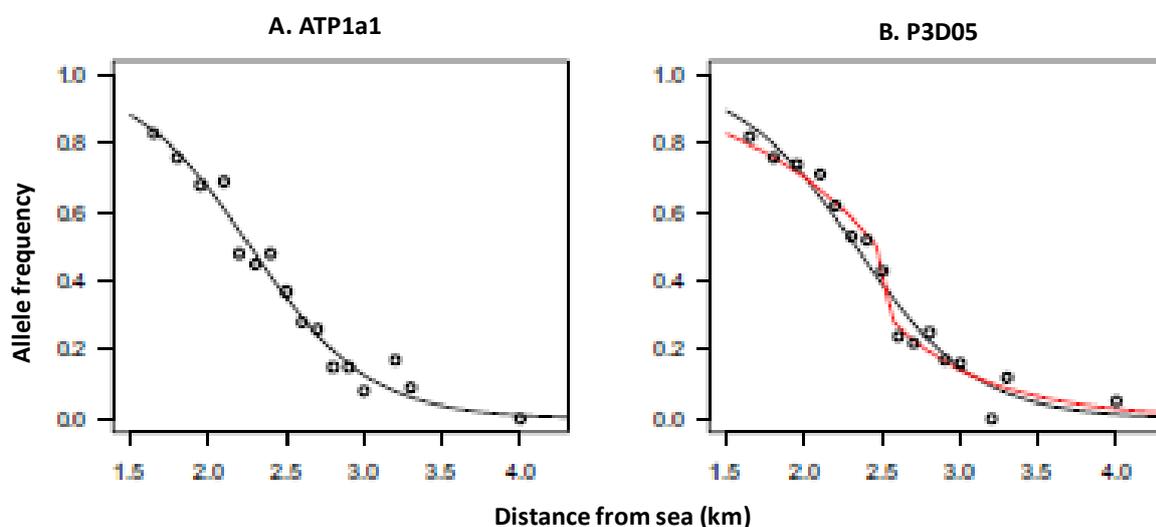
Fig. 1 Pattern of distribution of a sigmoid (black) and a stepped (red) cline for allele frequencies of two loci (ATP1a1 and P3D05 as for example) in three-spined stickleback (*Gasterosteus aculeatus*) at a geographical scale (km) from saltwater to freshwater habitat. Adapted from Vines et al., 2016.

The genetic basis of parasite-mediated selection is also an important study objective to understand adaptive evolution of a host species in response to a single agent of selection (Little, 2002). Parasites as an agent of evolutionary change under selection are known to

influence the life-history, behavioural, reproductive and fitness (defence) traits of their host, which may lead to genetic variation in parasite resistance (Minchella, 1985; Hart, 1994; Møller, 1997; Penn and Potts, 1998; Wilson et al., 2001; Rigby et al., 2002; Hedrick, 2002; Thomas, et al., 2005; Fredensborg and Poulin, 2006). The genetic variation in parasite resistance has demonstrated the occurrence of different types of selection such as directional, disruptive and stabilizing selection in several host populations (Carius et al., 2001; Clayton et al., 2005; Duncan and Little, 2007; Duffy et al., 2008;). For instance, Blanchet, et al, (2009) found a strong correlation between host genetic diversity and ectoparasitic load (*Tracheliastes polycolpus*), by observing disruptive selection with a higher infection rate in the intermediate level of heterozygosity of the fish host (Rostrum dace, *Leuciscus leuciscus*). A few studies based mainly on variation in gene expression levels of immune system of fish provide baseline evidence for genetic associations in response to parasitic infection (Robertson et al., 2017). For example, *Gyrodactylus* sp. a monogenean ectoparasite of freshwater and saltwater fish, were found to be associated with changes in the expression of immune-related genes in several species of fish including guppies (Cable and Van Oosterhout, 2007), rainbow trout (*O. mykiss*) (Lindenstrøm et al., 2004) and Atlantic salmon (Lindenstrøm et al., 2006).

The immune-related genes of major histocompatibility complex (MHC) are the most commonly studied example of the maintenance of genetic variation of the host population to parasite resistance or susceptibility (Hedrick, 2002; Penn et al., 2002; Wegner et al., 2003). The different alleles of these polymorphic genes can recognize antigens from a pathogen and play a major role in determining the level of resistance to various infectious agents including parasites (Paterson et al., 1998; Langefors et a., 2001; Hedrick et al., 2001; Blanchet et al., 2009). However, it seems difficult to identify specific genes for complex parasite resistance traits, which may be due to variable innate and adaptive immune response of the host (especially fish) towards a diverse range of parasite (Bernatchez and Landry, 2003; Tort et al., 2003). Most studies of genetic diversity of the host in response to a parasite infection are conducted in domesticated and laboratory populations under controlled conditions and targeting a specific parasite (Kloosterman et al., 1992; Sorci et al., 1997; Ebert et al., 1998; Carius et al., 2001; Capaul and Ebert, 2003). It has been assumed that the observation of genetic diversity in parasite resistance in a wild host population cannot reflect the direct response due to the unevenness and variety of parasite exposures to the host (Henter and Via, 1995; Little and Ebert, 2001). For this, observational natural and experimental studies of parasite resistance in association with specific candidate genes need to be conducted to understand the adaptive evolution of a species under selection (Wade and Kalisz, 1990).

16

1.5 Molecular approaches to study adaptive evolution

The recent advancement of molecular approaches including high-throughput genotyping using whole-genome sequencing, has facilitated the identification of genes associated with complex traitsthat have evolved through natural selection in various vertebrate species (Shimada et al., 2011; Jones et al., 2012a). In a bottom-up approach, the signature of selection is screened from a genome-scale sequence to estimate the proportion of the genome under selection, or to pick out putative loci with future functional importance. In comparison, a top-down approach involves selecting a candidate gene or mutation that is known to be associated with the specific phenotype under selection (Nielsen, 2005; Eyre-Walker, 2006; Olsen et al., 2007; Sella et al., 2009; Linnen et al., 2009; Jones et al., 2012a).

Several molecular techniques have been developed based on various genetic markers to observe the genetic variability under selection in various taxa, including humans (Hancock et al., 2008), birds (Li et al., 2021), reptiles (Barata et al., 2012; Hague and Routman, 2016), amphibians (Bonin et al., 2005), fish (Lamichhaney et al., 2012) and invertebrates (Wilding et al., 2001). As an approach, genome-wide associations studies (GWAS) are the most advanced, as well as expensive (>US$1,000 per sample) method to identify associations between loci and traits of interest, such as disease resistance (Bush and Moore, 2012; Tam et al., 2019). In addition, Quantitative Trait Locus (QTL) mapping and candidate gene approaches are the most common molecular techniques for the identification of the genetic architecture of an organism (Liu and Corde, 2004; Das and Sahoo, 2014; Yanez et al., 2014). The candidate gene approach has been applied to elucidate the genetic basis of both model organisms and humans where obvious 'candidates' for phenotypic variation are known from other studies of their physiological effects (Rothschild and Soller, 1997; Linnen and Hoekstra, 2009).

The most common nuclear genetic markers to characterize genetic diversity in animals includes allozymes, RAPD (random amplification polymorphic DNA), RFLP (restriction fragment length polymorphism), AFLP (amplified fragment length polymorphism) as early stage markers to advanced level microsatellite and SNP (single nucleotide polymorphism) (Freeland, 2005; Marsjan and Oldenbroek, 2007; Xu et al., 2016). Among mitochondrial markers, D-loop or control region, Cytochrome b gene, COI, 16S rRNA are common for intra-specific phylogenetic study which can be utilized to identify the evolutionary history of recently-diverged lineages descended from the ancestor based on their mitochondrial clades (Wilson et al., 1985; Avise et al., 1992; Orti, et al., 1994). SNPs are known to be major contributors to genetic variation, and generally

exhibit fixed differences between species caused by the addition or substitution of a single nucleotide segregating within a set of alleles (Eyre-Walker, 2006; Komar, 2009). The mapping of SNPs as a molecular marker has been utilized as a reliable and powerful tool for determining the chromosomal position of a mutant gene through chromosomal mapping and interval mapping to finely map the gene (Davis et al., 2005).

SNPs can be detected usually by direct sequencing of a PCR (polymerase chain reaction) product spanning the polymorphism, or by developing a secondary assay based on an RFLP method targeting specific SNPs (Donis-Keller et al., 1987; Syvänen, 2001; Kwok and Chen, 2003). The PCR-RFLP method has some advantages over other techniques (such as allele-specific oligonucleotide hybridization, allele-specific PCR, primer extension, oligonucleotide ligation, and endonuclease cleavage) as it allows simple, convenient and inexpensive detection of SNPs within the sequences of PCR products (Komar, 2009; Ota et al., 2009). The key advantages of using this polymorphism are that SNPs are codominant markers (enabling the identification of all alleles present at a particular locus) and usually have no associated phenotype. Therefore, it is easy to map a specific genetic region of a candidate gene with fix differences for targeted traits (Zipperlen et al., 2005; Hohenlohe et al., 2010). However, identifying candidate genes involved with a number of adaptive traits in animals is still apromising topicto compare adaptive differentiation among local populations (Merilä and Crnokrak, 2001; Kocher, 2004). Therefore, it is fundamental to understand the role of standing genetic variation occurred by specific mutations during environmental adaptation under natural selection in organisms (Linnen et al., 2013).

1.6 Study species: *Gasterosteus aculeatus* (Three-spined stickleback)

The three-spined stickleback (*Gasterosteus aculeatus,* Linnaeus, 1758) (hereafter 'stickleback') is an interesting and popular model fish across many fields in biology, including parasitology, ethology, evolutionary biology and ecology (Barber, 2013). The common name 'three-spined stickleback' is coined for the three dorsal spines that lie in front of the dorsal fin (Wootton, 1984). It is a small, oviparous (egg-laying) teleost, usually between 3 and 8 cm in length (average length about 5 cm) and has average life span varying from 1 to 3.6 years (Wootton, 1976; DeFaveri and Merilä, 2013). It is widely distributed throughout the Northern Hemisphere, from the open ocean to freshwater lakes and streams of North America, Europe, Asia and numerous islands (McPhail, 1994; Colosimo et al., 2005).

There are two distinct varieties (morphs) of the species: marine 'anadromous (relatively larger, broader and heavily armoured), and freshwater (typically small, slender

morphology with reduced body armour and spines) (MacColl et al., 2013) (Fig. 2). The anadromous form inhabits seawater but migrates to breed in fresh or brackish water and has a high salinity tolerance, while the freshwater form lives in freshwater all year round and occupies a wide variety of water bodies including ponds, lakes, streams and rivers (Bell and Foster, 1994). Anadromous stickleback are considered as ancestral to the newly-derived freshwater form, with a divergence in morphological, behavioural and life history traits since the Pleistocene ice sheet melted between 10,000 to 20,000 years ago (Wootton, 1976; Bell and Foster, 1994; Schluter, 1995; Boughman, 2001; McKinnon and Rundle, 2002; Ostlund-Nilsson et al., 2006). The stickleback's anti-predator adaptations, sensory physiology, reproductive physiology and endocrinology are well-studied, yet relatively few studies have investigatedtheir phenotypic and genotypic adaptation in relation to different environmental variables of natural selection (Barber, 2013; Hendry et al., 2013). Therefore, this fish, stickleback (*G. aculeatus*), was chosen for the work in this thesis, in order to study and understand its morphological, genetic, parasitological and microbial divergence in response to environmental variables.
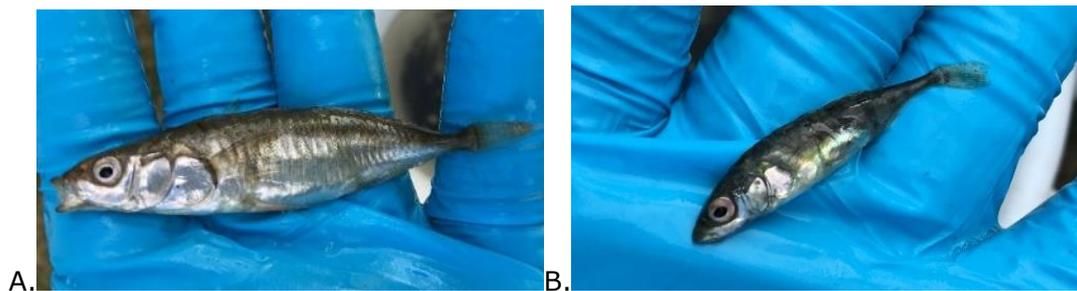
Fig 2. Stickleback fish (*Gasterosteus aculeatus*) showing two distinct morphs: A. Marine anadromous (approximately 70 mm) B. Freshwater resident (approximately 45 mm) collected from North Uist, Scotland.

Adaptive divergence suggestive of natural selection was first identified from variation in morphological traits between marine anadromous and freshwater stickleback populations (Bertin, 1925; Heuts, 1947; Bell, 2001). Variation in bony armour (Klepaker, 1993; Bell et al., 2004; Barrett, 2010; Reimchen et a., 2013), body size (McKinnon et al., 2004; MacColl et al., 2013; Taugbøl et al., 2014), body shape (Walker and Bell, 2000; Schluter et al. 2004; Aguirre, 2009), anti-predator defensive morphological structures (Hoogland et al., 1956; Reimchen, 1983, 1994; Marchinko, 2009), sexual dimorphism (Kitano et al., 2007, 2012), colouration (Boughman, 2001) and trophic traits (Berner et al., 2008; Matthews et al., 2010) are notable among stickleback populations of different environments such as freshwater vs. marine, benthic vs. limnetic, lake vs. stream often referred to as 'ecotypes' for their distinct form adapted to local environments (Schluter, 1993; McPhail, 1994; McKinnon and Rundle, 2002). The most striking polymorphism due to selection observed at a molecular level is in relation to the lateral bony armour plates

in freshwater populations, which are severely reduced or even missing, compared to the fully-plated ancestral anadromous populations (Bell and Foster, 1994; Colosimo et al., 2004, 2005; Chan et al., 2010). Phenotypic adaptative evolution in response to environmental selection pressures in divergent habitats, the body length, shape, lateral plate count, length of first dorsal spine, second dorsal spine and pelvic spine, and pelvis length and height, are studied in chapters 2 and 3 of this thesis.

Publication of the stickleback genome by the Broad Institute (2007) in the Ensembl data base, as well as the latest whole-genome sequencing by Jones et al. (2012a), have reached a milestone in the investigation of the genomic and molecular basis of adaptive divergence between freshwater and marine sticklebacks (Hendry et al., 2013; Roesti et al., 2015). Now, it has been established that marine and freshwater ecotypes exhibit remarkable genetic divergence in some genomic regions (at least in 81 different genomic loci) which allows stickleback to adapt to freshwater habitats through the action of natural selection (Hohenlohe et al., 2010; Jones et al., 2012a; Terekhanova et al., 2014; Ferchaud et al., 2014; Kusakabe et al., 2017; Haenel et al., 2019). Yet, the development of molecular tools with nuclear markers targeting specific candidate genes having fixed nucleotide differences and their association with adaptive phenotypic traits of stickleback is still largely insufficient (Jones et al., 2012b). The identification of the Ectodysplasin (Eda) locus as being responsible for the variation of the lateral plate morphs initiated this process of fine-scale exploration of adaptively-divergent genes under selection driving the evolution of stickleback (Avise, 1976; Schluter et al., 2004; Colosimo et al., 2004, 2005; Barrett et al., 2008). The study of the phenotypic-genotypic basis of environmental adaptation in a hybrid zone of this fish will enrich the knowledge of eco-evolutionary aspects relating to different populations, habitat conditions and abiotic-biotic agents of natural selection.

Stickleback are also a very common model for parasitological studies because they harbour a well-documented and experimentally-malleable parasitic fauna (Protista, Fungi, Cnidaria, Platyhelminthes, Annelida, Mollusca, Aschelminthes and Arthropoda), and are well suited to both laboratory and field parasitological investigation (Barber, 2007; MacColl, 2009; MacColl and Chapman, 2010). The role of parasites as agents of selection and diversification in different populations of stickleback has been documented in several studies (Rauch et al., 2006; Scharsack et al., 2007; Eizaguirre and Lenz, 2010; De Roij et al., 2011; De Roij and MacColl., 2012; Young and MacColl, 2017) as well as the likely connection to genetic factors (Rauch et al., 2006; El Nagar and MacColl, 2016; Robertson et al., 2017). Thus far, immune-related MHC genes have been studied specifically to observe pathogen-driven selection in stickleback, but this seems insufficient to understand the genetic basis of susceptibility to infection in a host

(Summers et al., 2003; Wegner et al., 2003; Kurtz et al., 2004; Janeway et al., 2005; Wegner et al., 2008; Eizaguirre and Lenz, 2010; Eizaguirre et al., 2012). Research in chapter 4 of this thesis incorporates a candidate gene approach based on SNPs in some selected loci to examine the effects of ectoparasite and endoparasite abundance on genotypes of stickleback in natural and experimental conditions.

To date, few studies have investigatedthe microbiota (especially on gut microbiota) of stickleback (Bolnick, et al., 2014; Milligan-Myhre et al., 2016). It has been suggested that there is an interaction between the skin mucosal surface of fish and microbial community due to direct contact with the surrounding environment (Austin, 2006). Yet, the studies on the effect of multiple factors such as abiotic ecological factors (e.g. pH, salinity) and genetic influences (fish population) on the skin microbiome composition of stickleback are comparatively limited for eco-evolutionary purposes (Smith et al., 2015). Strong variation in the relative frequency of bacterial communities among different groups of sticklebacks (such as freshwater-saltwater, high pH-low pH, anadromous-resident of saltwater and freshwater) will supplement the understanding of environmental adaptation of fish (chapter 5).

1.7 The study system: Outer Hebrides, Scotland

All stickleback samples examined in this thesis were collected from two neighbouring islands (North Uist and South Uist) in the Outer Hebrides or Western Isles (57°46'39"N;7°01'13"W) of Scotland, UK (Fig. 3). These two islands are connected through a small island called Benbecula and share similar aquatic environments, with a high number of freshwater and saltwater lochs (the Scottish term for lake). Both islands are enriched with numerous nationally and internationally important floral and faunal species, such as the aquatic plant Slender Naiad (*Najas flexilis*) (in North Uist), and many wading birds including the common redshank (*Tringa totanus*), ringed plover (*Charadrius hiaticula*) and dunlin (*Calidris alpina*), lapwing (*Vanellus vanellus*). Most importantly, the lochs of these islands are inhabited by several species of euryhaline fish including salmon (*S. salar*), sea and brown trout (*Salmo trutta*), eels (*Anguilla anguilla*), nine-spined stickleback (*Pungitius pungitius*) and three-spined stickleback (*G. aculeatus*), which has resulted in the designation of the lochs as sites of scientific interest for many years (Waterston et al., 1979).
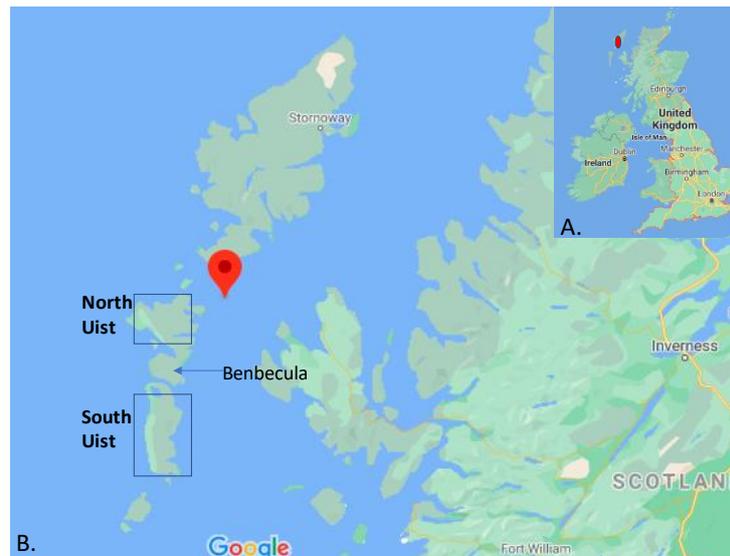
Fig. 3 A. Map of the United Kingdom B. A map of the Outer Hebrides showing the position of North Uist and South Uist islands in Scotland.

North Uist has long been the focus of studies for its diverse ecology which significantly impacts the adaptive radiation of stickleback (Giles, 1983; Campbell, 1985; MacColl et al., 2013; Magalhaes et al., 2016). There is a sharp transition in the surface geology of this island which results ina pH gradient in freshwater across the island: there are acidic, oligotrophic lochs on the east side, and basic, meso to eutrophic lochs on the west side (Waterston et al., 1979; MacColl et al., 2013; Haenel et al., 2019). Several studies have observed morphological variation of stickleback in relation to both abiotic and biotic factors of the loch water, specifically pH and Calcium ion concentration (Giles, 1983, Spence et al., 2013; MacColl and Aucott, 2014; Smith et al., 2014), temperature, conductivity, depth and predation (MacColl et al., 2013; Magalhaes et al., 2016), as well as phytoplankton and zooplankton (Chitheer, Rahman and MacColl, unpublished data). Parasitic communities of stickleback across the lochs have also been studied well in relation totemporal and spatial diversity (MacColl, 2009; MacColl and Chapman, 2010; De Roij and MacColl., 2012; Young and MacColl, 2017), variation in resistance or virulence (De Roij et al., 2011; Mahmud et al., 2017) and genetic responses to parasitic infection based on microsatellite and gene expression (El Nagar and MacColl, 2016; Robertson et al., 2017). Recent studies are mostly based on the genetic analysis of adaptive radiation in different populations of stickleback to observe ecological variation (Magalhaes et al., 2016; Rahn et al., 2016; Haenel et al., 2019; Magalhaes et al., 2020), life-history traits (Whiting et al., 2018) and speciation (Dean et al., 2019). In this thesis, I exclusively examine a natural putative hybrid zone between a freshwater loch (Loch Hosta) and the Atlantic Ocean on North Uist (Fig. 4) with the aim of observingthe effects of selection on phenotypic and genotypicvariation in a geographical gradation system (detail in chapter 3). Stickleback samples for observing natural infection status of

parasites and microbial diversity were also collected from different populations of North Uist.

Fig 4. Location of the Hosta hybrid zone running from (A) Loch Hosta, (B) through the upstream to (C) the downstream section of the Hosta stream in North Uist, Scotland.

In comparison, despite being the second largest island in theOuter Hebrides, South Uist has been undervalued for scientific research on stickleback. This island comprises high, rugged hills, and is more mountainous with the peaks of Beinn Mhor (620 metres) and Hecla (606 metres) mountains near the east coast, and low-lying croft land and machair with a continuous sandy beach at the west coast. Like North Uist, the geological features of this island might also have effects on the water chemistry of lochs, and therefore on adaptive divergence of the stickleback populations. Therefore, another aim of this thesis was to collect stickleback samples from South Uist to study the comparative morphological diversification with North Uist.

1.8 Thesis outline

The overall aim of this thesis was to understand the adaptive evolution of stickleback, including both morphological and genetic aspects, in association with interactions between different environmental selection agents, including salinity, pH, parasitic abundance and bacterial diversity. Each chapter addresses specific research questions regarding evolutionary perspectives of environmental adaptation. The remaining chapter outlines of this thesis are as follows:

Chapter 2: A comparative morphometric study of North Uist and South Uist stickleback

In this chapter, I examine the morphological traits of North Uist and South Uist stickleback to observe why adaptive radiation does or does not occur in closely-related lineages and how environmental selection agents contribute to this diversification.

Chapter 3: Strong genetic differentiation at a local scale in a three-spined stickleback hybrid zone.

This chapter describes the structure of a hybrid zone through the observation of phenotypic and genotypic variation. I use specifically develop SNPs-assay to investigate the process of environmental adaptation across a geographical gradation.

Chapter 4: Association of some genetic markers with parasitic infection in stickleback

I observe the divergence in parasite resistance through the targeting of selected candidate genes (Eda, PPARA, WNT7B and NLRC5) in both wild and laboratory-bred sticklebacks. For this, I conducted specific parasite infection experiments on a common ectoparasite (*G. gasterostei*) and endoparasite (*Diplostomum pseudospathaceum*) of stickleback to examine their association with host genotypes.

Chapter 5: Variation in the skin microbiomes of stickleback

In this chapter, I investigate the diversity of bacterial communities in freshwater and saltwater populations of stickleback to observe any obvious effects of major environmental differences or fish genetics.

Chapter 6: General discussion

In conclusion, I discuss the key findings from all the thesis chapters, and summarise general trends, significance and implications of this study, as well as future directions for further research.

References

Agrawal, A.A., 2001. Phenotypic plasticity in the interactions and evolution of species. *Science*, *294*(5541), pp.321-326.

Aguirre, W.E., 2009. Microgeographical diversification of threespine stickleback: body shape–habitat correlations in a small, ecologically diverse Alaskan drainage. *Biological Journal of the Linnean Society*, *98*(1), pp.139-151.

Albertson, R.C., Streelman, J.T. and Kocher, T.D., 2003. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proceedings of the National Academy of Sciences*, *100*(9), pp.5252-5257.

Alexandrino, J., Baird, S.J., Lawson, L., Macey, J.R., Moritz, C. and Wake, D.B., 2005. Strong selection against hybrids at a hybrid zone in the Ensatina ring species complex and its evolutinary implications. *Evolution*, *59*(6), pp.1334-1347.

Amundson, R., 1996. Historical development of the concept of adaptation. *Adaptation*, (ed.) Rose, M.R. and Lauder, G.V.,*1*, p.511. Academic press Inc.

Anon, 2020. Stabilizing, Directional, and Diversifying Selection. Available at: https://chem.libretexts.org/@go/page/13488.

Arnold, S.J. and Wade, M.J., 1984. On the measurement of natural and sexual selection: theory. *Evolution*, pp.709-719.

Austin, B., 2006. The bacterial microflora of fish, revised. *The Scientific World Journal*, *6*, pp.931-945.

Avise, J.C., 1976. Genetics of plate morphology in an unusual population of threespine sticklebacks (Gasterosteus aculeatus). *Genetics Research*, *27*(1), pp.33-46.

Avise, J.C., Bowen, B.W., Lamb, T., Meylan, A.B. and Bermingham, E., 1992. Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular biology and evolution*, *9*(3), pp.457-473.

Barata, M., Carranza, S. and Harris, D.J., 2012. Extreme genetic diversity in the lizard Atlantolacerta andreanskyi (Werner, 1929): A montane cryptic species complex. *BMC Evolutionary Biology*, *12*(1), pp.1-19.

Barber, I., 2007. Host-parasite interactions of the three-spined stickleback. In: *Biology of the Three-spined Stickleback*, eds. Östlund-Nilsson, S., Mayer, I. and Huntingford, F.A., CRC press. pp. 271-317.

Barber, I., 2013. Sticklebacks as model hosts in ecological and evolutionary parasitology. *Trends in parasitology*, *29*(11), pp.556-566.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2008. Natural selection on a major armor gene in threespine stickleback. *Science*, *322*(5899), pp.255-257.

Barrett, R.D.H., 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: A case study in functional analysis of natural variation. *Journal of Fish Biology*, *77*(2), pp.311-328.

Barton, N.H., 1979. The dynamics of hybrid zones. *Heredity*, *43*(3), pp.341-359.

Barton, N.H. and Hewitt, G.M., 1985. Analysis of hybrid zones. *Annual review of Ecology and Systematics*, *16*(1), pp.113-148.

Barton, N.H., Briggs, D.E.G., Eisen, J.A., Goldstein, D.B. and Patel, N.H., 2007. Random drift of allele frequencies. *Evolution. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press*, *415*.

Bell, G., 1997. *The basics of selection*. Springer Science & Business Media.

Bell, M.A. and Foster, S.A. eds., 1994. *The evolutionary biology of the threespine stickleback*. Oxford University Press.

Bell, M.A., 2001. Lateral plate evolution in the threespine stickleback: getting nowhere fast. In *Microevolution Rate, Pattern, Process* (pp. 445-461). Springer, Dordrecht.

Bell, M.A., Aguirre, W.E. and Buck, N.J., 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution*, *58*(4), pp.814-824.

Bernatchez, L. and Landry, C., 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years?. *Journal of evolutionary biology*, *16*(3), pp.363-377.

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.

Bertin, L., 1925. Recherches bionomique, biometrique et systematique sur les Epinoche. Annales de l'Institut Oceanographie Tome II 1–209.

Blanchet, S., Rey, O., Berthier, P., Lek, S. and Loot, G., 2009. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. *Molecular Ecology*, *18*(6), pp.1112-1123.

Boag, P.T. and Grant, P.R., 1981. Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galapagos. *Science*, *214*(4516), pp.82-85.

Bolnick, D.I., Snowberg, L.K., Hirsch, P.E., Lauber, C.L., Knight, R., Caporaso, J.G. and Svanbäck, R., 2014. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology letters*, *17*(8), pp.979-987.

Bonin, A., Taberlet, P., Miaud, C. and Pompanon, F., 2005. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). *Molecular biology and evolution*, *23*(4), pp.773-783.

Boughman, J.W., 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, *411*(6840), pp.944-948.

Brodie III, E.D., 1993. Differential avoidance of coral snake banded patterns by free-ranging avian predators in Costa Rica. *Evolution*, *47*(1), pp.227-235.

Bush, W.S. and Moore, J.H., 2012. Genome-wide association studies. *PLoS computational biology*, *8*(12), p.e1002822.

Cable, J. and Van Oosterhout, C., 2007. The role of innate and acquired resistance in two natural populations of guppies (Poecilia reticulata) infected with the ectoparasite Gyrodactylus turnbulli. *Biological Journal of the Linnean Society*, *90*(4), pp.647-655.

Calsbeek, R. and Cox, R.M., 2010. Experimentally assessing the relative importance of predation and competition as agents of selection. *Nature*, *465*(7298), pp.613-616.

Calsbeek, R. and Smith, T.B., 2007. Probing the adaptive landscape using experimental islands: density-dependent natural selection on lizard body size. *Evolution: International Journal of Organic Evolution*, *61*(5), pp.1052-1061.

Campbell, D.R., Faidiga, A. and Trujillo, G., 2018. Clines in traits compared over two decades in a plant hybrid zone. *Annals of botany*, *122*(2), pp.315-324.

Campbell, R.N., 1985. Morphological variation in the three-spined stickleback (Gasterosteus aculeatus) in Scotland. *Behaviour*, *93*(1-4), pp.161-168.

Capaul, M. and Ebert, D., 2003. Parasite-mediated selection in experimental Daphnia magna populations. *Evolution*, *57*(2), pp.249-260.

Carius, H.J., Little, T.J. and Ebert, D., 2001. Genetic variation in a host-parasite association: potential for coevolution and frequency-dependent selection. *Evolution*, *55*(6), pp.1136-1145.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J. and Myers, R.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *science*, *327*(5963), pp.302-305.

Cheplick, G.P., 2015. *Approaches to plant evolutionary ecology*. Oxford University Press, USA.

Cheplick, G.P., 2020. The semantics of evolutionary adaptation: clarification and evaluation. *Journal of Biological Education*, pp.1-14.

Clayton, D.H., Moyer, B.R., Bush, S.E., Jones, T.G., Gardiner, D.W., Rhodes, B.B. and Goller, F., 2005. Adaptive significance of avian beak morphology for ectoparasite control. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1565), pp.811-817.

Coleman, M.L. and Chisholm, S.W., 2010. Ecosystem-specific selection pressures revealed through comparative population genomics. *Proceedings of the National Academy of Sciences*, *107*(43), pp.18634-18639.

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.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. and Kingsley, D.M., 2004. The genetic architecture of parallel armor plate reduction in three spine sticklebacks. *PLoS biology*, *2*(5), p.e109.

Darwin, C. and Wallace, A.R., 1858. On the tendency of species to form varieties: and on the perpetuation of varieties and species by natural means of selection. Journal of the Proceedings of the Linnean Society, 3(9):45–62.

Darwin, C., 1859. On the origin of species by means of natural selection, or. *The Preservation of Favoured Races in the Struggle for Life, London/Die Entstehung der Arten durch natürliche Zuchtwahl, Leipzig oJ*.

Das, S. and Sahoo, P.K., 2014. Markers for selection of disease resistance in fish: a review. *Aquaculture international*, *22*(6), pp.1793-1812.

Davis, M.W., Hammarlund, M., Harrach, T., Hullett, P., Olsen, S. and Jorgensen, E.M., 2005. Rapid single nucleotide polymorphism mapping in C. elegans. *BMC genomics*, *6*(1), pp.1-11.

De La Torre, A., Ingvarsson, P.K. and Aitken, S.N., 2015. Genetic architecture and genomic patterns of gene flow between hybridizing species of Picea. *Heredity*, *115*(2), pp.153-164.

De Roij, J. and MacColl, A.D., 2012. Consistent differences in macroparasite community composition among populations of three-spined sticklebacks, Gasterosteus aculeatus L. *Parasitology*, *139*(11), p.1478.

De Roij, J., Harris, P.D. and MacColl, A.D., 2011. Divergent resistance to a monogenean flatworm among three-spined stickleback populations. *Functional Ecology*, *25*(1), pp.217-226.

Dean, L.L., Magalhaes, I.S., Foote, A., D'Agostino, D., McGowan, S. and MacColl, A.D., 2019. Admixture between ancient lineages, selection, and the formation of sympatric stickleback species-pairs. *Molecular biology and evolution*, *36*(11), pp.2481-2497.

DeFaveri, J. and Merilä, J., 2013. Variation in age and size in Fennoscandian three-spined sticklebacks (Gasterosteus aculeatus). *PLoS One*, *8*(11), p.e80866.

Dobzhansky, T., 1951. Genetics and the origin of Species: Columbia University. *New York*.

Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T.P., Bowden, D.W., Smith, D.R., Lander, E.S. and Botstein, D., 1987. A genetic linkage map of the human genome. *Cell*, *51*(2), pp.319-337.

Duffy, M.A., Brassil, C.E., Hall, S.R., Tessier, A.J., Cáceres, C.E. and Conner, J.K., 2008. Parasite-mediated disruptive selection in a natural Daphnia population. *BMC Evolutionary Biology*, *8*(1), pp.1-9.

Duncan, A.B. and Little, T.J., 2007. Parasite-driven genetic change in a natural population of Daphnia. *Evolution: International Journal of Organic Evolution*, *61*(4), pp.796-803.

Ebert, D., Zschokke-Rohringer, C.D. and Carius, H.J., 1998. Within–and between–population variation for resistance of Daphnia magna to the bacterial endoparasite Pasteuria ramosa. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1410), pp.2127-2134.

Eizaguirre, C. and Lenz, T.L., 2010. Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of fish biology*, *77*(9), pp.2023-2047.

Eizaguirre, C., Lenz, T.L., Kalbe, M. and Milinski, M., 2012. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology letters*, *15*(7), pp.723-731.

El Nagar, A. and MacColl, A.D., 2016. Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1836), p.20160691.

Endler, J.A., 1977. Geographic variation, speciation, and clines Princeton University Press. *Princeton, New Jersey, USA*.

Endler, J.A., 1986. *Natural selection in the wild*. Princeton University Press.

Eyre-Walker, A., 2006. The genomic rate of adaptive evolution. *Trends in ecology & evolution*, *21*(10), pp.569-575.

Ferchaud, A.L., Pedersen, S.H., Bekkevold, D., Jian, J., Niu, Y. and Hansen, M.M., 2014. A low-density SNP array for analyzing differential selection in freshwater and marine populations of threespine stickleback (Gasterosteus aculeatus). *BMC genomics*, *15*(1), pp.1-11.

Figueredo, A.J., Sefcek, J.A., Vasquez, G., Brumbach, B.H., King, J.E. and Jacobs, W.J., 2005.Evolutionary Personality Psychology. In: Buss, D.M. ed., 2005. *The handbook of evolutionary psychology*. John Wiley & Sons, pp.851-877.

Forrester, G.E., 1991. Social rank, individual size and group composition as determinants of food consumption by humbug damselfish, Dascyllus aruanus. *Animal Behaviour*, *42*(5), pp.701-711.

Fraterrigo, J.M., Wagner, S. and Warren, R.J., 2014. Local-scale biotic interactions embedded in macroscale climate drivers suggest Eltonian noise hypothesis distribution patterns for an invasive grass. *Ecology letters*, *17*(11), pp.1447-1454.

Fredensborg, B.L. and Poulin, R., 2006. Parasitism shaping host life-history evolution: adaptive responses in a marine gastropod to infection by trematodes. *Journal of animal ecology*, *75*(1), pp.44-53.

Freeland, J. R., 2005. Molecular Ecology. John Wiley and Sons, Ltd. PP. 388.

Futuyma, D.J., 2005. Evolution. Sinauer & Associates. *Inc., Sunderland, Massachusetts*, pp.226-243.

Garrido, E., Andraca-Gómez, G. and Fornoni, J., 2012. Local adaptation: simultaneously considering herbivores and their host plants. *New Phytologist*, *193*(2), pp.445-453.

Gay, L., Crochet, P.A., Bell, D.A. and Lenormand, T., 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution: International Journal of Organic Evolution*, *62*(11), pp.2789-2806.

Gibbs, H.L. and Grant, P.R., 1987. Oscillating selection on Darwin's finches. *Nature*, *327*(6122), pp.511-513.

Giles, N., 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *Journal of zoology*, *199*(4), pp.535-544.

Gowaty, P.A., 2005. Developmental Plasticity and Evolution, Mary Jane West-Eberhard (Ed.), Oxford University Press, Pp. xi+ 794.

Grant, P.R. and Grant, B.R., 1995. Predicting microevolutionary responses to directional selection on heritable variation. *Evolution*, *49*(2), pp.241-251.

Grant, P.R. and Grant, B.R., 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *science*, *296*(5568), pp.707-711.

Grant, P.R. and Grant, B.R., 2006. Evolution of character displacement in Darwin's finches. *science*, *313*(5784), pp.224-226.

Grant, P.R. and Grant, B.R., 2011. *How and why species multiply: the radiation of Darwin's finches*. Princeton University Press.

Gregory, T.R., 2009. Understanding natural selection: essential concepts and common misconceptions. *Evolution: Education and outreach*, *2*(2), pp.156-175.

Griffith, M.P., Noblick, L.R., Dowe, J.L., Husby, C.E. and Calonje, M.A., 2008. Cyclone tolerance in New World Arecaceae: biogeographic variation and abiotic natural selection. *Annals of Botany*, *102*(4), pp.591-598.

Griffiths, J.F., Griffiths, A.J., Wessler, S.R., Lewontin, R.C., Gelbart, W.M., Suzuki, D.T. and Miller, J.H., 2005. *An introduction to genetic analysis*. Macmillan.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D. and Berner, D., 2019. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution letters*, *3*(1), pp.28-42.

Hague, M.T.J. and Routman, E.J., 2016. Does population size affect genetic diversity? A test with sympatric lizard species. *Heredity*, *116*(1), pp.92-98.

Hairston Sr, N.G., Wiley, R.H., Smith, C.K. and Kneidel, K.A., 1992. The dynamics of two hybrid zones in Appalachian salamanders of the genus Plethodon. *Evolution*, *46*(4), pp.930-938.

Haldane, J.B.S., 1948. The theory of a cline. *Journal of genetics*, *48*(3), pp.277-284.

Hancock, A.M., Witonsky, D.B., Gordon, A.S., Eshel, G., Pritchard, J.K., Coop, G. and Di Rienzo, A., 2008. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet*, *4*(2), p.e32.

Hart, B.L., 1994. Behavioural defense against parasites: interaction with parasite invasiveness. *Parasitology*, *109*(S1), pp. S139-S151.

Hedrick, P.W., 2002. Pathogen resistance and genetic variation at MHC loci. *Evolution*, *56*(10), pp.1902-1908.

Hedrick, P.W., 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annu. Rev. Ecol. Evol. Syst.*, *37*, pp.67-93.

Hedrick, P.W., Kim, T.J. and Parker, K.M., 2001. Parasite resistance and genetic variation in the endangered Gila topminnow. *Animal Conservation*, *4*(2), pp.103-109.

31

Hendry, A.P., Peichel, C.L., Matthews, B., Boughman, J.W. and Nosil, P., 2013. Stickleback research: the now and the next. *Evolutionary Ecology Research*, *15*(2), pp.111-141.

Henter, H.J. and Via, S., 1995. The potential for coevolution in a host-parasitoid system. I. Genetic variation within an aphid population in susceptibility to a parasitic wasp. *Evolution*, *49*(3), pp.427-438.

Heuts, M.J., 1947. Experimental studies on adaptive evolution in Gasterosteus aculeatus L. *Evolution*, pp.89-102.

Hoekstra, H.E., Drumm, K.E. and Nachman, M.W., 2004. Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. *Evolution*, *58*(6), pp.1329-1341.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genet*, *6*(2), p.e1000862.

Hoogland, R., Morris, D. and Tinbergen, N., 1956. The spines of sticklebacks (Gasterosteus and Pygosteus) as means of defence against predators (Perca and Esox). *Behaviour*, pp.205-236.

Horth, L., 2004. Predation and the persistence of melanic male mosquitofish (Gambusia holbrooki). *Journal of Evolutionary Biology*, *17*(3), pp.672-679.

Huxley, J.S., 1939. Clines: an auxiliary method in taxonomy. *Bijdragen tot de Dierkunde*, *27*(1), pp.491-520.

Janeway, C.A., Travers, P., Walport, M. and Schlomchik, M.J., 2005.Immunobiology: The Immune System in Health and Disease, 6 Edn. Garland Science Publishing, New York.

Johnson, J.B. and Belk, M.C., 2001. Predation environment predicts divergent life-history phenotypes among populations of the livebearing fish Brachyrhaphis rhabdophora. *Oecologia*, *126*(1), pp.142-149.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M., Absher, D.M., Myers, R.M., Reimchen, T.E., Deagle, B.E. and Schluter, D., 2012b. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current biology*, *22*(1), pp.83-90.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012a. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), p.55.

Jutila, E., Jokikokko, E. and Julkunen, M., 2006. Long-term changes in the smolt size and age of Atlantic salmon, Salmo salar L., in a Northern Baltic River related to parr density, growth opportunity and postsmolt survival. *Ecology of Freshwater Fish*, *15*(3), pp.321-330.

Karlsson, L. and Karlström, Ö., 1994. The Baltic salmon (Salmo salar L.): its history, present situation and future. *Dana*, *10*, pp.61-85.

Karvonen, A. and Seehausen, O., 2012. The role of parasitism in adaptive radiations—when might parasites promote and when might they constrain ecological speciation?. *International Journal of Ecology*, *2012*.

King, R.B. and Lawson, R., 1997. Microevolution in island water snakes. *BioScience*, *47*(5), pp.279-286.

Kingsolver, J.G. and Diamond, S.E., 2011. Phenotypic selection in natural populations: what limits directional selection?. *The American Naturalist*, *177*(3), pp.346-357.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. and Beerli, P., 2001. The strength of phenotypic selection in natural populations. *The American Naturalist*, *157*(3), pp.245-261.

Kitano, J., Mori, S. and Peichel, C.L., 2007. Sexual dimorphism in the external morphology of the threespine stickleback (Gasterosteus aculeatus). *Copeia*, *2007*(2), pp.336-349.

Kitano, J., Mori, S. and Peichel, C.L., 2012. Reduction of sexual dimorphism in stream-resident forms of three-spined stickleback Gasterosteus aculeatus. *Journal of fish biology*, *80*(1), pp.131-146.

Klepaker, T., 1993. Morphological changes in a marine population of threespined stickleback, Gasterosteus aculeatus, recently isolated in fresh water. *Canadian Journal of Zoology*, *71*(6), pp.1251-1258.

Kloosterman, A., Parmentier, H.K. and Ploeger, H.W., 1992. Breeding cattle and sheep for resistance to gastrointestinal nematodes. *Parasitology Today*, *8*(10), pp.330-335.

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

Komar, A.A., 2009. Single nucleotide polymorphisms. *Methods in Molecular Biology*, *578*.

Kuchta, S.R., 2005. Experimental support for aposematic coloration in the salamander Ensatina eschscholtzii xanthoptica: implications for mimicry of Pacific newts. *Copeia*, *2005*(2), pp.265-271.

Kusakabe, M., Ishikawa, A., Ravinet, M., Yoshida, K., Makino, T., Toyoda, A., Fujiyama, A. and Kitano, J., 2017. Genetic basis for variation in salinity tolerance between stickleback ecotypes. *Molecular ecology*, *26*(1), pp.304-319.

Kwok, P.Y. and Chen, X., 2003. Detection of single nucleotide polymorphisms. *Current issues in molecular biology*, *5*, pp.43-60.

Lamichhaney, S., Barrio, A.M., Rafati, N., Sundström, G., Rubin, C.J., Gilbert, E.R., Berglund, J., Wetterbom, A., Laikre, L., Webster, M.T. and Grabherr, M., 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. *Proceedings of the National Academy of Sciences*, *109*(47), pp.19345-19350.

Lande, R. and Arnold, S.J., 1983. The measurement of selection on correlated characters. *Evolution*, pp.1210-1226.

Langefors, Å., Lohm, J., Grahn, M., Andersen, Ø. and Schantz, T.V., 2001. Association between major histocompatibility complex class IIB alleles and resistance to Aeromonas salmonicida in Atlantic salmon. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1466), pp.479-485.

Langerhans, R.B., Layman, C.A., Shokrollahi, A.M. and DeWitt, T.J., 2004. Predator-driven phenotypic diversification in Gambusia affinis. *Evolution*, *58*(10), pp.2305-2318.

Lenormand, T., 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), pp.183-189.

Lewis, J.S., Farnsworth, M.L., Burdett, C.L., Theobald, D.M., Gray, M. and Miller, R.S., 2017. Biotic and abiotic factors predicting the global distribution and population density of an invasive large mammal. *Scientific Reports*, *7*(1), pp.1-12.

Li, S., Li, D., Zhang, L., Shang, W., Qin, B. and Jiang, Y., 2021. High levels of genetic diversity and an absence of genetic structure among breeding populations of the endangered Rufous-backed Bunting in Inner Mongolia, China: implications for conservation. *Avian Research*, *12*(1), pp.1-11.

Li, W., 1997. *Molecular evolution*. Sinauer associates incorporated.

Liancourt, P., Spence, L.A., Song, D.S., Lkhagva, A., Sharkhuu, A., Boldgiv, B., Helliker, B.R., Petraitis, P.S. and Casper, B.B., 2013. Plant response to climate change varies with topography, interactions with neighbors, and ecotype. *Ecology*, *94*(2), pp.444-453.

Lindenstrøm, T., Secombes, C.J. and Buchmann, K., 2004. Expression of immune response genes in rainbow trout skin induced by Gyrodactylus derjavini infections. *Veterinary immunology and immunopathology*, *97*(3-4), pp.137-148.

Lindenstrøm, T., Sigh, J., Dalgaard, M.B. and Buchmann, K., 2006. Skin expression of IL-1β in East Atlantic salmon, Salmo salar L., highly susceptible to Gyrodactylus salaris infection is enhanced compared to a low susceptibility Baltic stock. *Journal of fish diseases*, *29*(2), pp.123-128.

Linnen, C.R. and Hoekstra, H.E., 2009, January. Measuring natural selection on genotypes and phenotypes in the wild. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 74, pp. 155-168). Cold Spring Harbor Laboratory Press.

Linnen, C.R., Kingsley, E.P., Jensen, J.D. and Hoekstra, H.E., 2009. On the origin and spread of an adaptive allele in deer mice. *Science*, *325*(5944), pp.1095-1098.

Linnen, C.R., Poh, Y.P., Peterson, B.K., Barrett, R.D., Larson, J.G., Jensen, J.D. and Hoekstra, H.E., 2013. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, *339*(6125), pp.1312-1316.

Little, E.E. and Fabacher, D.L., 1994.Comparative sensitivity of rainbow trout and two threatened salmonids, Apache trout and Lahontan cutthroat trout, to Ultraviolet-B radiation. *Advances in limnology [ERGEB. LIMNOL. /ADV. LIMNOL.]. 43:*217-226*.

Little, T.J. and Ebert, D., 2001. Temporal patterns of genetic variation for resistance and infectivity in a Daphnia-microparasite system. *Evolution*, *55*(6), pp.1146-1152.

Little, T.J., 2002. The evolutionary significance of parasitism: do parasite-driven genetic dynamics occur ex silico?. *Journal of Evolutionary Biology*, *15*(1), pp.1-9.

Littlejohn, M.J. and Watson, G.F., 1973. Mating-call variation across a narrow hybrid zone between Crinia laevis and C. victoriana (Anura: Leptodactylidae). *Australian Journal of Zoology*, *21*(2), pp.277-284.

Liu, Z.J. and Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, *238*(1-4), pp.1-37.

Lobón-Cerviá, J., 2005.Spatial and temporal variation in the influence of density dependence on growth of stream-living brown trout (Salmo trutta). *Canadian Journal of Fisheries and Aquatic Sciences*, *62*(6), pp.1231-1242.

Loewe, L., 2008. Negative selection. *Nature education*, *1*(1), p.59.

Losos, J.B., 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism: American Society of Naturalists EO Wilson Award address. *The American Naturalist*, *175*(6), pp.623-639.

MacColl, A.D. and Aucott, B., 2014. Inappropriate analysis does not reveal the ecological causes of evolution of stickleback armour: a critique of Spence et al. 2013. *Ecology and Evolution*, *4*(18), pp.3509-3513.

MacColl, A.D. and Chapman, S.M., 2010. Parasites can cause selection against migrants following dispersal between environments. *Functional ecology*, *24*(4), pp.847-856.

MacColl, A.D., 2009. Parasite burdens differ between sympatric three-spined stickleback species. *Ecography*, *32*(1), pp.153-160.

MacColl, A.D., 2011. The ecological causes of evolution. *Trends in Ecology & Evolution*, *26*(10), pp.514-522.

MacColl, A.D., Nagar, A.E. and de Roij, J., 2013. The evolutionary ecology of dwarfism in three-spined sticklebacks. *Journal of Animal Ecology*, *82*(3), pp.642-652.

Maes, S.L., De Frenne, P., Brunet, J., de la Peña, E., Chabrerie, O., Cousins, S.A., Decocq, G., Diekmann, M., Gruwez, R., Hermy, M. and Kolb, A., 2014. Effects of enhanced nitrogen inputs and climate warming on a forest understorey plant assessed by transplant experiments along a latitudinal gradient. *Plant Ecology*, *215*(8), pp.899-910.

Magalhaes, I.S., Whiting, J.R., D'Agostino, D., Hohenlohe, P.A., Mahmud, M., Bell, M.A., Skúlason, S. and MacColl, A.D., 2020. Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nature Ecology & Evolution*, pp.1-11.

Magurran, A.E. and Phillip, D.A., 2001. Evolutionary implications of large-scale patterns in the ecology of Trinidadian guppies, *Poecilia reticulata*. *Biological Journal of the Linnean Society*, *73*(1), pp.1-9.

Mahmud, M.A., Bradley, J.E. and MacColl, A.D., 2017. Abiotic environmental variation drives virulence evolution in a fish host-parasite geographic mosaic. *Functional Ecology*.pp.1-7.

Mallet, J., Barton, N., Lamas, G., Santisteban, J., Muedas, M. and Eeley, H., 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in Heliconius hybrid zones. *Genetics*, *124*(4), pp.921-936.

Malmquist, H.J., 1992. Phenotype-specific feeding behaviour of two arctic charr Salvelinus alpinus morphs. *Oecologia*, *92*(3), pp.354-361.

Marchinko, K.B., 2009. Predation's role in repeated phenotypic and genetic divergence of armor in three spine sticklebacks. *Evolution: International Journal of Organic Evolution*, *63*(1), pp.127-138.

Marsjan, P.A. and Oldenbroek, J.K., 2007. Molecular markers, a tool for exploring genetic diversity (Section C in part 4). In *The State of the World's Animal Genetic Resources for Food and Agriculture*. FAO.

Matthews, B., Marchinko, K.B., Bolnick, D.I. and Mazumder, A., 2010. Specialization of trophic position and habitat use by sticklebacks in an adaptive radiation. *Ecology*, *91*(4), pp.1025-1034.

McCarthy, I.D., 2001. Competitive ability is related to metabolic asymmetry in juvenile rainbow trout. *Journal of Fish Biology*, *59*(4), pp.1002-1014.

McCarthy, I.D., Houlihan, D.F. and Carter, C.G., 1994. Individual variation in protein turnover and growth efficiency in rainbow trout, Oncorhynchus mykiss (Walbaum). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *257*(1349), pp.141-147.

McCluskey, S., Mather, P.B. and Hughes, J.M., 1993. The relationship between behavioural responses to temperature and genotype at a PGI locus in the terrestrial isopod Porcellio laevis. *Biochemical systematics and ecology*, *21*(2), pp.171-179.

McKinnon, J.S. and Rundle, H.D., 2002. Speciation in nature: the threespine stickleback model systems. *Trends in ecology & evolution*, *17*(10), pp.480-488.

McKinnon, J.S., Mori, S., Blackman, B.K., David, L., Kingsley, D.M., Jamieson, L., Chou, J. and Schluter, D., 2004. Evidence for ecology's role in speciation. *Nature*, *429*(6989), pp.294-298.

McPeek, M.A., 1996. Linking local species interactions to rates of speciation in communities. *Ecology*, *77*(5), pp.1355-1366.

McPhail, J.D., 1994. Speciation and the evolution of reproductive isolation in the sticklebacks Gasterosteus) of southwestern British Columbia. *In*: M.A. Bell, S.A. Foster (Eds.), *The evolutionary biology of the threespine stickleback*, Oxford University Press (1994), pp. 399-437.

Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L. and Seehausen, O., 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature communications*, *8*(1), pp.1-11.

Merilä, J. and Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, *14*(6), pp.892-903.

Milligan-Myhre, K., Small, C.M., Mittge, E.K., Agarwal, M., Currey, M., Cresko, W.A. and Guillemin, K., 2016. Innate immune responses to gut microbiota differ between oceanic and freshwater threespine stickleback populations. *Disease models & mechanisms*, *9*(2), pp.187-198.

Mills, M.G., Greenwood, A.K. and Peichel, C.L., 2014. Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo*, *5*(1), p.5.

Minchella, D.J., 1985. Host life-history variation in response to parasitism. *Parasitology*, *90*(1), pp.205-216.

Møller, A.P., 1997. Parasitism and the evolution of host life history. *Host-parasite evolution: general principles and avian models.*, pp.105-127.

Moyaho, A., Garcia, C.M. and Manjarrez, J., 2004. Predation risk is associated with the geographic variation of a sexually selected trait in a viviparous fish (Xenotoca variata). *Journal of Zoology*, *262*(3), pp.265-270.

Mullen, L.M. and Hoekstra, H.E., 2008. Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution: International Journal of Organic Evolution*, *62*(7), pp.1555-1570.

Nielsen, R., 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.*, *39*, pp.197-218.

Noonan, B.P. and Comeault, A.A., 2009. The role of predator selection on polymorphic aposematic poison frogs. *Biology letters*, *5*(1), pp.51-54.

Nosil, P. and Crespi, B.J., 2006. Experimental evidence that predation promotes divergence in adaptive radiation. *Proceedings of the National Academy of Sciences*, *103*(24), pp.9090-9095.

Nürnberger, B., Barton, N., MacCallum, C., Gilchrist, J. and Appleby, M., 1995. Natural selection on quantitative traits in the Bombina hybrid zone. *Evolution*, *49*(6), pp.1224-1238.

Olsen, K.M., Sutherland, B.L. and Small, L.L., 2007. Molecular evolution of the Li/li chemical defence polymorphism in white clover (Trifolium repens L.). *Molecular Ecology*, *16*(19), pp.4180-4193.

Orti, G., Bell, M.A., Reimchen, T.E. and Meyer, A., 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution*, *48*(3), pp.608-622.

Ostlund-Nilsson, S., Mayer, I. and Huntingford, F.A. eds., 2006. *Biology of the three-spined stickleback*. CRC press.

Ota, M., Asamura, H., Oki, T. and Sada, M., 2009. Restriction enzyme analysis of PCR products. In *Single nucleotide polymorphisms* (pp. 405-414). Humana Press, Totowa, NJ.

Paterson, S., Wilson, K. and Pemberton, J.M., 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (Ovis aries L.). *Proceedings of the National Academy of Sciences*, *95*(7), pp.3714-3719.

Pedersen, S.H., Ferchaud, A.L., Bertelsen, M.S., Bekkevold, D. and Hansen, M.M., 2017. Low genetic and phenotypic divergence in a contact zone between freshwater and marine sticklebacks: gene flow constrains adaptation. *BMC evolutionary biology*, *17*(1), pp.1-16.

Penn, D. and Potts, W.K., 1998. Chemical signals and parasite-mediated sexual selection. *Trends in Ecology & Evolution*, *13*(10), pp.391-396.

Penn, D.J., Damjanovich, K. and Potts, W.K., 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences*, *99*(17), pp.11260-11264.

Pfennig, D.W., Harcombe, W.R. and Pfennig, K.S., 2001. Frequency-dependent Batesian mimicry. *Nature*, *410*(6826), pp.323-323.

Pfennig, D.W., Rice, A.M. and Martin, R.A., 2007. Field and experimental evidence for competition's role in phenotypic divergence. *Evolution*, *61*(2), pp.257-271.

Pigliucci, M. and Preston, K. eds., 2004. *Phenotypic integration: studying the ecology and evolution of complex phenotypes*. Oxford University Press.

Post, J.R., Parkinson, E.A. and Johnston, N.T., 1999. Density-dependent processes in structured fish populations: interaction strengths in whole-lake experiments. *Ecological Monographs*, *69*(2), pp.155-175.

Price, P.W., Westoby, M., Rice, B., Atsatt, P.R., Fritz, R.S., Thompson, J.N. and Mobley, K., 1986. Parasite mediation in ecological interactions. *Annual review of ecology and systematics*, *17*(1), pp.487-505.

Price, T.D., Grant, P.R., Gibbs, H.L. and Boag, P.T., 1984. Recurrent patterns of natural selection in a population of Darwin's finches. *Nature*, *309*(5971), pp.787-789.

Rahn, A.K., Krassmann, J., Tsobanidis, K., MacColl, A.D. and Bakker, T.C., 2016. Strong neutral genetic differentiation in a host, but not in its parasite. *Infection, Genetics and Evolution*, *44*, pp.261-271.

Rauch, G., Kalbe, M. and Reusch, T.B., 2006. Relative importance of MHC and genetic background for parasite load in a field experiment. *Evolutionary Ecology Research*, *8*(2), pp.373-386.

Reimchen, T.E., 1983. Structural relationships between spines and lateral plates in threespine stickleback (Gasterosteus aculeatus). *Evolution*, pp.931-946.

Reimchen, T.E., 1994. Predators and morphological evolution in threespine stickleback. *The evolutionary biology of the threespine stickleback*, pp.240-276.

Reimchen, T.E., Bergstrom, C. and Nosil, P., 2013. Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evol. Ecol. Res*, *15*, pp.241-269.

Reznick, D. and Endler, J.A., 1982. The impact of predation on life history evolution in Trinidadian guppies (Poecilia reticulata). *Evolution*, pp.160-177.

Reznick, D.N., Butler IV, M.J., Rodd, F.H. and Ross, P., 1996. Life-history evolution in guppies (Poecilia reticulata) 6. Differential mortality as a mechanism for natural selection. *Evolution*, *50*(4), pp.1651-1660.

Reznick, D.N., Shaw, F.H., Rodd, F.H. and Shaw, R.G., 1997. Evaluation of the rate of evolution in natural populations of guppies (Poecilia reticulata). *Science*, *275*(5308), pp.1934-1937.

Rigby, M.C., Hechinger, R.F. and Stevens, L., 2002. Why should parasite resistance be costly?. *Trends in parasitology*, *18*(3), pp.116-120.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific reports*, *7*, p.42677.

Roesti, M., Kueng, B., Moser, D. and Berner, D., 2015. The genomics of ecological vicariance in threespine stickleback fish. *Nature communications*, *6*(1), pp.1-16.

Rolhauser, A.G. and Pucheta, E., 2017. Directional, stabilizing, and disruptive trait selection as alternative mechanisms for plant community assembly. *Ecology*, *98*(3), pp.668-677.

Ross, C.L. and Harrison, R.G., 2002. A fine-scale spatial analysis of the mosaic hybrid zone between Gryllus firmus and Gryllus pennsylvanicus. *Evolution*, *56*(11), pp.2296-2312.

Rothschild, M.F. and Soller, M., 1997. Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. *Probe*, *8*(2), pp.13-20.

Sanjak, J.S., Sidorenko, J., Robinson, M.R., Thornton, K.R. and Visscher, P.M., 2018. Evidence of directional and stabilizing selection in contemporary humans. *Proceedings of the National Academy of Sciences*, *115*(1), pp.151-156.

Schaffner, S.F.P.C.S. and Sabeti, P., 2008. Evolutionary adaptation in the human lineage. *Nat. Educ*, *1*, p.14.

Scharsack, J.P., Kalbe, M., Harrod, C. and Rauch, G., 2007. Habitat-specific adaptation of immune responses of stickleback (Gasterosteus aculeatus) lake and river ecotypes. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1617), pp.1523-1532.

Schluter, D., 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology*, *74*(3), pp.699-709.

Schluter, D., 2000. *The ecology of adaptive radiation*. OUP Oxford.

Schluter, D., Clifford, E.A., Nemethy, M. and McKinnon, J.S., 2004. Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, *163*(6), pp.809-822.

Sella, G., Petrov, D.A., Przeworski, M. and Andolfatto, P., 2009. Pervasive natural selection in the Drosophila genome?. *PLoS Genet*, *5*(6), p.e1000495.

Shepard, D.B., 2007. Habitat but not body shape affects predator attack frequency on lizard models in the Brazilian Cerrado. *Herpetologica*, *63*(2), pp.193-202.

Shimada, Y., Shikano, T. and Merilä, J., 2010. A high incidence of selection on physiologically important genes in the three-spined stickleback, Gasterosteus aculeatus. *Molecular Biology and Evolution*, *28*(1), pp.181-193.

Slarkin, M., 1985. Gene flow in natural populations. *Annual review of ecology and systematics*, *16*(1), pp.393-430.

Slatkin, M., 1973. Gene flow and selection in a cline. *Genetics*, *75*(4), pp.733-756.

Smith, C., Spence, R., Barber, I., Przybylski, M. and Wootton, R.J., 2014. The role of calcium and predation on plate morph evolution in the three-spined stickleback (Gasterosteus aculeatus). *Ecology and evolution*, *4*(18), pp.3550-3554.

Smith, C.C., Snowberg, L.K., Caporaso, J.G., Knight, R. and Bolnick, D.I., 2015. Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME journal*, *9*(11), pp.2515-2526.

Sorci, G., Møller, A.P. and Boulinier, T., 1997. Genetics of host-parasite interactions. *Trends in Ecology & Evolution*, *12*(5), pp.196-200.

Sotka, E.E., Wares, J.P., Barth, J.A., Grosberg, R.K. and Palumbi, S.R., 2004. Strong genetic clines and geographical variation in gene flow in the rocky intertidal barnacle Balanus glandula. *Molecular ecology*, *13*(8), pp.2143-2156.

Spence, R., Wootton, R.J., Barber, I., Przybylski, M. and Smith, C., 2013. Ecological causes of morphological evolution in the three-spined stickleback. *Ecology and evolution*, *3*(6), pp.1717-1726.

Summers, K., McKeon, S.E.A., Sellars, J.O.N., Keusenkothen, M., Morris, J., Gloeckner, D., Pressley, C., Price, B. and Snow, H., 2003. Parasitic exploitation as an engine of diversity. *Biological Reviews*, *78*(4), pp.639-675.

Syvänen, A.C., 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nature Reviews Genetics*, *2*(12), pp.930-942.

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G. and Meyre, D., 2019. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, *20*(8), pp.467-484.

Taugbøl, A., Junge, C., Quinn, T.P., Herland, A. and Vøllestad, L.A., 2014. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. *Ecology and Evolution*, *4*(2), pp.144-156.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. and Mugue, N.S., 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. *PLoS Genet*, *10*(10), p.e1004696.

Thomas, F., Adamo, S. and Moore, J., 2005. Parasitic manipulation: where are we and where should we go?. *Behavioural processes*, *68*(3), pp.185-199.

Tort, L., Balasch, J.C. and Mackenzie, S., 2003. Fish immune system. A crossroads between innate and adaptive responses. *Inmunología*, *22*(3), pp.277-286.

Vamosi, S.M., 2005. On the role of enemies in divergence and diversification of prey: a review and synthesis. *Canadian Journal of Zoology*, *83*(7), pp.894-910.

Vignieri, S.N., Larson, J.G. and Hoekstra, H.E., 2010. The selective advantage of crypsis in mice. *Evolution: International Journal of Organic Evolution*, *64*(7), pp.2153-2158.

Vines, T.H., Dalziel, A.C., Albert, A.Y., Veen, T., Schulte, P.M. and Schluter, D., 2016. Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution*, *70*(5), pp.1023-1038.

Wade, M.J. and Kalisz, S., 1990. The causes of natural selection. *Evolution*, *44*(8), pp.1947-1955.

Wadgymar, S.M., Lowry, D.B., Gould, B.A., Byron, C.N., Mactavish, R.M. and Anderson, J.T., 2017. Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation. *Methods in Ecology and Evolution*, *8*(6), pp.738-749.

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, J.A. and Bell, M.A., 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of three-spine sticklebacks (Gasterosteus aculeatus). *Journal of Zoology*, *252*(3), pp.293-302.

Walton, W.E., Easter Jr, S.S., Malinoski, C. and Hairston Jr, N.G., 1994. Size-related change in the visual resolution of sunfish (Lepomis spp.). *Canadian Journal of Fisheries and Aquatic Sciences*, *51*(9), pp.2017-2026.

Waterston, A.R., Holden, A.V., Campbell, R.N. and Maitland, P.S., 1979. The inland waters of the Outer Hebrides. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *77*, pp.329-351.

Wegner, K.M., Kalbe, M., Kurtz, J., Reusch, T.B. and Milinski, M., 2003. Parasite selection for immunogenetic optimality. *Science*, *301*(5638), pp.1343-1343.

Wegner, K.M., Kalbe, M., Milinski, M. and Reusch, T.B., 2008. Mortality selection during the 2003 European heat wave in three-spined sticklebacks: effects of parasites and MHC genotype. *BMC Evolutionary Biology*, *8*(1), p.124.

Whiting, J.R., Magalhaes, I.S., Singkam, A.R., Robertson, S., D'Agostino, D., Bradley, J.E. and MacColl, A.D., 2018. A genetics‐based approach confirms immune associations with life history across multiple populations of an aquatic vertebrate (Gasterosteus aculeatus). *Molecular ecology*, *27*(15), pp.3174-3191.

Wilding, C.S., Butlin, R.K. and Grahame, J., 2001. Differential gene exchange between parapatric morphs of Littorina saxatilis detected using AFLP markers. *Journal of Evolutionary Biology*, *14*(4), pp.611-619.

Williamson, C.E., Metzgar, S.L., Lovera, P.A. and Moeller, R.E., 1997. Solar ultraviolet radiation and the spawning habitat of yellow perch, Perca flavescens. *Ecological Applications*, *7*(3), pp.1017-1023.

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M., 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, *26*(4), pp.375-400.

Wilson, K., Cotter, S.C., Reeson, A.F. and Pell, J.K., 2001. Melanism and disease resistance in insects. *Ecology letters*, *4*(6), pp.637-649.

Wootton, R.J., 1976. *Biology of the sticklebacks*. Academic Press.

Wootton, R.J., 1984. *A functional biology of sticklebacks*. Univ of California Press.

Xu, P., Jiang, Y., Xu, J., Li, J. and Sun, X., 2016. Genomics in the common carp. In *Genomics in Aquaculture* (pp. 247-274). Academic Press.

Yáñez, J.M., Houston, R.D. and Newman, S., 2014. Genetics and genomics of disease resistance in salmonid species.

Young, R.E. and MacColl, A.D., 2017. Spatial and temporal variation in macroparasite communities of three-spined stickleback. *Parasitology*.

Zhang, J., 2008. Positive selection, not negative selection, in the pseudogenization of rcsA in Yersinia pestis. *Proceedings of the National Academy of Sciences*, *105*(42), pp.E69-E69.

Zipperlen, P., Nairz, K., Rimann, I., Basler, K., Hafen, E., Hengartner, M. and Hajnal, A., 2005. A universal method for automated gene mapping. *Genome biology*, *6*(2), pp.1-10.

Chapter 2: A comparative morphometric study of North Uist and South Uist stickleback

Abstract

Adaptive radiations play a major role in understanding the process of adaptive evolution and speciation by showing bursts of diversification of many evolutionary lineages. The context and cause of these has been widely described and explored but why rapid evolutionary diversification does not occur in related evolutionary lineages has yet to be understood. One possible answer to this is simply that evolutionary diversification is provoked by environmental diversity, and that some lineages do not encounter the necessary environmental diversity. Stickleback on North Uist show enormous diversification, which seems to be associated with the diversity of aquatic habitats, but stickleback on the neighbouring island of South Uist have not been reported to show anything like the same level of evolutionary diversity, despite levels of environmental variation that we might expect to be similar to North Uist. In this study, I compared patterns of morphological diversity on North and South Uist and also examined variation in some important environmental parameters (such as pH, conductivity and salinity). I found significant associations between the pH of the water and variation in measured armour traits of stickleback populations collected from North Uist and South Uist. Ancestral anadromous stickleback of both islands showed similar morphological features including size and bony 'armour'. Freshwater and resident saltwater fish showed significant variation in all armour traits (lateral plate count, first dorsal spine, second dorsal spine, pelvic spine, pelvis length and pelvis height) in relation to pH of water. However, North Uist stickleback exhibited greater diversity of morphological traits than South Uist and this was associated with greater diversity in pH of the loch water on North Uist. Highly acidic and highly alkaline freshwater habitats are missing, or uncommon, on South Uist. Thus, pH appears to act as a causal factor driving the evolutionary diversification of stickleback in local adaptation in North and South Uist, Scotland.

1. Introduction

In evolutionary biology, adaptive radiation is a phenomenon in which an evolutionary lineage undergoes diversification into a variety of types adapted to specialised modes of life. These results in the origin of multiple related species with features that are phenotypically distinct from the ancestor and that confer adaptation to divergent ecological niches (Schluter, 2000). The observation of phenotypic divergence among populations of a species due to local adaptation is the first step in understanding the process of adaptive radiation in nature (DeWoody et al., 2015). According to naturalists, adaptive radiation of phenotypic characters of populations is the consequence of three ecological processes: 1) resource-based divergent natural selection, 2) resource competition and 3) ecological speciation (Schluter, 1996). In essence, differences in ecological conditions among local populations are the reason for divergence leading to reproductive isolation and speciation (Schluter, 2000; MacColl, 2009). Some classic examples of adaptive radiation through phenotypic changes are: beak shape and size of Darwin's Galapagos finches (Grant and Grant, 2006), body shape, feeding apparatus, colour patterns and behaviour of East African cichlid fishes (Kocher, 2004; Koblmuller et al., 2011), limb size and structure of Caribbean Anolis Lizards (Losos et al., 1998) and leaf shape and branch tips of the Hawaiian silverwords (Baldwin and Sanderson, 1998). These examples provide strong evidence for the process of natural selection in evolution with prominent diversity in form and function among a group of closely-related taxa (Schluter, 2000).

Following Darwin's theory of natural selection, explicit research on adaptive evolution has been conducted to understand how adaptation, divergence and speciation occur in natural populations (Nosil et al., 2009; Yoder et al., 2010; Wiens et al., 2010). However, understanding why some evolutionary lineages exhibit much morerapid evolutionary diversification than closely-related lineages still needs to be understood (Wagner et al., 2012). Generally, hybridization among members of divergent lineages may seed the onset of adaptive radiation and speciation, but the role of biological and environmental factors underlying this variation remains to be tested (Seehausen, 2004; Meier et al., 2017). Primarily, any environmental factor (agent of selection) and its interaction across taxa resulting in differential fitness among phenotypes is considered as the cause of natural selection and adaptive divergence (MacColl, 2011). Furthermore, a combination of several factors such as intrinsic (lineage-specific traits such as morphological, behavioural or genetical architecture), extrinsic abiotic (e.g., light, temperature, water chemistry, latitude, altitude etc.) or biotic factors (e.g. predators, competitors, parasites) including non-ecological explanations (e.g. genetic drift) can be responsible for a

particular adaptive radiation (Wagner et al., 2012; Magalhaes et al., 2016; Magalhaes et al., 2020). Therefore, it is important to investigate why adaptive radiation does or does not occur in closely-related lineages and how environmental selection agents contribute to morphological diversification among natural populations ofadiversifying taxa.

The three-spined stickleback (*Gasterosteus aculeatus*) has served as a model organism for the study of evolution due in part to its parallel evolution of freshwater adaptative radiations from bony-armoured marine ancestors, that occurred around 10,000 years ago (Bell and Foster, 1994; Colosimo et al., 2004). As a result, a diverse range of stickleback populations have evolved in new freshwater environments with conspicuous differences in morphology, physiology and behaviour (Bell and Foster, 1994; Ostlund-Nilsson et al., 2006). The divergence of morphological traits through major changes in the bony armour and spines that have repeatedly evolved in various locations, are common examples of adaptive radiation of stickleback (Colosimo et al., 2004; Chan et al., 2010). Ancestral marine anadromous sticklebacks (live in the sea and migrate to freshwater to spawn) are heavily armoured with a continuous row of 30-36 bony lateral plates running from head to tail on each side (known as a complete morph) (Colosimo et al., 2005; Barrett et al., 2008). On the other hand, derived freshwater (live in freshwater all time) and saltwater-resident (inhabit coastal saltwater all year round without migration to the open sea) stickleback exhibit substantial reduction in the total plate number with either a discontinuous row of 9-28 plates (partial morph) or with 0-9 lateral plates at the anterior end (low morph) (Colosimo et al., 2004; Colosimo et al., 2005). In addition, other phenotypic traits such as body size, shape, pelvic girdle and spines of freshwater and saltwater-resident stickleback show morphological transformation from the ancestral anadromous form (Moodie and Reimchen, 1976; Schluter, 1993; Bell et al., 1993; Schluter et al., 2004; Shapiro et al., 2004; Ravinet et al., 2015). Marine sticklebacks have a well-developed pelvic girdle with long, bilateral pelvic spines and more slender, spindle-shaped body than the freshwater form (Walker and Bell 2000; Shapiro et al., 2004; Schluter et al., 2004). In contrast, many populations of derived freshwater stickleback exhibit complete or partial loss of the pelvic girdle along with a shorter and more posterior positioned first dorsal spine (Bell et al., 1993; Walker and Bell, 2000; Shapiro et al., 2004; Ravinet et al., 2015).

To date, several abiotic and biotic environmental factors have been considered as causes for these major phenotypic changes including variations in salinity, temperature, nutrient availability, calcium deficiency, stream gradient, predators and parasites (Giles, 1983; Bergstrom, 2002; Barrett et al., 2009; Myhre and Klepaker, 2009; Marchinko, 2009; Wark and Peichel, 2010). For example, predation-driven selection has been reported to

influence the evolution of bony armour structures such as the lateral plates and spines within marine and freshwater stickleback populations (Marchinko, 2009). Many marine and lake stickleback populations were reported to possess large numbers of lateral plates and long spines which provide mechanical protection against piscivorous fish and birds (Moodie et al., 1973; Reimchen, 1992, 1994, 2000). In contrast, reduction or loss of lateral plates, pelvic girdle and associated spines in many freshwater populations might provide an advantage against invertebrate predators (e.g. aquatic insects) and pursuit predators (e.g. fish) (Reimchen, 1980; Bell et al., 1993; Marchinko Bergstrom, 2002, 2009). Moreover, other factors such as reduced nutrients or salinity, and calcium ion deficiency have also been reported to have an association with the loss of lateral plates in freshwater populations (Giles, 1983; Bourgeois et al., 1994; Barrett et al., 2009; Myhre and Klepaker, 2009; Spence et al., 2012). The genetic basis of some phenotypic traits has also been investigated in relation to their strong contribution to natural selection which provides additional support in identifying the role of the selective agents in the adaptive evolution of stickleback (Colosimo et al., 2004; 2005; Barrett et al., 2008; Chan et al., 2010; Jones et al., 2012). For example, the Eda gene which is involved in the variation (approximately 80%) of bony lateral plates between marine and freshwater conditions has associations with immunity and parasite resistance in stickleback (Colosimo et al., 2004, 2005; Robertson et al., 2017). However, this present study was designed to identify the role of environmental variables which cause selection on the morphological diversification of stickleback in response to diverse ecological conditions of two islands.

The island of North Uist in the Western Isles of Scotland comprises a mosaic of interconnected freshwater and brackish water lochs and lagoons which have exceptional variation in their water chemistry; there are high pH alkaline lochs in the west and low pH acidic lochs in the east, which are associated with variation in the concentrations of alkaline metals (sodium, potassium, magnesium, calcium etc.) (Waterston et al., 1979). This variation has been shown to correlate with the evolution of high diversity of stickleback populations across the island (Giles, 1983; MacColl et al., 2013; Magalhaes et al., 2016) and offers a unique opportunity to study the variation in phenotypes in relation to adaptation under different environmental conditions. In contrast, South Uist is the second largest of the Western Isles (~320km^2) with quite different geographical features than North Uist. This Island comprises high, rugged hills and mountainous areas, with the peaks of Beinn Mhor (620 metres) and Hecla (606 metres) at the east coast and low-lying croft land and machair with a continuous sandy beach on the west coast. In contrast to North Uist, where freshwater catchments form a radial pattern around the island, many of the catchments on South Uist drain from the eastern hills

across the more fertile machair. This might facilitate gene flow, even between divergent environments in different parts of catchments, and hence constrain diversification.

North Uist has been a common area for many studies on stickleback including looking at systematic variation in natural parasite burdens between isolated natural populations (De Roij and MacColl, 2012), resistance to infections (De Roij et al., 2011; El Nagar and MacColl, 2016), immune system variation (Robertson et al., 2017) and the virulence of parasites (Mahmud et al., 2017). In addition, the evolution of body size and armour in North Uist stickleback was reported to be due to variation in pH and calcium ions (Giles, 1983; MacColl et al., 2013; Spence et al., 2013; MacColl and Aucott, 2014; Smith et al., 2014). Recent studies on the adaptive radiation of stickleback have shown wide ranging phenotypic and genotypic variation in association with environmental factors (biotic and abiotic) in North Uist (Magalhaes et al., 2016, Haenel et al., 2019). However, to my knowledge no studies have been carried out on the South Uist stickleback to observe their morphological divergence. A comparison with the ecologically-diverse North Uist populations would help understand the role of environmental factors on selection in adaptive radiation. Therefore, in this study I collected samples from ten different lochs in North Uist and eight lochs in South Uist with different water chemistry (freshwater with acidic or alkaline condition and saltwater) and compared the morphological variation among stickleback populations.

I hypothesised that morphological traits (standard length, lateral plate count, first dorsal spine, second dorsal spine, pelvic spine, pelvis length and pelvis height) of three-spined stickleback would show variation between South Uist and North Uist due to differences in the environmental diversity of the lochs. I expect that there will be more morphological variation among different populations of North Uist than the South Uist due to the influence of the highly diverse water chemistry (such as pH and salinity) in local environmental conditions.

2. Materials and methods

2.1 Study area

Two neighbouring islands with apparently similar aquatic environments were selected to make a comparative study of morphological variation in three-spined stickleback. North Uist (57°31'12"N; 7°27'42"W) is in the centre of the Outer Hebrides or Western Isles of Scotland (Isles of Lewis, Harris, Scalpay, Beneray, North Uist, Benbecula, South Uist, Eriskay, Barra and Vatersay) and is approximately 303 km^2 in total area (Fig. 1A-C). It consists of a series of complex interconnected freshwater and saline lochs and coastal lagoons that show exceptional variation in their water chemistry, for example alkaline high pH in the western lochs and acidic low pH in the eastern lochs. North Uist encompasses a mosaic of peat bogs, heathland and low hills which makes it different from South Uist. South Uist (57°13'54"N; 7°02'38"W) is the second largest island of the Outer Hebrides, connected with North Uist through Benbecula (Fig. 1B). It is around 320.3 km^2 in total area and differs greatly between its west and east sides. Before this study, its inland aquatic habitats have not been characterised in detail in relation to stickleback morphological evolution.

2.2 Sample collection

Fish samples were collected from 10 lochs on North Uist and 8 lochs on South Uist between the dates of 06th – 16th May 2019 (Fig. 1). Locations on both islands were selected randomly to maximise likely variation in surface geology using satellite imagery (Google Earth, version 7.3) and hence water chemistry. Fish were caught in unbaited Gee's Minnow Traps (Gee traps, Dynamic Aqua, Vancouver, Canada), set overnight (approximately 16 hrs.) in all sites. The geolocation of each sample site (GPSmap 60CSx, Germin, UK) and water quality parameters including temperature, absolute conductivity, salinity and pH (multi-parameter probe - Multi340/set, WTW, Germany) were recorded from all sites (Table 1-2). The loch was grouped based on the pH of water: high pH >7.5, neutral pH = >7.0 - 7.4, low pH< 7.0 (Table 1 - 2).

Across all ten sampling sites of North Uist, 135 live fish samples were haphazardly selected and transported immediately to the laboratory in loch water in darkened boxes with continuous aeration (Fig. 1C). From South Uist, 128 fish were collected in the same way from 11 sites (3 sites from loch Aroa) of the eight lochs (Fig. 1D). Of the total 263 samples, 45 anadromous fish were sampled from both islands to estimate the ancestral state of stickleback in the Uist lochs. All animal work in this thesis was reviewed by the University of Nottingham AWERB, and carried out under Home Office licence 3003415.

No licence is required to collect stickleback in Scotland. Fish were euthanized with an overdose of MS222 (400mgL^{-l}) following Schedule 1 techniques (death was confirmed by destruction of the brain using a sharp needle) according to UK Home Office regulations. All fish were preserved with a tag in 70% ethanol for the morphometric study.

Fig.1 Sites for the collection of samples from ten lochs in North Uist and eight lochs in South Uist, Scotland (A. Map of United Kingdom B. North Uist and South Uist Island C. Loch position in North Uist D. Loch position in South Uist with road map).

Table 1. Sampling locations, habitat types, sample number (N) and physicochemical parameters of eight lochs at South Uist, Scotland (FW = Fresh water, SW = Saltwater, Anad = Anadromous, Resi = Resident).

Sl.	Island	Loch	Location	Habitat	N	pH	Temperature (°C)	Conductivity (µS/cm)	Salinity (ppt)	Date of sample collection
1.		a'Mhoil (Mhoi)	57°17'11"N;7°25'06"W	FW high pH	11 FW	7.5	10.9	63.9	0.04	08.05.2019
2.		Grogarry Loch (GroS)	57°19'59"N; 7°22'55"W	FW high pH	5 FW	7.7	9.4	1167	0.746	08.05.2019
3.		Eadaray (Eada)	57°15'39"N; 7°22'14"W	FW high pH	5 FW	7.7	10.3	60.5	0.038	08.05.2019
4.	South Uist	West Loch Ollay (OllW)	57°15'58"N; 7°24'01"W	FW neutral pH	30 FW	7.4	10.0	138.2	0.088	08.05.2019
5.		Stilligarry (Stil)	57°19'09"N; 7°22'12"W	FW neutral pH	3 FW	7.4	10.2	116.5	0.074	08.05.2019
6.		a' Phuirt-ruaidh (Phui)	57°17'48"N; 7°21'54"W	FW neutral pH	1 Anad 11 Resi	7.3	11.4	48.2	0.038	08.05.2019
7.		Druidibeg (Drui)	57°19'26"N; 7°19'38"W	FW neutral pH	10 FW	7.3	12.2	40.9	0.026	08.05.2019
8.		Abhainn Roag (Aroa 1,2,3)	57°17'17"N; 7°21'58"W 57°17'27"N; 7°22'33"W 57°17'30"N; 7°22'47"W	FW neutral pH Brackish water	13 Anad 38 Resi	7.4 7.0 7.4	9.0 10.6 9.5	52.2 106.8 2630	0.033 0.068 1.683	06.05.2019 08.05.2019

Table 2. Sampling locations, habitat types, sample number (N) and physicochemical parameters of ten lochs at North Uist, Scotland (FW = Fresh water, SW = Saltwater, Anad = Anadromous, Resi = Resident).

Sl.	Island	Loch	Location	Habitat	N	pH	Temper-ature (^{0}C)	Conduc-tivity (μS/cm)	Salinity (ppt)	Date of sample collection
1.		Chadha Ruaidh (Chru)	57°35'37"N; 7°11'44"W	FW low pH	15 FW	6.6	12.9	148	0.094	14.05.2019
2.		Scadavay (Scad)	57°35'6"N; 7°14'10"W	FW low pH	12 FW	6.5	13.2	130.9	0.083	15.05.2019
3.		Tormasad (Torm)	57°33'45"N; 7°19'1"W	FW low pH	14 FW	7.0	13.8	162.5	0.104	13.05.2019
4.	North Uist	Trosavat (Tros)	57°35'3"N; 7°24'45"W	FW low pH	5 Anad 9 Resi	6.6	16.4	165.9	0.106	14.05.2019
5.		Hosta (Hosta)	57°37'40"N; 7°29'18"W	FW high pH	11 FW	8.5	12.2	432	0.280	13.05.2019
6.		Grogary (Grog)	57°36'54"N; 7°30'40"W	FW high pH	5 Anad 10 Resi	8.3	14.6	340	0.220	16.05.2019
7.		naReival (Reiv)	57°36'39"N; 7°30'50"W	FW high pH	12 FW	9.0	14.7	439	0.280	14.05.2019
8.		Ard heiskir (Ardh)	57°34'48"N; 7°24'48"W	SW / Brackish water)	7 Anad 7 Resi	8.3	13.2	47,400	30.33	14.05.2019
9.		Fairy Knoll (Faik)	57°38'7"N; 7°12'54"W	SW / Brackish water)	7 Anad 7 Resi	8.5	13.2	45,100	28.86	16.05.2019
10.		Loch Duin (Duin)	57°38'35"N; 7°12'40"W	SW / Brackish water)	7 Anad 7 Resi	8.4	13.3	27,460	13.73	15.05.2019

2.3. Body armour and spine data collection for morphometric analysis

To collect data on external bony skeletal structures (armour), all ethanol-preserved samples were stained with Alizarin Red solution following a standard staining method of rehydration (70% EtOH /30% H20 → 50%/50% → 30/70 → 100% distilled water), fixation (10% buffered neutral formalin, pH 7), washing (distilled water), maceration (1% KOH solution), bleaching (5ml of 33% H_2O_2 in 1L KOH), staining (Alizarin red S stock solution to 1% KOH, ratio - 1:50) and subsequent washing with water and 1% KOH solution (Peichel et al., 2001). Finally, samples were stored in 40% isopropyl alcohol (propan-2-ol) which was changed weekly to remove purple dye. After confirming the appearance of bony parts, a digital photograph of the right side of every fish was captured using a digital SLR camera (Nikon D5200, Thailand) with macro lens (60 mm), digital ring flash and a tripod (to set a fixed distance). All photographs included a graph paper as a scale (cm) and the measurements of standard length (from the tip of the snout to the end of caudal peduncle), first and second dorsal spine, pelvic spine (from insertion point to the tip), pelvis height and pelvis length were recorded using tpsDig2 version 2.31 (Rohlf, 2010). The total number of lateral plates was also counted from the right side of the stained photograph.
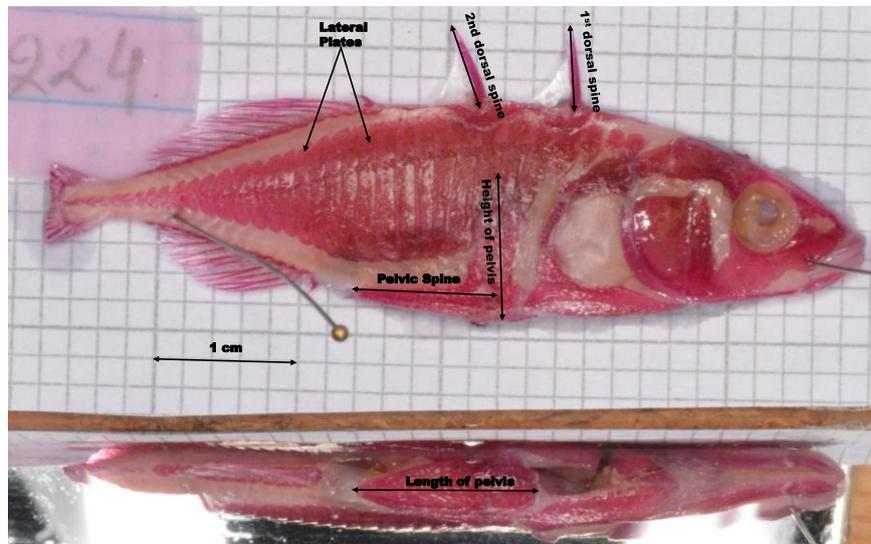
Fig. 2 Measurement of standard length (SL), dorsal spines (1st and 2nd), pelvic spine, pelvis (height and length) and lateral plates in a fully plated stained fish. Pelvis length was collected from the reflection on mirror.

2.4 Data analysis

All data were collated in Excel (Microsoft) and statistical analyses were conducted using R, version 3.6.3 (R Core Team, 2020).

Variation in morphological traits between North Uist and South Uist

All measured armour traits (except lateral plates) were size-standardized by calculating the residuals of a regression of each trait against standard length to obtain their allometric relationship with body size. The resident (including freshwater) (hereafter 'resident') and anadromous fish were analysed separately to quantify variation in morphological data [lateral plate count (hereafter 'plate count'), standard length, residuals of 1st dorsal spine length (first dorsal spine), 2nd dorsal spine length (second dorsal spine), pelvic spine length (pelvic spine), length of pelvis (pelvis length), height of pelvis (pelvis height)]. Principal component analysis (PCA) was performed using the singular value decomposition method to explore the axis of greatest variation in the measured armour data. All variables were scaled and centred for PCA and grouped for visualisation purposes according to three variables: location (North Uist and South Uist), pH (high pH, neutral pH and low pH) and salinity (freshwater and saltwater).

A one-wayanalysis of variance (ANOVA) was performed to test for differences in each morphological trait between the two different locations (North Uist and South Uist). Modified signed-likelihood ratio tests (MSLRT) (Krishnamoorthy and Lee, 2014) were also performed to quantify the differences in coefficient of variations (CV) between North Uist and South Uist for each morphological trait. Pearson's correlations were performed for all measured armour traits to observe their relationship with each other. To compare the mean fish length of the two ecotypes (anadromous and resident stickleback), a Wilcoxon rank sum t-test was performed for all samples of North Uist and South Uist. Finally, a Generalised Linear Model (GLM) with a Gaussian distribution and identity link function was fitted individually to observe the association of standard length of anadromous and resident stickleback with all armour traits of both islands. GLMs of plate count were also fitted using a negative binomial distribution and logit-link function with all measured traits as predictor factors in the two locations. Stepwise regression with a combination of forward and backward selection based on likelihood-ratio tests was conducted for all GLM models. The best fitting model was then selected based on Akaike's Information Criterion (AIC). All models were checked for goodness-of-fit using Quantile-Quantile (Q-Q) plots of the residuals and the significance of each component was tested using ANOVA tests.

Environmental factors affecting the variation of morphological traits in North Uist and South Uist

To observe associations between morphological traits of resident stickleback and different environmental factors of the loch water in North Uist and South Uist, GLMs were fitted with a Gaussian distribution and identity link function. These included location, pH, salinity and conductivity as predictor variables and PC1 and PC2 from the PCA of armour traits as response variable. GLMs using the mean of each morphological trait (standard length, 1st dorsal spine length, 2nd dorsal spine length, pelvic spine length, length of pelvis, height of pelvis and plate count) of each population as the response variable were fitted in relation to pH value, conductivity and salinity of loch water. A Scheffe post-hoc test (Scheffe, 1951) for multiple comparisions of mean with unequal sample sizes was performed to observe the differences in the size of measured armour traits among four low pH lochs in North Uist.

3. Results

3.1 Variation in morphological traits between North Uist and South Uist

The first two PCs from the PCA of the armour traits for the anadromous stickleback from both North and South Uist accounted for 58% of the total variation. PC3 and PC4 of armour traits explained around 21% and 11% of the variation respectively. All measured armour traits had positive loadings for PC1 and PC2 except pelvic spine length and length of pelvis (Fig. 3). Overall, anadromous fish from North and South Uist were morphologically similar, with some exceptions. The mean height of the pelvis of South Uist anadromous fish was significantly greater than North Uist (one-way ANOVA: $F = 5.19$, $df = 1, 44$, $p = 0.02$; Table 3). The standard length of South Uist anadromous stickleback showed significantly less variation than North Uist fish (MSLRT = 4.15, $p = 0.04$) (Fig. 4). In general, standard length of the anadromous fish showed significant associations with the length of the second dorsal spine ($F_{1, 44} = 6.0$, $p = 0.02$) and pelvis ($F_{1, 43} = 271.7$, $p<0.001$). Plate count of the anadromous fish did not show any associations with any measured armour traits.
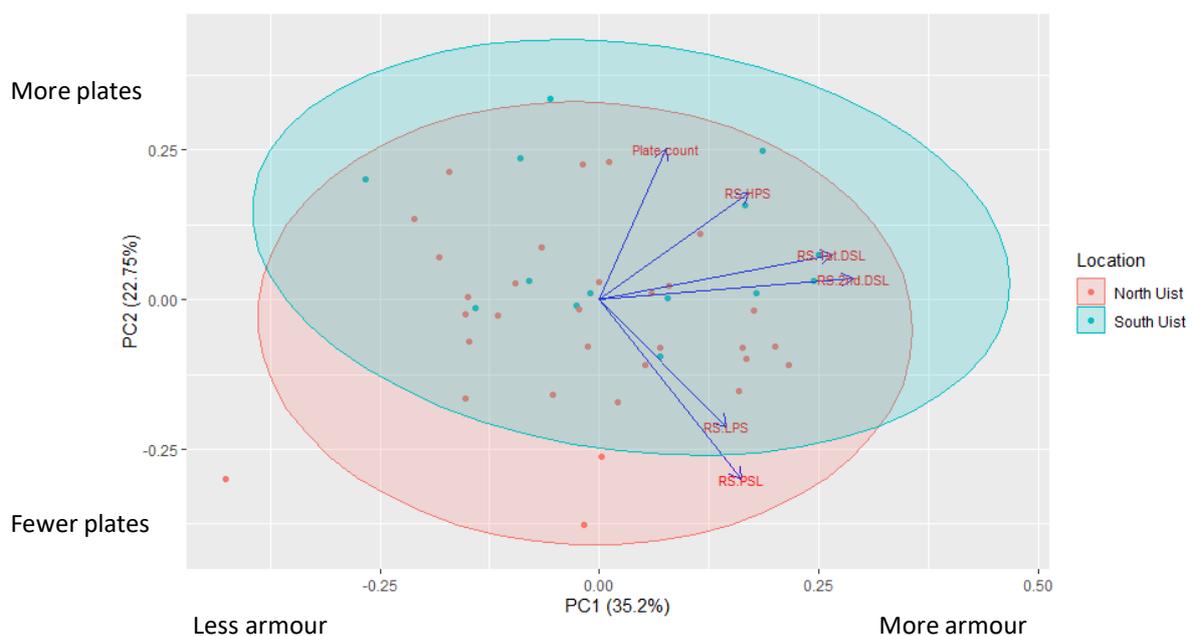
Fig. 3 The first two PCs of body armour traits - 1st dorsal spine length (DSL), 2nd dorsal spine length (DSL), pelvic spine length (PSL), length of pelvis (LPS), height of pelvis (HPS) with their residual value (RS) and plate count explaining 57.95% of total variation within the data of anadromous sticklebacks collected from North and South Uist. PC1 describing the positive loading of all variables with increasing in size whereas PC2 describing negative loading for LPS and PSL and an increase in size for plate count. Ellipses represent 95% confidence levels within each data set which shows overlapping clusters for anadromous fish indicating similar variation on North (red) and South Uist (blue).

Resident populations of stickleback from North and South Uist were similar in size and significantly smaller (smaller mean standard length) than the anadromous populations (W = 9742, p < 0.001; Fig. 4). Mean standard length of the North Uist resident stickleback (mean ± S.E. 36.44 ± 0.64 mm) was significantly greater than South Uist fish (34.55 ± 0.53 mm) (one-way ANOVA: F= 5.21, df = 1, 216, p= 0.02; Fig. 4). A GLM revealed that the standard length of resident fish of North Uist were significantly associated with some measured armour traits including length of 1st dorsal spine ($F_{1, 103}$ = 11.7, p<0.001), 2nd dorsal spine ($F_{1, 102}$ = 6.92, p = 0.009) and pelvis ($F_{1, 101}$ = 152.8, p<0.001). On the other hand, standard length of the South Uist resident fish showed a significant association only with length of pelvis ($F_{1, 112}$ = 304.4, p<0.001). Plate count of resident fish also showed a significant relationship with standard length ($F_{1, 103}$ = 7.66, p = 0.006) and height of pelvis ($F_{1, 102}$ = 62.6, p<0.001) in North Uist, compared to with length of second dorsal spine ($F_{1, 112}$ = 14.26, p<0.001) in South Uist stickleback.
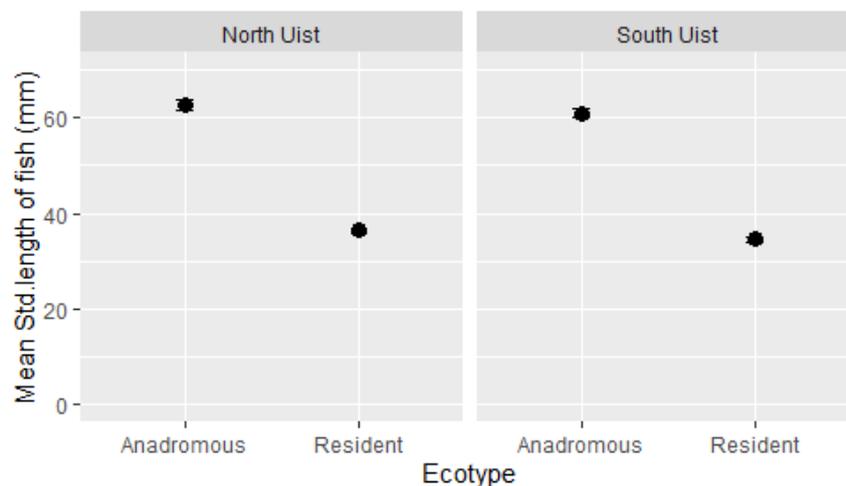
Fig. 4 Mean standard length (± S.E.) of anadromous and resident fish shows little overall variation in size between North and South Uist sticklebacks. Resident stickleback of North and South uist were significantly smaller than the anadromous fish.

The first two principal components of a PCA of armour traits for resident fish collected from the North and South Uist explained approximately 85% of the variance among individuals (Fig. 5A). PC1 was strongly correlated with all measures of armour traits so that high values of PC1 are associated with more armour. Plate count was highly correlated with PC2 and had a stronger influence on the variation in North Uist fish than South Uist. All measured armour traits (except plate count) were strongly correlated with each other Pearson's correlation (r)>0.5 (Fig. 5B).

For resident fish, despite no significant differences in trait means, there was significantly more variation (CV) in all traits (except standard length) in North Uist than South Uist (Table 3).
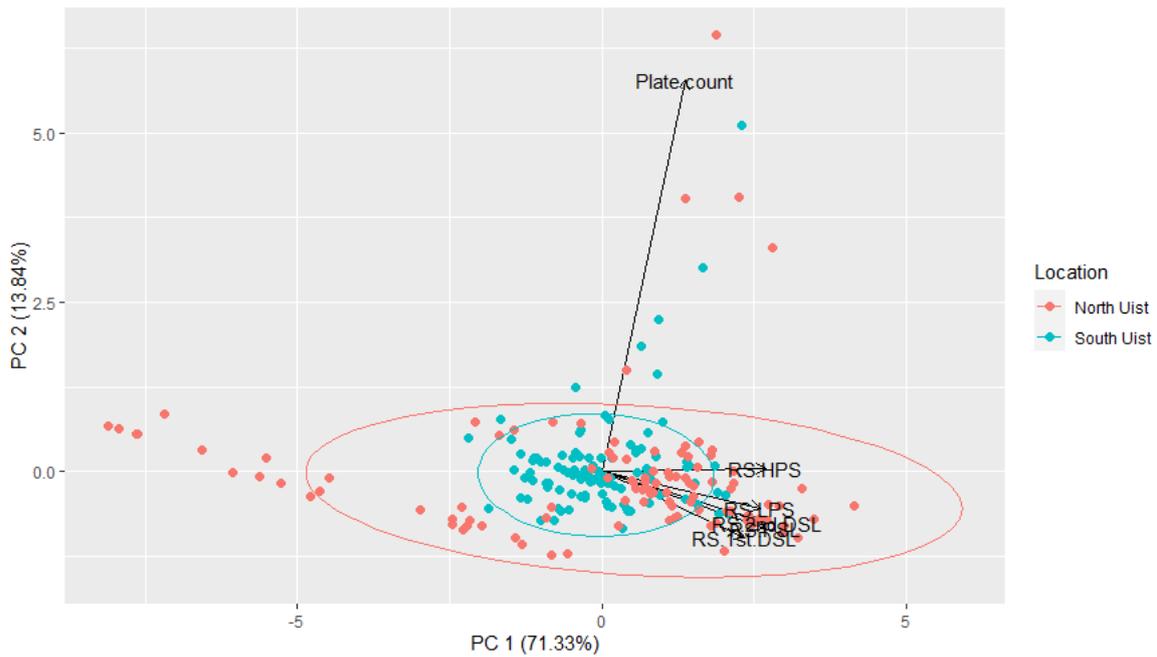
A.
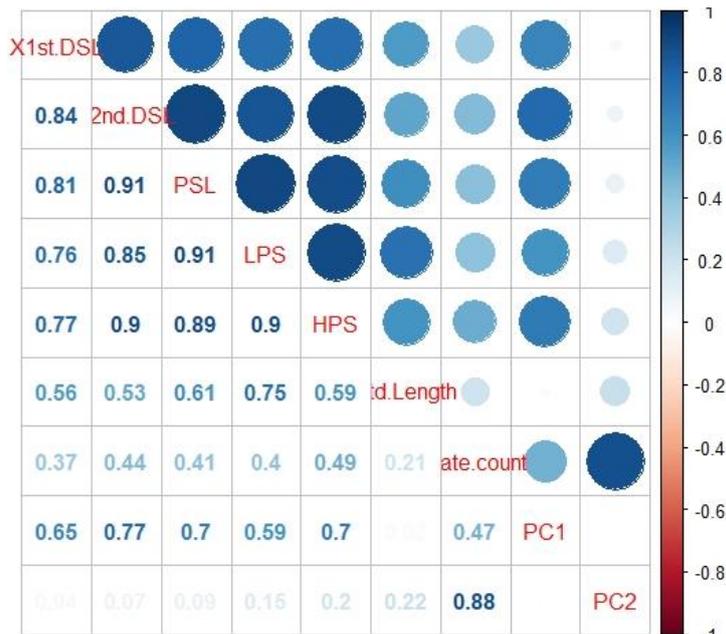
B.

Fig. 5 A The first two PCs of six body armour traits: 1st dorsal spine length (DSL), 2nd dorsal spine length (DSL), pelvic spine length (PSL), length of pelvis (LPS), height of pelvis (HPS) with their residual value and plate count of all resident sticklebacks explaining 85.17% of the total variability within the data set. Plate count shows the largest effect on PC2 than all other highly correlated traits. Ellipses represent 95% confidence levels separating the larger cluster of the data values of North Uist than the South Uist. B. Correlogram showing correlations between all measured armour traits including plate count, PC1, PC2 and standard length of North and South Uist resident fish. Blue circles denote for positive correlation and red circles denotes for negative correlation where size and shade of the ball indicates the strength of the correlation.

Table 3: One-way analysis of variance (ANOVA) for mean value and modified signed-likelihood ratio test (MSLRT) for coefficients of variation (CV) of each morphological trait [standard length, 1st dorsal spine length (DSL), 2nd dorsal spine length (DSL), pelvic spine length (PSL), length of pelvis (LPS), height of pelvis (HPS) and plate count] of stickleback sample fish (217 residents and 45 anadromous) collected from two locations (North Uist and South Uist, Scotland).

Morphological traits	North and South Uist									
	Anadromous					Resident				
	Mean			CV		Mean			CV	
	F	df	p	MSLRT	p	F	df	p	MSLRT	p
Std. Length	1.03	1, 44	0.31	4.15	**0.04***	5.21	1, 216	**0.02**	0.89	0.35
1st DSL	0.13	1, 44	0.72	0.45	0.50	3.13	1, 216	0.08	23.44	**<0.001**
2nd DSL	0.26	1, 44	0.61	2.25	0.13	0.22	1, 216	0.64	112.11	**<0.001**
PSL	3.99	1, 44	0.05	0.31	0.58	0.73	1, 216	0.40	84.46	**<0.001**
LPS	0.59	1, 44	0.45	1.05	0.30	0.64	1, 216	0.43	50.90	**<0.001**
HPS	5.19	1, 44	**0.03**	0.39	0.53	3.67	1, 216	0.06	100.76	**<0.001**
Plate count	2.64	1, 44	0.11	0.17	0.68	0.92	1, 216	0.34	12.46	**<0.001**

(Significance at p<0.05 is denoted in bold)

3.2 Factors affecting the variation of morphological traits in North Uist and South Uist

There was wide variation in many of the environmental factors including salinity and pH across the 8 lochs of South Uist and 10 lochs of North Uist: loch pH ranged from extremely low (acidic) to high (alkaline) pH of water (one-way ANOVA: $F = 10.28$, df = 1, 216, $p = 0.001$) in the freshwater lochs of North Uist compared to mostly neutral to high pH lochs of South Uist (Table 1 and 2, Fig. 6).
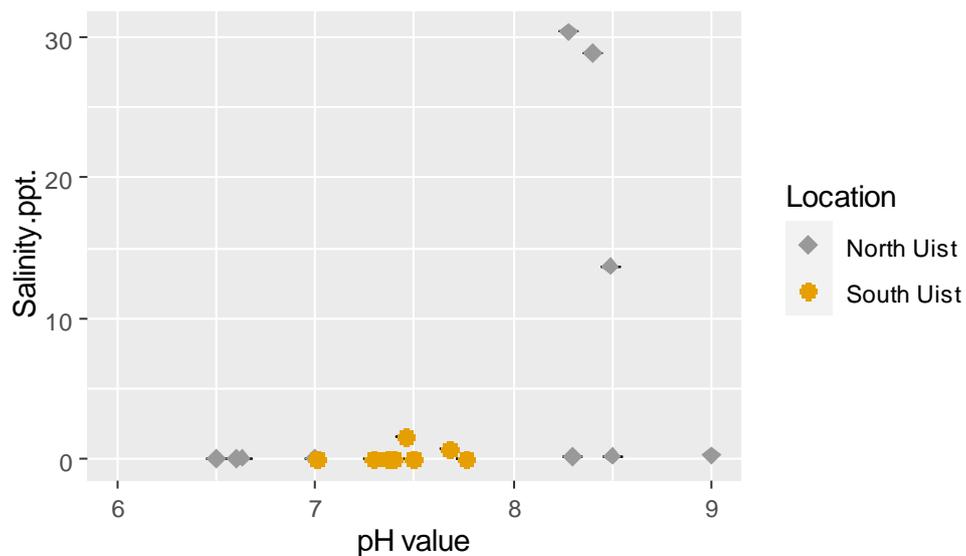
Fig. 6 Environmental factors (salinity and pH) of the sample collection sites (loch) shows more variation in freshwater lochs (with high range of pH at 0 ppt salinity) of North Uist compared to lochs of South Uist with neutral to slightly alkaline pH level.

The pH level of the lochs showed significant associations with the first two PCs of armour traits of the 18 populations of resident fish collected from North and South Uist (Table 4): there was more morphological variation in fish found in freshwater low pH lochs than other groups (Fig. 7).
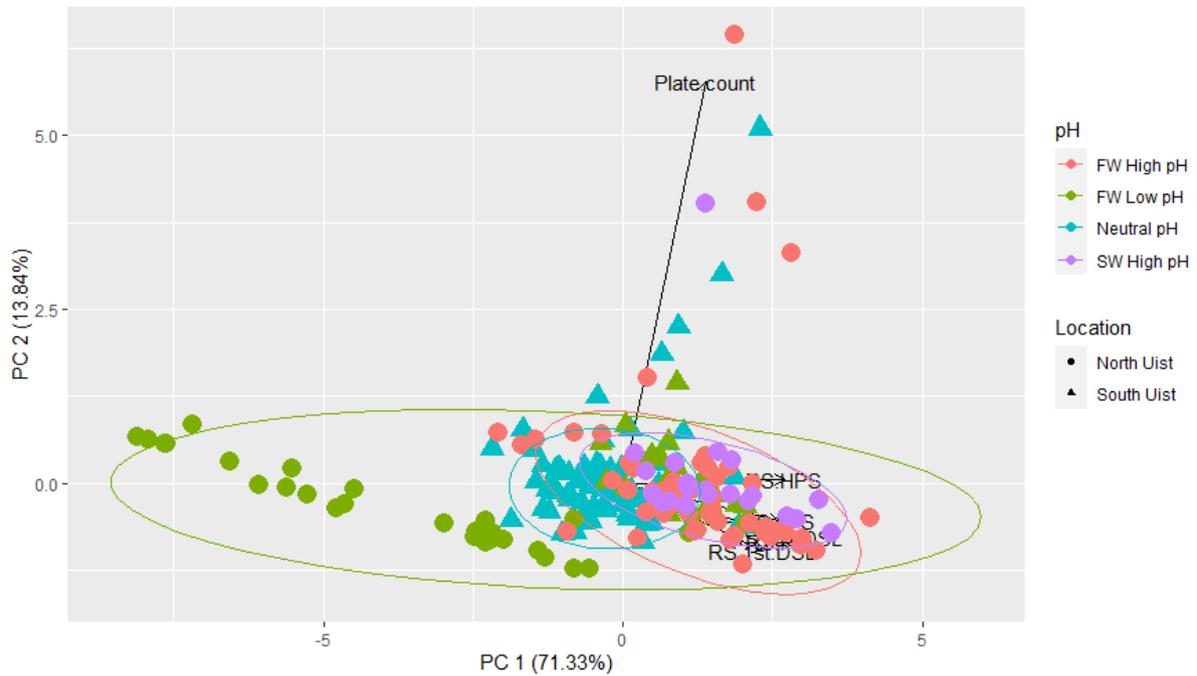
Fig. 7 Variation in body armour traits among resident sticklebacks of North and South Uist across the first two PCs on six measured variables: 1st dorsal spine length (DSL), 2nd dorsal spine length (DSL), pelvic spine length (PSL), length of pelvis (LPS), height of pelvis (HPS) and plate count. PC1 represents approximately 71% of total variation and PC2 represents approximately 14% of the variability within the data. Ellipses with 95% confidence intervals for each data set describing variation in four pH groups of loch water, separating the larger cluster for samples in freshwater low pH fish than the other groups.

Table 4: GLMs of minimum adequate models with significant variables fitted for morphological traits including PCs of armour traits of resident fish collected from North Uist and South Uist in association with environmental factor (pH value).

Response variable	Predictor variable	Family	Population	F	df	p
PC1	pH value	Gaussian	18	4.88	1,17	**0.042**
PC2	pH value	Gaussian	18	5.25	1,17	**0.034**
Avg. standard length	pH value	Gaussian	18	4.85	1,17	**0.044**
Avg. plate count	pH value	Gaussian	18	11.3	1,17	**0.003**
Avg. first dorsal spine	pH value	Gaussian	18	4.46	1,17	0.050.
Avg. second dorsal spine	pH value	Gaussian	18	4.83	1,17	**0.043**
Avg. pelvic spine	pH value	Gaussian	18	7.42	1,17	**0.015**
Avg. length of pelvis	pH value	Gaussian	18	11.3	1,17	**0.004**
Avg. height of pelvis	pH value	Gaussian	18	10.4	1,17	**0.005**

(Significance at p<0.05 is denoted in bold)

All morphological traits of the resident fish showed significant associations with pH levels in both North and South Uist (Table 4). The mean standard length of resident fish was highest in freshwater high pH populations (37.02 ± 0.97 mm) and lowest in freshwater low pH (33.49 ± 0.92 mm) ($F_{1, 17}$ = 4.85, p = 0.044; Fig. 8A). Total plate count showed different associations with pH of water across the two locations with significantly lower number of plates in higher pH on South Uist, although overall across both North and South Uist the mean plate count was highest in the high pH populations (6.58 ± 1.41 mm) and lowest in freshwater low pH lochs of North Uist (1.86 ± 0.27 mm) ($F_{1, 17}$ = 11.3, p<0.001; Fig. 8B). Among the 18 populations of resident fish of both islands, the mean length of all measured armour variables (first dorsal spine, second dorsal spine, pelvic spine, pelvis length and height) showed an increasing trend with increasing pH level having reduced size in freshwater low pH (except population from Tros) than the freshwater high pH populations (Fig. 8C-G). Post-hoc tests (ScheffeTest) revealed that all measured armour traits of stickleback collected from loch Tros (pH - 6.63, Table 2) were significantly larger in length than the three other freshwater low pH populations of North Uist (p < 0.001).
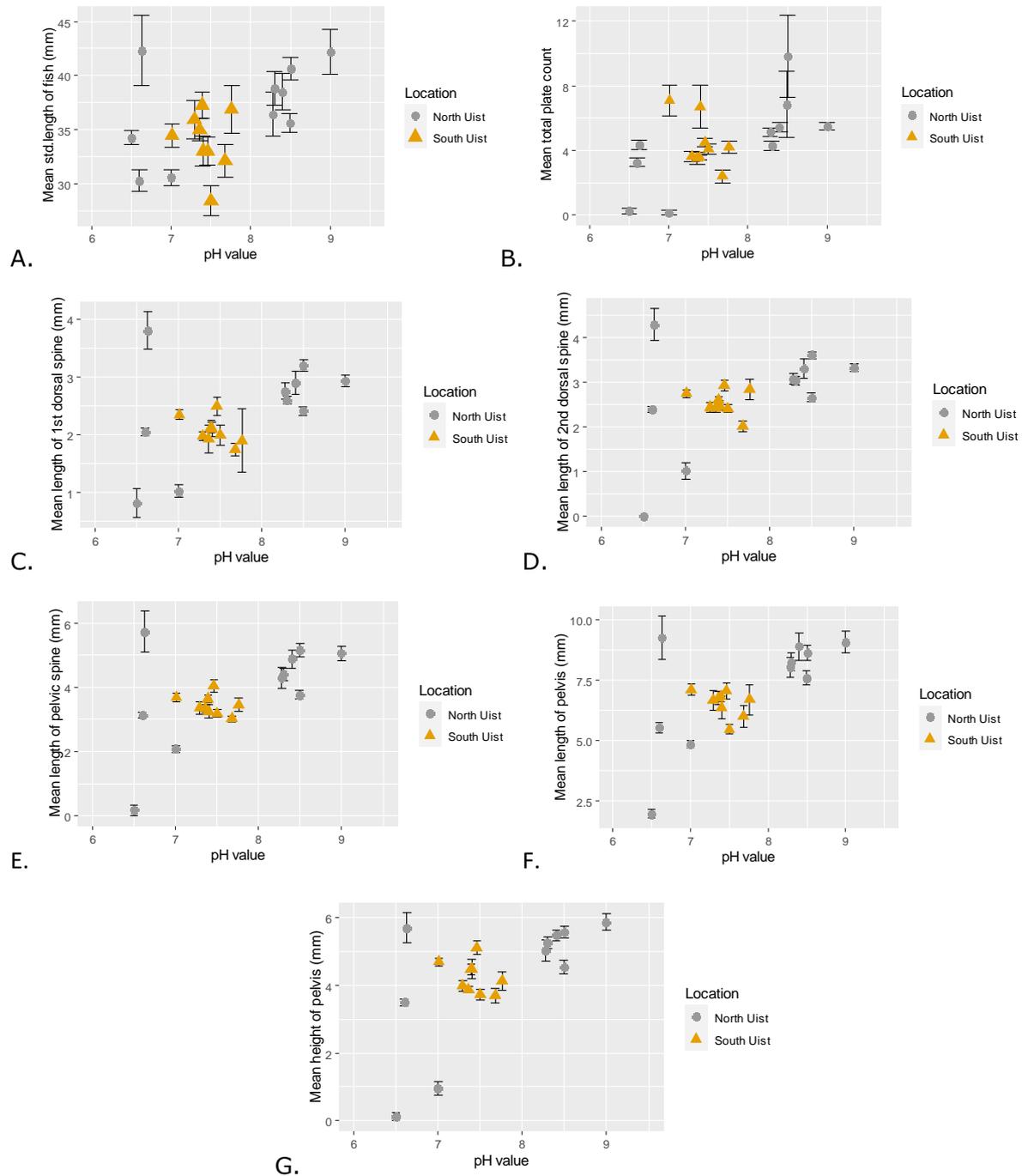
Fig. 8 A-G. Mean (± S.E.) standard length and six armour traits of 18 populations of North and South Uist resident fish shows significantly different pattern in length, total plate number, 1st dorsal spine length, 2nd dorsal spine length, pelvic spine length, length of pelvis and height of pelvis on the basis of pH of loch water having increasing trends in high pH levels of the sample collection sites except one loch of low pH in North Uist.

4. Discussion

The ecological theory of adaptive radiation suggests that phenotypic variance and speciation are generated by divergent natural selection emerging from differences in the environment and competition between species (Schluter, 2000). The present study was designed to investigate the phenotypic diversity of closely-related lineages of three-spined stickleback collected from North Uist and South Uist in Scotland, and to understand the role of some environmental variables as agents of natural selection. In this study, there was substantial variation in the observed morphological traits of stickleback between the two neighbouring islands which supports the hypothesis that underlying environmental factors may cause adaptive diversity. From the previous studies on stickleback ecotypes, it has been observed that anadromous and freshwater-resident fish are adaptively divergent and have morphologically and ecologically distinct forms (Jones et al., 2006; Raeymaekers et al., 2007; Hendry et al., 2009; Ravinet et al., 2015). This study also observed less variation among measured phenotypic traits (standard length, first dorsal spine, second dorsal spine, pelvic spine, pelvis length and pelvis height) and lateral plate count of anadromous stickleback than the freshwater-resident form among individuals of the two islands. Although, sample size was small (mostly 7 - 12) but I found significant differences (Type I error) from the mean of morphological variables to estimate the differences between North and South Uist.

Anadromous stickleback showed a similar type of variation of armour traits (except height of pelvis) between North Uist and South Uist consistent with close common ancestry in the marine environment. Bell and Foster (1994) described the typical structure of a three-spined stickleback which is represented by this ancestral anadromous fish having a large, fully armoured and well-developed pelvic girdle and long spines. The standard length of South Uist anadromous stickleback showed little overall variation compared to North Uist but were significantly larger in size than the resident populations of both islands. The reason for this variation among populations of the two islands might be a balance between natural selection in the local environment. Previous studies have suggested multiple causes of body size evolution of animals in island populations such as resource environment, competition and predation (Case, 1978; Raia et al., 2003; Raia and Meiri, 2006; Lomolino et al., 2012). The present finding is consistent with a previous study of stickleback body size evolution which reported the reduction of size in resident-freshwater ecotype compared to anadromous fish (McKinnon et al., 2004; MacColl et al., 2013; Taugbøl et al., 2014). However, the association of body length with the length of second dorsal spine and pelvis of anadromous stickleback may be related to the increased growth potential in the marine

environment for defence morphology against predators (Moodie and Reimchen, 1976; Reimchen, 1983; Taugbøl et al., 2014).

Resident fish collected from North and South Uist showed variance in all measured armour traits within and between populations which indicates there is adaptive radiation of phenotypic characters among populations as a result of colonization and adaptation to the freshwater environment (Bell and Foster, 1994; Nosil and Reimchen, 2005). There was significantly more variation in armour traits in resident fish of North Uist than South Uist which indicates the rapid evolutionary diversification of closely-related lineages. It has previously been observed that there was a large amount of phenotypic variation among freshwater-resident populations of North Uist stickleback, especially in body size, shape, armour and spine traits which supports the result of present study (Giles, 1983; MacColl et al., 2013; Spence et al., 2013; MacColl and Aucott, 2014; Smith et al., 2014; Magalhaes et al., 2016, Haenel et al., 2019). All measured armour traits were strongly correlated with each other across populations which suggests that variation among these phenotypic traits is the result of environmentally-driven selection varying across populations of North Uist and South Uist. Another reason might be the genetic correlation of traits controlling skeletal elements of the evolving populations of stickleback (Miller et al., 2014).

The overall body length of North Uist resident stickleback was significantly greater than South Uist fish which might be the result of greater variation of the environmental variables in North Uist (MacColl et al., 2013). The variation in the association of standard length with different armour traits in North Uist (1st dorsal spine, 2nd dorsal spine and pelvis length) and South Uist (length of pelvis) clearly indicates the effect of selection on the evolutionary diversification provoked by environmental diversity. Total plate number also showed different associations with measured traits in North Uist (standard length and height of pelvis) and South Uist (second dorsal spine) because of the enormous diversification of these island stickleback, which seems to be associated with the diversity of aquatic habitats (Haenel et al., 2019).

Environmental factors such as pH of the lochs showed strong associations with the armour traits of resident fish collected from North and South Uist which is supported by several studies of stickleback adaptive evolution in freshwater environments (Giles, 1983; Bell and Foster, 1994; Spence et al., 2012; MacColl et al., 2013; Magalhaes et al., 2016). In the present study, the armour traits of resident stickleback showed substantial variation in the first dorsal spine, second dorsal spine, pelvic spine and pelvis length in the freshwater populations of both islands at different levels of pH compared to the

saltwater populations of North Uist. The effect of pH on such morphological variation between freshwater-and saltwater-resident populations may indicate the role of divergent selection in maintaining armour traits for physiological adaptation (osmoregulation) or origin of populations (Heuts, 1947; Marchinko and Schluter, 2007; Spence et al., 2012).

North Uist stickleback exhibited much more morphological variation than the stickleback of South Uist and this variation was directly associated with greater diversity in pH of the loch water on North Uist. This is consistent with the prediction of this study since the pH levels in the lochs of South Uist were much less variable than those of North Uist, varying only from neutral to slightly alkaline compared to moderately acidic to moderately alkaline on North Uist. The surface geology of both islands might also be responsible for the differences in the pH level and productivity of the loch water (Waterston et al., 1979; Haenel et al., 2019). Catchments on North Uist are radial, draining areas of more consistent surface geology and resulting in higher variance between catchments than on South Uist, where the headwaters of catchments are in the hilly east, but streams drain west across the machair. The latter pattern results in a kind of 'environmental flow' that may reduce variation in water chemistry between lochs. This may also facilitate gene flow between lochs on South Uist that is absent on North Uist, but the restricted environmental variance alone on South Uist appears sufficient to explain the reduced morphological diversification.

Body length and armour traits of North Uist resident stickleback showed significant associations with pH and had shorter lengths in freshwater low pH populations than the freshwater high pH populations. In a similar study, Magalhaes et al., (2016) identified a strong association between stickleback phenotypes (particularly in armour traits) and environmental variation (abiotic and biotic) including pH across lakes of North Uist and also demonstrated the extreme reduction of body length, dorsal and pelvic spines including armour in acidic, deep, low ion composition and oligotrophic lakes compared to more alkaline eutrophic lakes.

Among resident fish, all armour traits were larger at higher pH levels, including saltwater-resident populations, which might be influenced by the alternate ion concentrations and resource environment (nutrient content). Another possible reason might be the genomic responses of high pH-resident sticklebacks to selection due to ecological similarities of habitat shared by the marine ancestor (Haenel et al., 2019). In contrast, there were significantly different patterns of measured traits in freshwater low pH populations than the high pH stickleback on North Uist. Body size of some low pH populations was extremely short which mirrored the previous finding of dwarfism in

stickleback in association with low pH indicative of poor resource conditions (Giles, 1983; MacColl et al., 2013). The dorsal spines and pelvic complex with spine were also either rudimentary or missing in most of the populations from low pH lochs (except Tros) in North Uist. Previously, several studies suggested this striking variation was due to calcium ion limitations in relation to extreme acidic conditions (Giles, 1983; Bourgeois et al., 1994; Schluter, 1995; Barrett, 2010; Magalhaes et al., 2016). Another reason might be the effect of local adaptation influenced by other confounding factors such as resource predation (Marchinko, 2009; MacColl et al., 2013). This latter explanation is supported by the exceptional outlier population of this study collected from Tros (Loch Trosavat), which possess larger body sizes and armour traits compared to other low pH populations of Uist. This loch has a direct connection to the sea, and as a consequence is inhabited by large salmonids (Atlantic salmon, *Salmo salar* and sea trout, *Salmo trutta*) and possibly other large predatory fishes (Magalhaes et al., 2016). It is known that long dorsal and pelvic spines with strong pelvic girdles increase protection from vertebrate predators (Reimchen, 1983, 1994; Bell and Foster, 1994; Marchinko, 2009). However, South Uist stickleback populations showed a common pattern for all measured traits in neutral to slightly alkaline loch water indicating less diversity of this lineage which do not seem to be provoked by the environmental diversity.

Lateral armour plate variation has been considered as the most common example of adaptive radiation of stickleback: there is substantial reduction of the total plate number in freshwater-and saltwater-resident fish (0-28 plates) compared to their ancestral anadromous fish (30-36 plates) (Colosimo et al., 2005; Barrett et al., 2009). Previously, several agents of selection have been suggested as causes of this loss of lateral plates in freshwater populations, including predators (Reimchen 1980; Bell et al., 1993; Bergstrom, 2002; Marchinko, 2009), reduced nutrients or salinity (Heuts, 1947; Marchinko and Schluter, 2007; Barrett et al., 2009; Myhre and Klepaker, 2009; Spence et al., 2012) and calcium ion deficiency (Giles, 1983; Bourgeois et al., 1994). In the present study, pH level of the loch water showed strong influences on the lateral armour plate variation by having rudimentary or missing lateral plates in some freshwater low pH populations and large number of plates in high pH populations of North Uist. South Uist populations had a medium number of plates compared to the extremely low or high numbers of plates in acidic and highly alkaline populations of North Uist which strongly indicates pH of water is a selective factor for plate number variation.

The divergent natural selection between environments originating from differences in resources and habitat structure (including biotic and abiotic factors) can be the main driver of morphological differentiation of these closely-related lineages (Schluter, 2000).

Although variation in phenotypes of North Uist stickleback has been more explained by the biotic variables (such as presence of predators and density of predator and competitors) than abiotic factors (Reimchen et al., 2013; Magalhaes et al., 2016), this study suggested strong influences of abiotic factors, especially the pH of the loch water. Additionally, as pH determines the productivity (nutrient content) and ion concentration of the loch, it can also influence the presence or absence of predators or competitors in connection with morphological adaptation of stickleback (Reimchen, 1980; Muniz, 1990; MacColl et al., 2013). There might be slight seasonal fluactuation in freshwater lochs to get more acid in winter time for organic matter but water chemistry including pH value of North and South Uist have been known to be stable for the last 15 years (MacColl et al., 2013; Spence et al., 2013; MacColl and Aucott, 2014; Smith et al., 2014) and correlated with the record of 1970s (Waterston et al., 1979; Giles, 1983). Therefore, from the comparison of phenotypic diversity of the two islands, it has been clear that pH acts as a causal factor of selection on the morphological diversification of stickleback in response to diverse ecological conditions.

To conclude, the increased evolutionary variation of North Uist stickleback populations compared to South Uist populations can be explained by greater environmental variation on North Uist, especially linked to variation in the pH of loch water. Although recent molecular studies through genome-wide linkage mapping approaches have identified genotype-phenotype associations, further study is required focusing on stickleback genetic divergence consistent with variation in selection due to environmental differences. However, other biotic and abiotic factors characterising the aquatic environments need to be identified in order to understand adaptive evolution through morphological diversification among natural populations of a diversifying taxa.

References

Baldwin, B.G. and Sanderson, M.J., 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences*, *95*(16), pp.9402-9406.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2008. Natural selection on a major armour gene in threespine stickleback. *Science*, *322*(5899), pp.255-257.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2009. Environment specific pleiotropy facilitates divergence at the Ectodysplasin locus in threespine stickleback. *Evolution: International Journal of Organic Evolution*, *63*(11), pp.2831-2837.

Barrett, R.D.H., 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: A case study in functional analysis of natural variation. *Journal of Fish Biology*, *77*(2), pp.311-328.

Bell, M.A. and Foster, S.A. eds., 1994. *The evolutionary biology of the three-spine stickleback*. Oxford University Press.

Bell, M.A., Ortí, G., Walker, J.A. and Koenings, J.P., 1993. Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. *Evolution*, *47*(3), pp.906-914.

Bergstrom, C.A., 2002. Fast-start swimming performance and reduction in lateral plate number in threespine stickleback. *Canadian Journal of Zoology*, *80*(2), pp.207-213.

Bourgeois, J.F., Blouw, D.M., Koenings, J.P. and Bell, M.A., 1994. Multivariate analysis of geographic covariance between phenotypes and environments in the threespine stickleback, Gasterosteus aculeatus, from the Cook Inlet area, Alaska. *Canadian Journal of Zoology*, *72*(8), pp.1497-1509.

Case, T.J., 1978. A general explanation for insular body size trends in terrestrial vertebrates. *Ecology*, *59*(1), pp.1-18.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J. and Myers, R.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *science*, *327*(5963), pp.302-305.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. and Kingsley, D.M., 2004. The genetic architecture of parallel armour plate reduction in three spine sticklebacks. *PLoS biology*, *2*(5), p.e109.

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.

De Roij, J. and MacColl, A.D., 2012. Consistent differences in macroparasite community composition among populations of three-spined sticklebacks, Gasterosteus aculeatus L. *Parasitology*, *139*(11), p.1478.

De Roij, J., Harris, P.D. and MacColl, A.D., 2011. Divergent resistance to a monogenean flatworm among three-spined stickleback populations. *Functional Ecology*, *25*(1), pp.217-226.

DeWoody, J., Trewin, H. and Taylor, G., 2015. Genetic and morphological differentiation in P opulus nigra L.: isolation by colonization or isolation by adaptation?. *Molecular Ecology*, *24*(11), pp.2641-2655.

Giles, N., 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *Journal of zoology*, *199*(4), pp.535-544.

Grant, P.R. and Grant, B.R., 2006. Evolution of character displacement in Darwin's finches. *science*, *313*(5784), pp.224-226.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D. and Berner, D., 2019. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution letters*, *3*(1), pp.28-42.

Hendry, A.P., Bolnick, D.I., Berner, D. and Peichel, C.L., 2009. Along the speciation continuum in sticklebacks. *Journal of fish biology*, *75*(8), pp.2000-2036.

Heuts, M.J., 1947. Experimental studies on adaptive evolution in Gasterosteus aculeatus L. *Evolution*, pp.89-102.

Jones, F.C., Brown, C., Pemberton, J.M. and Braithwaite, V.A., 2006. Reproductive isolation in a threespine stickleback hybrid zone. *Journal of evolutionary biology*, *19*(5), pp.1531-1544.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012a. The genomic basis of adaptive evolution in three-spine sticklebacks. *Nature*, *484*(7392), pp.55-61.

Koblmüller, S., Albertson, R.C., Genner, M.J., Sefc, K.M. and Takahashi, T., 2011. Cichlid evolution: lessons in diversification. *International journal of evolutionary biology*, *2011*.

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

Krishnamoorthy, K. and Lee, M., 2014.Improved tests for the equality of normal coefficients of variation. *Computational Statistics*, *29*(1), pp.215-232.

Lomolino, M.V., Sax, D.F., Palombo, M.R. and van der Geer, A.A., 2012. Of mice and mammoths: evaluations of causal explanations for body size evolution in insular mammals. *Journal of Biogeography*, *39*(5), pp.842-854.

Losos, J.B., Jackman, T.R., Larson, A., de Queiroz, K. and Rodríguez-Schettino, L., 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, *279*(5359), pp.2115-2118.

MacColl, A.D. and Aucott, B., 2014. Inappropriate analysis does not reveal the ecological causes of evolution of stickleback armour: a critique of Spence et al. 2013. *Ecology and Evolution*, *4*(18), pp.3509-3513.

MacColl, A.D., 2009. Parasite burdens differ between sympatric three-spined stickleback species. *Ecography*, *32*(1), pp.153-160.

MacColl, A.D., 2011. The ecological causes of evolution. *Trends in Ecology & Evolution*, *26*(10), pp.514-522.

MacColl, A.D., Nagar, A.E. and de Roij, J., 2013. The evolutionary ecology of dwarfism in three-spined sticklebacks. *Journal of Animal Ecology*, *82*(3), pp.642-652.

Magalhaes, I.S., D'Agostino, D., Hohenlohe, P.A. and MacColl, A.D., 2016. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular Ecology*, *25*(17), pp.4319-4336.

Magalhaes, I.S., Whiting, J.R., D'Agostino, D., Hohenlohe, P.A., Mahmud, M., Bell, M.A., Skúlason, S. and MacColl, A.D., 2020. Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nature Ecology & Evolution*, pp.1-11.

Mahmud, M.A., Bradley, J.E. and MacColl, A.D., 2017. Abiotic environmental variation drives virulence evolution in a fish host-parasite geographic mosaic. *Functional Ecology*.pp.1-7.

Marchinko, K.B. and Schluter, D., 2007. Parallel evolution by correlated response: lateral plate reduction in threespine stickleback. *Evolution: International Journal of Organic Evolution*, *61*(5), pp.1084-1090.

Marchinko, K.B., 2009. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution: International Journal of Organic Evolution*, *63*(1), pp.127-138.

Marwick, B. and Krishnamoorthy K. 2019. Cvequality: tests for the equality of coefficients of variation from multiple groups. *R package version 0.1*, *3*.

McKinnon, J.S., Mori, S., Blackman, B.K., David, L., Kingsley, D.M., Jamieson, L., Chou, J. and Schluter, D., 2004. Evidence for ecology's role in speciation. *Nature*, *429*(6989), pp.294-298.

Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L. and Seehausen, O., 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature communications*, *8*(1), pp.1-11.

Miller, C.T., Glazer, A.M., Summers, B.R., Blackman, B.K., Norman, A.R., Shapiro, M.D., Cole, B.L., Peichel, C.L., Schluter, D. and Kingsley, D.M., 2014. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics*, *197*(1), pp.405-420.

Moodie, G.E.E. and Reimchen, T.E., 1976. Phenetic variation and habitat differences in Gasterosteus populations of the Queen Charlotte Islands. *Systematic Zoology*, *25*(1), pp.49-61.

Moodie, G.E.E., McPhail, J.D. and Hagen, D.W. 1973. Experimental demonstration of selective predation on Gasterosteus aculeatus. *Behaviour*, *47*(1-2), pp.95-105.

Muniz, I.P., 1990. Freshwater acidification: its effects on species and communities of freshwater microbes, plants and animals. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *97*, pp.227-254.

Myhre, F. and Klepaker, T., 2009. Body armour and lateral-plate reduction in freshwater three-spined stickleback Gasterosteus aculeatus: adaptations to a different buoyancy regime?. *Journal of Fish Biology*, *75*(8), pp.2062-2074.

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.

Ostlund-Nilsson, S., Mayer, I. and Huntingford, F.A. eds., 2006. *Biology of the three-spined stickleback*. CRC press.

Peichel, C.L. and Marques, D.A., 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1713), p.20150486.

R Core Team, 2020. R: a language and environment for statistical computing. Version 3.6.3 [computer program].R Foundation for Statistical Computing, Vienna, Austria.

Raeymaekers, J.A., Van Houdt, J.K., Larmuseau, M.H., Geldof, S. and Volckaert, F.A., 2007. Divergent selection as revealed by PST and QTL-based FST in three-spined stickleback (Gasterosteus aculeatus) populations along a coastal-inland gradient. *Molecular ecology*, *16*(4), pp.891-905.

Raia, P. and Meiri, S., 2006. The island rule in large mammals: paleontology meets ecology. *Evolution*, *60*(8), pp.1731-1742.

Raia, P., Barbera, C. and Conte, M., 2003. The fast life of a dwarfed giant. *Evolutionary Ecology*, *17*(3), pp.293-312.

Ravinet, M., Hynes, R., Poole, R., Cross, T.F., McGinnity, P., Harrod, C. and Prodöhl, P.A., 2015. Where the lake meets the sea: strong reproductive isolation is associated with adaptive divergence between lake resident and anadromous three-spined sticklebacks. *PloS one*, *10*(4), p.e0122825.

Reimchen, T.E., 1980. Spine deficiency and polymorphism in a population of Gasterosteus aculeatus: an adaptation to predators?. *Canadian Journal of Zoology*, *58*(7), pp.1232-1244.

Reimchen, T.E., 1983. Structural relationships between spines and lateral plates in threespine stickleback (Gasterosteus aculeatus). *Evolution*, pp.931-946.

Reimchen, T.E., 1992. Injuries on stickleback from attacks by a toothed predator (Oncorhynchus) and implications for the evolution of lateral plates. *Evolution*, *46*(4), pp.1224-1230.

Reimchen, T.E., 1994. Predators and morphological evolution in three-spine stickleback. *The evolutionary biology of the threespine stickleback*, pp.240-276 in M. A. Bell and S. A. Foster, eds. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford.

Reimchen, T.E., 2000. Predator handling failures of lateral plate morphs in Gasterosteus aculeatus: functional implications for the ancestral plate condition. *Behaviour*, *137*(7-8), pp.1081-1096.

Reimchen, T.E., Bergstrom, C. and Nosil, P., 2013. Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evol. Ecol. Res*, *15*, pp.241-269.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific reports*, *7*, p.42677.

Rohlf, F. J, 2010. tpsDig, digitize landmarks and outlines, ver. 2.3. *Department of Ecology and Evolution, State University of New York at Stony Brook (computer program and documentation)*.

Scheffe, H., 1951. The analysis of variance, New York, 1959. *J. Wiley & Sons. Schild, HO, Fitzpatrick, RJ, and Nixon WCW: Lancet*, *1*, p.250.

Schluter, D., 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology*, *74*(3), pp.699-709.

Schluter, D., 1995. Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology*, *76*(1), pp.82-90.

Schluter, D., 1996. Ecological causes of adaptive radiation. *The American Naturalist*, *148*, pp. S40-S64.

Schluter, D., 2000. *The ecology of adaptive radiation*. OUP Oxford.

Schluter, D., Clifford, E.A., Nemethy, M. and McKinnon, J.S., 2004. Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, *163*(6), pp.809-822.

Seehausen, O., 2004. Hybridization and adaptive radiation. *Trends in ecology & evolution*, *19*(4), pp.198-207.

Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D. and Kingsley, D.M., 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, *428*(6984), pp.717-723.

Smith, C., Spence, R., Barber, I., Przybylski, M. and Wootton, R.J., 2014. The role of calcium and predation on plate morph evolution in the three-spined stickleback (Gasterosteus aculeatus). *Ecology and evolution*, *4*(18), pp.3550-3554.

Spence, R., Wootton, R.J., Barber, I., Przybylski, M. and Smith, C., 2013. Ecological causes of morphological evolution in the three-spined stickleback. *Ecology and evolution*, *3*(6), pp.1717-1726.

Spence, R., Wootton, R.J., Przybylski, M., Zięba, G., Macdonald, K. and Smith, C., 2012. Calcium and salinity as selective factors in plate morph evolution of the three-spined stickleback (G asterosteus aculeatus). *Journal of evolutionary biology*, *25*(10), pp.1965-1974.

Taugbøl, A., Junge, C., Quinn, T.P., Herland, A. and Vøllestad, L.A., 2014. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. *Ecology and Evolution*, *4*(2), pp.144-156.

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, J.A. and Bell, M.A., 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of three-spine sticklebacks (Gasterosteus aculeatus). *Journal of Zoology*, *252*(3), pp.293-302.

Wark, A.R. and Peichel, C.L. 2010. Lateral line diversity among ecologically divergent threespine stickleback populations. J. Exp. Biol., 213: 108–117.

Waterston, A.R., Holden, A.V., Campbell, R.N. and Maitland, P.S., 1979. The inland waters of the Outer Hebrides. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *77*, pp.329-351.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P. and Hawkins, B.A., 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology letters*, *13*(10), pp.1310-1324.

Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.P., Robertson, J. and Sarver, B.A.J., 2010. Ecological opportunity and the origin of adaptive radiations. *Journal of evolutionary biology*, *23*(8), pp.1581-1596.

Chapter 3: Strong genetic differentiation at a local scale in a three-spined stickleback hybrid zone

Abstract

Hybrid zones serve as natural laboratories that can show how the process of selection plays an influential role in population divergence and speciation. The present study investigated the structure of a naturally-occurring hybrid zone between marine and freshwater stickleback in a small stream (~1 m wide and approximately 1km long) running from a loch (Loch Hosta) to the sea on North Uist, an island in the Western Isles of Scotland. The major aims of this study were to examine the pattern of morphological and genetical variation through the zone and to identify genetic association with phenotypes among eleven populations of three micro-geographical groupings: one anadromous fish representative of one parental type, freshwater loch fish representative of the other, and eight samples from the stream, representing the hybrid zone. There was substantial evidence for admixture of phenotypic and genotypic traits consistent with variation in selection due to environmental differences. Morphological traits showed substantial variation across the hybrid zone due to adaptive divergence between parental populations with mosaic patterns in length, total plate count, parasite abundance and clinal patterns in body shape, spines and pelvic structure. Genetic diversity varied among populations of the hybrid zone, highlighting the strength of selection across the geographical gradient depending on intrinsic factors with high level of adaptive genetic differentiation between ancestral marine and diverged freshwater stickleback. Mitochondrial marker (Cytochrome b gene) showed a unique pattern with two major clades present in the parental populations, but only one in the stream, suggesting strong selection and/or habitat choice associated with mitochondrial haplotype. Allele frequencies of four candidate adaptive genetic loci assayed from SNPs: Eda, PPARA, WNT7B and NLRC5, showed great variation across the hybrid zone with a clinal pattern from sea to loch, which emphasises the strong genetic structure of the Hosta stream population. The deviation from Hardy-Weinberg equilibrium and linkage disequilibrium among candidate loci suggested the existence of heterozygote deficit due to strong selection maintaining reproductive isolation among populations. Body size, armour, shape and parasite of fish showed associations with Eda, WNT7B and PPARA loci which indicated their significant impact on phenotypic traits due to strong selection in divergent environments. Despite its small size, the Hosta hybrid zone provides a significant insight into the fine-scale patterns of phenotypic and genotypic variation under divergent selection of ecological adaptations in stickleback.

1. Introduction

A hybrid zone is a narrow geographical area between two biomes where two genetically-distinct populations or species meet, mate and produce offspring with a mixed genome (Harrison, 1990; Jiggins and Mallet, 2000). The structure of the hybrid zone between species, sub-species, races or lineages can be in several forms of variable size depending on the populations and environment involved (Gay et al., 2008). A primary hybrid zone forms between adjacent populations of a previously homogeneous species developing parapatric divergence, whereas a secondary hybrid zone originates from secondary contact between two populations after a period of geographic isolation (allopatric speciation) (Barton and Hewitt, 1985). A hybrid zone between divergent populations may be persistent or ephemeral depending on the process of reproductive isolation and its influencing factors (exogenous such as environmental factors, or endogenous such as genetic incompatibilities) leading to speciation (Jones et al., 2006; Kawakami and Butlin, 2012).

The occurrence of hybrid zones has been documented extensively in both plants and animals, including a wide range of vertebrates (such as mammals, birds, reptiles, amphibians, fish etc.) and invertebrates (snails, insects etc.), which offers a unique opportunity to understand the nature and origin of species (Barton and Hewitt, 1985; Sage et al., 1986; Szymura, 1993; Guiller et al., 1996; Seddon et al., 2001; Riginos and Cunningham, 2005; Teeter et al., 2008; Johannesson et al., 2010; Morales-Rozo et al., 2017). Natural hybrid zones also serve as an evolutionary filter through the spreading of advantageous alleles from one species to another, which provides information for the identification of significant adaptive traits in species (Barton and Gale, 1993; De La Torre et al., 2015). Studies on hybrid zones are therefore indispensable for gaining insights about traits and the processes of population divergence and speciation (Jones et al., 2006, Morales-Rozo et al., 2017).

The structure of a hybrid zone provides valuable information about ecological and evolutionary interactions between divergent lineages (Gompert et al., 2017). It contains a wide range of phenotypes and genotypes which can be utilized to analyse the genetic differences and selective processes separating the taxa involved (Barton and Hewitt, 1985). Generally, a hybrid zone between two distinct taxa or populations comprises a set of clines or gradation in the value of phenotypic traits or the frequency of alleles that distinguish them (Barton and Hewitt, 1985; Kawakami and Butlin, 2012). A typical cline is a spatial gradation of a phenotypic trait or allele frequency across a hybrid zone showing a smooth sigmoid shape with a steep transition near the centre and shallow tail

of introgression (the incorporation of genetic materials from one species to another by hybridization and repeated backcrossing with one of its parent species) towards the edges (Kawakami and Butlin, 2012). It can be formed by gene flow from both parental populations and maintained by a balance between dispersal (geographical distance between the reproductive sites of parents and their offspring) and selection against hybrids (Barton and Hewitt, 1985).

The dynamics of the cline such as shape and movement are maintained either by intrinsic (endogenous) selection independent from the environment (e.g. heterozygote disadvantage or epistasis), or extrinsic (exogenous) selection dependent on the environment (e.g. geographic gradient) across the hybrid zone, which determines the role of adaptation in the mechanism of speciation (Endler, 1977; Mallet and Barton, 1989; Kruuk et al., 1999; Alexandrino et al., 2005). In addition, the width of the cline and linkage disequilibrium (gametic correlation or non-random associations between alleles at different loci) between genes across the hybrid zone can both be used to estimate the strength of natural selection and gene flow (Mallet et al., 1990; Barton and Gale, 1993). When linkage disequilibrium (LD) among genetic loci is strongest, clines of all clustered loci will be held together (coincidence and concordance of the cline centre and widths, respectively) by strong indirect selection resulting in much narrower and steeper clines in the centre of the hybrid zone (Barton and Hewitt, 1985; Szymura and Barton, 1991). Therefore, it is important to assess the structure of a natural hybrid zone to understand how dispersal and selection shape the hybrid zone.

The distribution and pattern of phenotypic and genotypic clines in a hybrid zone also provide information to assess the extent of reproductive isolation occurring within the zone (McKenzie et al., 2015). The estimation of a hybrid index (the calculation of similarity of each hybrid individual to the parental types) of multiple loci is another approach for the analysis of hybrid zone including the evaluation of reproductive isolation (Jiggins and Mallet, 2000; McKenzie et al., 2015). If there is excessive introgression between the hybridizing populations (limited reproductive isolation), the frequency distribution of hybrid indices or traits tends towards a unimodal pattern indicative of a hybrid swarm (Gay et al., 2008) (Fig. 1A). If there is high dispersal and strong selection against hybrids resulting from substantial reproductive isolation or strong assortative mating behaviour between the taxa, a bimodal distribution of hybrid indices or traits will be observed across the hybrid zone (Jiggins and Mallet, 2000) (Fig. 1B). Typically, a unimodal zone is characterized by the presence of Hardy-Weinberg equilibrium (HWE) and linkage equilibrium at multiple loci due to heterozygote advantage, whereas a bimodal zone represents the existence of heterozygote deficit and

77

LD due to prezygotic isolation or local adaptation of either parental population to divergent habitats (Harrison and Bogdanowicz, 1997; McKenzie et al., 2015). In addition, a trimodal distribution is also observed in many hybrid zones due to varying levels of introgression, where hybrids form a distinct group with intermediate allele frequencies and extensive genetic variance (Gay et al., 2008) (Fig. 1C).
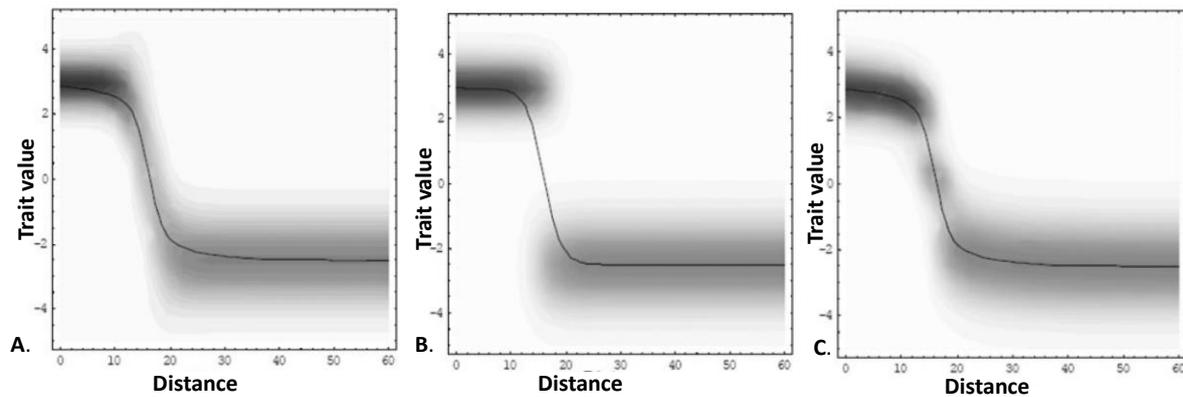
Fig.1 Distribution of individual phenotypic trait based on modelling of morphological clines at a given distance across a contact zone showing example of a A. unimodal, B. bimodal and C. trimodal distribution (adapted from Gay et al., 2008).

Hybrid zones serve as a window for evolutionary study, providing insight into the identification of significant adaptive species traits through the transference of advantageous alleles from one species to another (Harrison, 1990; Barton and Gale, 1993). The pattern of phenotypic or genotypic clines also provide insight into the traits or genes under selection that are responsible for specific morphological, physiological and reproductive changes among taxa (Teeter et al., 2010). Depending on the strength of selection or linkage acting on a specific locus or quantitative trait, the centre and steepness of clines may vary among markers with fixed differences adapted to divergent environments (Jiggins and Mallet, 2000; McKenzie et al., 2015). For example, if a phenotypic trait or genetic locus is under positive selection in freshwater, there will be an asymmetrically broad cline usually with high frequencies of the freshwater allele of such genes or traits toward the marine water. When some genetic markers exhibit narrow cline widths, it indicates an association of such genomic regions with reproductive isolation, including behavioural and physiological responses (Teeter et al., 2008). However, it is possible to detect genomic regions underlying divergently-selected traits by investigating the association between phenotypic and genotypic characteristics of a hybrid zone (Nadeau et al., 2014; Hollander et al., 2015; Westram et al., 2018). It has been established that the phenotype of an organism depends on the combined action of many genes, with the hybrid zones acting as a sieve for the rapid detection of adaptive loci underlying phenotypic variation (Barton, 1983; Rieseberg and Buerkle, 2002; Lindtke et al., 2013; Bresadola et al., 2019). Therefore, identification of differences in

the pattern of phenotypic variants and the genes under natural selection is the important step in identifying the functional mechanism and adaptive evolution of a species.

The three-spined stickleback (*Gasterosteus aculeatus*) (hereafter 'stickleback') is a model organism in eco-evolutionary research due to its wide distribution across different habitat types and as a host to a variety of parasites (Scharsack et al., 2016). Since the last glaciation period (~10,000 years ago), the colonization and adaptation of marine stickleback within different freshwater environments like rivers, lochs and streams, provides an unusual opportunity to identify the genetic basis of repeated evolutionary change in natural populations. Many studies have shown that marine forms of stickleback undergo repeated changes in body shape, skeletal armour, pigmentation, salt tolerance, feeding habits, mating preferences and life histories when evolving within different types of freshwater habitats (Bell and Foster, 1994; McKinnon and Rundle 2002; Hendry et al., 2013). However, the genetic and molecular changes that underlie the phenotypic diversity and adaptation of stickleback to divergent environments are still under investigation (Peichel and Marques, 2017).

A genome-wide linkage map was developed initially for stickleback to analyse the recently-evolved phenotypic changes (skeletal armour and feeding morphologies) between populations that have adapted to different environments (Piechel et al., 2001). Following this, a number of additional genomic resources have been developed for this system such as large-insert genomic libraries and physical maps (Kingsley et al., 2004; Kingsley et al., 2007), restriction site-associated DNA sequencing (RAD-seq) (Miller et al., 2007b; Baird et al., 2008; Hohenlohe et al., 2010), genome-wide SNPs genotyping arrays (Jones et al., 2012b) including QTL mapping of many different phenotypes (e.g. plate morph, pelvic loss, lateral line patterning, schooling body position, pigmentation, tooth number etc.) (Shapiro et al., 2004; Colosimo et al., 2004, 2005; Miller et al., 2007a, 2014; Chan, 2010; Wark, 2012; Greenwood et al., 2011, 2013; 2016; Cleves et al., 2014; Mills et al., 2014; Ellis et al., 2015).

Gene expression assays have also been conducted to identify genes that show differential expression between marine and freshwater environments (Kitano et al., 2010; Lenz et al., 2013; Nikinmaa et al., 2013; Greenwood et al., 2016; Robertson et al., 2017), but are unable to discriminate between the direct target of selection and the downstream consequence of such selected changes (Hendry et al., 2013). The most notable work on the genomic basis of adaptive evolution of stickleback is the development of a high-quality reference genome assembly revealing 81 genomic regions that are consistently associated with marine-freshwater divergence (Jones et al., 2012a). These have become the focus for fine-scale mapping of candidate genes to better

explore links between divergence in genotypes, phenotypes and adaptation. To understand the dynamics of evolutionary changes in nature, it is important to know about the candidate genes underlying phenotypic diversity to demonstrate that variation in the gene is responsible for variation in the phenotype of interest (Peichel and Marques, 2017). But the identification of the genes associated with specific phenotypic traits is still a major challenge (Hendry et al., 2013; Pallares et al., 2014).

To date, only a few genes have been identified for specific phenotypes of stickleback: 1. Eda gene for plate morph, lateral line patterning, schooling body position (Colosimo et al., 2004, 2005; Wark, 2012; Greenwood et al., 2016), 2. Pitx1 gene for pelvic reduction (Shapiro et al., 2004; Chan, 2010), 3. Gdf6 gene for plate size (Colosimo et al., 2004, 2005), 4. Kitlg gene for pigmentation (Miller et al., 2007a), 5. Bmp6 gene for tooth number (Cleves et al., 2014; Ellis et al., 2015; Mills et al., 2016) and 6. MSX2A gene for dorsal spine reduction (Howes et al., 2017). However, a candidate gene approach can be applied to identify SNPs associated with specific phenotypic traits such as salinity, plate morph and immunity in naturally-occurring hybrid fish. Recently, some studies have researched natural hybrid zones that occur between marine and freshwater stickleback to observe their adaptive divergence into different habitats (Jones, et al., 2006; Taugbøl et al., 2014; Ravinet et al., 2015; Vines et al., 2016; Pedersen et al., 2017). Most of these studies found strong differentiation among two divergent lineages concerning the variation of lateral plates due to its clear genetic basis associated with the Eda gene. In this study, I focused on a natural putative hybrid zone in a stream that runs between a freshwater loch (Loch Hosta) and the Atlantic Ocean in North Uist, Scotland. This is the first time this site has been studied in order to observe the strength of selection on the changes in genetic variation of selected diagnostic loci along with different phenotypic traits.

Firstly, the Eda gene located on chromosome IV was examined which is known to be responsible for major morphological differences between freshwater and marine stickleback (Colosimo et al., 2004; Barrett et al., 2008). It determines the lateral armour plates (hereafter 'plate morph') in stickleback based on a polymorphism in the second intron of this gene (Hagen 1967; Colosimo et al., 2004, 2005). The major aim of genotyping the Eda locus was to confirm the genetic evidence of the 3 types of plate morphs: 1. Completely-plated morph (CM) with 30-34 plates in a continuous row, 2. Partially-plated morph (PM) with discontious row of plates and 3. Low-plated morph (LM) ususally with 4-8 plates present in the Hosta stream as a hybrid zone (Jones, et al., 2006). Secondly, the distribution of Mitochondrial Cytochrome b (Cyt b) was observed across the hybrid zone to find out the genetic origin of stickleback based on trans-

Atlantic (At) and European (EU) lineages. Previous research on mitochondrial DNA has confirmed these two major evolutionary lineages among populations of sticklebacks in the European distribution range including the Black Sea lineage (Orti, et al., 1994; Mäkinen and Merilä, 2008; Dean et al., 2019). I predicted that the hybrid individuals of the hybrid zone will be admixed of both lineages due to temporal dispersion of parental anadromous stickleback from the Atlantic Ocean.

Finally, three SNP assays were developed for specific genes (PPARA, WNT7B and NLRC5) by using the PCR-RFLP (Restriction fragment length polymorphism) method from a panel of 10 candidate genes selected from previous workon stickleback utilizing genome scans and QTL analyses of adaptation related to salinity and immunity (Hohenlohe et al., 2010; Jones et al 2012a; Terekhanova et al., 2014; Ferchaud et al., 2014; Haenel et al., 2019). Although PPARA (lipid and glucose metabolism), WNT7B (paracrine signalling pathway of kidney) and NLRC5 (immunity response toward microbial invasion) are known for specific phenotypes, previous work on the stickleback genome assumed that some of these genomic regions are the result of adaptation to differences in salinity (Yu et al., 2009; Neerincx et al., 2010; Jones et al., 2012a; Wafer et al., 2017). I predicted that unlike the well-known standing genetic variation in the Eda locus (Colosimo et al., 2004, 2005), the point mutations (SNPs) of PPARA, WNT7B and NLRC5 loci will separate marine anadromous and freshwater stickleback with fixed differences of homozygous genotype along the at geographic gradient of the Hosta hybrid zone.

In this study, I aimed to address two main research questions: 1) how do phenotypes and genotypes vary across a small-scale hybrid zone? and 2) what does this reveal about adaptation? I quantified variation in phenotypic traits (such as body size and shape, armour plate and spines, parasites etc.) and genotypes of selected genes (Eda, PPARA, WNT7B and NLRC5) along with clinal patterns across the hybrid zone of North Uist. Through observing associations between phenotypic and genotypic variables in sticklebacks (stream-resident and anadromous), I predicted that different phenotypes might have strong relationships with selected candidate gene loci due to similar selection processes in the environmental gradient of the hybrid zone. Therefore, the present study may provide significant evidence to better explore the links between divergence in genotypes, phenotypes and adaptation under natural selection and the mechanism of speciation.

2. Materials and methods

2.1 Study area and sample collection

Fish samples were collected from a putative hybrid zone between divergent stickleback ecotypes (marine and freshwater): the Hosta stream, North Uist, Scotland (Fig. 2). This stream is the outflow of the freshwater Loch Hosta (Hosta) (57°37'40"N; 7°29'18"W) to the Atlantic Ocean and is approximately 1.05 km in length. Samples were collected from eight sites that were on average 108 m apart along the stream (from sea to loch) and also at two sites inside the loch (Fig. 2; Table 1). A small number of samples (N = 20) were also collected from another stream running from the freshwater Loch Grogarry (GROG) to the sea as the nearest known source of pure anadromous fish. Both freshwater and anadromous fish occur at GROG, but there is no evidence of hybridisation.

Fig. 2 Sites for the collection of fish from ten points in the Hosta stream hybrid zone, North Uist, Scotland [A. Map of North Uist Island B. Position of loch Hosta and GROG with streams towards Atlantic Ocean and C. Hosta stream with 10 sites with an average of 100m (approx.) distance from each other].

All fish were collected between the dates of 08-16th May 2018. Fish were caught in Gee's Minnow Traps (Gee traps, Dynamic Aqua, Vancouver, Canada), set overnight (approximately 16h) at ten sites in the hybrid zone. The grid reference of each sample site (GPSmap 60CSx, Germin, UK) and water quality parameters including temperature and pH (multi-parameter probe - Multi340/set, WTW, Germany) were recorded from all sites (Table 2). Across all ten sampling sites, 216 live fish samples were haphazardly selected and transported immediately to the laboratory in stream water in darkened

boxes with continuous aeration. From GROG, 20 fish were collected in the same way as a reference group from nearest source of pure anadromous fish (apparently pure anadromous fish do occur in the Hosta stream, at least sporadically, but none were caught during this sampling period) (A. D. C. MacColl, personal communication). All animal work in this thesis was reviewed by the University of Nottingham AWERB, and carried out under Home Office licence 3003415. No licence is required to collect stickleback in Scotland. Fish were euthanized with an overdose of MS222 (400mgL^{-l}) following Schedule 1 techniques according to UK Home Office regulations and death was confirmed by destruction of the brain.

Table 1: Sampling locations, number of sample fish and physicochemical parameters of the 10 sites of HOSTA and loch GROG at North Uist, England.

Site	Distance from the sea (m)	Location	Sample size (N)	pH	Temperature (^{0}C)	Date of sample collection
GROG	0	57°36′54″N; 7°30′40″W	20	7.93	12.1	15.05.2018
Hybrid zone 1	150	57°62′72″N; 7°49′91″W	34	7.66	9.3	08.05.2018
Hybrid zone 2	220	57°62′71″N; 7°49′86″W	20	7.68	9.2	09.05.2018
Hybrid zone 3	280	57°62′68″N; 7°49′82″W	19	7.78	9.7	10.05.2018
Hybrid zone 4	434	57°62′62″N; 7°49′70″W	20	7.73	11.7	11.05.2018
Hybrid zone 5	537	57°62′59″N; 7°49′59″W	20	7.79	10.4	14.05.2018
Hybrid zone 6	762	57°62′54″N; 7°49′36″W	13	7.63	9.7	10.05.2018
Hybrid zone 7	820	57°62′56″N; 7°49′30″W	20	7.76	9.8	13.05.2018
Hybrid zone 8	925	57°62′53″N; 7°49′20″W	20	8.57	14.5	14.05.2018
Hybrid zone 9 (inside the loch)	1090	57°62′62″N; 7°49′09″W	30	8.05	10.6	08.05.2018
Hybrid zone 10 (inside the loch)	1230	57°62′72″N; 7°49′03″W	20	7.85	13.6	16.05.2018

For the preliminary identification of ecotypes (anadromous, hybrid or resident) of fish, the standard length (mm), weight (gm) and presence of lateral plates and keel were recorded, alongside a digital photograph of the left side for further morphometric study. Reproductive status was recorded by macroscopic inspection of the size or appearance of the gonads following dissection using a binocular microscope (Euromax, Holland). For females, a four-stage classification of the reproductive status was observed based on the macroscopic appearance of the ovaries outlined by Schultz, et al., (2006). For male

reproductive status, a three-stage classification was followed based on the qualitative scale on size of testes and kidney (Whiting et al., 2018).

Samples were also examined for their natural infection status by observing nine different macroparasites: *Gyrodactylus* spp. and *Dermocystidium* spp. from the skin surface including fins, *Apatemon* spp., *Diplostomum spathaceum* (in the lens) and *Diplostomum gasterostei* (in the retina) from the eyes, *Schistocephalus solidus* from the body cavity, *Proteocephalus filicollis*, *Diphyllobothrium* spp. and Nematode from the intestinal cavity following the method described in El Nagar and MacColl (2016). For genetic analysis, two pectoral fins and the caudal fin were collected in 100% ethanol and stored at -20^0C, and the rest of the fish was preserved with a tag in 70% ethanol for further morphometric study.

2.2 Morphometric analysis

To quantify geometric body shape variation, a total of 25 landmarks (Fig. 3) were placed on all 236 fish photographs based on Svanback's landmark system (Svanbäck and Eklöv, 2003) using tpsDig232 (Rohlf, 2010). The digitised landmark positions were then analysed using the centroid variation of each landmark in MorphoJ (Klingenberg, 2011) following the procedure of Magalhaes et al., (2016). Principal component analysis (PCA) was performed using the covariance matrix of the log centroid size of the landmarks to explore shape variation in hybrid zone sticklebacks. To make a comparison based on location, the sampling sites were grouped into 1. Hybrid zone fish between the bottom (BS) (sites 1-4) and top of stream (TS) (sites 5-8), 2. loch fish (LF) (sites 9-10) and 3. the GROG (GF) with marine fish.
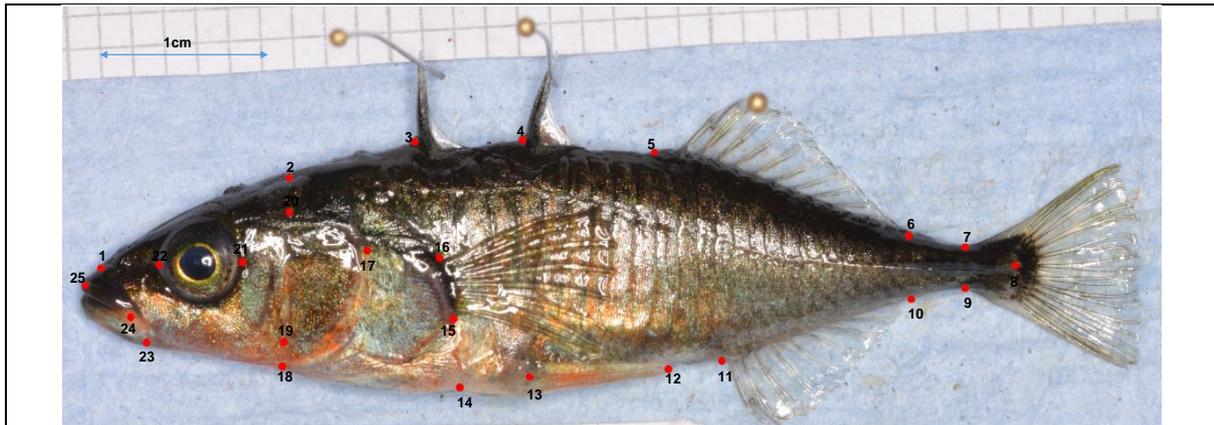
Fig. 3 Location of 25 landmarks on stickleback sample for morphometric analysis of shape variation. The landmarks refers to 1) Top of upper lip; 2) Posterior extent of supra-occipital; 3) Anterior insertion point of primary dorsal spine; 4)Anterior insertion point of secondary dorsal spine; 5) Anterior insertion point of tertiary dorsal spine; 6) Posterior insertion point of dorsal fin; 7) The dorsal narrowest depth of the caudal peduncle together with point 9; 8) Posterior insertion point of anal fin; 9) The ventral narrowest depth of the caudal peduncle together with point 7; 10) Posterior insertion point of anal fin; 11) Anterior insertion point of anal fin; 12) Posterior tip of pelvic fin; 13) Anterior insertion point of pelvic spine; 14) Anterior insertion point of pelvic girdle; 15) Lower insertion point of pectoral fin; 16) Upper insertion point of pectoral fin 17) Posterior extent of operculum 18) Intersection of operculum with the line of the belly; 19) Lower corner of facial plate 20) Upper corner of facial plate 21) Posterior extent of the eye; 22) Anterior extent of the eye; 23) Posterior corner of angled jaw region; 24) Posterior extent of the gape; 25) Tip of the snout (Source: Svanbäck and Eklöv, 2003).

To collect data on external bony skeletal structures (armour), all ethanol-preserved samples were stained with Alizarin Red solution following a standard staining method described in 2.3 in chapter 2 (Peichel et al., 2001). Finally, samples were stored in 40% isopropyl alcohol (propan-2-ol) which was changed weekly to remove purple dye. After confirming the appearance of bony parts, a digital photograph of the right side of every fish was captured and the length of the first dorsal, second dorsal, pelvic spine (from insertion point to the tip), pelvis height and pelvis length were recorded using tpsDig232.

The total number of lateral plates was counted from the stained fish photograph and used to classify fish plate morphology (plate morph) into three groups representing Eda phenotypes: completely plated with 30-34 continuous row of plates (CM), partially plated with discontinuous row of plates (PM) and low (4-8) plated (LM) (Hagen, 1967) (Fig. 4).
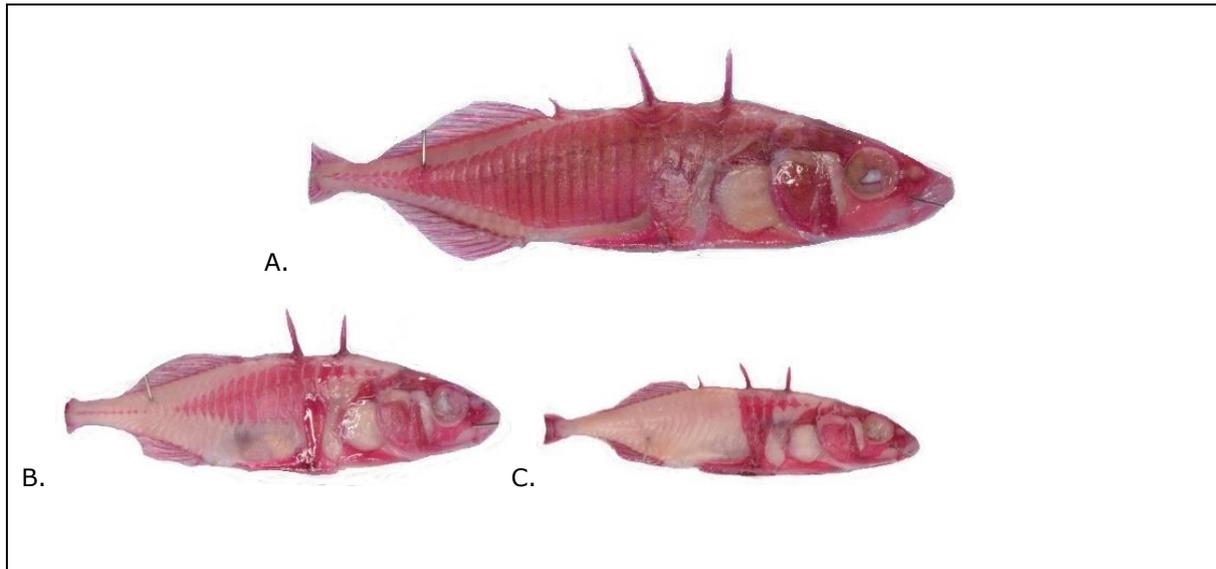
Fig. 4 Example of the different armour phenotypes to describe three ecotypes found in the hybrid zone: A) Completely plated fish with continuous row of plates (CM) B) Partially plated fish with discontinuous row of plates (PM) and 3) Low plated (LM) fish.

2.3 Genetic analysis

Genomic DNA was extracted from pectoral fins of all 236 fish using Quanta Bio-ExtractaTM DNA prep for PCR Tissue kit (UK) following a proteinase K and ethanol precipitation procedure (Goldenberger et al., 1995). DNA concentration was measured with a Bio-drop spectrophotometer (Biochrom, USA) following the manufacturer's protocol and then preserved at -20^0C for further analysis. The extracted DNA was used for the confirmation of phenotypic characteristics (Eda genotyping), mitochondrial cytochrome b genotyping and genotype of SNPs-based candidate gene mapping assay.

2.3.1 Ectodysplasin A (Eda) genotyping

The Ectodysplasin (Eda) gene is responsible for determining the plate morph in stickleback and is well-known as a unique marker with standing genetic variation between maine and freshwater populations across the world (Colosimo et al., 2004; 2005; Kitano et al., 2008). All 236 samples were genotyped for the Eda gene including 20 samples from GROG as a source of anadromous fish. For Eda genotyping, a previously identified microsatellite marker Stn382 (flanks a 60bp indel polymorphism in the 1st intron of Eda gene) was used (Colosimo et al., 2005). This genotyping produces a single band of either 218 bp for the Eda^C allele or 158 bp product for Eda^L allele and double two bands for heterozygotes.

2.3.2 Mitochondrial Cytochrome b genotyping

Mitochondrial Cytochrome b (Cyt b) was genotyped for 80 samples as a subset of 236 fish collected in 2018 to observe the frequency of anciently diverged trans-Atlantic (At) and European (EU) mitochondrial lineages across the Hosta hybrid zone. The Cyt b gene of 1000bp region was amplified from genomic DNA following the method described in Makinen and Merila, (2008). The 80 representative samples were: 20 samples from GROG, 40 samples from four sites in the HOSTA hybrid zone (10 from each of site 1, 3, 5 and 6) and 20 samples from loch HOSTA (site 9 and 10) to limit the cost of sequencing. To confirm the observe genetic pattern based on the mitochondrial lineage of these samples, another 50 stickleback samples were collected and sequenced from the same location in 2019. After amplification and sequencing (Source Biosciences, UK), DNA sequences were aligned using BioEdit version 7.2.5 (Hall, 1999) to identify lineages based on nucleotide differences at diagnostic SNPs.

Following the results obtained above, an experiment was conducted in the Hosta hybrid zone in May, 2019 to determine whether there was any association between mitochondrial haplotypes (At or EU) of individuals (based on Cyt b sequence) and their movement patterns. One hundred anadromous sticklebacks were collected from GROG as a nearest source of pure marine migratory fish and released into the Hosta stream after collection of small fin clips (tip of caudal fin). These were preserved in 96% Ethanol for further analysis. The fish were released at the middle of the stream (57°37'32.8"N 7°29'45.4"W) after setting minnow traps at 4 upstream (+25 m, +50 m, +100 m, +200 m) and 4 downstream (-25 m, -50 m, -100 m, -200 m) locations. Traps were checked every 24hrs until the 5th day after release and all recaptures were euthanised with an overdose of MS222 (400mgL^{-l}) and preserved in 70% ethanol after confirmation of death by destruction of the brain. All recaptured fish (N = 50) were processed and sequenced using the same method to identify the trans-Atlantic lineage (At) and European (EU) mitochondrial lineage.

2.3.3 Development of PCR-RFLP-based SNPs assay

The development of SNP assays for candidate allele polymorphisms was based on PCR-restriction fragment length polymorphism (RFLP) including the following steps: 1. selection of candidate genes from the related literature (Hohenlohe et al., 2010; Jones et al 2012a; Terekhanova et al., 2014; Ferchaud et al., 2014), 2. filtering of SNPs from previously sequenced whole genome data using Unix (Jones et al., 2012a; A.D.C MacColl unpublished data), 3. identification of blocks of 600-1000bp of sequence in the target gene from the Ensembl genome browser 97 (Hubbard et al., 2002), 4. the design of

primers targeting specific restriction enzymes sites using Primer3 (Untergasser et al., 2012), BioEdit (Hall, 1999) and SNP2CAPS (Thiel et al., 2004), 5. the amplification and treatment of target segment with specific restriction enzyme (following Komar, 2009) and finally, 6. the sequencing of 4-8 fish per gene to check differences between marine and freshwater fish. Three SNP-based assays of three genes (PPARA, WNT7B and NLRC5) were developed out of 10 targeted genes (described in Appendix S1) and 236 samples were genotyped for each gene assay.

2.3.4 Primer and PCR condition

The PCR reactions were preliminarily fixed in 20µL volumes. Each reaction contained 3µL DNA template, 10 μl of PCR master Mix (Quanta Bio- Accustart TM II Gel^{TM} Super Mix), 2μl of forward primer (1pmol/μl), 2μl of reverse primer (1pmol/μl) (Sigma) and 3 µL molecular biology grade water. The primers and their properties used in the present study are described in Table 2. Usually, the PCR reaction was performed in a thermocycler (BIO Rad, Mexico) under the following conditions: heated lid at 105°C and initial denaturation for 3 minutes at 94^{0}C, followed by a total of 35/40 cycles of denaturation at 94°C for 15 seconds, annealing at 58°C (depending on primer annealing temperature, Tm) for 15 seconds, and extension at 72°C for 30 seconds, and a final extension at 72°C for 4 min and hold at 4°C. PCR products were separated on a 1% agarose / TAE (2 M tris base,1 M Glacial Acetic Acid, 0.05 M EDTA, pH 8.0) gel at 100 Volts, 110 mAMP for 50 mins to 1 hour. To determine the approximate size of the amplified products, 5 µl of 100bp or 1Kb Hyper ladder (Bio line) was loaded as a molecular weight marker. After treatment of targeted segment with specific restriction enzyme (according to manufacturer protocol, NEB, UK), the PCR product was separated on a 2% agarose gel at 70-100 Volts, 110 mAMP for 50 mins to 1.5 hr. and observed using UV-illumination (Photo Doc-II-imaging system, UVP, UK) for each SNPs assay. After purification [by ExoSAP-IT PCR product clean-up kits (Thermo-Fisher Scientific)] and quantification by UV-spectrophotometer (Bio-drop, USA), DNA was diluted to 15 µg/ml where possible and sent to Source Biosciences (UK) for DNA sequence determination.

Table 2: The primers and their properties, used in the present study for genetic analysis of Hosta hybrid zone, Scotland.

Primer Name	Restriction enzyme (RE)	Primer Sequence (5'-3')	Tm	Product Size (bp)	Fragment length after RE treatment/ alternative alleles	Source
Eda FD Eda ReV	--	5' CCCTTAGAGAATTTCCTAGCAG 3' 5' CTTGTCCCGGATCATACGC 3'	58^0C	218bp/ 158bp	--	Colosimo et al., 2005
Cytb FD Cytb ReV	--	5' ATGAAACTTTGGTTCCCTCC 3' 5' CGCTGAGCTACTTTTGCATGT 3'	60^0C	1000bp	--	Makinen and Merila, 2008.
PPARA FD (P2) PPARA ReV (P2)	BstUI	5'-TCATGACTCTTCTCTGCTGC-3' 5'-CAGCATGTATGTGGACTTCTTC-3'	60^0C	331bp	273bp 58bp	developed in the present study
WNT7B FD (P1) WNT7B ReV (P1)	Msal	5'-GTTCCCCTCTCACTGATCCC-3' 5'-ATAACGGCGCAGAAATTCCC-3'	60^0C	429bp	333bp 96bp	developed in the present study
NLRC5 FD (P1) NLRC5 ReV (P1)	AluI	5'-AGCGGGACTACAAAGGGATC-3' 5'-CAGGATTCAGCACACCACAC-3'	60^0C	460bp	355bp 125bp	developed in the present study

2.4. Statistical analysis

All data were collated in Microsoft Excel and statistical analyses were conducted using R, version 3.6.3 (R Core Team, 2020). All phenotypic and genotypic data were analysed based on the geographical distance from the sea (0 m) towards the loch (~1km) to quantify spatial variation of each trait. The sampling sites included GROG (taken to be 0 m from the sea) as a control site for marine anadromous fish, eight sites from the stream Hosta and two sites (site 9 and 10) from the loch Hosta and considered as three geographic groups: GROG, hybrid zone and loch.

Phenotypic variation across the hybrid zone

Standard length, total lateral plate count, length of body armour and spines, body shape and parasite abundance were analysed according to distance from the sea (Appendix Fig. S1). All measured armour traits including 1st dorsal spine length (first dorsal spine), 2nd dorsal spine length (second dorsal spine), pelvic spine length (pelvic spine), length of pelvis (pelvis length) and height of pelvis (pelvis height) were size-standardized by calculating the residuals of a regression against standard length. Principal component analysis (PCA) was performed to explore variation in all armour regression residuals described previously in Chapter 2 as well as total lateral plate number (hereafter 'total plate count') and to determine the axis of highest variation in bony armour among three morphs of fish (completely plated, partially plated and low plated). The PCA for shape variation from the morphometric data was analysed and grouped according to four categoric variables: site (loch, top of stream, bottom of stream and GROG), sex, and plate morphology (completely plated, partially plated and low plated). A one-way analysis of variance (ANOVA) and t-test were also performed to make comparisons between different morphological traits.

Genotypic variation across the hybrid zone

The allele frequencies of the four genes (Eda, PPARA, WNT7B and NLRC5) were calculated for the 11 populations (sites), from a total of 236 samples, to observe their distribution across the hybrid zone (Appendix Table S1). Genetic diversity wasestimated with the test of Hardy-Weinberg Equilibrium (HWE), observed heterozygosity (Ho), expected heterozygosity (He), Shannon information index, fixation index (Fst) and Fis (Inbreeding coefficient within individuals) for each locus of four genes in each population using GenAlEx version 6.5 (Peakall and Smouse, 2012). The fixation index or Fst value ranged from 0 to 1 where 0 means no differentiation between populations and 1 means complete diffentiation or no sharing of genetic materials (Bird et al., 2017). A hybrid index (HI) for four loci was calculated by counting the number of anadromous / marine alleles across the hybrid zone to observe the allele frequency difference between marine

and freshwater populations. The program GENEPOP version 4.1.4 (Raymond and Rousset, 1995) was used to test for linkage disequilibrium (LD) among four loci for 236 samples collected from all populations (GROG, Hybrid zone site 1 to 8 and loch zone site 9 and 10). Pearson's chi-square test was performed for the Cyt b gene sequence data to observe the differences in the frequency and movement of At and EU lineages across the hybrid zone.

Cline analysis for phenotypic and genotypic traits

To observe the effect of dispersal, clinal patterns were observed for gradual changes in the phenotypic traits and genotypic frequencies across the Hosta hybrid zone from the sea (0 m) towards the loch (1230 m). Seven morphological traits including standard length, total number of lateral plates, armour PC1 and PC3, shape PC1 and PC3 and parasite abundance (*Gyrodactylus* sp.) as well as allele frequencies of four loci (Eda, PPARA, WNT7B and NLRC5) were analysed to fit clines using the program HZAR (Hybrid zone analysis for R) following Derryberry et al., (2014). The lateral plate number and *Gyrodactylus* parasite abundance were log-transformed before running the program. The classic equilibrium models were fitted using the Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithm with likelihood functions, and the best fitting model selected based on Akaike's Information Criterion (AIC). The cline parameters included centre, width, muL, muR, varL and varH for morphological traits and Pmin and Pmax for the allele frequency of candidate loci. Pearson's correlations were performed between allele frequency of marine genotypes and distance from the sea to observe their clinal pattern across the hybrid zone.

Phenotype-genotype association across the hybrid zone

Out of the 216 fish samples from the Hosta stream and loch (without GROG), nine samples were removed due to missing data [one sample in morphometric data for shape, 5 samples in WNT7B genotype data and three samples in NLRC5 genotype data]. Generalised Linear Models (GLMs) with a Gaussian distribution and identity link function were fitted individually to observe the association of genotype-phenotypic traits (Eda genotype, WNT7B genotype, PPARA genotype, NLRC5 genotype, location (hybrid zone and loch), plate morph and sex of fish) using standard length, armour PC1 and PC3 without plate count, shape PC1 and PC3 as response variables. A GLM was also fitted for total plate count using a Poisson distribution and log-link function in relation to the genotypes of the four loci, location, standard length, plate morph, armour PC1 and PC3, shape PC1 and PC3 and sex. Asymptotic tests for the equality of coefficients of variation (CV) (Feltz and Miller, 1996) were also performed to quantify the differences in CV among genotypes and phenotypes for standard length and total plate count of fish. A

GLM of the most abundant parasite (*Gyrodactylus*) was also fitted using a negative binomial distribution and logit-link function with all four loci genotypes, sex, armour PC1 and PC3, shape PC1 and PC3, plate morph and location as predictor variables.

To quantify any association of the EU mitochondrial lineage with location, genotype (Eda, WNT7B, PPARA and NLRC5) and phenotype (sex, length, armour PC1 and PC3, shape PC1 and PC3 and parasites including *Gyrodactylus, Diplostomum* in lens, *Schistocephalus*, Apatemon and Nematode), a GLM was fitted with a binomial distribution and logit-link function.

Stepwise regression with a combination of forward and backward selection based on likelihood ratio tests was conducted for all GLM models. The best fitting model was then selected based on AIC. All models were checked for goodness-of-fit using Quantile-Quantile (Q-Q) plots of the residuals and the significance of each component was tested using ANOVA tests.

3. Results

3.1 Physio-chemical properties of the hybrid zone

Properties of the eleven sampling sites are shown in Table 1. The pH and temperature of the sampling sites ranged from 7.63 to 8.57 and 9.2°C to 14.5°C respectively.

3.2 Phenotypic variation across the hybrid zone
3.2.1 Fish length and plate number variation

There were significant differences in the length of fish (standard length) and total plate count across the ten sampling sites of the Hosta stream. Mean standard length of stickleback fish shows no specific pattern across the stream (150 m to 925 m) with highest length (67.84 ± 1.25 mm) in the GROG site (0 m) representing marine anadromous fish (Fig. 5A). Furthermore, the mean total plate count of fish in the loch (1090 m and 1230 m) was significantly lower than fish in the stream sites (one-way ANOVA: F = 27.4, df = 1, 214, p< 0.001) (Fig. 5B). The mean standard length and total plate count of GROG fish (0 m) was significantly higher than the fish from all other sites [two sample t (224) = 18.11, p<0.001 and 19.84, p<0.001 respectively)] (Fig. 5 A-B).

A.

B.
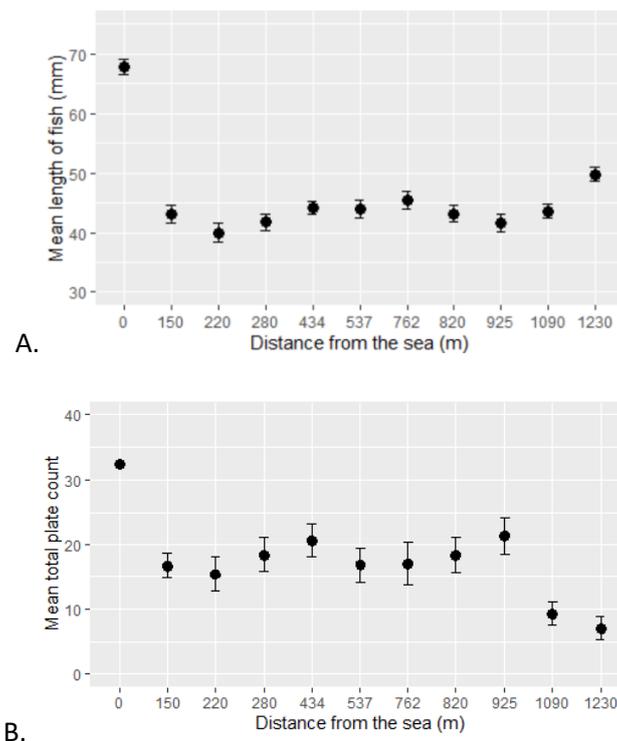
Fig. 5 A. Mean fish length (± S.E.) across the ten sites (from sea to loch) shows no specific pattern or gradient in standard length of fish B. Mean total plate count (± SE) across ten sites (from sea to loch) shows a lower number of plates in the loch (1090 m and 1230 m) and GROG (0 m) with the highest average number of fully-plated stickleback fish.

93

3.2.2 Body shape variation

Geometric-morphometric analysis of body shape using a PCA of 25 digitized morphological landmarks of 235 sticklebacks showed significant variation across the sample sites ($F = 29.39$, $df = 10$, 224, $p < 0.001$) (Fig. 6A-B), plate morph ($F = 30.06$, $df = 2$, 232, $p < 0.001$), location [loch fish (LF), stream fish: four sites in the bottom of stream (BS) and four sites in the top of stream (TS) and GROG (GF)] ($F = 88.05$, $df = 3$, 231, $p < 0.001$) and sex ($F = 8.69$, $df = 1$, 233, $p = 0.003$). The first Principle Component (PC) of body shape explained 27.3% of the total variation with changes in the forward position of pelvic girdle, pectoral fin base and backward position of snout, mouthparts and operculum, especially in the loch fish (LF) (Fig. 6C). PC1 also exhibited prominent differences in body shape with slightly elongated profile including caudal peduncle (Fig. 6D). The second PC (16.9%) was ignored as it represents malformation (twisting) of the specimens during preservation. PC3 accounted for 9.70% of variation and corresponded to prominent shape changes with a deep, wider body profile at the dorsal and pelvic region (Fig. 6C, E).
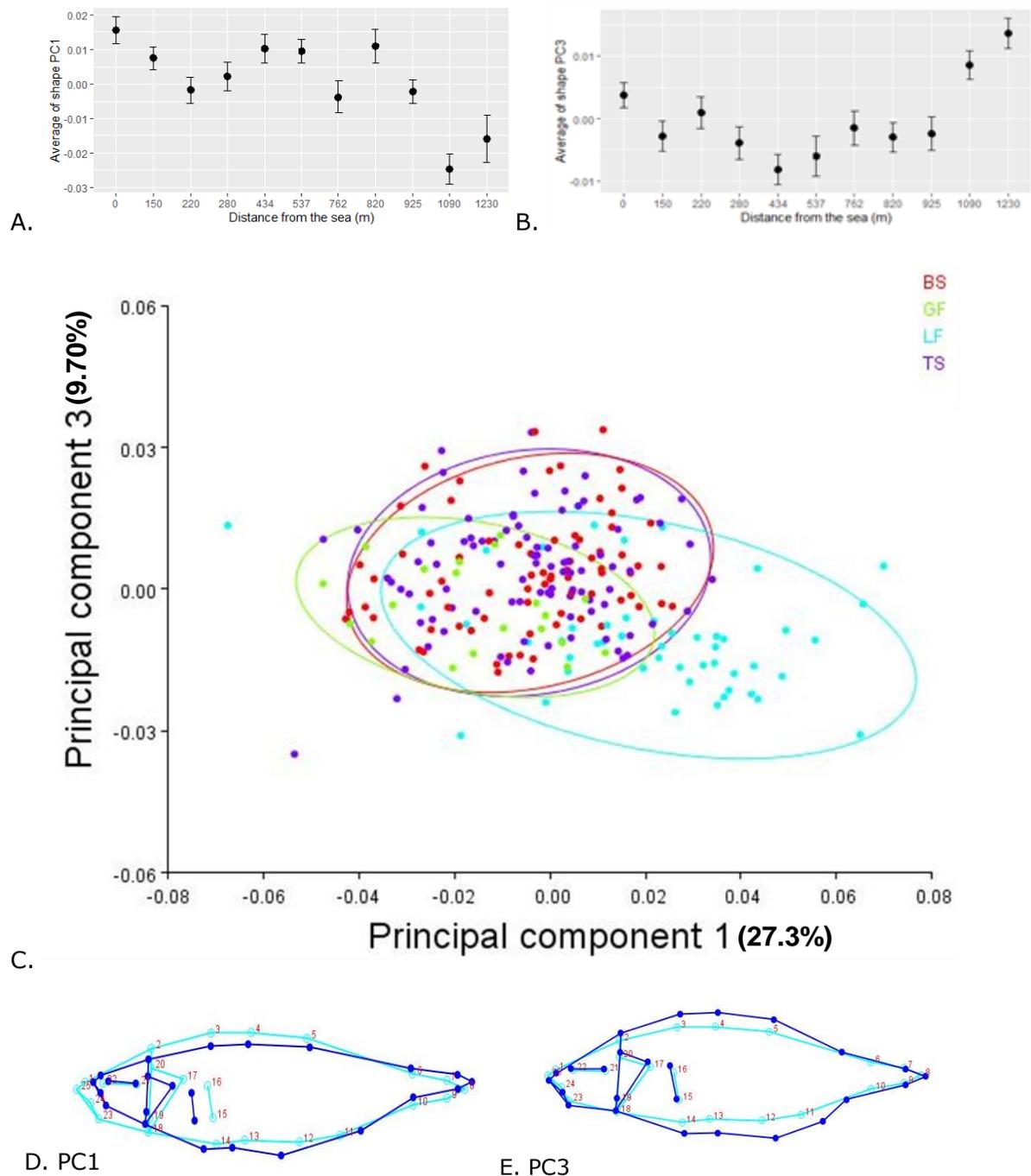
Fig. 6 A) PC1 and B) PC3 of body shape (mean ± S.E.) showing variation at different sites across the hybrid zone and GROG (0 m) as control group. C) The first and third body shape PCs derived from 25 landmarks data explaining 37% of total variation within the data. The distribution of shape phenotype and their associated 95% confidence ellipses (based on four groups of location) across two PCs showing significant differences among loch fish (LF) (Turquoise), stream fish [from the bottom of stream (BS) (Red) and top of the stream (TS) (purple)] and anadromous fish from GROG (GF) (light green). D) Warped outline drawings of PC1 and E) PC3 showing direction of shape change among sample fish (Dark blue = observed, Turquoise = standard).

3.2.3 Body armour and spine variation

All five measured traits of bony armour (not including total plates) showed variation across the 10 sites of the hybrid zone, specifically a general trend of decreasing size from the sea towards the loch (1090 m and 1230 m) (PC1, $F_{1, 235}$ = 39.83, p< 0.001 and PC3, $F_{1, 235}$ = 0.04, p> 0.05, Fig. 7A). Samples from GROG represented a different pattern in pure anadromous fish, with these fish having larger structures except for the pelvic spine (Fig. 6A). The first and third PCs of body armour integrating length of 1st dorsal spine, 2nd dorsal spine, pelvic spine, length and height of pelvis accounted for 59.9% (PC1) and 11.0% (PC3) of the total variation among all five phenotypes (Fig. 7B).

A.

B.
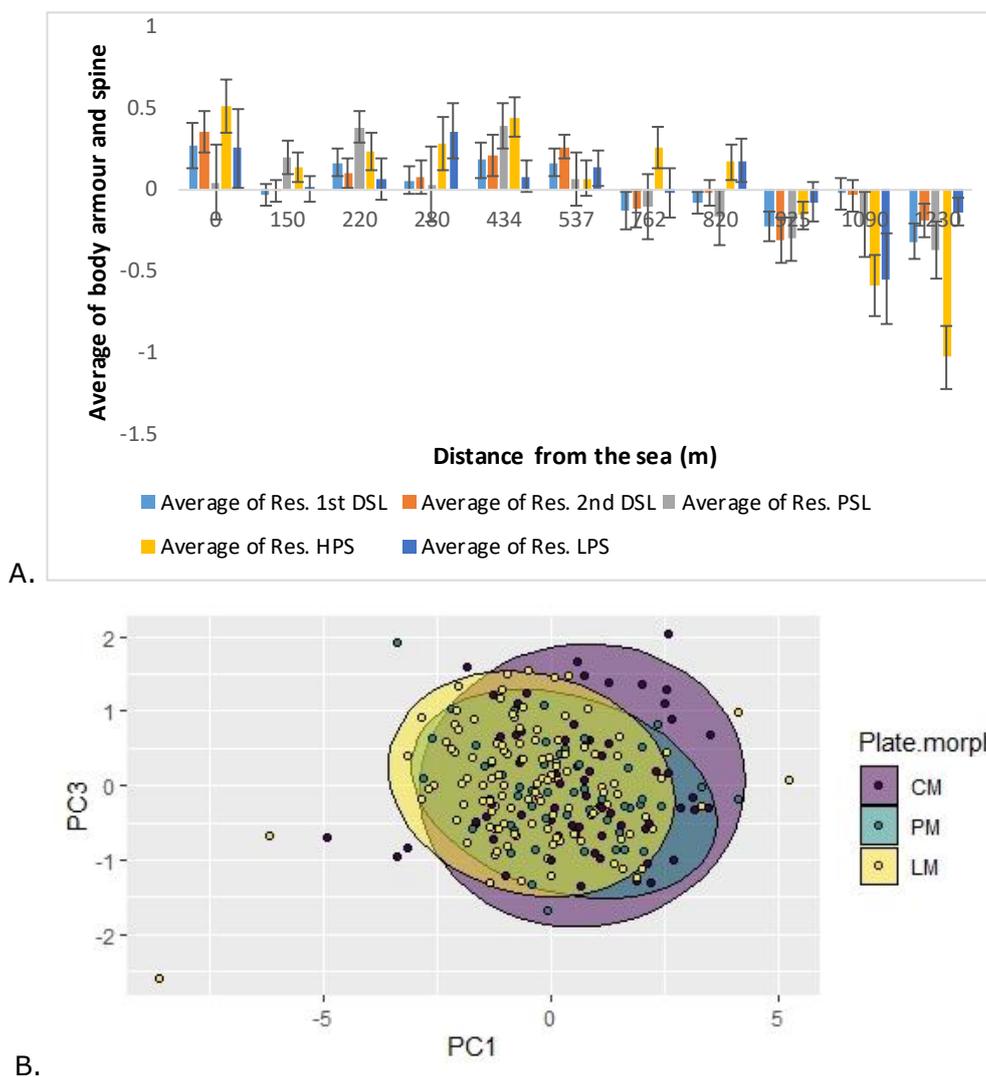
Fig.7 A. Average (± SE) of five armour traits: 1st dorsal spine length, 2nd dorsal spine length, pelvic spine length, height of pelvis and length of pelvis with their residual values showing variation in size across the hybrid zone with gradual decreasing trend towards the loch B. The first and third PCs of body armour traits with residuals showed 70.9% of total variation among five measured armour elements and total plate count in three morph types - completely plated (CM), partially platted (PM) and low plated (LM).

3.2.4 Variation in parasite abundance

Parasite abundance from the sampled fish varied across the hybrid zone, with increasing diversity towards the loch (762-1090 m) (Fig. 8A). Among them, *Gyrodactylus arcuatus* was the most frequent in all sample sites with mean abundance significantly highest in the loch (1090 m and 1230 m) compared to the stream sites and GROG (one-way ANOVA: $F = 4.85$, $df = 2, 233$, $p = 0.008$) (Fig. 8B).

A.

B.
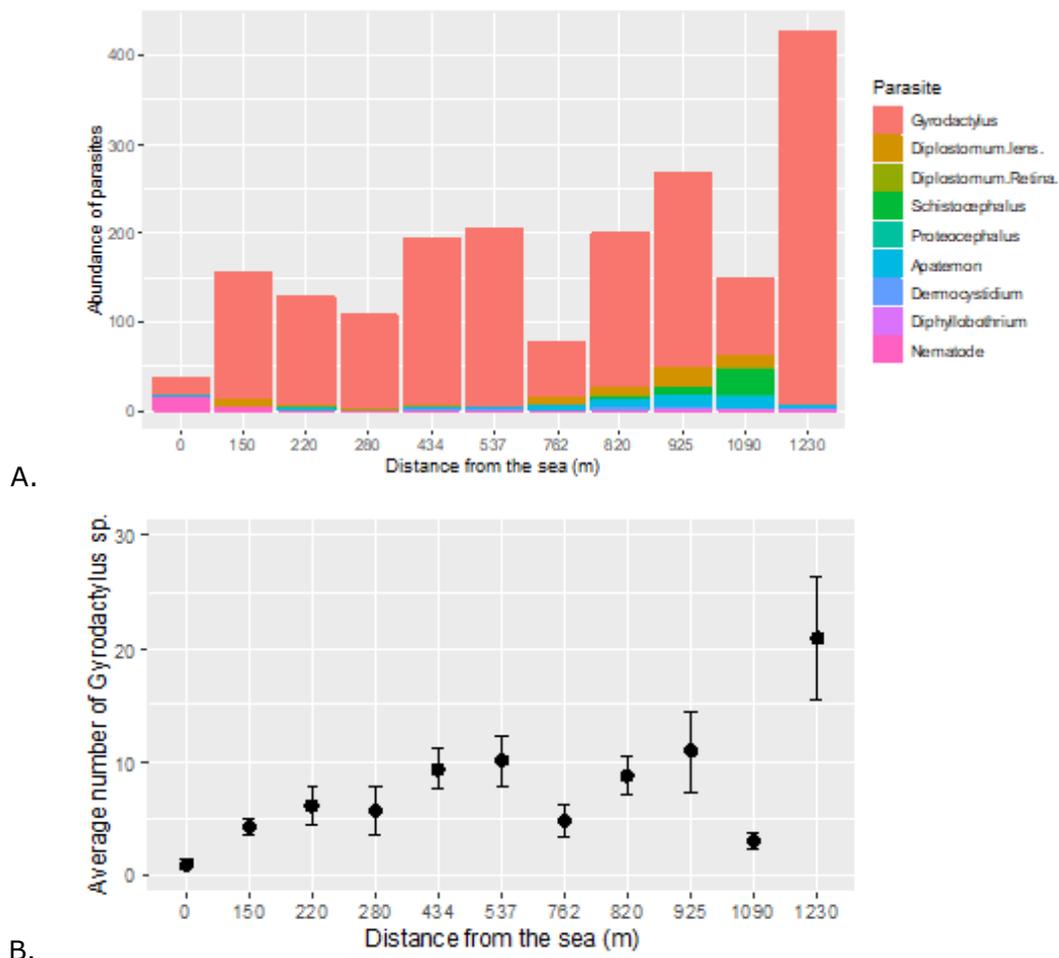
Fig. 8A. The abundance of nine types of macroparasites: *Gyrodactylus* spp. *Dermocystidium* spp., *Apatemon* spp., *Diplostomum spathaceum* (lens), *D. gasterostei* (retina), *Schistocephalus solidus*, *Proteocephalus filicollis*, *Diphyllobothrium* spp. and Nematode identified from the external body surface, eyes, body cavity and intestine of stickleback sample fish across the hybrid zone. GROG (0 m distance from the sea) with fully-plated fish showing the lowest abundance of different groups of parasites than the hybrid zone and loch sticklebacks. B. *Gyrodactylus* sp. was the most abundant group of parasites than all others showing an increasing trend towards loch Hosta for their mean abundance (± S.E.).

97

3.3 Genotypic variation across the hybrid zone
3.3.1 Eda genotyping

Genotypic variation at the Eda locus for 236 fish showed there was a higher frequency of the heterozygous genotype (Eda^{CL}) across the hybrid zone compared to the frequency in both the loch (1090 m and 1230 m) and the sea (0 m) (χ^2 = 130.33, df = 4, p<0.001) (Fig. 9). The allele frequency of marine C homozygotes (complete morph Eda^{C}) showed almost a similar pattern to the phenotypic data (total plate count in Fig. 5B). The only discrepancies were found in sites 5 (537 m) and 9 (1090 m) where there were no CC genotypes found, but phenotypic data showed 3 fully-plated fish in each site.
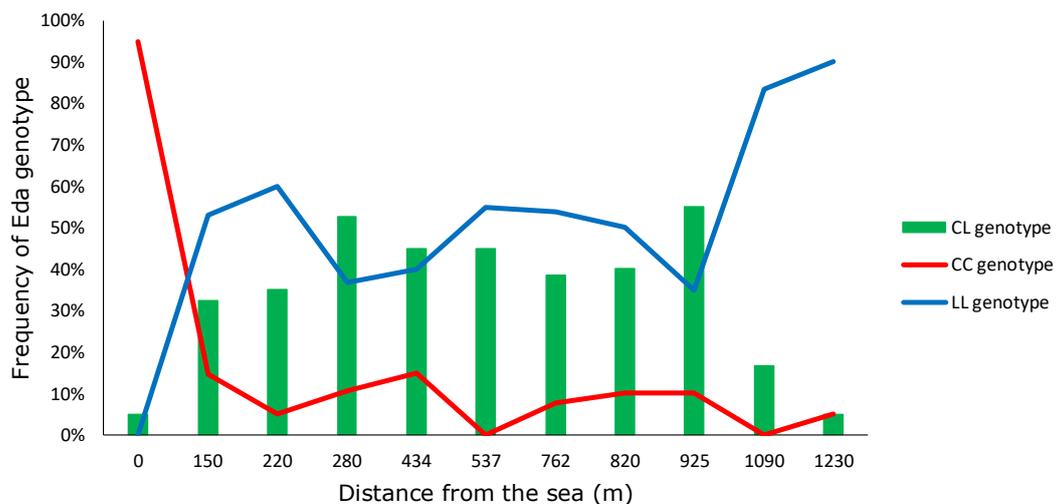
Fig. 9 The genotype frequency of Eda locus for plate morph across the sampling sites showing high frequency of heterozygous CL genotype (green) inside the hybrid zone than the loch (1090 m and 1230 m) and the sea (0 m). Eda genotype for homozygous marine (red) and freshwater alleles (blue) showing an opposite direction to each other across the hybrid zone with a decreasing (CC) and increasing (LL) trends towards the loch, repectively.

3.3.2 Mitochondrial Cytochrome b genotyping

There were three diagnostic SNPs in three positions (approx. 195, 276, 843 bp of 1000 bp) of the mitochondrial Cyt b gene sequence that differentiated the At lineage from EU. There were no At lineage fish in the hybrid zone (n = 40), but 40% of stickleback in Loch Hosta were At (n = 20) and 50% of marine fish from GROG (n = 20) (Fig. 10 A). Therefore, there were significant differences between the distribution of At and EU across the hybrid zone (χ^2 = 270.18, df = 6, p <0.001). However, the experiment on the movement of anadromous fish (n = 100) to observe the effect of mitochondrial lineage on the direction of migration showed no significant differences between upstream (+ 200 m toward the loch) and downstream (- 200 m towards the sea) dispersion of At and EU lineages (χ^2 = 0.058, df = 2, p > 0.05, Fig. 10B) across the hybrid zone.

98
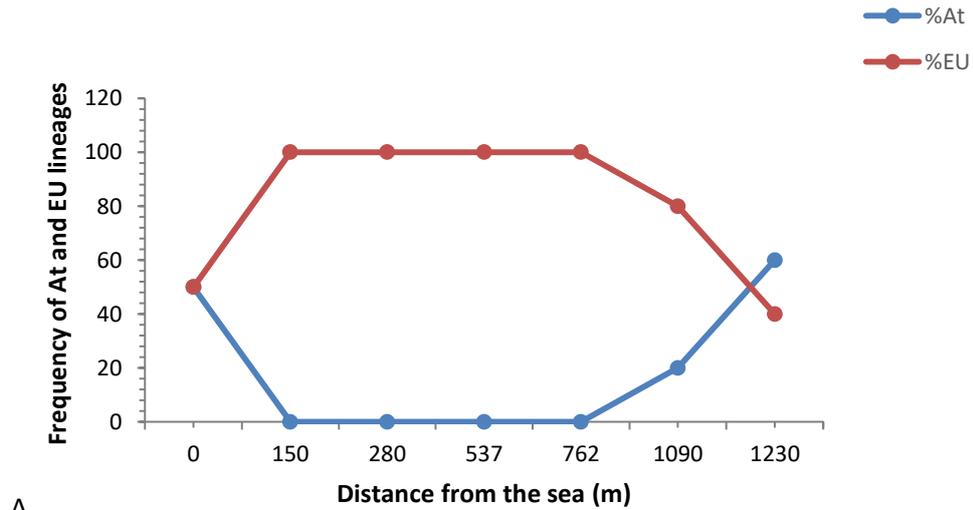
A.
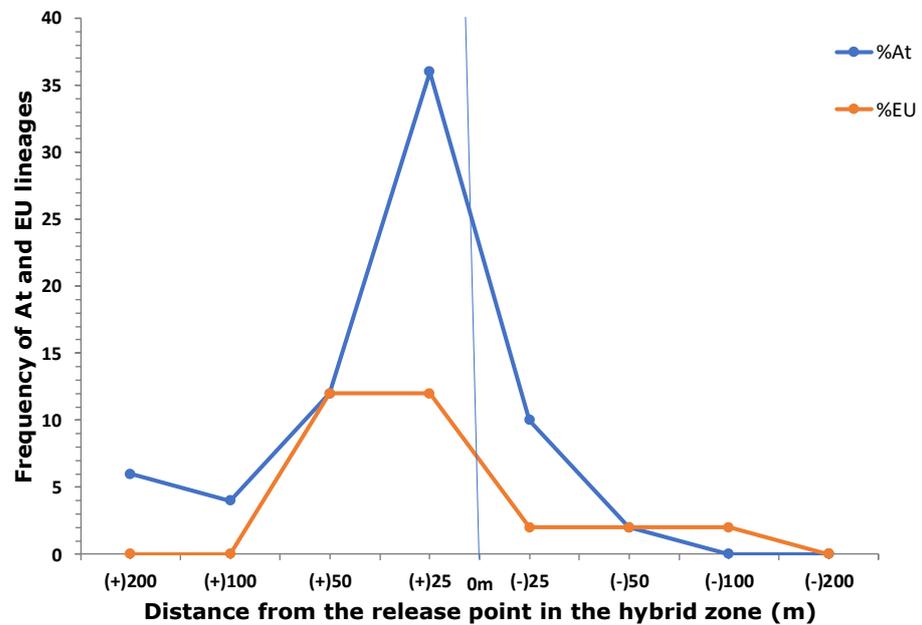
B.

Fig. 10 A Frequency of trans-Atlantic (At) and European (EU) mitochondrial lineages based on SNPs differences in the Cyto b gene showing with only European lineage in the samples among seven sites of Hosta hybrid zone including GROG (0 m) and loch (1090 m and 1230 m) as a representative of freshwater and anadromous stickleback, respectively. B. Frequency of trans-Atlantic (At) and European (EU) mitochondrial lineages in the experimental release of anadromous fish showing no specific pattern of dispersion towards upstream (+) and downstream (-) direction from the release point (0 m) across the Hosta hybrid zone.

3.3.3 SNP assay analyses of PPARA, WNT7B and NLRC5 gene

The SNP assays of three genes (PPARA, WNT7B and NLRC5) showed different patterns of allele frequency between marine and freshwater populations (Appendix Table S1). The outcome of the gel electrophoresis of the PPARA locus showed the heterozygotes (three bands for PR genotype), homozygotes of freshwater (two bandsfor RR genotype) and marine (one band for PP genotype) allele (Fig. 11 A). The amplified segment (429 bp length) of the WNT7B locus showed two bands for the homozygous marine allele (WW genotype), three bands for the heterozygous (WT genotype) and one band for homozygous freshwater allele (TT genotype) (Fig. 11 B). The NLRC5 locus (460 bp length) also displayed heterozygotes with three bands (NL genotype) and homozygotes of marine (two bands for NN genotype) and freshwater (one band for LL genotype) allele (Fig. 11 C).
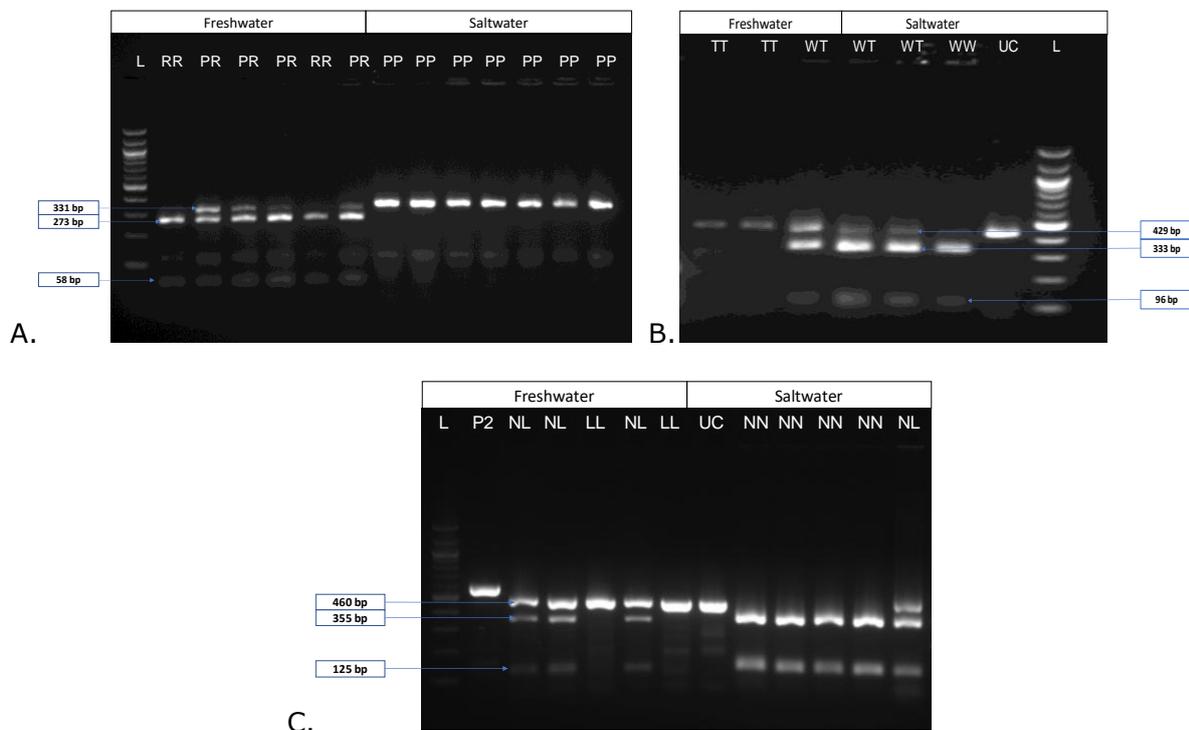
Fig. 11 A. PPARA gene treated with BstUI restriction enzyme showing a single uncut band for 331bp product of homozygous marine PP genotype or a cut band of 273 + 58 bp for homozygous freshwater RR genotype and triple bands for heterozygous PR genotype having copy of each allele B. WNT7B gene treated with restriction enzyme (Msal) showing a uncut (UC) band for 429 bp product of homozygous freshwater TT genotype and a cut band of 333 + 96 bp for homozygous marine WW genotype C. NLRC5 gene treated with restriction enzyme (AluI) showing a uncut band for 460 bp product of homozygous freshwater LL genotype and a cut band of 355 + 125 bp for homozygous marine NN genotype.

3.3.4 Estimation of genetic diversity based on four loci

Frequency-based genetic diversity across the hybrid zone showed variation in allelic diversity, heterozygosity and high level of polymorphism at all four loci (Eda, PPARA, WNT7B, NLRC5), except for the control group, GROG (Table 3). The overall observed heterozygosity was significantly lower than the expected for all populations of four loci (t-test, $t = 2.073$, df = 20, $p = 0.025$). Most populations especially at the downstream end of the Hosta hybrid zone such as HB zone 1, 2, 4, 5 and loch zone 9, 10 showed departure from HWE at Eda, WNT7B and NLRC5 loci (Fisher's method). On the basis of locus, Eda ($\chi^2 = 8.18$, df = 1, $p = 0.004$), PPARA ($\chi^2 = 7.39$, df = 1, $p = 0.007$) and NLRC5 ($\chi^2 = 5.49$, df = 1, $p = 0.019$ showed significant deviation from the HWE. Pairwise population Fst estimated among populations of hybrid zone showed less differentiation (close to each other) at the centre of the stream compared to the two extreme sites (GROG and loch) (Table 4). The overall Fis, Fit and Fst for each genetic locus also showed differentiation among all populations such as Eda: -0.005, 0.234, 0.238, PPARA: 0.102, 0.390, 0.321, WNT7B: 0.483, 0.624, 0.273 and NLRC5: 0.442, 0.561, 0.213. The Fst values evaluated by permutation test showed significant deviation from zero at all loci ($p<0.001$).

Linkage disequilibrium was significant ($p<0.05$) for Eda, NLRC5, PPARA and WNT7B loci in several populations of sticklebacks collected from HB zone 3, 4, 7 and loch zone 9 and 10. Overall, LD was significant for Eda and PPARA ($\chi^2 = 38.7$, df = 20, $p = 0.007$) and PPARA and WNT7B ($\chi^2 = >59.65$, df = 20, $p <0.001$) loci in all populations corresponded to markers located on the same chromosome (Chromosome IV). The mean LD (D bar) estimated from HI of four loci was 0.134 for the central eight sites of the hybrid zone (without GROG and loch) indicating maximum value of LD was around 13% in these sites.

Table 3: Populations, sample size (N), mean marine allele frequencies (AFM), observed heterozygosity (Ho), expected heterozygosity (He), Shannon information index (SI), fixation index (FI) and chi-square estimation of HWE for Eda, PPARA, WNT7B and NLRC5 loci including mean and variance of the Hybrid Index (HI) collected from 236 samples of 11 sites of Hosta hybrid zone, North Uist, Scotland (Highest FI value in bold, positive FI indicates heterozygote deficit and negative value indicates heterozygote excess)**.**

Populations/site	Locus	N	AFM	Ho	He	SI	FI	HWE ChiSq.	HWE P value	HI for 4 loci	Var (HI) for 4 loci
	EDA		0.975	0.05	0.049	0.117	-0.026	0.013	0.909		
	NLRC5		0.900	0.1	0.18	0.325	**0.444**	3.951	**0.047**		
	WNT7B		1.000	0	0	0	#N/A	Monomorphic			
GROG	PPARA	20	1.000	0	0	0	#N/A	Monomorphic		0.97	0.005
	EDA		0.309	0.324	0.427	0.618	0.242	1.994	0.158		
	NLRC5		0.500	0.118	0.5	0.693	**0.765**	19.882	**0.0**		
	WNT7B		0.576	0.121	0.489	0.682	**0.752**	18.656	**0.0**		
HB Zone 1	PPARA	34	0.809	0.382	0.309	0.488	-0.236	1.9	0.168	0.55	0.043
	EDA		0.225	0.35	0.349	0.533	-0.004	0	0.987		
	NLRC5		0.275	0.15	0.399	0.588	**0.624**	7.783	**0.005**		
	WNT7B		0.675	0.15	0.439	0.631	**0.658**	8.662	**0.003**		
HB Zone 2	PPARA	20	0.850	0.3	0.255	0.423	-0.176	0.623	0.43	0.61	0.039
	EDA		0.368	0.526	0.465	0.658	-0.131	0.326	0.568		
	NLRC5		0.684	0.316	0.432	0.624	**0.269**	1.377	0.241		
	WNT7B		0.789	0.211	0.332	0.515	**0.367**	2.554	0.11		
HB Zone 3	PPARA	19	0.895	0.211	0.188	0.336	-0.118	0.263	0.608	0.5	0.035
	EDA		0.375	0.45	0.469	0.662	0.04	0.032	0.858		
	NLRC5		0.775	0.05	0.349	0.533	**0.857**	14.676	**0.0**		
	WNT7B		0.750	0.2	0.375	0.562	**0.467**	4.356	**0.037**		
HB Zone 4	PPARA	20	0.800	0.2	0.32	0.5	0.375	2.813	0.094	0.68	0.05

Table 3: (cont.)

Populations/site	Locus	N	AFM	Ho	He	SI	FI	HWE ChiSq.	HWE P value	HI for 4 loci	Var (HI) for 4 loci
	EDA		0.225	0.45	0.349	0.533	-0.29	1.686	0.194		
	NLRC5		0.350	0.1	0.455	0.647	**0.78**	12.175	**0.0**		
	WNT7B		0.842	0.211	0.266	0.436	**0.208**	0.825	0.364		
HB Zone 5	PPARA	20	0.800	0.4	0.32	0.5	-0.25	1.25	0.264	0.54	0.025
	EDA		0.269	0.385	0.393	0.582	0.023	0.007	0.935		
	NLRC5		0.577	0.538	0.488	0.681	-0.103	0.138	0.71		
	WNT7B		0.654	0.385	0.453	0.645	**0.15**	0.294	0.588		
HB Zone 6	PPARA	13	0.808	0.231	0.311	0.49	**0.257**	0.86	0.354	0.58	0.051
	EDA		0.300	0.4	0.42	0.611	0.048	0.045	0.831		
	NLRC5		0.184	0.263	0.301	0.478	0.124	0.294	0.588		
	WNT7B		0.579	0.211	0.488	0.681	**0.568**	6.134	**0.013**		
HB Zone 7	PPARA	20	0.875	0.15	0.219	0.377	**0.314**	1.976	0.16	0.48	0.038
	EDA		0.375	0.55	0.469	0.662	-0.173	0.601	0.438		
	NLRC5		0.300	0.3	0.42	0.611	**0.286**	1.633	0.201		
	WNT7B		0.600	0.3	0.48	0.673	**0.375**	2.813	0.094		
HB Zone 8	PPARA	20	0.850	0.3	0.255	0.423	-0.176	0.623	0.43	0.53	0.074
	EDA		0.083	0.167	0.153	0.287	-0.091	0.248	0.619		
	NLRC5		0.233	0.267	0.358	0.543	0.255	1.946	0.163		
	WNT7B		0.172	0.069	0.285	0.46	**0.758**	16.677	**0.0**		
Loch Zone 9	PPARA	30	0.317	0.3	0.433	0.624	**0.307**	2.824	0.093	0.2	0.055
	EDA		0.075	0.05	0.139	0.266	0.64	8.183	**0.004**		
	NLRC5		0.300	0.2	0.42	0.611	**0.524**	5.488	**0.019**		
	WNT7B		0.105	0.105	0.188	0.336	0.441	3.698	0.054		
Loch Zone 10	PPARA	20	0.150	0.1	0.255	0.423	**0.608**	7.389	**0.007**	0.16	0.057

(Significance at p<0.05 is denoted in bold)

Table 4: Results of pairwise population Fst analysis using GenAlEx (version 6.5) calculated via frequency option from 236 samples collected from 11 sites of Hosta hybrid zone of North Uist (Highest Fst in bold indicated a considerable degree of differentiation of loch populations from GROG and stream populations).

Populations/ site	GROG	HB Zone 1	HB Zone 2	HB Zone 3	HB Zone 4	HB Zone 5	HB Zone 6	HB Zone 7	HB Zone 8	Loch Zone 9	Loch Zone 10
GROG	0.000										
HB Zone 1	0.262	0.000									
HB Zone 2	0.316	0.019	0.000								
HB Zone 3	0.165	0.027	0.053	0.000							
HB Zone 4	0.173	0.030	0.072	0.008	0.000						
HB Zone 5	0.276	0.030	0.012	0.040	0.056	0.000					
HB Zone 6	0.245	0.004	0.025	0.015	0.017	0.025	0.000				
HB Zone 7	0.336	0.030	0.008	0.078	0.100	0.034	0.045	0.000			
HB Zone 8	0.279	0.013	0.008	0.049	0.064	0.027	0.024	0.007	0.000		
Loch Zone 9	**0.619**	0.144	0.148	0.263	0.247	0.185	0.166	0.145	0.153	0.000	
Loch Zone 10	**0.684**	**0.203**	**0.219**	**0.325**	**0.301**	**0.254**	**0.224**	**0.219**	**0.222**	**0.014**	0.000

*An Fst value greater than 0.15 can be considered as significant in differentiating populations (Frankham et al., 2002).

3. 4 Cline analysis for phenotypic and genotypic traits
3.4.1 Clines for phenotypic traits

All morphological traits exhibited significant differences in clinal pattern, showing significant variation in the length of the centre (χ^2 = 2970.5, df = 6, p <0.001) and width (χ^2 = 569.19, df = 6, p <0.001) across the hybrid zone (Table 5). The stepped cline (representing mosaic pattern) of standard length (r = -0.28, p = 0.06) and total plate count (r = -0.71, p = 0.06) showed non-significant but negative correlations with distance from the sea (Fig. 12 A-B). PCs for armour traits and shape of fish showed a small-scale geographic cline with a decreasing trend towards the loch except PC3 for shape (Fig. 12 C and D). *Gyrodactylus* sp. abundance showed a specific clinal pattern with non-significant but positive correlation (r = 0.61, p = 0.13) in relation to distance towards the loch (Fig. 12 G).
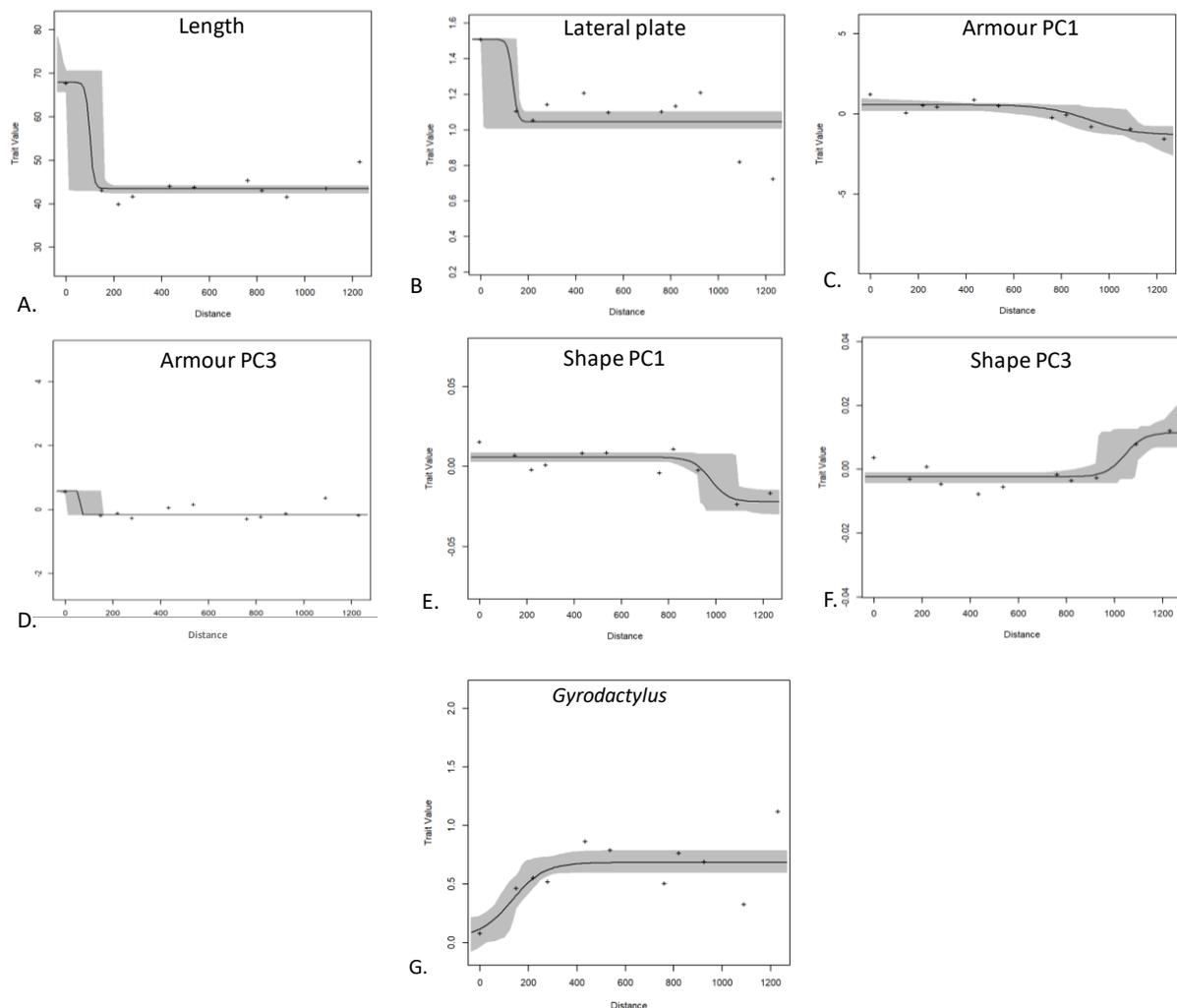
Fig. 12 The maximum likelihood clines fitted for seven morphological traits: A. standard length, B. total number of lateral plates, C. armour PC1, D. armour PC3, E. shape PC1, F. shape PC3 and G. parasite (*Gyrodactylus*) showing variation in clinal patterns from 0 m distance of sea towards the loch up to 1230 m.

Table 5: Summary of cline analysis with cline parameters obtained from likelihood models for seven phenotypic and four loci genotypic traits of sticklebacks collected from Hosta hybrid zone, Scotland. The best fit model with lowest AIC value in bold.

Phenotype											
Morphological traits	**Centre (m)**	**Width (m)**	**muL**	**muR**	**varL**	**varR**	**varH**	**AIC Model I**	**AIC Model II**	**AIC Model III**	**In L**
Length	116.89	19.75	67.51	43.73	29.15	44.95	7484.09	1821.62	**1591.84**	1600.51	-788.67
Lateral plates	109.86	57.30	32.39	6.07	0.06	0.90	166.33	1854.84	1842.60	**1749.62**	-863.22
Armour PC1	1069.20	255.30	-0.36	1.97	2.16	1.40	3.69	921.62	**908.57**	916.93	-447.04
Armour PC3	61.75	3.64	-0.58	0.17	1.36	0.74	13.80	**629.81**	634.34	642.96	-311.86
ShapePC1	967.17	159.04	-0.01	0.02	0.00	0.00	0.00	-1144.66	**-1162.36**	-1153.06	588.42
ShapePC3	1074.15	89.15	0.00	-0.01	0.00	0.00	0.00	-1317.91	**-1392.84**	-1383.36	703.67
Gyrodactylus	131.84	263.70	0.04	0.68	0.04	0.24	0.01	352.30	**308.91**	316.77	-147.21

Genotype										
Locus name	**Centre (m)**	**Width (m)**	**Pmin**	**Pmax**		**AIC Null model**	**AIC Model I**	**AIC Model II**	**AIC Model III**	**In L**
EDA	130.40	32.46	0.25	0.97		61.63	40.55	**22.63**	29.44	-7.23
PPARA	1063.16	85.96	0.15	0.85		77.83	34.61	**16.42**	24.87	-4.12
WNT7B	967.83	310.94	0.08	0.74		76.41	41.13	**27.73**	35.67	-9.78
NLRC5	327.33	1228.37	0.18	0.90		53.65	**34.34**	36.80	43.48	-15.15

3.4.2 Clines for genotypic traits

For the Eda locus, there was a sharp, stepped decrease in the frequency of the allele between marine fish and those in the hybrid zone, but the overall shallow decrease towards the loch was not significant (r = -0.65, p = 0.08, Fig. 13 A). PPARA and WNT7B loci (Fig. 13 B-C) showed a significant negative correlation for the marine P allele (r = -0.75, p = 0.008) and W allele (-0.808, p = 0.002) with the distance from the sea. NLRC5 locus showed a shallower slope with a significant negative correlation (r = -0.66, p = 0.02) between the distance from the sea and the frequency of the marine N allele (Fig. 13 D). All loci had significantly different length of cline centre (χ^2 = 1033, df = 3, p <0.001) and width (χ^2 = 2236.8, df = 3, p <0.001). Hybrid index (HI) data for each fish across all four loci showed an overall gradual decrease in the frequency of marine alleles between the sea and the loch (r = - 0.811, p = 0.002) (Table 3, Fig. 13 E).
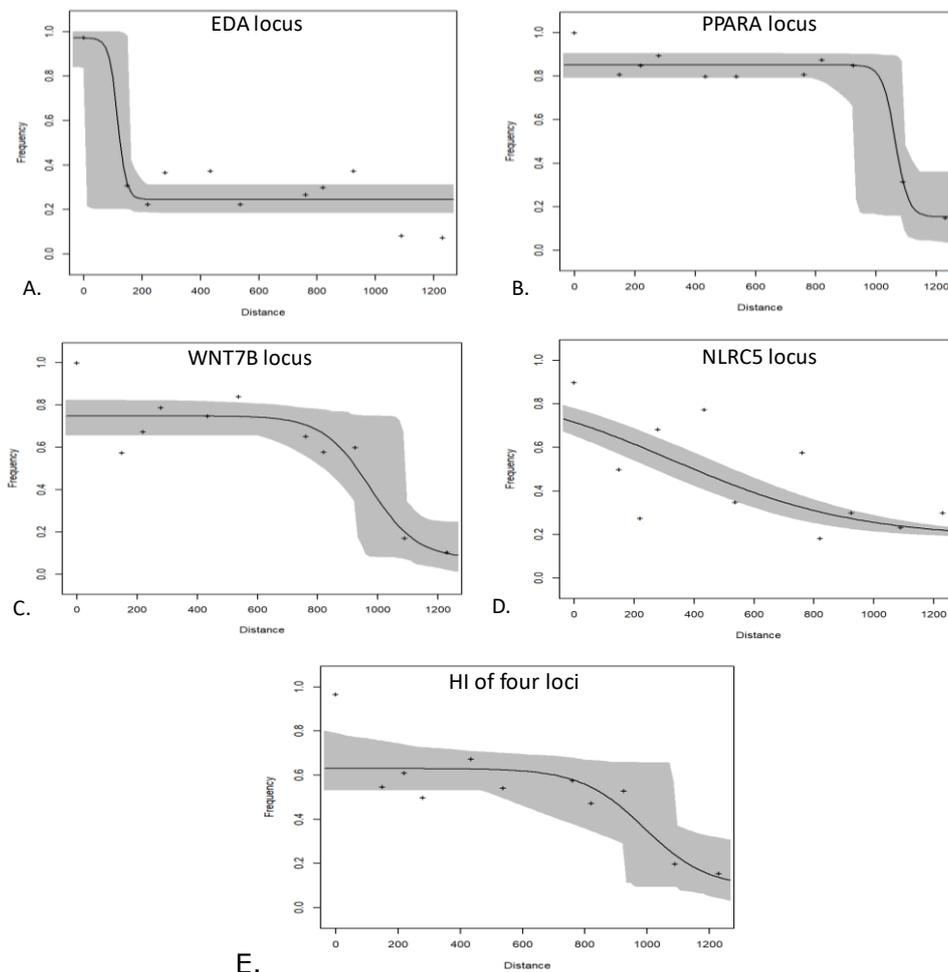
Fig. 13 Clines fitted for the allele frequency of marine genotype of four loci showing variation in pattern from the distance of the sea toward the loch (0 m to 1230 m) through the hybrid zone. A. Frequency of the Eda C allele, B. Frequency of PPARA P allele, C. Frequency of the WNT7B W allele and D. Frequency of the NLRC5 N allele. E. Frequency of hybrid index (HI) of four loci showing a clinal pattern across the hybrid zone. Grey shading indicates 95% credible cline region.

3.5. Phenotype-genotype association across the hybrid zone

For fish sampled in the Hosta stream and loch (without GROG), genotypes of WNT7B gene and location, plate morph, shape PC1, sex of fish were significantly associated with the size of fish (Table 6). Asymptotic tests for coefficient of variation (CV) showed non-significant variation in the pattern of standard length among plate morphs, genotypes of WNT7B and sex of fish ($p>0.05$), however post-hoc comparisons revealed significant differences ($p<0.05$) between heterozygous and homozygous freshwater genotypes of WNT7B (Fig. 14 A-C) including male and female fish. There was a significant association of total plate count among the three plate morph phenotypes (CM, PM, LM) and genotype of Eda and PPARA loci (Table 6). Asymptotic tests revealed that fish carrying the homozygous freshwater genotypes of Eda (D_AD = 22.61, $p< 0.001$) and PPARA (D_AD = 22.33, $p< 0.001$) had a significantly lower mean number of plates than the two other genotypes, consistent with the pattern for plate morph (D_AD = 80.06, $p< 0.001$) (Fig. 15. A-C).
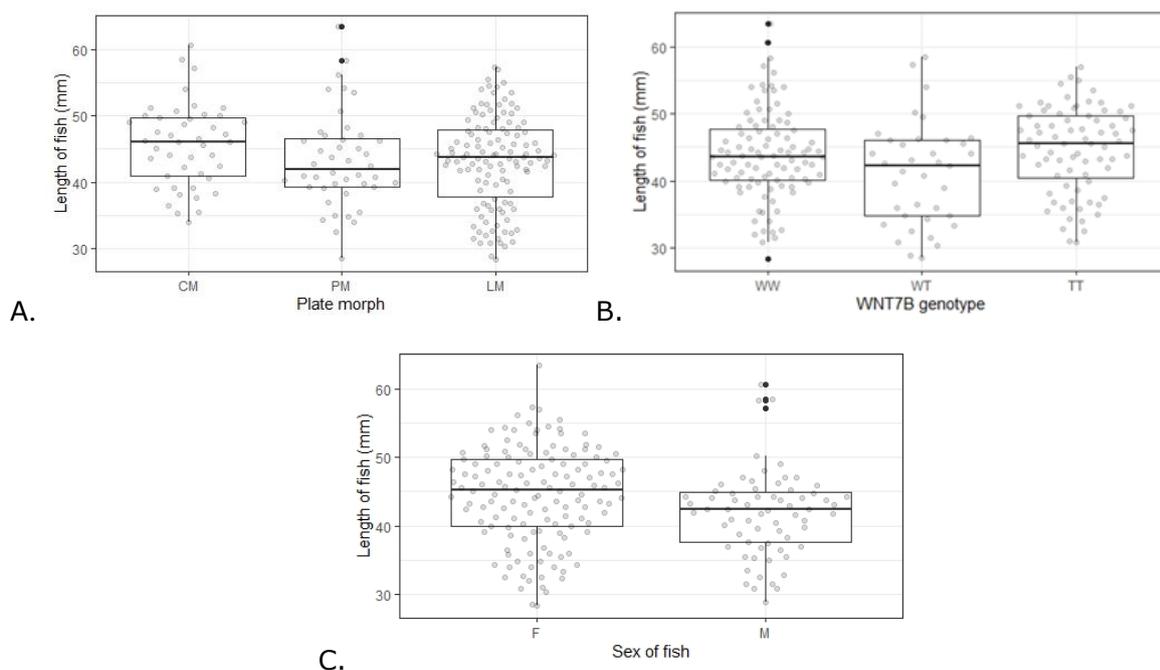
Fig. 14 A. Standard length of fish (mm) showing different patterns among three plate morphs with comparatively larger in size of completely-plated (CM) stickleback than two other morphs (partialy-plated and low-plated morph), (B) Stickleback carrying homogenous freshwater genotype (WW) of WNT7B locus showing the highest length than heterozygous and marine homozygous genotypes, (C) Male was larger in size than female stickleback.
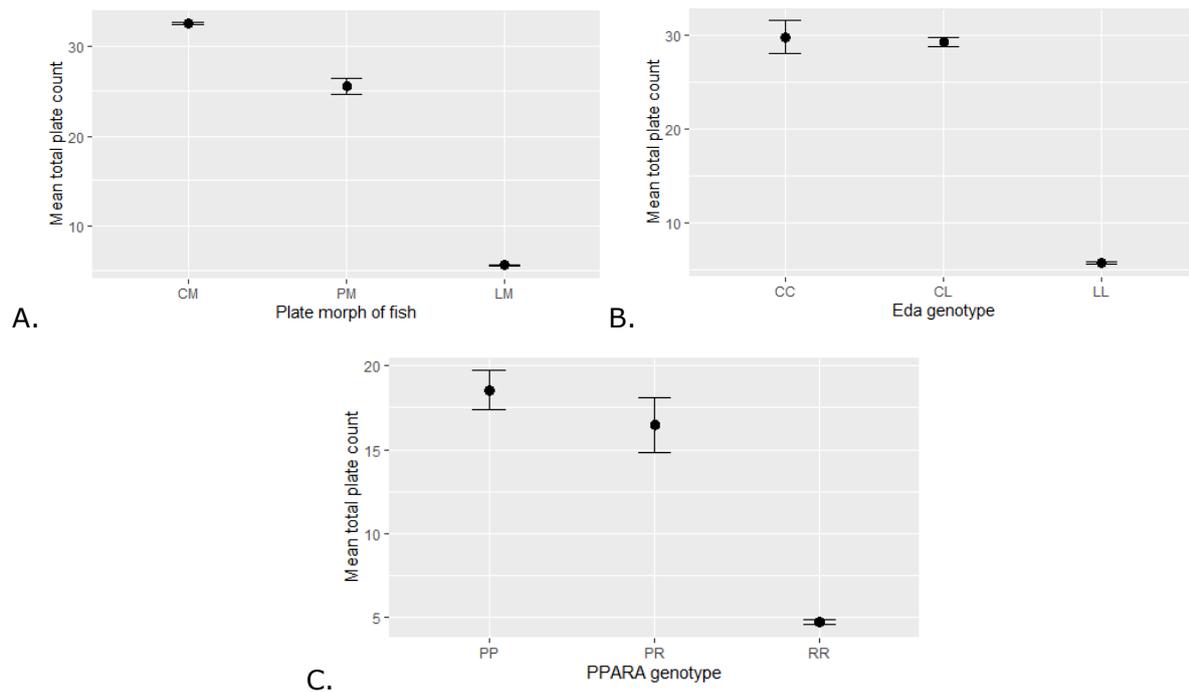
Fig. 15 A. Mean total count of plates (± S.E.) showing similar pattern among three plate morphs (CM, PM, LM) as well as genotypes of B. Eda locus and C. PPARA locus with significantly lowest number of plates in the homozygous freshwater genotype (LL and RR, respectively).

The genotypes of Eda, PPARA and WNT7B and shape of fish showed significant associations with the first and third PCs of armour traits (without plate count) of Hosta hybrid zone fish (Table 6): morphological variation was higher in fish with marine CC genotype of Eda and heterozygous genotypes of PPARA and WNT7B (Fig. 16 A-C). Body shape PC1 and PC3 showed an association with location, length, armour PC1 phenotypes and genotypes of PPARA with more variation in the heterozygote genotype across the hybrid zone (Table 6; Fig. 17). Sex of fish also showed a significant association with armour traits and shape of fish.

Table 6: GLMs of minimum adequate models with significant variables fitted for morphological traits including standard length, total plate count, PCs of armour traits and body shape, parasite to observe phenotype-genotypes association in sticklebacks of the Hosta hybrid zone collected from 2 sites of loch Hosta and 8 sites of stream Hosta, North Uist, Scotland. There were four nuclear genes (Eda, PPARA, WNT7B and NLRC5 loci) for genotypes and one mitochondrial Cyt b gene for the frequency of EU lineage.

Characteristics	Response variable	Predictor variable	Family	n	F	df	p
Size of fish	Standard length (mm)	Location	Gaussian	208	9.25	1, 207	**0.002**
		Sex			6.17	1, 206	**0.013**
		Plate morph			5.00	2, 204	**0.007**
		Shape PC1			10.99	1, 203	**0.001**
		WNT7Bgenotype			3.12	2, 201	**0.045**
Body armour	Total count of lateral plate	Plate morph	Poisson	208	103.60	2, 206	**0.000**
		Eda genotype			43.87	2, 204	**0.000**
		PPARA genotype			14.37	2, 202	**0.000**
Armour girdle and spines	ArmourPC1	Shape PC1	Gaussian	208	36.78	1, 207	**0.000**
		Shape PC3			4.36	1, 206	**0.038**
		WNT7Bgenotype			6.16	2, 204	**0.002**
		Eda genotype			2.61	2, 202	0.07.
	ArmourPC3	Sex	Gaussian	208	11.66	1, 207	**0.000**
		PPARA genotype			4.45	2, 205	**0.012**
Shape of fish	ShapePC1	Location	Gaussian	208	87.23	1, 207	**0.000**
		Sex			192.25	1, 206	**0.000**
		Length			12.46	1, 205	**0.000**
		Armour PC1			15.93	1, 204	**0.000**
		PPARA genotype			3.78	2, 202	**0.001**
	ShapePC3	Sex	Gaussian	208	6.23	1, 207	**0.013**
		Armour PC1			6.24	1, 206	**0.013**
		PPARA genotype			26.25	2, 204	**0.000**

Table 6: (cont.)

Characteristics	Response variable	Predictor variable	Family	n	F	df	p
Parasite abundance	Count of *Gyrodactylus* sp.	Length ShapePC1 Plate morph PPARA genotype	Negative binomial	208	15.59 16.97 3.75 4.22	1, 207 1, 206 1, 205 2, 204	**0.000** **0.000** **0.025** **0.015**
Cytochrome b gene	Count of EU lineage	Location Length	Binomial	57	38.68 10.86	1, 56 1, 55	**0.000** **0.002**

(Significance at p<0.05 is denoted in bold)

A.

B.

A.
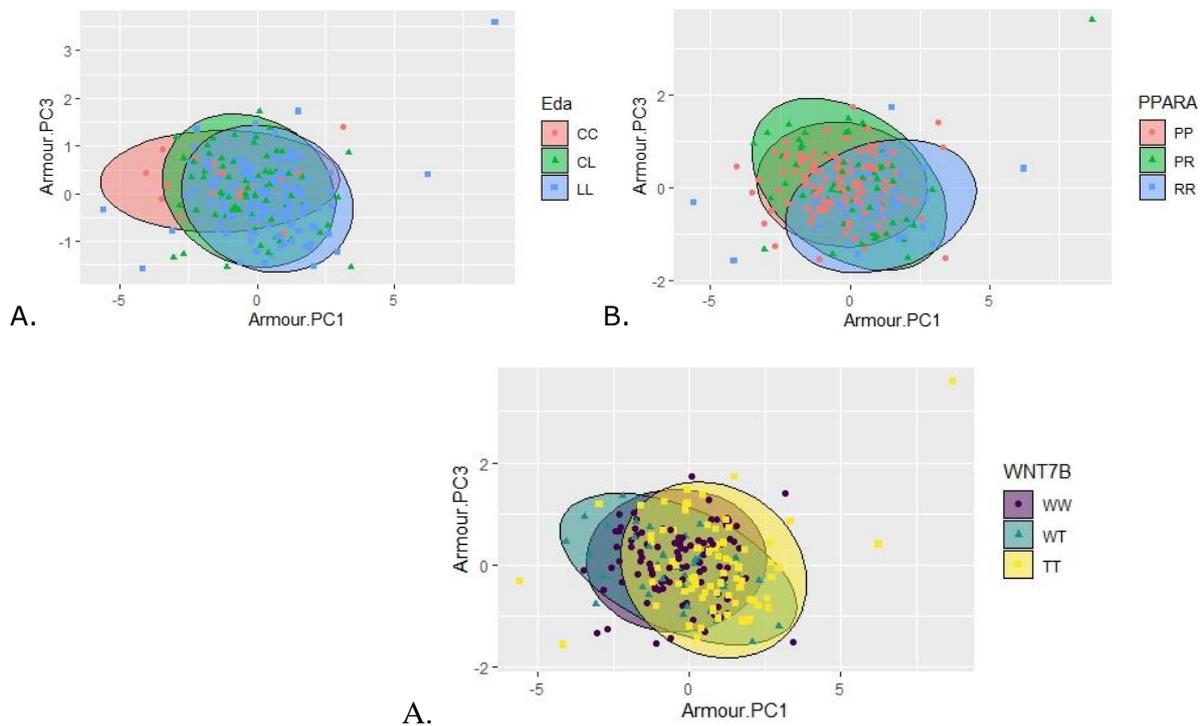
Fig. 16 The first and third PCs of body armour traits (1st dorsal spine length, 2nd dorsal spine length, pelvic spine length, length of pelvis, height of pelvis with their residual value and plate count) of all sticklebacks shows comparatively more significantly different variation in A. homozygous marine genotype (CC) than homozygous freshwater (LL) and heterozygous (CL) of Eda locus B. heterozygous genotype (PR) of PPARA locus and C. WT genotype of WNT7B locus.
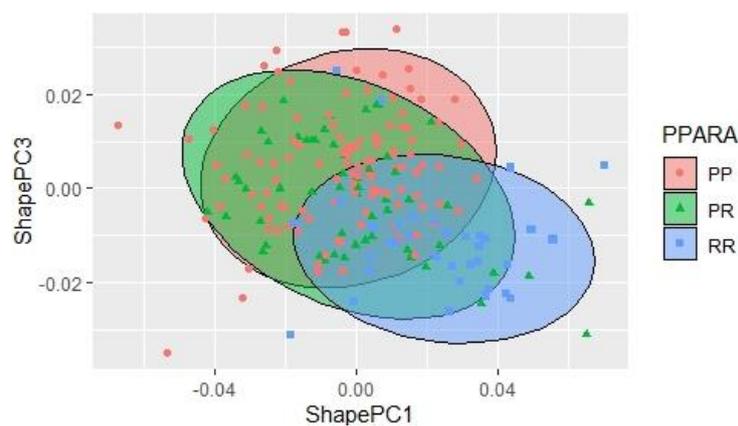
Fig. 17 The first and third PCs of body shape of hybrid zone fish showed comparatively more variation in heterozygous genotype (PR) than homozygous PP and RR genotypes of PPARA locus.

Body length, shape, plate morph phenotypes and PPARA genotype were significantly associated with *Gyrodactylus* abundance across the hybrid zone (Table 6, detail of plate morph and PPARA in chapter 4). The frequency of the EU mitochondrial lineage of Cyt b gene also showed a significant association with location and length of Hosta hybrid zone sticklebacks (Table 6, Fig. 18). The fish carrying At mitochondrial lineage were significantly larger than EU mitochondrial lineage and were absent from sticklebacks within the hybrid zone ($F_{1, 56} = 6.57$, $p = 0.013$).
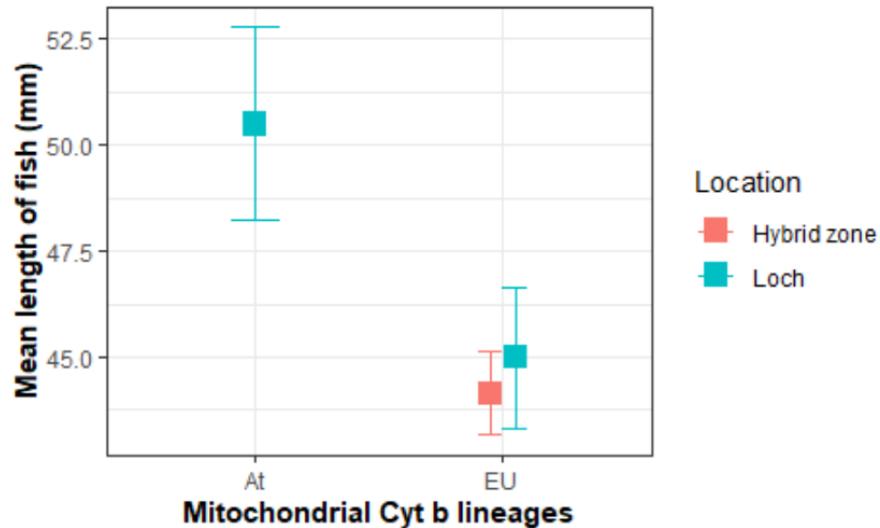
Fig. 18 Mean length (mm) (± S.E.) of sticklebacks carrying trans-Atlantic mitochondrial lineage (At) were bigger than fish carrying EU lineage (EU) across the Hosta hybrid zone and loch. The At lineage was absent from the hybrid zone.

4. Discussion

The study of evolutionary genetics is generally focused on the mechanisms of genetic and molecular change underlying phenotypic diversity, as well as the selective processes responsible for such changes (Romero et al., 2012). This includes investigating the evolution of genome structure, the genetic basis of adaptation and speciation, and genetic changes associated with phenotypic differences in response to selection within populations (Barrett, 2010). In this study, I examined the structure of a putative hybrid zone between a freshwater loch and the sea with the aim of observing phenotypic and genotypic variation among hybrids of ancient marine anadromous and newly-adapted freshwater stickleback populations. Overall, I found substantial variation among morphological (body size, body shape, plate number of body armour, spine structure, parasite abundance etc.) and molecular (polymorphism in Cyt b, Eda, PPARA, WNT7B and NLRC5) traits of these ecotypes across the hybrid zone.

Stickleback morphology across the hybrid zone of Hosta, North Uist varied greatly: this hybrid zone has formed between morphologically-and genetically-distinct populations of stickleback exhibiting no specific pattern for the length of fish which indicated substantial admixture or hybridization among two morph types (Jones et al., 2006). Body shape showed clear differentiation among sea (GROG), hybrid zone and loch fish, especially inside the loch (1090 m and 1230 m), which might be the reason for adaption into freshwater: freshwater stickleback are known to have adapted from their ancestral marine form through divergence of many phenotypic traits including body size (McPhail, 1994; Jones et al., 2012a), shape (Walker and Bell, 2000; Schluter et al., 2004; Aguirre, 2009) and reduced number of lateral plates (Hagen, 1967; Klepaker, 1996). However, studies on the divergent body shape between lake and stream sticklebacks have been shown to have comparatively smaller and deeper body in stream population corresponded to variation in water depth and predators (presence of benthic macroinvertebrates) which also echos in the present study (Lavin and McPhail, 1993; Walker, 1997; Hendry et al., 2002).

Three plate morphs (completely plated fish, hybrids and low plated fish) were identified in Hosta based on their total lateral plate count. This classification mirrored morphological variation in other studies of adaptive divergence in stickleback (Hagen, 1967; Colosimo et al., 2004, 2005; Jones et al., 2006). Measurements of five body armour traits (length of first dorsal spine, second dorsal spine, pelvic spine, pelvis and pelvis height) showed a clinal pattern with comparatively long spines and large pelvic structure in the sea that gradually decreased in size towards the loch (Bell and Ortí, 1994; Marchinko, 2009). This might be a function of change in selection with distance

from a saltwater environment and the gradual change to a freshwater environment. Anadromous fish collected from GROG had bigger spines and pelvic structures supporting several previous studies on stickleback (Bell et al., 1993; Shapiro et al., 2004; Chan, 2010; Vines et al., 2016). However, it has also been suggested that positive selection involving either biotic (e.g. predation) or abiotic (e.g. pH or ion concentration) mechanisms may be involved in the evolution of reduced armour of freshwater stickleback populations (Giles, 1983; Muniz, 1990; Bell et al., 1993; Reimchen, 1994; Schluter, 1995; Barrett, 2010; Smith et al., 2020).

Parasites are thought to be a potent selective agent for the adaptive evolution of animals (Combes, 1996). I found that there was substantial variation in parasite abundance across the hybrid zone and GROG sites. The most common ectoparasite, *Gyrodactylus*, was observed throughout the hybrid zone but was more abundant in loch fish compared to all other sites. The reason might be that infected marine stickleback in the sea (where this parasite tends to be more common), transfer the *Gyrodactylus* into freshwater fish, which have low resistance to this parasite (Mahmud et al., 2017). This result is also supported by previous studies where low plated fish experienced a higher abundance of *Gyrodactylus* than fully plated fish (Robertson et al., 2017). This also indicates that the association between plate phenotype and susceptibility to *Gyrodactylus* infection might be mediated by water salinity.

The structure of the Hosta hybrid zone showing sigmoid clines in some phenotypic traits (body shape, body armour), and mosaic patterns in other traits (length, total plate, parasite abundance) indicates the role of selection in adaptive evolution of stickleback. Different patterns with substantially discordant cline centres (average 504.41 m) and variable widths (average 121.13 m) for seven phenotypic traits (length, total plate count, armour PC1 and PC3, shape PC1 and PC3 and *Gyrodactylus* abundance) were seen. This might be due to the dynamic nature of a hybrid zone for different selection regimes experienced by selected adaptive traits with a diverse range of functions including defence, swimming performance, competition and immune response (Vines et al., 2016; Morales-Rozo et al., 2017). The length of fish and total plate count exhibited steep clines with a narrow width at the centre, indicating their bimodal distribution due to low dispersal and strong disruptive selection (Jiggins and Mallet, 2000; Pedersen et al., 2017). It has been suggested that if there is strong selection against hybrids, it will result in substantial reproductive isolation or strong assortative mating between the taxa (Jiggins and Mallet, 2000). On the other hand, body armour, shape and parasite abundance clines were shallower across the narrow geographical width of the hybrid zone, indicating a unimodal distribution with high dispersal and weak selection due to

excessive introgression (limited reproductive isolation) among the hybrids (Gay et al., 2008).

The Eda gene in stickleback is one of the best studied examples of the genetic control of an adaptive phenotypic change (Colosimo et al., 2005; Jones et al., 2012a). In the present study, variation in allele frequency at the Eda locus closely-mirrored the observed phenotypic data for total plate number across the hybrid zone, echoing previous findings by Jones et al. (2006) and Vines et al. (2016). The presence of a high frequency of the heterozygous Eda^{CL} genotype in the PM fish confirmed the presence of hybrid fish in the hybrid zone which presumably results from hybridization between parental CM and LM fish (Buerkle and Lexer, 2008). The allele frequency of homozygous L was comparatively higher than that of homozygous C across the hybrid zone, but did not show any specific patterns which is quite different from the finding of Vines et al., (2016), for the river Bonsall Creek, a Canada hybrid zone, but similar to the findings of Pedersen et al., (2017) for the Odder river in Denmark.

For the mitochondrial Cyt b gene, stickleback in the Hosta stream showed an interesting pattern of At and EU lineages: the At lineage was absent from among hybrid zone fish but present in the sea and inside the loch. To confirm this finding, another fifty individuals were collected from the stream in the following year (2019) and the same pattern was observed among the hybrid zone fish. It is highly unlikely that this pattern is random, but there are a number of possible explanations. The pattern could arise as a straightforward result of differences in behaviour, if fish of the At lineage somehow attempt to avoid the stream environment by moving through it as rapidly as possible. Alternatively, the pattern may be evolutionary, arising from strong selection against the At haplotype. Experiments on movement pattern of pure anadromous stickleback across the hybrid zone did not show any specific direction for the dispersal of two lineages, suggesting that the first of the two explanations are not correct. This prominent genetic differentiation therefore suggests the effect of selection on the mitochondrial divergence and adaptive colonization to new ecological niches (Orti, et al., 1994; Mäkinen and Merilä, 2008). There are several possibilities for this unique pattern for At lineage related to physical, abiotic or biotic properties of the stream: 1) differences in the oxygen saturation levels between loch/sea and the stream due to continuous flow of water (ranged from 0.07 to 0.35 m/S), 2) the higher temperature of the stream water due to shallow in depth (max. approximately 1 m), 3) specific physiological or genetic aspects of marine stickleback or 4) differences in ion or nutrient concentrations between stream and loch/sea. Regarding the last possibility, it may be that the At haplotype is associated with differences in nutrient requirements, either directly because of differences in the operation of the mitochondria, or indirectly because of linkage. The latter is suggested

116

by a recent study on Japan Sea and Pacific Ocean lineages of stickleback which revealed that deficiency in docosahexaenoic acid (DHA) controlled by a key metabolic gene Fatty acid desaturase 2 (Fads2), can constrain freshwater colonization and the radiation of marine fish with lower physiological ability (Ishikawa et al., 2019).

To observe the genetic variation across the hybrid zone, candidate genes with fixed differences between freshwater and marine stickleback populations, and that are likely to be linked to adaptive traits, were selected. Development of PCR-RFLP-based SNP assays for a gene with specific SNPs differences was challenging, given the large gene size, large number of SNPs, specific restriction enzymes and time constraint. Initially, I tried with several previously identified SNPs with reference genes (ATP1A1, CSMD1, LEMD3, EPX, MFAP1, PRL2) but found no differences in the amplified target regions. Finally, three SNP assays were developed for three candidate genes: PPARA, WNT7B and NLRC5, all putatively related to salinity and immunity of stickleback. From previous genome scan studies of stickleback, it is evident that PPARA and WNT7B genes are present on Chromosome IV and show similar marine-freshwater divergence patterns to Eda (Jones et al., 2012a). The successful genotyping of these three genes in the natural hybrid zone confirmed their fixed differences between marine and freshwater with heterozygous genotype for polymorphism in that loci. In this study, the population genetic clusters did not test by using STRUCTURE (Falush et al., 2003) due to utilization of only adaptive markers (SNPs) in the selected candidate genes.

Genetic diversity of a hybrid zone can be interpreted in several ways, mainly based on the allele frequency of molecular markers and their behaviour in the mode of selection (Barton and Hewitt, 1985; Mallet et al., 1990; Barton and Gale, 1993; Bert and Arnold, 1995). In a hybrid zone study, it is fundamental to evaluate the mode of selection on a trait or locus by observing deficiencies of hybrids or fitness of hybrids across the zone by measuring the deviation from HWE, linkage disequilibrium and genetic differentiation (Fst) among diagnostic markers (Barton and Hewitt, 1985; Bert and Arnold, 1995; McKenzie et al., 2015). In general, observed heterozygosity was significantly lower than expected for all populations at four loci indicating an overall heterozygote deficit across the hybrid zone. WNT7B and NLRC5 loci showed significant departure from HWE in the downstream sticklebacks of the hybrid zone (site 1, 2, 4, 5) but all in loch fish which indicated strong selection in these populations. Similarly, positive Fst value for all the populations for WNT7B and NLRC5 loci (except site 6) and loch fish for all loci indicated heterozygote deficiencies due to strong selection on that loci. The average Fst value for all loci was 0.261 suggesting a medium level of inbreeding among populations. However, the overall genetic differentiation between pairwise populations was highest between loch and GROG populations, which echoes the adaptive divergence of marine and

freshwater stickleback. In addition, strong LD was also observed among Eda-PPARA and PPARA-WNT7B loci, which had significant disequilibrium in several populations across the hybrid zone (such as site 3, 4, 7) and loch zone (site 9 and 10) indicating an indirect selection leading to possible prezygotic reproductive isolation among stream and loch stickleback. Therefore, the deviation from HWE, significant LD among loci, steps clines in phenotypic and genotypic traits and strong genetic differentiation in Fst value among populations suggests that selection influences this hybrid zone differentially with strong endogenous selection (independently of the environment) for the population of downstream and loch fish as well as exogenous selection (in response to environment) in some populations in the middle of hybrid zone depending on the adaptive loci.

The pattern of genotypic clines of this hybrid zone was also assessed to understand the extent of reproductive isolation occurring within the zone (McKenzie et al., 2015). The clines fitted for Eda, PPARA and WNT7B loci exhibited stepped clines of marine allele frequencies along the geographical range from sea to loch, corresponding to a bimodal distribution similar to that for plate morph and length of fish. The centre of these clines was discordant with a narrow width indicative of strong direct selection on these loci and the existence of reproductive isolation (Barton and Hewitt, 1985; Szymura and Barton, 1991; Gay et al., 2008; Teeter et al., 2008). The bimodal distribution of the Eda locus and plate morph has been observed in several stickleback populations corresponding to strong differentiation between marine and freshwater parental populations (Jones, et al., 2006; Taugbøl et al., 2014; Ravinet et al., 2015; Vines et al., 2016; Pedersen et al., 2017). A bimodal distribution in a locus also represents the existence of a heterozygote deficit and LD in some populations due to prezygotic isolation or local adaptation of either parental population to divergent habitats (Harrison and Bogdanowicz, 1997; McKenzie et al., 2015).

The NLRC5 locus displayed a shallow cline with a wide width, representing a unimodal distribution which may be due to weak selection and excessive introgression between the hybridizing populations (Mallet et al., 1990; Barton and Gale, 1993; Gay et al., 2008). Nevertheless, the NLRC5 gene showed fixed differences between freshwater and marine fish, which suggests it plays a role in adaptation to divergent environments, but it is also known to be responsible for anti-viral responses in humans (Neerincx et al., 2010) and this may also be true for stickleback (Terekhanova et al., 2014). However, the combined hybrid index of the four loci displayed a sigmoid clinal pattern with a gradual decrease of marine allele frequencies towards the loch, suggesting that weak environmental selection may be occurring on such loci for divergence with limited reproductive isolation (Barton and Hewitt, 1985; Pedersen et al., 2017). Therefore, PPARA, WNT7B and NLRC5 loci can be considered as important candidate genes connected to salinity: this is an

important factor for selection in nature, as well as influencing the structure of the hybrid zone based on geographical dispersal range.

A hybrid zone also serves as a natural laboratory for investigating phenotype-genotype association and detection of genomic regions underlying divergent traits under selection (Nadeau et al., 2014; Ravinet et al., 2015). In the Hosta hybrid zone, most of the phenotypic traits including body size, shape, armour and parasite showed a significant relationship with selected loci (Eda, PPARA, WNT7B and NLRC5) indicating the importance of these putative genomic regions for divergent adaptation (Westram et al., 2018). The similar cline centres and width of lateral plates and Eda reconfirmed their strong association in the divergent evolution of plate morph under selection between freshwater and saltwater sticklebacks (Colosimo et al., 2004, 2005; Jones et al., 2012a). The association of WNT7B locus with fish size might be the function of the Wnt signalling pathway for the development of the fish body, similar to the development of the swim bladder of zebrafish (Yin et al., 2011) and lungs of mammals (Rajagopal et al., 2008). Body armour and shape of fish also showed strong associations with WNT7B and PPARA loci, which were supported by their coinciding cline centres at the upstream point (around 1028 m) of the Hosta stream. The PPARA locus also had a significant impact on other phenotypic traits including plate count and *Gyrodactylus* parasite abundance. This might be due to its physiological homeostasis roles in lipids, fatty acids and glucose metabolism as a member of the peroxisome proliferator-activated receptor (PPAR) family of nuclear receptors observed in fish and mammals (Michalik and Wahli., 1999; Leaver et al., 2005). However, the mitochondrial Cyt b gene was significantly associated with the length of fish based on the dispersal of the EU lineage found in the populations of hybrid zone and loch. The stickleback carrying At mitochondrial lineage were larger in size than EU lineage fish, which clearly indicates the effect of origin in the adaptation of stickleback (Mäkinen and Merilä, 2008; Ishikawa et al., 2019).

In conclusion, the characterization of the stickleback Hosta hybrid zone represented an interesting area of eco-evolutionary research considering its small size (~1km) but diverse phenotypic and genotypic variation among hybrid populations compared to other studied hybrid zone (Jones et al., 2006; Vines et al. 2016; Pedersen et al., 2017). These patterns have been seemed to be stable across the Hosta stream since 2007 with almost similar morphological variation in stickleback samples collected in the different time of the year (September, 2007) over distances of 100-200 m (Gostick, 2008). The observed clinal patterns of the phenotypes and genotypes across the hybrid zone didn't completely fulfil the criteria of hybrid zone theory (gradual pattern) which is probably due to large-scale dispersal of fish in terms of hybrid zone size. Besides, there was no ecotone

variation or salinity gradation across the Hosta stream because tidal flow does not enter the stream directly due to wide shallow shore at the bottom. Nevertheless, gene flow among populations of this zone still existed but the overall strong selection against hybrids suggested reproductive isolation to some extent enabling distinct population structure of the stream in the transition between freshwater stickleback of the loch and anadromous stickleback of the sea. It has been observed that there were two founder populations of marine and freshwater origin and a third hybrid population which was less connected to the parental populations across the stream. This suggests that this may not a true hybrid zone but a contact zone between two divergent lineages similar to Burrishoole catchment in western Ireland (Ravinet et al., 2015) or may be an on-going hybrid zone with the formation of a stable hybrid population for many generations (Gostick, 2008). However, the admixed population of this zone helps to identify the influence of point mutations in a specific gene associated with divergence phenotypic traits that contributes to adaptation in stickleback. Further research using large-scale molecular scanning will confirm the dynamics of this hybrid zone to further understand the process of adaptive divergence and speciation in stickleback.

References

Aguirre, W.E., 2009. Microgeographical diversification of three spine sticklebacks: body shape–habitat correlations in a small, ecologically diverse Alaskan drainage. *Biological Journal of the Linnean Society*, *98*(1), pp.139-151.

Alexandrino, J., Baird, S.J., Lawson, L., Macey, J.R., Moritz, C. and Wake, D.B., 2005. Strong selection against hybrids at a hybrid zone in the Ensatina ring species complex and its evolutinary implications. *Evolution*, *59*(6), pp.1334-1347.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, *3*(10), p.e3376.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2008. Natural selection on a major armor gene in threespine stickleback. *Science*, *322*(5899), pp.255-257.

Barrett, R.D.H., 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: A case study in functional analysis of natural variation. *Journal of Fish Biology*, *77*(2), pp.311-328.

Barton, N.H., 1983. Multilocus clines. *Evolution*, pp.454-471.

Barton, N.H. and Gale, K.S., 1993. Genetic analysis of hybrid zones. *Hybrid zones and the evolutionary process*, pp.13-45.

Barton, N.H. and Hewitt, G.M., 1985. Analysis of hybrid zones. *Annual review of Ecology and Systematics*, *16*(1), pp.113-148.

Bell, M. A. and Foster, S.A. 1994. *The evolutionary biology of the threespine stickleback* (pp. 381-398). Oxford: Oxford University Press.

Bell, M.A. and Ortí, G., 1994. Pelvic reduction in three spine sticklebacks from Cook Inlet lakes: geographical distribution and intra-population variation. *Copeia*, pp.314-325.

Bell, M.A., Orti, G., Walker, J.A. and Koenings, J.P., 1993. Evolution of pelvic reduction in three spine stickleback fish: a test of competing hypotheses. *Evolution*, *47*(3), pp.906-914.

Bert, T.M. and Arnold, W.S., 1995. An empirical test of predictions of two competing models for the maintenance and fate of hybrid zones: both models are supported in a hard-clam hybrid zone. *Evolution*, *49*(2), pp.276-289.

Bird, K.A., An, H., Gazave, E., Gore, M.A., Pires, J.C., Robertson, L.D. and Labate, J.A., 2017. Population structure and phylogenetic relationships in a diverse panel of Brassica rapa L. *Frontiers in plant science*, *8*, p.321.

Bresadola, L., Caseys, C., Castiglione, S., Buerkle, C.A., Wegmann, D. and Lexer, C., 2019. Admixture mapping in interspecific Populus hybrids identifies classes of genomic architectures for phytochemical, morphological and growth traits. *New Phytologist*, *223*(4), pp.2076-2089.

Buerkle, C.A. and Lexer, C., 2008. Admixture as the basis for genetic mapping. *Trends in ecology & evolution*, *23*(12), pp.686-694.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J. and Myers, R.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *science*, *327*(5963), pp.302-305.

Cleves, P.A., Ellis, N.A., Jimenez, M.T., Nunez, S.M., Schluter, D., Kingsley, D.M. and Miller, C.T., 2014. Evolved tooth gain in sticklebacks is associated with a cis-regulatory allele of Bmp6. *Proceedings of the National Academy of Sciences*, *111*(38), pp.13912-13917.

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.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. and Kingsley, D.M., 2004. The genetic architecture of parallel armor plate reduction in three spine sticklebacks. *PLoS biology*, *2*(5), p.e109.

Combes, C., 1996. Parasites, biodiversity and ecosystem stability. *Biodiversity & Conservation*, *5*(8), pp.953-962.

Dean, L.L., Magalhaes, I.S., Foote, A., D'Agostino, D., McGowan, S. and MacColl, A.D., 2019. Admixture between ancient lineages, selection, and the formation of sympatric stickleback species-pairs. *Molecular biology and evolution*, *36*(11), pp.2481-2497.

De La Torre, A., Ingvarsson, P.K. and Aitken, S.N., 2015. Genetic architecture and genomic patterns of gene flow between hybridizing species of Picea. *Heredity*, *115*(2), pp.153-164.

Derryberry, E.P., Derryberry, G.E., Maley, J.M. and Brumfield, R.T., 2014. HZAR: hybrid zone analysis using an R software package. *Molecular ecology resources*, *14*(3), pp.652-663.

El Nagar, A. and MacColl, A.D., 2016, August. Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. In *Proc. R. Soc. B*. 283(1836), p. 20160691. The Royal Society.

Ellis, N.A., Glazer, A.M., Donde, N.N., Cleves, P.A., Agoglia, R.M. and Miller, C.T., 2015. Distinct developmental genetic mechanisms underlie convergently evolved tooth gain in sticklebacks. *Development*, *142*(14), pp.2442-2451.

Endler, J.A., 1977. *Geographic variation, speciation, and clines* (No. 10). Princeton University Press.

Falush, D., Stephens, M. and Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*(4), pp.1567-1587.

Feltz, C.J. and Miller, G.E., 1996. An asymptotic test for the equality of coefficients of variation from k populations. *Statistics in medicine*, *15*(6), pp.647-658.

Ferchaud, A.L., Pedersen, S.H., Bekkevold, D., Jian, J., Niu, Y. and Hansen, M.M., 2014. A low-density SNP array for analyzing differential selection in freshwater and marine populations of threespine stickleback (Gasterosteus aculeatus). *BMC genomics*, *15*(1), pp.1-11.

Frankham, R., Ballou, S.E.J.D., Briscoe, D.A. and Ballou, J.D., 2002. *Introduction to conservation genetics*. Cambridge university press.

Gay, L., Crochet, P.A., Bell, D.A. and Lenormand, T., 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution: International Journal of Organic Evolution*, *62*(11), pp.2789-2806.

Giles, N., 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *Journal of zoology*, *199*(4), pp.535-544.

Goldenberger, D., Perschil, I., Ritzler, M. and Altwegg, M. (1995). A simple universal DNA extraction procedure using sds and proteinase-k is compatible with direct pcr amplification. *Pcr-Methods and Applications* **4**(6): 368-370.

Gompert, Z., Mandeville, E.G. and Buerkle, C.A., 2017. Analysis of population genomic data from hybrid zones. *Annual Review of Ecology, Evolution, and Systematics*, *48*.

Gostick D., 2008. *MS thesis*: Morphological variation and selection within a putative hybrid zone of the three-spined stickleback *Gasterosteus aculeatus*, University of Nottingam, UK.

Greenwood, A.K., Jones, F.C., Chan, Y.F., Brady, S.D., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Kingsley, D.M. and Peichel, C.L., 2011. The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. *Heredity*, *107*(2), p.155.

Greenwood, A.K., Mills, M.G., Wark, A.R., Archambeault, S.L. and Peichel, C.L., 2016. Evolution of schooling behavior in threespine sticklebacks is shaped by the Eda gene. *Genetics*, *203*(2), pp.677-681.

Greenwood, A.K., Wark, A.R., Yoshida, K. and Peichel, C.L., 2013. Genetic and neural modularity underlie the evolution of schooling behavior in threespine sticklebacks. *Current Biology*, *23*(19), pp.1884-1888.

Guiller, A., Coutellec-Vreto, M.A. and Madec, L., 1996. Genetic relationships among suspected contact zone populations of Helix aspersa (Gastropoda: Pulmonata) in Algeria. *Heredity*, *77*(2), pp.113-129.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D. and Berner, D., 2019. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution letters*, *3*(1), pp.28-42.

Hagen, D. W. 1967. Isolating mechanisms in threespine sticklebacks (Gasterosteus). J. Fish. Res. Board Can. 24:1637–1692.

Hall, T.A., 1999, January. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* (Vol. 41, No. 41, pp. 95-98). [London]: Information Retrieval Ltd., c1979-c2000.

Harrison, R.G. and Bogdanowicz, S.M., 1997. Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution*, *51*(2), pp.493-505.

Harrison, R.G., 1990. Hybrid zones: windows on evolutionary process. *Oxford surveys in evolutionary biology*, *7*, pp.69-128.

Hendry, A.P., Peichel, C.L., Matthews, B., Boughman, J.W. and Nosil, P., 2013. Stickleback research: the now and the next. *Evolutionary Ecology Research*, *15*(2), pp.111-141.

Hendry, A.P., Taylor, E.B. and McPhail, J.D., 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution*, *56*(6), pp.1199-1216.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genet*, *6*(2), p.e1000862.

Hollander, J., Galindo, J. and Butlin, R.K., 2015. Selection on outlier loci and their association with adaptive phenotypes in Littorina saxatilis contact zones. *Journal of Evolutionary Biology*, *28*(2), pp.328-337.

Howes, T.R., Summers, B.R. and Kingsley, D.M., 2017. Dorsal spine evolution in threespine sticklebacks via a splicing change in MSX2A. *BMC biology*, *15*(1), p.115.

Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T. and Durbin, R., 2002. The Ensembl genome database project. *Nucleic acids research*, *30*(1), pp.38-41.

Ishikawa, A., Kabeya, N., Ikeya, K., Kakioka, R., Cech, J.N., Osada, N., Leal, M.C., Inoue, J., Kume, M., Toyoda, A. and Tezuka, A., 2019. A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science*, *364*(6443), pp.886-889.

Jiggins, C.D. and Mallet, J., 2000. Bimodal hybrid zones and speciation. *Trends in ecology & evolution*, *15*(6), pp.250-255.

Johannesson, K., Panova, M., Kemppainen, P., André, C., Rolan-Alvarez, E. and Butlin, R.K., 2010. Repeated evolution of reproductive isolation in a marine snail: unveiling mechanisms of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1547), pp.1735-1747.

Jones, F.C., Brown, C., Pemberton, J.M. and Braithwaite, V.A., 2006. Reproductive isolation in a three-spine stickleback hybrid zone. *Journal of evolutionary biology*, *19*(5), pp.1531-1544.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M., Absher, D.M., Myers, R.M., Reimchen, T.E., Deagle, B.E. and Schluter, D., 2012b. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current biology*, *22*(1), pp.83-90.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012a. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), p.55.

Kawakami, T. and Butlin, R.K., 2001. Hybrid zones*In*: eLS. John Wiley, New York. http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.

Kingsley, D.M., Peichel, C.L., Ostlund-Nilsson, S. and Mayer, I., 2007. The molecular genetics of evolutionary change in sticklebacks. *Biology of the Three-spined Stickleback*, pp.41-81.

Kingsley, D.M., Zhu, B., Osoegawa, K., De Jong, P.J., Schein, J., Marra, M., Peichel, C., Amemiya, C., Schluter, D., Balabhadra, S. and Friedlander, B., 2004. New genomic tools for molecular studies of evolutionary change in threespine sticklebacks. *Behaviour*, pp.1331-1344.

Kitano, J., Bolnick, D.I., Beauchamp, D.A., Mazur, M.M., Mori, S., Nakano, T. and Peichel, C.L., 2008. Reverse evolution of armor plates in the threespine stickleback. *Current Biology*, *18*(10), pp.769-774.

Kitano, J., Lema, S.C., Luckenbach, J.A., Mori, S., Kawagishi, Y., Kusakabe, M., Swanson, P. and Peichel, C.L., 2010. Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. *Current Biology*, *20*(23), pp.2124-2130.

Klepaker, T., 1996. Lateral plate polymorphism in marine and estuarine populations of the three-spine stickleback (Gasterosteus aculeatus) along the coast of Norway. *Copeia*, pp.832-838.

Klingenberg, C.P., 2011. MorphoJ: an integrated software package for geometric morphometrics. *Molecular ecology resources*, *11*(2), pp.353-357.

Komar, A.A., 2009. Single nucleotide polymorphisms. *Methods in Molecular Biology*, *578*.

Kruuk, L.E.B., Baird, S.J.E., Gale, K.S. and Barton, N.H., 1999. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics*, *153*(4), pp.1959-1971.

Lavin, P.A. and McPhail, J.D., 1993. Parapatric lake and stream sticklebacks on northern Vancouver Island: disjunct distribution or parallel evolution?. *Canadian Journal of Zoology*, *71*(1), pp.11-17.

Leaver, M.J., Boukouvala, E., Antonopoulou, E., Diez, A., Favre-Krey, L., Ezaz, M.T., Bautista, J.M., Tocher, D.R. and Krey, G., 2005. Three peroxisome proliferator-activated receptor isotypes from each of two species of marine fish. *Endocrinology*, *146*(7), pp.3150-3162.

Lenz, T.L., Eizaguirre, C., Rotter, B., Kalbe, M. and Milinski, M., 2013. Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Molecular ecology*, *22*(3), pp.774-786.

Lindtke, D., González-Martínez, S.C., Macaya-Sanz, D. and Lexer, C., 2013. Admixture mapping of quantitative traits in Populus hybrid zones: power and limitations. *Heredity*, *111*(6), pp.474-485.

Magalhaes, I.S., D'Agostino, D., Hohenlohe, P.A. and MacColl, A.D., 2016. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular ecology*, *25*(17), pp.4319-4336.

Mahmud, M.A., Bradley, J.E. and MacColl, A.D., 2017. Abiotic environmental variation drives virulence evolution in a fish host-parasite geographic mosaic. *Functional Ecology*.pp.1-7.

Mallet, J. and Barton, N.H., 1989. Strong natural selection in a warning-color hybrid zone. *Evolution*, *43*(2), pp.421-431.

Mallet, J., Barton, N., Lamas, G., Santisteban, J., Muedas, M. and Eeley, H., 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in Heliconius hybrid zones. *Genetics*, *124*(4), pp.921-936.

Mäkinen, H.S. and Merilä, J., 2008. Mitochondrial DNA phylogeography of the three-spined stickleback (Gasterosteus aculeatus) in Europe—evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, *46*(1), pp.167-182.

Marchinko, K.B., 2009. Predation's role in repeated phenotypic and genetic divergence of armor in three spine stickleback. *Evolution: International Journal of Organic Evolution*, *63*(1), pp.127-138.

McKenzie, J.L., Dhillon, R.S. and Schulte, P.M., 2015. Evidence for a bimodal distribution of hybrid indices in a hybrid zone with high admixture. *Royal Society open science*, *2*(12), p.150285.

McKinnon, J.S. and Rundle, H.D., 2002. Speciation in nature: the threespine stickleback model systems. *Trends in ecology & evolution*, *17*(10), pp.480-488.

McPhail, J.D., 1994. Speciation and the evolution of reproductive isolation in the sticklebacks Gasterosteus) of southwestern British Columbia. *In*: M.A. Bell, S.A. Foster (Eds.), *The evolutionary biology of the threespine stickleback*, Oxford University Press (1994), pp. 399-437.

Michalik, L. and Wahli, W., 1999. Peroxisome proliferator-activated receptors: three isotypes for a multitude of functions. *Current opinion in biotechnology*, *10*(6), pp.564-570.

Miller, C.T., Beleza, S., Pollen, A.A., Schluter, D., Kittles, R.A., Shriver, M.D. and Kingsley, D.M., 2007a. cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*, *131*(6), pp.1179-1189.

Miller, C.T., Glazer, A.M., Summers, B.R., Blackman, B.K., Norman, A.R., Shapiro, M.D., Cole, B.L., Peichel, C.L., Schluter, D. and Kingsley, D.M., 2014. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics*, *197*(1), pp.405-420.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A. and Johnson, E.A., 2007b. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research*, *17*(2), pp.240-248.

Mills, M.G., Greenwood, A.K. and Peichel, C.L., 2014. Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo*, *5*(1), p.5.

Morales-Rozo, A., Tenorio, E.A., Carling, M.D. and Cadena, C.D., 2017. Origin and cross-century dynamics of an avian hybrid zone. *BMC evolutionary biology*, *17*(1), pp.1-18.

Muniz, I.P., 1990. Freshwater acidification: its effects on species and communities of freshwater microbes, plants and animals. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, *97*, pp.227-254.

Nadeau, N.J., Ruiz, M., Salazar, P., Counterman, B., Medina, J.A., Ortiz-Zuazaga, H., Morrison, A., McMillan, W.O., Jiggins, C.D. and Papa, R., 2014. Population genomics of parallel hybrid zones in the mimetic butterflies, H. melpomene and H. erato. *Genome research*, *24*(8), pp.1316-1333.

Neerincx, A., Lautz, K., Menning, M., Kremmer, E., Zigrino, P., Hösel, M., Büning, H., Schwarzenbacher, R. and Kufer, T.A., 2010. A role for the human nucleotide-binding domain, leucine-rich repeat-containing family member NLRC5 in antiviral responses. *Journal of Biological Chemistry*, *285*(34), pp.26223-26232.

Nikinmaa, M., McCairns, R.J.S., Nikinmaa, M.W., Vuori, K.A., Kanerva, M., Leinonen, T., Primmer, C.R., Merilä, J. and Leder, E.H., 2013. Transcription and redox enzyme activities: comparison of equilibrium and disequilibrium levels in the three-spined stickleback. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1755), p.20122974.

Orti, G., Bell, M.A., Reimchen, T.E. and Meyer, A., 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution*, *48*(3), pp.608-622.

Pallares, L.F., Harr, B., Turner, L.M. and Tautz, D., 2014. Use of a natural hybrid zone for genomewide association mapping of craniofacial traits in the house mouse. *Molecular Ecology*, *23*(23), pp.5756-5770.

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

Pedersen, S.H., Ferchaud, A.L., Bertelsen, M.S., Bekkevold, D. and Hansen, M.M., 2017. Low genetic and phenotypic divergence in a contact zone between freshwater and marine sticklebacks: gene flow constrains adaptation. *BMC evolutionary biology*, *17*(1), pp.1-16.

Peichel, C.L. and Marques, D.A., 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1713), p.20150486.

Peichel, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F., Buerkle, C.A., Schluter, D. and Kingsley, D.M., 2001. The genetic architecture of divergence between threespine stickleback species. *Nature*, *414*(6866), p.901.

R core Team, 2016. Vienna: R Foundation for Statistical Computing, 2016. *R: A language and environment for statistical computing*.

Rajagopal, Jayaraj, Thomas J. Carroll, J. Sawalla Guseh, Sam A. Bores, Leah J. Blank, William J. Anderson, Jing Yu, Qiao Zhou, Andrew P. McMahon, and Douglas A. Melton. "Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme." *Development* 135, no. 9 (2008): 1625-1634.

Ravinet, M., Hynes, R., Poole, R., Cross, T.F., McGinnity, P., Harrod, C. and Prodöhl, P.A., 2015. Where the lake meets the sea: strong reproductive isolation is associated with adaptive divergence between lake resident and anadromous three-spined sticklebacks. *PloS one*, *10*(4), p.e0122825.

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

Reimchen, T.E., 1994. Predators and morphological evolution in threespine stickleback. *The evolutionary biology of the threespine stickleback*, pp.240-276.

Rieseberg, L.H. and Buerkle, C.A., 2002. Genetic mapping in hybrid zones. *The American Naturalist*, *159*(S3), pp.S36-S50.

Riginos, C. and Cunningham, C.W., 2005. Invited review: local adaptation and species segregation in two mussel (Mytilus edulis× Mytilus trossulus) hybrid zones. *Molecular ecology*, *14*(2), pp.381-400.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific reports*, *7*, p.42677.

Rohlf, F.J., 2010. TpsDig, version 2.16, Software, Department of Ecology and Evolution, State. *University New York, Stony Brook. Available in http://life. bio. sunysb. edu/morph*.

Romero, I.G., Ruvinsky, I. and Gilad, Y., 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics*, *13*(7), pp.505-516.

Sage, R.D., Heyneman, D., Lim, K.C. and Wilson, A.C., 1986. Wormy mice in a hybrid zone. *Nature*, *324*(6092), pp.60-63.

Scharsack, J.P., Franke, F., Erin, N.I., Kuske, A., Büscher, J., Stolz, H., Samonte, I.E., Kurtz, J. and Kalbe, M., 2016. Effects of environmental variation on host–parasite interaction in three-spined sticklebacks (Gasterosteus aculeatus). *Zoology*, *119*(4), pp.375-383.

Schluter, D., 1995. Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology*, *76*(1), pp.82-90.

Schluter, D., Clifford, E.A., Nemethy, M. and McKinnon, J.S., 2004. Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, *163*(6), pp.809-822.

Schultz, E. T., Topper, M. and Heins, D. C. 2006. Decreased reproductive investment of female three spine stickleback Gasterosteus aculeatus infected with the cestode Schistocephalus solidus: parasite adaptation, host adaptation, or side effect? EEB Articles. 14 – Oikos, 114: 303-310.

Seddon, J.M., Santucci, F., Reeve, N.J. and Hewitt, G.M., 2001. DNA footprints of European hedgehogs, Erinaceus europaeus and E. concolor: Pleistocene refugia, postglacial expansion and colonization routes. *Molecular Ecology*, *10*(9), pp.2187-2198.

Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D. and Kingsley, D.M., 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, *428*(6984), p.717.

Shapiro, M.D., Summers, B.R., Balabhadra, S., Aldenhoven, J.T., Miller, A.L., Cunningham, C.B., Bell, M.A. and Kingsley, D.M., 2009. The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. *Current Biology*, *19*(13), pp.1140-1145.

Smith, C., Zięba, G., Spence, R., Klepaker, T. and Przybylski, M., 2020. Three-spined stickleback armour predicted by body size, minimum winter temperature and pH. *Journal of Zoology*, *311*(1), pp.13-22.

Svanbäck, R., and Eklöv, P. (2003). Morphology dependent foraging efficiency in perch: A trade-off for ecological specialization? *Oikos*, 102, 273–284.

Szymura, J.M. and Barton, N.H., 1991. The genetic structure of the hybrid zone between the fire-bellied toads Bombina bombina and B. variegata: comparisons between transects and between loci. *Evolution*, *45*(2), pp.237-261.

Szymura, J.M., 1993. Analysis of hybrid zones with Bombina. In: *Hybrid zones and the evolutionary process (ed. RG Harrison)*, pp.261-289. NewYork: Oxford University Press.

Taugbøl, A., Junge, C., Quinn, T.P., Herland, A. and Vøllestad, L.A., 2014. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. *Ecology and Evolution*, *4*(2), pp.144-156.

Teeter, K.C., Payseur, B.A., Harris, L.W., Bakewell, M.A., Thibodeau, L.M., O'Brien, J.E., Krenz, J.G., Sans-Fuentes, M.A., Nachman, M.W. and Tucker, P.K., 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome research*, *18*(1), pp.67-76.

Teeter, K.C., Thibodeau, L.M., Gompert, Z., Buerkle, C.A., Nachman, M.W. and Tucker, P.K., 2010. The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution: International Journal of Organic Evolution*, *64*(2), pp.472-485.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. and Mugue, N.S., 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. *PLoS Genet*, *10*(10), p.e1004696.

Thiel, T., Kota, R., Grosse, I., Stein, N. and Graner, A., 2004. SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Research*, *32*(1), pp.e5-e5.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012.Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15), pp.e115-e115.

Vines, T.H., Dalziel, A.C., Albert, A.Y., Veen, T., Schulte, P.M. and Schluter, D., 2016. Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution*, *70*(5), pp.1023-1038.

Wafer, R., Tandon, P. and Minchin, J.E., 2017. The role of peroxisome proliferator-activated receptor gamma (PPARG) in adipogenesis: Applying knowledge from the fish aquaculture industry to biomedical research. *Frontiers in endocrinology*, *8*, p.102.

Walker, J.A., 1997. Ecological morphology of lacustrine threespine stickleback Gasterosteus aculeatus L.(Gasterosteidae) body shape. *Biological Journal of the Linnean Society*, *61*(1), pp.3-50.

Walker, J.A. and Bell, M.A., 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks (Gasterosteus aculeatus). *Journal of Zoology*, *252*(3), pp.293-302.

Wark, A.R., Mills, M.G., Dang, L.H., Chan, Y.F., Jones, F.C., Brady, S.D., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M. and Kingsley, D.M., 2012. Genetic architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3: Genes, Genomes, Genetics*, *2*(9), pp.1047-1056.

Westram, A.M., Rafajlović, M., Chaube, P., Faria, R., Larsson, T., Panova, M., Ravinet, M., Blomberg, A., Mehlig, B., Johannesson, K. and Butlin, R., 2018. Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow. *Evolution letters*, *2*(4), pp.297-309.

Whiting, J.R., Magalhaes, I.S., Singkam, A.R., Robertson, S., D'Agostino, D., Bradley, J.E. and MacColl, A.D., 2018. A genetics-based approach confirms immune associations with life history across multiple populations of an aquatic vertebrate (Gasterosteus aculeatus). *Molecular ecology*, *27*(15), pp.3174-3191.

Yin, A., Korzh, S., Winata, C.L., Korzh, V. and Gong, Z., 2011. Wnt signaling is required for early development of zebrafish swimbladder. *PloS one*, *6*(3), p.e18431.

Yu, J., Carroll, T.J., Rajagopal, J., Kobayashi, A., Ren, Q. and McMahon, A.P., 2009. A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. *Development*, *136*(1), pp.161-171.

Chapter 4: Association of some genetic markers with parasitic infection in stickleback

Abstract

Parasites are thought to be potent biotic selective agents in the adaptive evolution of animals, but little is known about the genetic basis of susceptibility to infection in wild organisms. Three-spined stickleback provide a good model for improving our understanding in this area, because of their well characterised variation in susceptibility and availability of diverse genetic tools. In the present study, the quantification of parasite abundance in natural and experimental environments and their association with selected candidate genes (Eda, PPARA, WNT7B and NLRC5) has assisted in elucidating the genetic regions underlying adaptation to infection. Candidate genes were selected because of previous work showing putatively adaptive divergence in allele frequencies between marine and freshwater. There was a strong relationship between some alleles and the abundance of ectoparasites and endoparasites of stickleback. Under natural conditions, *Gyrodactylus* sp. infection burden was associated with PPARA locus, plate morph and sex of fish while *Diplostomum* sp. had a significant relationship with plate morph and WNT7B locus. For experimental infections with the ectoparasite *Gyrodactylus gasterostei*, there was a significant association of parasite abundance with sex, plate morph and Eda locus, whereas growth rate of fish was associated with plate morph and PPARA locus. For the endoparasite *Diplostomum pseudospathecum*, there was a significant association of Eda locus with parasite burden of the fish under experimental condition. The divergent effects of parasites on different stickleback genes suggest parasite-mediated selection in the ecological adaptation of fish.

1. Introduction

Modern genomic studies are revealing genes and genomic regions that differ in allele frequencies between even closely-related populations, in ways that suggest they may be involved in adaptation (Jones et al., 2012a). To understand the adaptive significance of these differences we need to identify the environmental factors, or 'agents of selection' that drive them, but we know rather little in general the relative importance of different agents of selection, or which ones are important in specific circumstances (MacColl, 2011). Examples of selective agents or environmental factors in nature that are responsible for adaptive differentiation among local populations of a given species, include intraspecific competition that is responsible for the changes of the size and shape of Galapagos finch bills (Grant and Grant, 2006), variation in body size of humbug damselfish (*D. aruanus*) (Forrester, 1991) and trophic morphology of arctic charr (*S. alpinus*) (Malmquist, 1992).

Natural enemies like parasites are also potential agents of selection which might affect traits or taxa to drive adaptive changes in natural populations (Colosimo et al., 2005; MacColl, 2011; Karvonen and Seehausen, 2012). Differences in infections within or between host populations, depending on different ecological or environmental conditions, provide the foundations for investigating parasite-mediated divergent selection (Karvonen and Seehausen, 2012). In freshwater fishes such as Arctic charr (*S. alpinus*), whitefish (*Coregonus lavaretus*), African cichlid fish (*Pundamilia pundamilia*) and three-spined stickleback (*Gasterosteus aculeatus*) have been reported for divergent parasite infections among ecotypes and species inhabiting mostly pelagic versus benthic habitats but knowledge gaps remain to identify the genetic basis of susceptibility to infection (Frandsen et al., 1989; Knudsen et al., 1997; Little, 2002, Knudsen et al., 2003; Maan et al., 2008; MacColl, 2009). The strength and prevalence of parasite-mediated selection in nature largely relies on two criteria: firstly, reduction of host fitness in terms of reproduction or survival and secondly, host genetic variation for resistance within populations (Tompkins and Begon, 1999; Little, 2002). Therefore, it is important to know about the genetic correlates of infection in a host-parasite system by observing the variation of susceptibility in host genotypes within populations in the wild and laboratory condition.

The three-spined stickleback (*G. aculeatus*) has recently-emerged as an important model fish in eco-evolutionary research (Kingsley et al., 2007). This is due in part to its parallel evolution of freshwater adaptative radiations from bony-armoured marine ancestors that occurred around 10,000 years ago (Bell and Foster, 1994; Colosimo et al., 2005). To

adapt to the new environmental conditions (new food sources, predators, parasites, salinity, water temperature etc.), freshwater stickleback show extensive divergence in morphological, physiological and behavioural characteristics (Ostlund-Nilsson et al., 2006; Miller et al., 2007). However, the genomic mechanisms underlying such adaptive evolution are still largely undetermined (Jones et al., 2012a). A recent study revealed that three-spined stickleback possess a genome-wide set of adaptive loci associated with repeated marine-freshwater evolution, but relatively few genes have been identified underlying the genetic basis of a specific phenotype undergoing natural selection and evolution (Jones et al., 2012a; Jones et al., 2012b). The genes contributing to the variation of armour plates (*Eda* gene), reduction of pelvic structure (*Pitx1* gene) and pigmentation (*Kitlg* gene) of stickleback have been identified using microsatellite markers, allele-specific expression and QTL mapping (Colosimo et al., 2005; Miller et al., 2007; Chan et al., 2010). In addition, advanced genome-scan approaches using RAD-data sequencing also allow us to investigate an array of genes with predicted functions related to osmoregulation, metabolism, immunity, behaviour and morphology, and are likely to be important for adaptation from marine to freshwater forms (Terekhanova et al., 2014; Haenel et al., 2019). Therefore, the identification of adaptive loci or genes involved in parallel divergence of stickleback into two different habitats is the prerequisite to map the genetic basis of an individual trait evolved by natural selection.

In the present study, a set of candidate genes (PPARA, WNT7B and NLRC5) was selected from among those showing strong (often fixed) SNP differences between freshwater and marine populations (detail in chapter 2) (Hohenlohe et al., 2010; Jones et al., 2012a; Terekhanova et al., 2014; Haenel et al., 2019). Support for these loci having a causal effect on a particular phenotype can be obtained through genotyping individuals only at those candidate loci, using SNPs as markers and testing statistically whether different alleles can predict the phenotype in a large sample of individuals. An implicit assumption is that these genomic regions are the result of adaptation to differences in salinity, but many other things differ between marine and freshwater environments which may act as selective agents causing ecological adaptation. Although PPARA (lipid metabolism), WNT7B (paracrine signalling pathway of kidney) and NLRC5 (immunity response toward microbial invasion) are known for specific phenotypes, experimental evidence demonstrates that multiple phenotypes can map to the same genomic region in association with specific factors (Yu et al., 2009; Neerincx et al., 2010; Mills et al., 2014; Wafer et al., 2017). For example, the Eda gene contributing to the variation of armour plates in freshwater and marine ecotypes due to strong environmental selection, it also associated with growth, behaviour, immunity and sensory system development of stickleback (Barrett et al., 2009a; Barrett et al., 2009b; Mills et al., 2014; Robertson et

al., 2017). It is not immediately obvious that these candidate genes should be involved in tolerating salinity variation per se. Other environmental factors, including parasitism (El Nagar and MacColl, 2016), differ strongly between marine and freshwater environments and are putative agents of selection.

Previous molecular research on stickleback has confirmed that the Eda gene located in chromosome IV is responsible for most of the variation (approximately 80%) in the pattern and number of lateral armour plates (hereafter 'plate morph') between marine and freshwater fish (Colosimo et al., 2004, 2005). A polymorphism in the second intron of this gene exhibits two major alleles which results in three different phenotypes : 1. fish that are homozygous for the ancestral Eda^C allele are completely plated (32-36 plates), 2. fish that are homozygous for the derived Eda^L allele are low plated (<10 plates), and 3. heterozygous fish exhibit an intermediate partially plated phenotype with a discontinuous row of lateral plates (10-25 plates) (Colosimo et al., 2005; Jones et al., 2012a; Jones et al., 2012b). Besides lateral armour plates, multiple phenotypes have been shown to be associated with a genomic region around the Eda gene due to potential pleiotropic effects of this gene or to multiple tightly-linked genes (Barrett et al., 2009a; Barrett et al., 2009b; Mills et al., 2014; Robertson et al., 2017). There is also evidence of association between the Eda haplotype block and the expression pattern of key immune system genes (Foxp3, Stat4, Stat6, Cmip, Il-1β, Tbet, Tnfα, Tgfβ) in ectoparasite-infected stickleback in semi-natural conditions which showed changes in expression levels of certain genes (increase or decrease) in linkage with Eda (Robertson et al., 2017). It is therefore possible that parasites might cause additional selection on the Eda locus, either as a direct result of pleiotropy, or through selection on linked genes involved in the immune response. Therefore, it is important to investigate the possibility that parasites are a selective agent in local adaptation to freshwater by identifying association with the genetic regions underlying adaptation to infection.

Differences in the patterns of infection in various populations and ecotypes of host species are the most important prerequisite to understand parasite-mediated divergent selection (Karvonen and Seehausen, 2012). Like Arctic charr (*S. alpinus*), whitefish (*Coregonus lavaretus*) and African cichlid fish (*Pundamilia pundamilia*), stickleback also show a pattern of divergent parasite burdens between lake and river populations in freshwaters as well as between freshwater and marine ecotypes (Frandsen et al., 1989; Knudsen et al., 2003; Rauch et al., 2006; Scharsack et al., 2007; Maan et al., 2008; Eizaguirre et al., 2010; McCairns et al., 2011; Scharsack and Kalbe, 2014). MacColl and Chapman (2010) also demonstrated that parasites can cause selection against marine migrants to freshwater, with anadromous stickleback having higher burdens of novel

parasites than resident freshwater populations and suffering a growth cost as a direct result. To date, there is limited study of the genetic basis of the pattern of divergent parasite resistance within and between populations of stickleback and this is confined largely to the major histocompatibility complex (MHC) (Rauch et al., 2006; El Nagar and MacColl, 2016; Robertson et al., 2017). MHC genes exhibited divergence selection in parasite-mediated local adaptation between river and lake stickleback where fish carrying lake MHC genotype experienced lower parasite burden in lakes (Eizaguirre et al., 2012). In another study, there was an affiliation of Eda alleles with the expression pattern of some immune genes with higher abundance of *Gyrodactylus* sp. in low plated than high plated fish grown in cages in saltwater (Robertson et al., 2017). The present study is an attempt to use a candidate gene approach to understand the genetic basis of parasite-mediated selection other than the extensively studied pathogen-driven selection of MHC gene in stickleback (Wegner et al., 2003; Kurtz et al., 2004; Janeway et al., 2005; Wegner et al., 2008; Eizaguirre and Lunz, 2010; Eizaguirre et al., 2012).

Study of stickleback has documented a diverse range of parasite fauna with varied selective roles in ecological divergence of phenotypic traits (Wegner et al., 2003; Barber, 2007; MacColl, 2009; MacColl and Chapman, 2010; De Roij and MacColl., 2012). *Gyrodactylus* sp. is a well-known monogenean flatworm and a dominant ectoparasite of stickleback having significant impact on fitness of infected fish through disruption of the surface layer of skin and fins and also resulting in secondary fungal infection (Bakke et al., 2007; Wegner et al., 2008; El Nagar and MacColl, 2016). The digenean trematode *Diplostomum* sp*.* (Eye fluke) is another common parasite of stickleback with a complex life cycle involving three hosts: lacustrine birds, fish and pond snails (Chappell, 1995). Stickleback are infected by the cercariae (larvae) of this endoparasite as a second intermediate host. The parasite is found in the lens and retina as metacercariae (next stage larvae) and affects the fitness negatively through impaired vision (Owen et al., 1993). Stickleback and their parasites found on North Uist, in the Outer Hebrides of Scotland, represent a well-studied system exhibiting a varying degree of resistance to infection among different populations of this fish (De Roij et al., 2011; De Roij and MacColl., 2012; El Nagar and MacColl, 2016). Previous study of North Uist stickleback confirmed that the infection level of *Gyrodactylus arcuatus*, was typically high in fishes of saltwater populations while *Diplostomum gasterostei* wasspecific to freshwater populations. Hence these two parasites are potential selective agents of local adaptation (De Roij et al., 2012; El Nagar and MacColl, 2016).

In this study, I hypothesised that parasite resistance would show divergence in selected candidate genes (Eda, PPARA, WNT7B and NLRC5) in both wild and laboratory-bred

stickleback. A putative hybrid zone (a freshwater stream around 1 km in length) located in North Uist, Scotland was selected with admixed populations of marine anadromous and freshwater stickleback to observe the pattern of natural abundance of *Gyrodactylus* sp. and *Diplostomum* sp. infection. A naturally-occurring hybrid zone (a transition area between two species or ecotypes) serves as a sieve to isolate the effects of individual loci on the phenotype of interest from admixture or recombinant organisms of divergent parental taxa (Buerkle and Lexer, 2008; Kawakami and Butlin, 2012). To compare the natural infection status, laboratory experiments were conducted to understand the genetic basis of susceptibility to infection of a single host. Two stickleback populations with contrasting types of parasite prevalence were selected for the experiments to examine the effects of parasite abundance on the genotypes of selected candidate genes. For experimental infections with the ectoparasite, *G. gasterostei* and endoparasite, *D. pseudospathaceum*, fish were raised in the lab from hybrid crosses between marine and freshwater parents, which also had contrasting Eda genotypes. This randomised the genetic background of each individual along with segregating the lateral plate phenotype variation and maternal effects to facilitate the screening of the impact of host genotypes on parasite burden.

I expect that the genetic effects will be different in both artificial and natural conditions for the same parasites due to the influence of confounding factors such as spatial variation, salinity of the hybrid zone etc. Besides, I assume that the patterns of ectoparasite (*Gyrodactylus* sp.) and endoparasite (*Diplostomum* sp.) infection and their association with genotypes at candidate loci would be different due to their mode of infections. Through artificial infection experiments, I predicted that plate morph and Eda in a haplotype block with other immune-related gene (Garp, Baff) might have similar impacts on parasite resistance due to their selection and adaptive role in divergent environment. Therefore, the possible findings of the present study might serve as a proof of genetic basis of susceptibility to infection and may establish parasites as an agent of selection in local adaptation.

2. Materials and method

2.1 Study design

Patterns of parasite infection and their association with genotypes at candidate loci were examined in both wild fish from a hybrid zone and in artificial infection experiments of laboratory-raised hybrid fish. To observe the effects of genotype on natural infection patterns of natural hybrids of marine anadromous and freshwater stickleback, 216 sticklebacks were collected from 10 sites of the Hosta stream, North Uist, Scotland (57^0 $36'0.$ N, $7^0 19'58.8.$ W). Abundance of a common ectoparasite (*Gyrodactylus arcuatus*) and an endoparasite (*Diplostomum* spp.) were investigated and all fish were genotyped at four candidate loci (Eda, PPARA, WNT7B and NLRC5). To examine infection patterns experimentally, three experiments were conducted – two for the ectoparasite (*G. gasterostei*) artificially infecting the F2 generation of two different marine x freshwater crosses (Faim x Torm and Obsm x Chru) from North Uist, Scotland and one for the endoparasite (*D. pseudospathecum)* artificially infecting lab-raised F1 fish from the already admixed Hosta hybrid zone on North Uist. For *G. gasterostei,* experiment 1 used 50 fish genotyped for only one candidate gene (Eda). Experiment 2 used 100 fish that were genotyped for all four candidate genes. Experiment 3 (*D. pseudospathaceum*) was conducted with 90 fish that were genotyped for all four candidate loci.

2.2 Natural infection of ectoparasite and endoparasites

All fish samples were caught from the freshwater Hosta stream with Gee's Minnow Traps, set overnight (approximately 16h) at ten sites in 08-16th May, 2018. Across all ten sampling sites, 216 live fish samples were haphazardly selected and transported immediately to the laboratory in stream water in darkened boxes with continuous aeration. Fish were euthanized with an overdose of MS222 (400mgL^{-l}) following Schedule 1 techniques (confirmation of death by destruction of the brain) according to UK Home Office regulations. For the preliminary identification of ecotypes (anadromous, hybrid or resident) of fish, the standard length (mm), weight (gm) were recorded, alongside a digital photograph of the left side for further morphometric study. Samples were then examined for natural infection status of *Gyrodactylus* spp. by scanning the whole body and all fins (dorsal, Pectoral, pelvic, anal and caudal fins) at 20x magnification under a stereoscope microscope. After that both eyes were removed from the fish to a petri dish with small amount of "tap water", the number of metacercariae larvae of *Diplostomum* spp. in the lens and retina of fish were recorded after observation under the microscope (El Nagar and MacColl, 2016). For genetic analysis, two pectoral fins and the caudal fin

were collected in 100% ethanol and stored at -20^0C, and the rest of the fish was preserved with a tag in 70% ethanol for further morphometric study.

2.3 Experimental infection ofectoparasite and endoparasite
2.3.1 Fish crossing for infection experiment

For experiment 1, F1 families of low armour-plated fish from a freshwater loch-Tormasad (Torm) (57°33"N; 7°19"W) with fully plated pure anadromous individuals from a saltwater loch - Fairy Knoll (Faik) (57°38"N; 7°12"W) in June 2016. F2 hybrids were created by crossing brother-sister F1 generation with 4 pairs of male and female [Cross ID: E7 $_{17x3}$ (25 offspring), E7 $_{17x5}$ (22 offspring), E7 $_{17x7}$ (33 offspring), E7 $_{17x8}$ (12 offspring)] produced from a single grandparent in August, 2017. Fifty-five fish (50 infected and 5 control) were used from 92 progenies in this pilot ectoparasite (*G. gasterostei*) infection experiment conducted in June 2018.

For experiment 2, F1 families were created by crossing individuals with low armour plates from a freshwater loch- Loch a Chadha Ruaidh ('Chru') with fully plated pure anadromous individuals from a saltwater lagoon - Loch Ob nan Stearnain ('Obse') in May 2017. These two populations have been shown to have a stable contrast in abundance of *G. gasterostei* infection with high burden in Obsm and zero in Chru (De Roij et al., 2011). In May 2018, an F2 generation was created by crossing F1 Obsm x Chru families following the same method from brother-sister crosses with 6 males and 6 females to produce a total of 150 offspring [Cross ID: Obsm x Chru- $_{18 \times 3}$ (14 offspring), $_{18 \times 5}$ (24 offspring), $_{18 \times 6}$ (30 offspring), $_{18 \times 9}$ (32 offspring), $_{18 \times 10}$ (40 offspring), $_{18 \times 15}$ (10 offspring)].

The crossing and artificial fertilization process followed De Roij et al., (2011). Briefly, fully gravid females and mature males (with bright blue-green eyes and orange-red fore belly and throat) were caught using minnow traps and kept in a container before the cross. Eggs from a female were stripped into a small petri dish with a small volume of sterile embryo medium (1% NaCl solution). Immediately, a male was killed by schedule 1 (overdose of anaesthesia using MS222 (400mg/L) and confirmed by destruction of brain) to collect the testes and released sperm thoroughly over the eggs. Fertilised clutches were incubated in well-aerated, saline (1-2 ppt) water for 8 to 10 days. Methylene blue (2.303% commercial solution, Kordon, USA) was used to prevent fungal infections at a rate of 3 drops / 10L water in each egg-incubation tank for 10 days. After hatching, all families were reared in an aquarium in a temperature- and light-controlled room (13.5 ± 1^0C natural photoperiod/14hrs light + 10hrs dark) at the University of Nottingham. Newly-hatched fry were fed with *Paramecium* spp. for the first five days and

then with freshly-hatched brine shrimp (*Artemia salina*) for around 60 days. After that the fingerlings were fed with chironomid larvae (blood worm) daily (morning) until maturity and then twice daily (morning and evening) for around a month to increase their size. For the F2 Obsm x Chru crosses, due to a shortage of males in the stock population, sperm of a male fish were cryopreserved in liquid nitrogen for the fertilization of eggs. One hundred and three fish (100 infected and 3 control) were used in the second ectoparasite (*Gyrodactylus* sp.) infection experiment conducted in June 2019.

For experiment 3 (*Diplostomum*), an F1 generation was created following the same procedure described above from the crosses of 11 pairs of admixed parents of marine anadromous and freshwater stickleback collected from the HOSTA hybrid zone, North Uist in May 2018. 199 offspring were produced from 11 males and 11 females [Cross ID: Hosta $_{18x1}$ (45 offspring), $_{18x2}$ (35 offspring), $_{18x3}$ (25 offspring), $_{18x4}$ (25 offspring), $_{18x5}$ (8 offspring), $_{18x6}$ (16 offspring), $_{18x7}$ (5 offspring), $_{18x8}$ (28 offspring), $_{18x9}$ (1 offspring), $_{18x11}$ (10 offspring), $_{18x12}$ (1 offspring)] and reared in the aquarium for about one year. Among them, 94 fish (90 infected and 4 control) were used in the *Diplostomum* infection experiment conducted in September 2019.

2.3.2 Experimental exposure of ectoparasite (*G. gasterostei*)

All experiments were conducted under UK Home Office licence (PPL-30/3415 and PPI-IDA022B8C) and approved by the Animal Welfare and Ethical Review Body (AWERB) committee at the University of Nottingham. The same procedure of exposure was followed for both experiment 1 and 2.

To infect the fish with *Gyrodactylus*, naturally-infected donor fish were collected, using minnow traps, from the local stream - Tottle brook (52°56'00.1"N 1°12'01.5"W) running across the University of Nottingham campus. The entire infection procedure followed De Roij et al., 2011. Briefly, *G. gasterostei* identification was confirmed under a dissecting microscope by their distinguishing anatomical features - short body shape and absence of excretory bladders. After observing the presence of parasites in the donor fish (N = 30), they were housed in a single 100L tank for a week to increase the number of parasites. For individual experimental fish, a 2L plastic tank was prepared with one litre of dechlorinated tap water two days prior to the start of the experiment. On the day of exposure, donor fish were killed with an overdose of MS222 (400mg L^{-1}). All fins with the parasite were removed and were then placed in a petri dish containing a small amount of tap water. When the worms detached from the fins (after 20-30 minutes) and started to

move in the water, the tail of the recipient fish was held close until two parasites transferred onto it. Prior to infection, the experimental fish was anaesthetized with a weak dose of MS222 (100mg L-1) and their length (standard length) and weight (dried with paper tissue) recorded. They were then released into the small tank after infection. All tanks were kept at a constant in a temperature-(12.73 ± 1^0C) and light-controlled room (14hrs light + 10hrs dark). On the following day, each fish was scanned under the microscope to confirm the presence of two worms and re-infected if any had been lost. Three to five fish were housed under the same conditions without infection as a control group.

Infections were then followed until the end of the experiment (40 days). The number of parasites was counted by scanning the whole-body including mouth, eyes, gill opening, flanks, girdles, cloaca, all spines and fins (dorsal, pectoral, pelvic, anal and caudal fin) under a stereomicroscope every 4/5 days. One third of the water from each tank was changed every three days and the fish were fed with whole blood worm (3-4 worms) once a day. Any fish showing poor performance in movement and feeding were killed by the schedule 1 method. Two artificial infection experiments were conducted with *Gyrodactylus*. In experiment 1 in June 2018 [F2 generation of Faim (male) + Torm (female)], 50 fish were infected and 5 were kept as controls. In experiment 2, the same infection protocol of the previous experiment was followed with 100 infected and 3 control fish in June 2019 [F2 generation of Obse (male) + Chru (female)]. On the 35th day, 70 fish that showed a decline in parasite abundance from the peak were killed by schedule 1 and 30 fish were kept for another 5 days (40th day) to observe the decline from the peak. At the end of the experiment, the final length (mm) and weight (gm) was measured. Tissues were collected (two pectoral and caudal fins) in 95% ethanol and preserved at -20^{0}C for future genetic study. The sex of the fish was also recorded by dissection of the abdominal region and a photo was captured for the record of identification. In experiment 1 four fish (on day 8, 15, 16 and 25), and in experiment 2 one fish (on day 30th), were killed by schedule 1.

2.3.3 Experimental exposure of endoparasite (*D. pseudospathecum*)

Pond snails (*Lymnea stagnalis*) were collected from a local lake (Martin's pond) ($52^057'23.06''$ N and $1^o13'02.60''$ W) located in Wollaton, Nottingham, UK in September 2019 (Fig. 1A-B). The presence of cercariae larvae of *Diplostomum* was confirmed by placing the snails in water exposed under bright, fluorescent light for an hour, which prompts the discharge cercariae from the snail (Fig. 1C). The cercariae larvae were identified as *D. pseudospathecum* from the Y-shaped body (bifurcate tail) and jerking

140

movement under a light microscope (De Roij et al., 2011; El Nagar and MacColl, 2016). After that the snails were kept in a 15 L tank for 5 days and fed with fresh soft lettuce daily. The snails were then re-exposed to the fluorescent light to encourage parasite shedding. Twenty cercariae were used to infect each fish and all larvae were collected from a single snail.
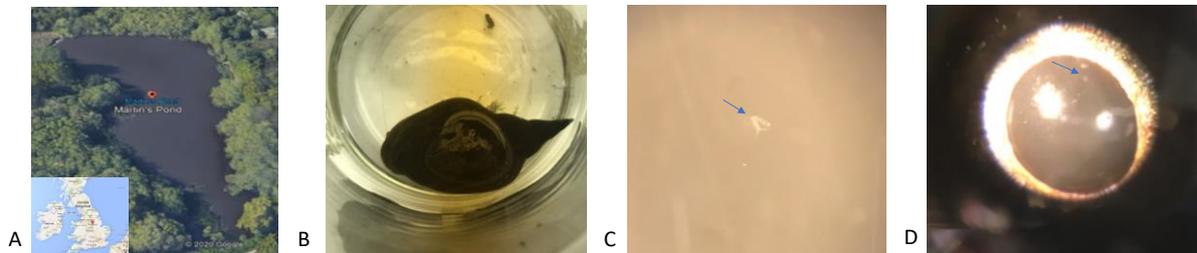
Fig. 1 Collection of *Diplostomum* infected pond snail from a local lake in Nottingham, UK. A. Martin's pond with UK map, B. Pond snail (*L. stagnalis*), C. *D. pseudospathecum* and D. Stickleback eye infected with metacercaria larvae *of D. pseudospathecum* observed under microscope.

For each individual fish, a 2L plastic tank was prepared with 1L of dechlorinated tap water two days prior to the start of the experiment. On the day of exposure, recipient fish were anaesthetized with a weak dose of MS222 (100mg L^{-1}) to measure their length and weight and were then released into a single holding tank with aeration. All fish were haphazardly selected based on their plate morphology in order to have completely plated, partially plated and low plated fish of equal number. After that a dose of 20 cercariae was added to each tank and infection was confirmed by checking the water for free moving cercaria after one hour of exposure.

In this experiment, 90 fish were infected with the larvae of *Diplostomum* and 4 fish were kept as a control without any exposure to parasites for the duration of 3 days. The fish were kept in a temperature and light controlled room and fed with whole bloodworm once a day. At the end of the experiment, the fish were killed by schedule 1, their weight recorded and tissue (two pectoral and caudal fins) collected in 96% ethanol and preserved at -20^{0}C for further genetic analysis. Length of fish was not recorded at the end because the short period (3 days only) of the experiment would have made changes difficult to detect. After that both eyes were removed from the fish into a petri dish with a small amount of tap water and the number of metacercaria larvae observed in the lens of the fish under a dissection microscope (Fig. 1D) was recorded.

2.4 Morphometric analysis

Armour plate number (plate morph) of sticklebacks distinguishes the fully plated marine form from the low plated freshwater ecotype. All natural and experimental fish were tagged with an identifying number and preserved in 70% ethanol for staining with Alizarin Red solution following a standard staining method described in chapter 2. After staining, the number of plates were counted from a photograph to separate them into three morpho-types: completely plated (CM), partially plated (PM) and low plated (LM) and compared with the corresponding Eda genotype (detail in chapter 2).

2.5 Genetic analysis

For genetic study of parasite-infected fish, a total 468 fish tissue samples were collected from the natural infection study (N = 216) and three infection experiments- ectoparasite infection with *Gyrodactylus* (N = 158) and endoparasite infection with *Diplostomum* (N = 94) including control fish. Genomic DNA was extracted from pectoral fins of all samples using Quanta Bio- Extracta^{TM} DNA prep for PCR Tissue kit (UK) and purified with proteinase K-ethanol precipitation procedure (Goldenberger et al., 1995). DNA concentration was measured with a Bio-drop spectrophotometer (Biochrom, USA) following the manufacturer's protocol and then preserved at -20^0C for gene specific analysis. The extracted DNA was used to genotype fish at the Eda locus and then at the remaining three candidate loci using a SNP-based assay (Appendix Table S2 and S3).

2.5.1 Eda genotyping

The genotyping of EDA locus was carried out to quantify the relationship with the three phenotypic morphs (CM, PM and LM). All fish were genotyped for the Eda gene using a previously identified microsatellite marker *Stn382* (flanks a 60bp indel polymorphism in the 1st intron of Eda gene) developed by Colosimo *et al.*, 2005. The marker *Stn382* produced a single band for either a 158bp product for the Eda^L allele (LL) or a 218bp for the Eda^C allele (CC) and a double band for heterozygous (Eda^{CL}) fish.

2.5.2 SNPs assay typing

Generally, the method of developing a SNP assay based on PCR-restriction fragment length polymorphism (RFLP) includes: a. selection of candidate genes from the related literature b. filtering of SNPs from previously sequenced whole genome data, c. collection of sequences of 600-1000bp of target gene from the Ensembl genome browser 97 (Hubbard et al., 2002), d. identification of restriction enzyme sites that differ between marine and freshwater sequences using Bio-edit (Hall, 1999), e. the design of primers

targeting specific polymorphic restriction enzymes sites (detail in chapter 3), and finally f. the amplification and treatment of target segments with specific restriction enzyme (following Komar, 2009) including the sequencing of 4-8 samples per gene to find out the differences between marine and freshwater fish. SNP-based assays of three candidate genes (PPARA, WNT7B and NLRC5) have been developed successfully (detail in chapter 2) and 197 samples (103 from experiment 2 and 94 from experiment 3) from the experiment and 216 samples from the Hosta hybrid zone have been genotyped for each SNP assay.

2.5.3 Primer and PCR condition

Generally, the PCR reactions were preliminary fixed in 20µL volumes. Each reaction contained 3µL DNA template, 10 μl of PCR master Mix (Quanta Bio- Accustart TM II Gel^{TM} Super Mix$^)$, 2μl of forward primer (1pmol/μl), 2μl of reverse primer (1pmol/μl) (Sigma) and 3µL molecular biology grade water. The properties of primers used for the amplification of each gene segment are described in chapter 2. Usually, the PCR reaction was performed in a thermocycler (BIO Rad, Mexico) under the following conditions: heated lid at 105°C and initial denaturation for 3 minutes at 94^{0}C, followed by a total of 40 cycles of denaturation at 94°C for 15 seconds, annealing at 58°C (depending on primer Tm) for 15 seconds and extension at 72°C for 30 seconds, and a final extension at 72°C for 4 min and hold at 4°C. PCR products were separated on a 1% agarose / TAE (2 M tris base, 1 M Glacial Acetic Acid, 0.05 M EDTA, pH 8.0) gel at 100Volts, 110mAMP for 50mins to 1 hour. To determine the approximate size of the amplified products, 5 µl of 100bp or 1Kb Hyper ladder (Bio line) was loaded as a molecular weight marker. For SNPs assay, the restriction enzyme-treated DNA was separated on a 2% agarose gel at 70-100 Volts, 110 mAMP for 50 mins to 1.5 hr. and observed using UV-illumination (Photo Doc-II-imaging system, UVP, UK) to record the genotypes.

2.6 Data analysis

Data of all experiments were collected in Excel (Microsoft) and statistical analyses were conducted using R, version 3.6.3 (R Core Team, 2020).

Natural infection dynamics of ectoparasite and endoparasite

The total number of *Gyrodactylus* and *Diplostomum* (lens) in individual fish of the Hosta hybrid zone were quantified as measures of natural infection. A generalized linear model (GLM) with negative binomial error distribution and log-link function was fitted for *Gyrodactylus* abundance. This included plate morph (CM, PM, LM), Eda genotype (CC, CL, LL), PPARA genotype (PP, PR, RR), WNT7B genotype (WW, WT, TT), NLRC5 genotype (NN, NL, LL) and sex of fish (2 levels) as predictor variables. For *Diplostomum*, fish were considered as infected or non-infected, because of low abundance. A GLM was fitted with a binomial distribution and logit-link function with all previous factors as predictor variables (Sex, Plate morph, Eda, PPARA, WNT7B and NLRC5). Stepwise regression with a combination of forward and backword selection based on likelihood ratio tests was conducted. The best fitting model was then evaluated based on Akaike's Information Criterion (AIC).

Experimental infection dynamics of ectoparasite and endoparasite

The abundance of *Gyrodactylus* (ectoparasite) in the infection experiment generally followed a dynamic pattern: increasing from the initial dose to a peak and then decreasing to zero (De Roij et al., 2011). Therefore, the maximum abundance of *Gyrodactylus* at the peak of the infection (Max. count) and the total number of parasites throughout the whole experiment, calculated as the sum of all daily counts ("area under curve", AUC) were considered as the response variables. To investigate the association of parasite abundance with morpho-genetic factors, generalized linear models (GLMs) were used with a negative binomial error distribution and log-link function. AUC (which was strongly correlated to max count) was the response variable fitted with predictor variables as before (sex, plate morph, Eda, PPARA, WNT7B, NLRC5 loci). For experiment 1 with 50 fish, only Eda genotype along with plate morph were recorded as explanatory variables.

To quantify the relationship between fish growth and all genotypes (Eda, PPARA, WNT7B and NLRC5) as well as plate morph, the mean daily percentage increase in fish length or specific growth rate (SGR) was calculated using the formula: SGR (%/day) = $100*[\ln$ (final length in mm) $- \ln$ (initial length in mm)]/t (in days) (Barber, 2005; De Roij et al.,

2011). A GLM (gaussian distribution with identity-link function) was fitted also using stepwise regression with SGR as the response variable and the same predictor variables as the second experiment. SGR of control fish were added to the model to compare the growth rate with exposed fish.

For experiment 3, the total number of metacercariae of *D. pseudospathecum* (endoparasite) present in the lens of fish was considered as the response variable. A GLM with negative binomial error distribution and log-link function was fitted with sex of fish, plate morph, Eda, PPARA, WNT7B, NLRC5 loci as predictor variables. The best fitting model was obtained by stepwise regression with top down selection of significant variables determined by AIC.

Pattern of natural and experimental infection dynamics of ectoparasite and endoparasite

To compare the level of parasite abundance between the experimental and natural data, *Gyrodactylus* count from experiment 2 (N = 100) and the Hosta hybrid zone (N = 216), a Wilcoxon rank sum test was performed. The *Diplostomum* infection experiment (N = 90) was also compared with the natural infection level of Hosta hybrid zone following the above method.

All models were checked for the goodness-of-fit using Quantile-Quantile (Q-Q) plots of the residuals and the significance of the components was tested using F statistics.

3. Results

3.1 Natural infection dynamics of ectoparasite and endoparasite

The mean abundance of the endoparasite (*Diplostomum pseudospathaceum*) in the lens of Hosta hybrid-zone fish was (mean ± S.E.) 0.30 ± 0.05. That of the ectoparasite (*Gyrodactylus* spp.) was 7.99 ± 0.78.

A GLM revealed that plate morph, PPARA locus and sex of fish were significantly associated with the abundance of *Gyrodactylus* of naturally-infected fish (Table 1). The mean abundance of *Gyrodactylus* was highest in completely plated fish (CM) (10.30 ± 1.41) and lowest in partially plated fish (PM) (4.42 ± 0.70) (Fig. 2A). The same pattern was seen across Eda genotypes. The highest *Gyrodactylus* abundance was observed in the homozygous freshwater genotype (12.26 ± 3.45) of PPARA locus, the lowest counts were in the homozygous marine genotype (7.03 ± 0.84) (Fig. 2B). Male fish had significantly higher abundance (10.75 ± 2.13) of *Gyrodactylus* than female fish (6.70 ± 0.62) (Fig. 2C).
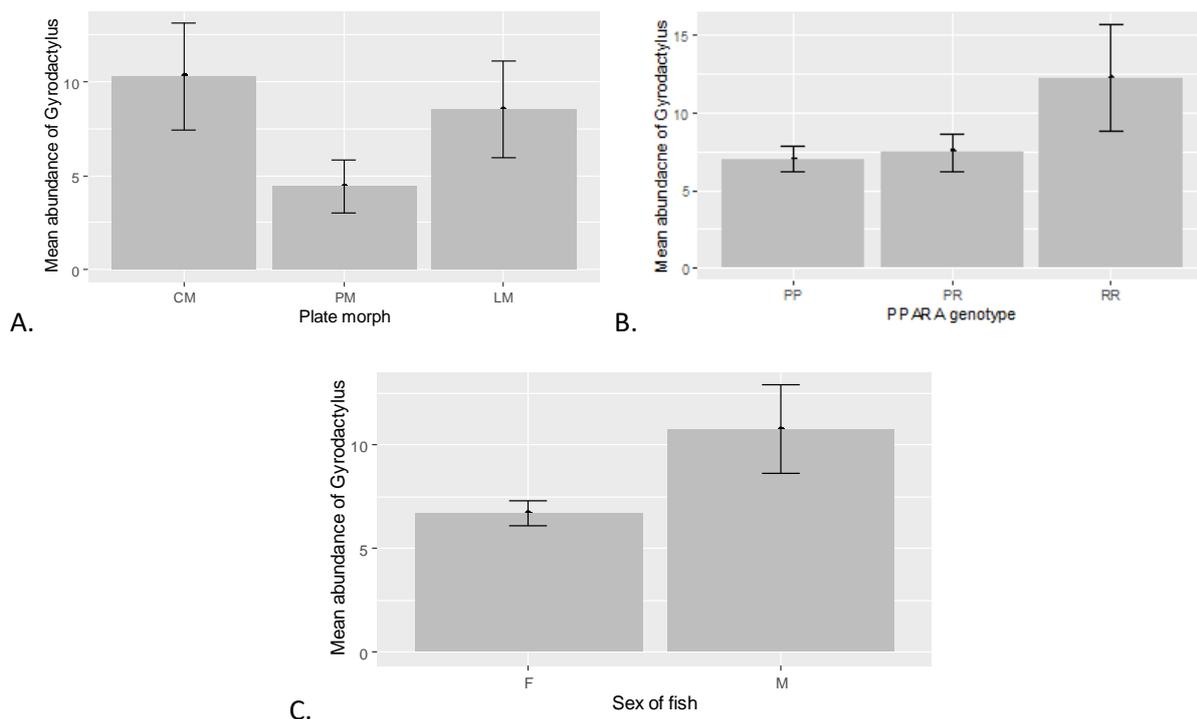
Fig. 2 Naturally-infected stickleback from the Hosta hybrid zone showing significantly higher *Gyrodactylus* spp. abundance (mean ± S.E.) in A. completely plated fish (CM) than partially plated (PM) and low plated fish (LM), B. homozygous freshwater genotype of PPARA than heterozygous and homozygous marine genotype and C. male (M) fish than female (F).

146

The prevalence of *D. pseudospathaceum* in the lens of fish had significant association with plate morph and WNT7B locus (Table 1). The completely plated fish had the highest prevalence of infection which was significantly higher than the partially-plated and low plated fish ($F_{2,\ 206}$ = 5.43, p = 0.005; Fig. 3A). The prevalence of *Diplostomum* was higher in fish that were heterozygous (WT) at the WNT7B locus ($F_{2,\ 204}$ = 4.80, p = 0.018; Fig. 3B).

A.

B.
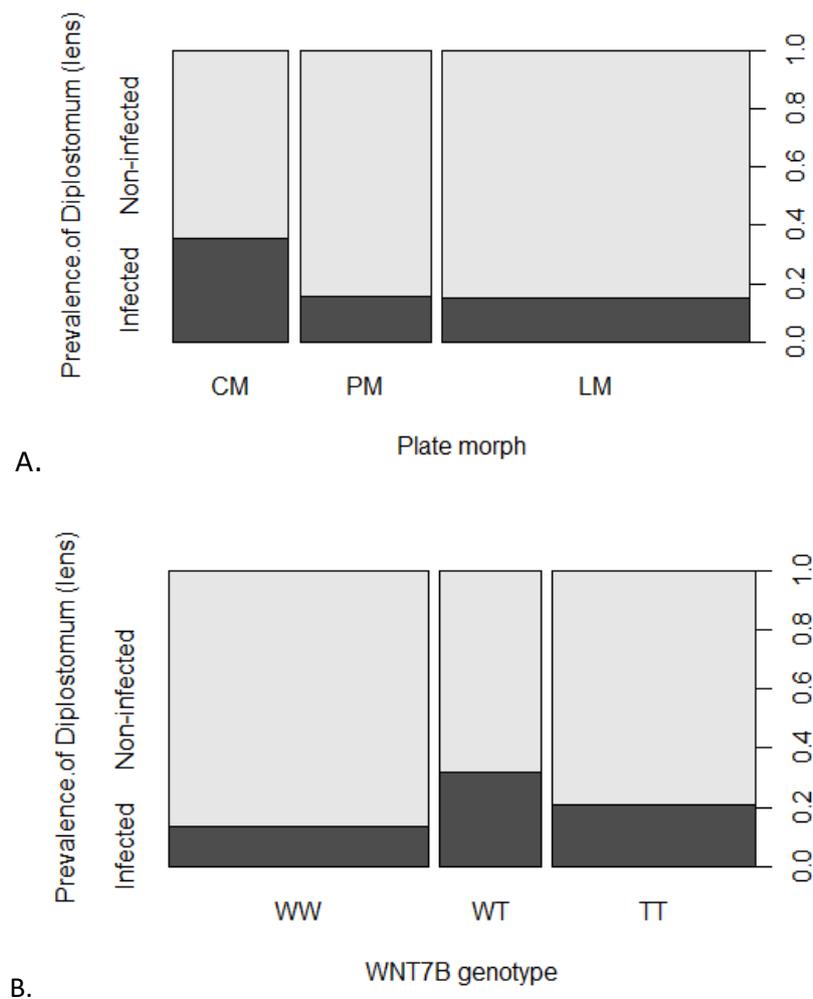
Fig. 3 *D. pseudospathaceum* infected in the lens of stickleback of natural source from the Hosta hybrid zone showing higher parasite prevalence in A. completely plated fish (CM) and B. heterozygous genotype (WT) of WNT7B locus.

147

Table 1: GLMs of minimum adequate models with significant variables fitted for phenotype and genotypes of parasite-infected stickleback in natural and experimental conditions. The natural ectoparasite and endoparasite abundance were recorded from the Hosta hybrid zone. Experiment 1 and 2 were conducted with *Gyrodactylus* (ectoparasite) for a period of 40 days and experiment 3 with *Diplostomum* (endoparasite) for 3 days.

Type	Parasite	Response variable	Predictor variable	Family	n	F	df	p
Natural (ectoparasite)	*Gyrodactlylus*	Total count	Plate morph	Negative binomial	208	4.61	2, 206	**0.010**
			PPARA genotype			3.21	2, 204	**0.042**
			Sex			6.21	1, 203	**0.013**
Natural (endo-parasite)	*Diplostomum*	Prevalence	Plate morph	Binomial	208	5.43	2, 206	**0.005**
			WNT7B genotype			4.08	2, 204	**0.018**
Experiment 1 (ectoparasite)	*Gyrodactylus*	AUC	Sex	Negative binomial	46	6.04	1, 45	**0.018**
Experiment 2 (ectoparasite)	*Gyrodactylus*	AUC	Plate morph	Negative binomial	100	6.41	2, 98	**0.002**
			Eda genotype			5.35	2, 96	**0.006**
			Sex			12.3	1, 95	**0.000**
		SGR	PPARA genotype	Gaussian	108	5.48	2, 106	**0.005**
			Group (Control/Exposed)			20.17	1, 105	**0.000**
			Initial Length			9.10	1, 104	**0.003**
			Plate morph			3.41	2, 102	**0.04**
			Initial Length: Plate morph			3.58	2, 100	**0.031**
Experiment 3 (endoparasite)	*Diplostomum*	Total count	Eda genotype	Negative binomial	90	4.46	2, 88	**0.014**
			WNT7B genotype			2.96	2, 86	0.057**.**

(Significance at p<0.05 is denoted in bold)

3.2 Experimental infection dynamics of ectoparasite and endoparasite

3.2.1 Ectoparasite infection dynamics (*G. gasterostei* experiment)

The maximum abundance (peak) and AUC (sum of count) of *Gyrodactylus* were strongly positively correlated (r = 0.95, p = 0.001) in both ectoparasite experiments, I therefore report only results for AUC. In the same way, I also only report results up to 35th day parasite counts, in order to include all samples in experiment 2.

In Experiment 1, only sex of fish had significant association with *Gyrodactylus* abundance (AUC) with higher infection burden in male fish than females ($F_{1, 45}$ = 6.04, p = 0.01) (Table 1). *G. gasterostei* tended to be most abundant on completely plated (CM) fish and had the lowest abundance on partially plated (PM) fish ($F_{2, 44}$ = 0.19, p = 0.83). This mirrored the pattern across Eda genotypes (CC>LL>CL) ($F_{2, 44}$ = 0.03, p = 0.97). These patterns were not significant but are reported here because of their similarity to those in the larger experiment 2 (see below).

In Experiment 2 the GLM modelling revealed that plate morph, Eda genotype and sex of fish had significant effects on total abundance of *Gyrodactylus* (Table 1). The highest *Gyrodactylus* abundance (AUC) was observed in completely plated fish and lowest in the partially plated fish (Fig. 4A). For Eda genotype the highest counts were in the homozygous marine genotype, consistent with the pattern for plate morph, but the lowest counts were in the heterozygous genotype (Fig. 4B). Male fish experienced higher abundance of *Gyrodactylus* than female fish (Fig. 4C). There were no significant differences among genotypes of NLRC5, WNT7B or PPARA in experiment 2.

A.

B.

C.
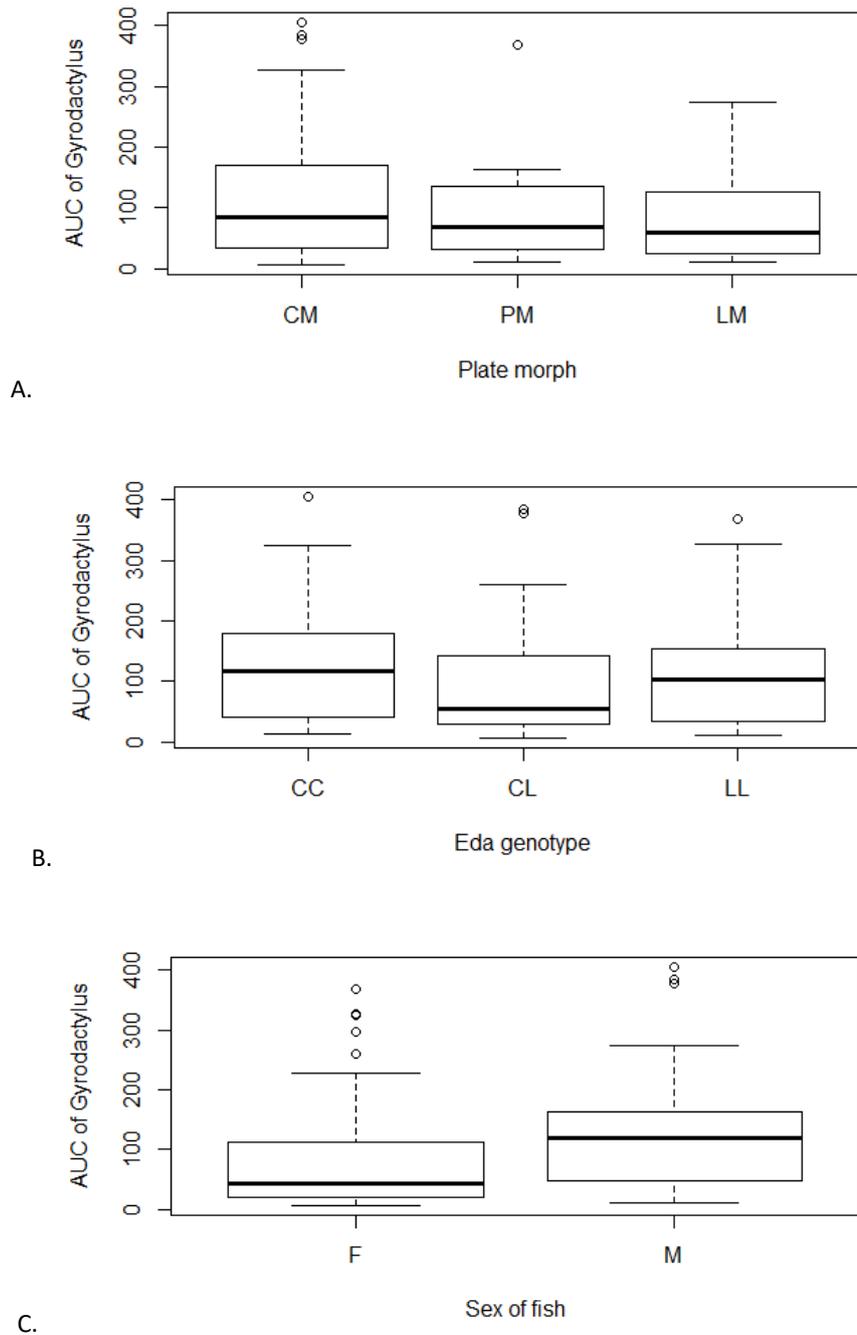
Fig. 4 A. AUC of *G. gasterostei* in experiment 2 shows different patterns of infection among three plate morphs (CM, PM, LM) and B. Eda genotype (CC, CL, LL) with significantly higher abundance in completely-plated fish and homozygous marine genotype of Eda locus in artificially-infected stickleback from an F2 cross between anadromous, completely-plated and freshwater, low-plated parental populations. C. Infected stickleback show significantly higher parasite burden in male fish than females.

The average growth rate of control fish was higher than those of the experimentally-infected fish ($F_{1, 105}$=20.17, p< 0.001). The overall growth rate (SGR) of fish in experiment 2 was greater in fish which began the experiment smaller (Fig. 5). This pattern was significantly more pronounced in partially plated fish (PM) than completely (CM) or low plated (LM) fish through out the experiment ($F_{2, 100}$ = 3.58, p< 0.05).
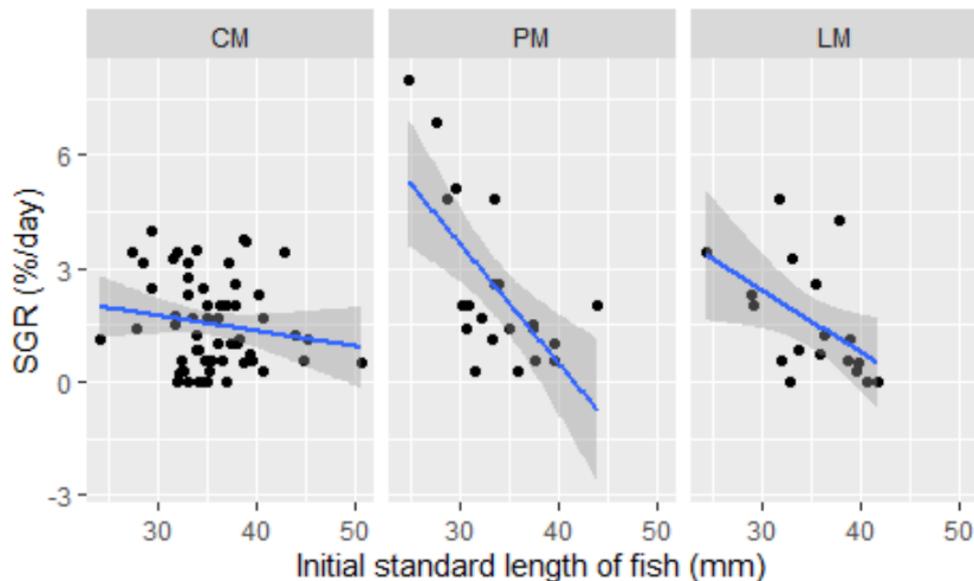
Fig. 5 Specific growth rate (%/day) shows variation among three plate morphs (CM, PM, LM) having higher growth in fish that were smaller size (standard length, SL) at the start of the experiment, especially in partially-plated fish (PM) in experiment 2.

As a result of the more rapid growth of smaller fish, SGR of partially plated fish (PM) was significantly greater (2.49 ± 0.47) than that of completely plated (CM) fish (1.52 ± 0.15, $F_{2, 102}$ = 3.41, p< 0.05; Fig. 6A). In contrast, fish that were heterozygous (PR) at the PPARA locus grew more slowly (1.41 ± 0.16) than the homozygous marine genotype (2.32 ± 0.43) ($F_{2, 106}$ = 5.48, p< 0.01; Fig. 6B). The genotype of Eda, WNT7B and NLRC5 gene showed no significant association with the growth rate of fish. There were no significance differences in the SGR between sexes.

A.

B.
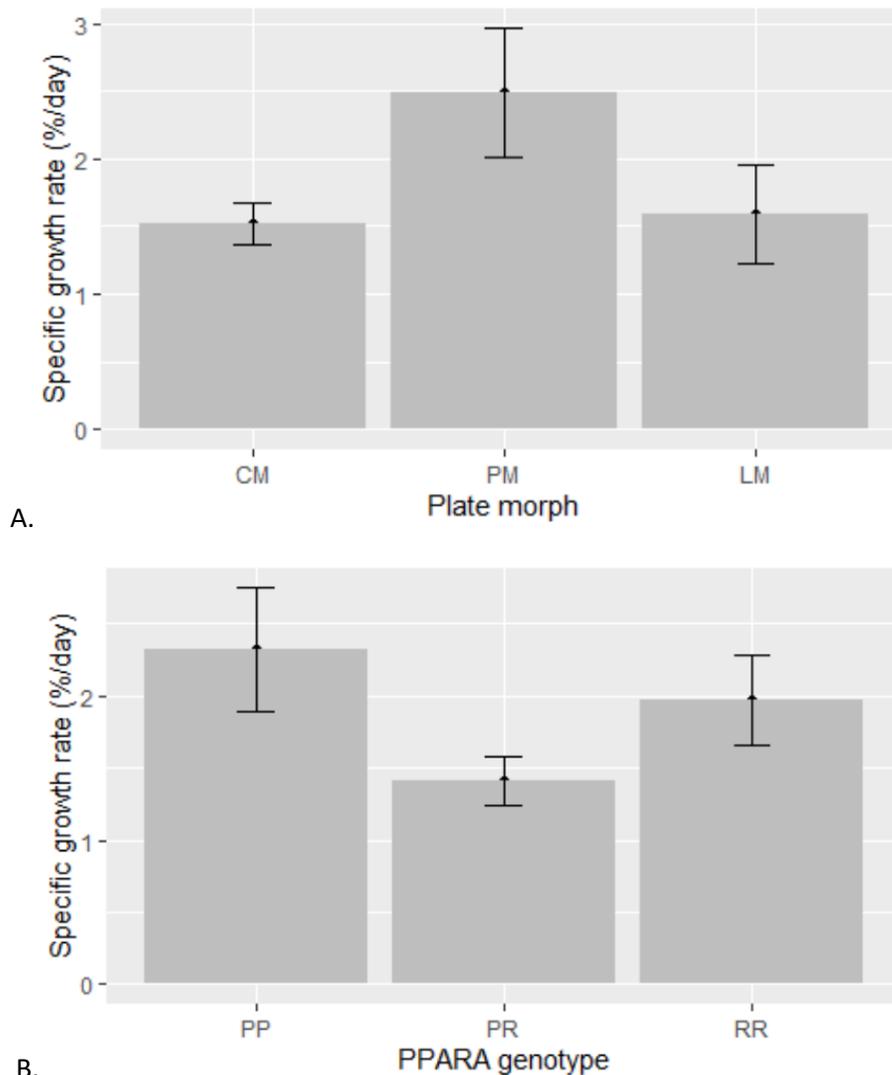
Fig. 6 A. Differences in the mean specific growth rate (%/day) (± S.E.) of *Gyrodactylus*-infected fish in experiment 2 shows highest growth rate in partially-plated (PM) fish among three plate morphs (CM, PM, LM) and B. the opposite pattern in the genotypes of PPARA locus with lowest growth rate in the heterozygous genotype (PR). The populations used was a pure anadromous by freshwater cross - Obse and Chru (North Uist, Scotland).

3.2.2 Endoparasite infection dynamics (*D. pseudospathecum* experiment)

In Experiment 3, F1 admixed fish from the Hosta hybrid zone (N = 90) infected with the endoparasite (*D. pseudospathecum*) showed different patterns of association than in the ectoparasite experiments for genotypes of four genes (Eda, PPARA, WNT7B and NLRC5) and the plate morph and sex of fish.

The abundance of metacercaria of *Diplostomum* observed in the lens of the fish showed a common pattern for all both Eda and WNT7B genotypes with highest counts in homozygous marine genotypes for both genes, intermediate in heterozygous and lowest in homozygous freshwater form (Fig. 7A-B). The differences between genotypes were

significant only for the EDA locus (Table 1). There pattern for the WNT7B locus was not quite formally significant ($F_{2, 88}$ = 2.96, p = 0.06; Fig. 7B). There were no significant differences among the plate morphs or between sexes.

A.

B.
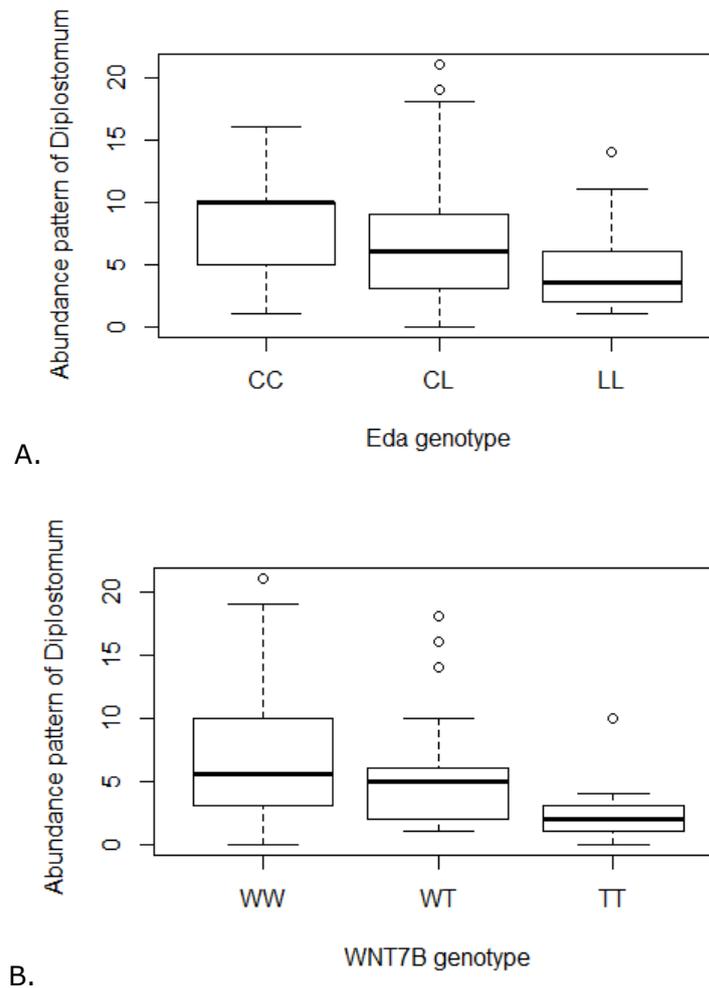
Fig. 7 *D. pseudospathecum* infected stickleback during the experimental period (experiment 3) showing similar pattern among genotypes with higher parasite burden in the homozygous marine genotypes A. CC genotype of Eda locus and B. WW genotype of WNT7B locus.

153

3.3 Pattern of natural and experimental infection dynamics of ectoparasite and endoparasite

Experimental infection patterns were similar to natural infection dynamics in relation to plate morph and Eda genotype in both ectoparasite and endoparasite, but patterns differed for other genes.

The highest infection level of *Gyrodactylus* and *Diplostomum* in both experimental and natural environments were observed in completely plated phenotype and CC genotype of Eda but were lowest in partially plated CL genotype for the ectoparasite and low plated LL genotype for the endoparasite, respectively. The mean abundance of the ectoparasite (*Gyrodactylus*) was significantly higher under experimental conditions than the natural abundance (W = 18613, p< 0.001; Fig. 8A). The same pattern was observed in the endoparasite (*Diplostomum*) infection burden between natural and experimental condition (W = 18697, p < 0.001; Fig. 8B).

A.

B.
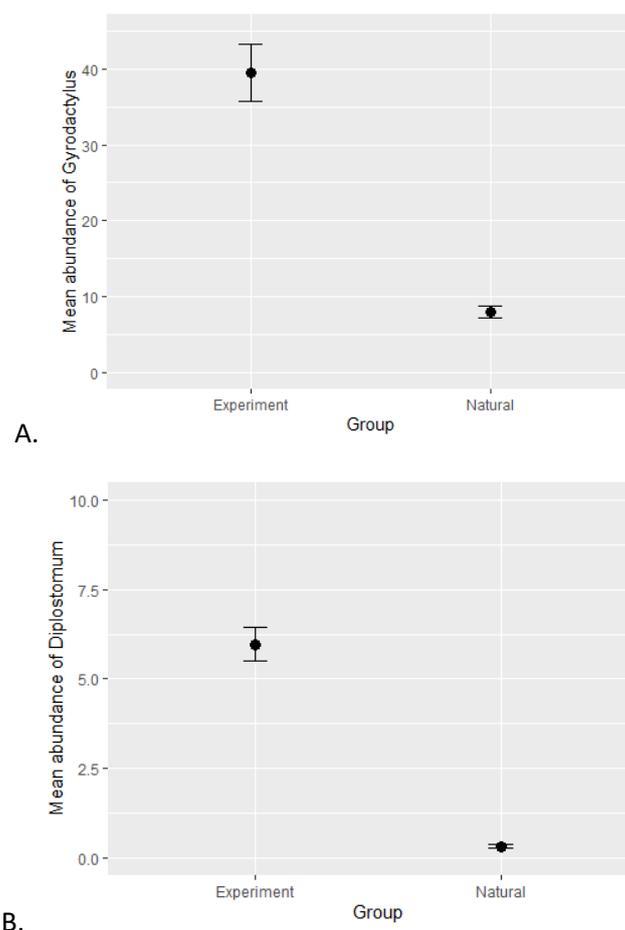
Fig. 8 A. Mean abundance (± S.E.) of *Gyrodactylus* and B. *Diplostomum* infection burden shows comparatively higher abundance in experimental stickleback than in natural infections.

4. Discussion

Adaptive radiation of three-spined stickleback from marine to freshwater represents an excellent system to understand the role of parasite-mediated selection in population divergence, speciation and evolution (Konijnendijk et al., 2013; El Nagar and MacColl, 2016). Previous studies of local adaptation of host-parasite coevolution revealed divergence in resistance to parasites in populations of North Uist stickleback (Obse and Chru) along with genetic associations with immune system function (De Roij et al., 2012; El Nagar and MacColl, 2016; Mahmud et al., 2017; Robertson et al., 2017). The present study was designed to investigate the effects of genotypically-distinct loci on abundance of two macro-parasites (*Gyrodactylus* sp. and *Diplostomum* sp.) which may contribute to local adaptation between marine and freshwater populations. Hence, three SNP assays were developed for three physiologically important genes (PPARA, WNT7B and NLRC5), with fixed differences between marine and freshwater ecotypes, and successfully genotyped for parasite-infected F1 and F2 generation hybrid stickleback. In addition, three distinct morpho-types of the hybrids (completely plated, partially plated and low plated) were also compared with corresponding Eda genotypes while controlling for background genetic variation and maternal effects (Robertson et al., 2017). This study explored for the first time the relationship between parasite susceptibility and genotypes of four candidate genes both in natural and experimental conditions using one ectoparasite (*G. gasterostei*) and one endoparasite (*D. pseudospathecum*).

The infection pattern of ectoparasites in natural and experimental conditions exhibited substantial variation in relation to plate morph phenotype and genotypes of selected candidate loci. Abundance of *Gyrodactylus* sp. in naturally-occurring hybrid stickleback showed a significant association with armour plate phenotype which was also consistent with experimental infections with this ectoparasite.

In both wild and experimental fish those with the completely plated phenotype experienced the highest parasite burden while partially plated had the lowest. This suggests either that *Gyrodactylus* finds it easier to attach to plated fish or that fish with the Eda C allele are somehow more susceptible. This is opposite to the pattern found by Robertson et al., (2017) who reported completely plated fish infected with a lower number of *G. arcuatus* than low plated fish in a common garden experiment carried out in saltwater. This suggests that the association between plate phenotype and susceptibility to *Gyrodactylus* infection might be mediated by water salinity or partly confounded by fish populations. Completely plated sticklebacks, which are ubiquitous in saltwater, may be stressed in freshwater, or struggle to maintain growth, consistent with

the pattern of SGR reported here. Generally, *Gyrodactylus* sp. are common in marine and less frequent in freshwater stickleback (El Nagar and MacColl, 2016) and salt water populations of North Uist are heavily infected with *G. arcuatus* while it was absent from some freshwater populations such as CHRU (De Roij et al., 2012; Mahmud et al., 2017). Here, the *G. gasterostei* were collected from freshwater stickleback and the infection experiment was conducted in freshwater with F2 hybrids. Thus, it is also possible that completely plated fish have lower resistance to freshwater *G. gasterostei* (MacColl and Chapman, 2010; El Nagar and MacColl, 2016). Lower burden of parasites in hybrids than parental forms were also evident in other animals like freshwater snail, *Melanopsis* spp. (Guttel and Ben-Ami, 2014) and house mice, *Mus musculus* (Baird et al., 2012) in response to local environmental conditions.

Eda genotype showed significant association with *Gyrodactylus* abundance in experimental conditions (experiment 2) which indicated its genetic effect on the ectoparasite burden of stickleback. Infection pattern of Eda locus mirrored the plate morph pattern with the highest counts in the homozygous marine genotype and the lowest in the heterozygous genotype. This might be due to its direct controls on bony plate architecture (Colosimo et al., 2004). The Eda haplotype block contains important immune genes (Baff, Garp and Dusp1) (Colosimo et al., 2005; Jones et al., 2012a; Jones et al., 2012b), and it has been speculated that Eda might show strong effects on parasite resistance due to their pleiotropy or linkage. However, the pattern was not evident in the naturally-infected stickleback of this study which may suggest the direct effect of plate number variation on ectoparasite infection is more important than the immunological effect of the Eda locus.

In experimental conditions (experiment 2), growth rate of uninfected, control fish was faster than that of the exposed fish, consistent with previous work (Robertson, 2016) and demonstrating that *Gyrodactylus* infection does have fitness consequences. Plate morph had a significant association with the growth of *Gyrodactylus*-infected fish. Growth was higher in fish which began the experiment smaller, particularly in partially plated fish. As a result, partially plated fish had higher growth rate than the completely and low plated fish which might due to lower parasite burden on the hybrid fish. This pattern is inconsistent with a previous study where *Gyrodactylus*-infected completely plated stickleback showed higher growth rates in saltwater (Robertson et al., 2017). Uninfected, low plated stickleback grew faster in freshwater (Marchinko and Schluter, 2007; Barrett et al., 2009a). This suggests that growth of fish can be influenced either by environmental factors such as salinity and parasites or by the genetic background of the F2 generation of exposed fish.

156

PPARA locus showed strong association with *Gyrodactylus* abundance in natural condition which indicated its genetic effect on the ectoparasite burden of stickleback. The higher abundance of this parasite in the homozygous freshwater genotype of PPARA indicated that both *Gyrodactylus* and RR genotype (homozygous freshwater) might be more common in the loch fish. PPARA genotype also showed significant association with growth of stickleback under experimental condition. Heterozygotes grew more slowly, the opposite pattern to plate morph. This might be caused by its regulatory role in the metabolic process and development of fish (Wafer et al., 2017). Outbreeding depression and fitness reduction of the F2 generation of comparatively large size anadromous stickleback and small size freshwater resident could be another reason.

The endoparasite, *D. pseudospathaceum* specific to the lens of fish, showed comparatively different patterns of infection for phenotype (plate morph) and genotypes of candidate loci than the ectoparasite burden both in wild and experimental conditions. The prevalence of *Diplostomum* recorded from the stickleback in the freshwater Hosta stream had a significant association with plate morph, with highest prevalence in the completely plated and lowest in the low plated fish. This result echoes several studies of stickleback which have demonstrated the high susceptibility of saltwater populations to *Diplostomum* infection in freshwater environments (Kalbe and Kurtz 2006; MacColl and Chapman, 2010; De Roij, 2011; El Nagar and MacColl, 2016).

Eda genotype was significantly associated with abundance of the endoparasite *Diplostomum* in artificial infections of admixed fish from the Hosta hybrid population. This may suggest some link between Eda and the host immune response of host such that this endoparasite may represent an additional source of selection on this adaptive locus. Previous work revealed an association between the Eda haplotype block and the expression pattern of key immune system genes in *Gyrodactylus*-infected stickleback (Robertson et al., 2017) but this result is the first to demonstrate its association with infection by the endoparasite (*Diplostomum*). The highest burden of the parasite was observed in fish that were homozygous for the complete allele (CC) and lowest in homozygous low allele fish (LL) indicating a similar effect on fitness as with the plate morph. This pattern was expected for this freshwater parasite experiment since homozygous complete fish are the ancestral form found in the marine environment and may be susceptible in the freshwater environment.

The association of WNT7B locus genotypes with endoparasite burdens was strong in wild, naturally-infected fish but weak in experimentally-infected stickleback, however the

pattern of association was different in the two groups. In wild fish, the prevalence of *Diplostomum* was higher in fish that were heterozygous at the WNT7B locus and lowest in the homozygous marine form. In experimentally-infected fish, the abundance of *Diplostomum* was highest in fish that were homozygous for the marine allele and lowest in those that were homozygous for the freshwater allele. The low prevalence of *Diplostomum* or the effect of extensive admixture of parental forms in the Hosta hybrid zone might be the reason for different patterns in wild condition and experimental infections. The parasite may act as a selective agent on this locus due to its role in developmental signalling of kidney function which is known to contribute to long-term adaptation of freshwater fish by producing copious hypotonic urine compared to marine fish (Marshall and Grosell, 2006; Yu et al., 2009; Jones et al., 2012b).

Sex of fish is also an important trait to understand the role of selection in adaption, especially in mate choice and reproductive isolation. Male fish were more susceptible to parasite infection than females, harbouring significantly higher ectoparasite (*Gyrodactylus* spp.) burden both in wild and experimental condition. Reimchen and Nosil (2001) found higher abundance in males of a different trematode parasite. It is common in vertebrates for males to have higher susceptibility to parasite infection than females, possibly as a result of costs of mating behaviour (Hamilton and Zuk, 1982; Folstad and Karter, 1992). Sex had no significant impact on *Diplostomum* infection in either wild or experimental conditions.

Parasite abundances were significantly higher in the experiment than natural infections for both ectoparasite and endoparasites which is analogous to the finding for experimental infections of North Uist stickleback with another parasite, *Schistocephalus solidus* (De Roij, 2011). This pattern probably arises because experimentally fish are being exposed for the first time and will have no adaptive immunity, whereas those in the wild are probably being infected repeatedly through their lifetimes and may have some acquired immunity that controls the infections.

The results reported here are consistent with the idea that major adaptive loci in stickleback, as indicated by their consistent, widely different allele frequencies in marine and freshwater populations, may be experiencing selection due to parasite infections. Eda genotype was constantly associated with both types of parasites in experimental condition which supports the idea that parasites might cause additional selection on the Eda locus. WNT7B genotype was also associated with endoparasite abundance in both natural and experimental conditions which provides evidence of a new genetic correlation with parasite infection.

The overall result of the present study suggests a role for parasites in ecological selection that may be separate from the obvious factor of salinity. The association of PPARA genotype with growth of stickleback requires further research to understand the mechanism underlying the observed variation, especially in non-infected fish. This study has shown that important genetic regions that exhibit adaptive divergence may do so partially in response to parasite infection. Furthermore, other genes with allele frequency differences needs to be identified to understand the genetic basis of susceptibility to infection and local adaptation.

References

Baird, S.J., Ribas, A., Macholán, M., Albrecht, T., Piálek, J. and Goüy de Bellocq, J., 2012. Where are the wormy mice? A re-examination of hybrid parasitism in the European house mouse hybrid zone. *Evolution: International Journal of Organic Evolution*, *66*(9), pp.2757-2772.

Bakke, T.A., Cable, J. and Harris, P.D., 2007. The biology of gyrodactylid monogeneans: the "Russian-doll killers". *Advances in parasitology*, *64*, pp.161-460.

Barber, I., 2005. Parasites grow larger in faster growing fish hosts. *International journal for parasitology*, *35*(2), pp.137-143.

Barber, I., 2007. Host-parasite interactions of the three-spined stickleback. In: *Biology of the Three-spined Stickleback*, eds. Östlund-Nilsson, S., Mayer, I. and Huntingford, F.A., CRC press. pp.271-317.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2009a. Environment specific pleiotropy facilitates divergence at the Ectodysplasin locus in three-spine stickleback. *Evolution: International Journal of Organic Evolution*, *63*(11), pp.2831-2837.

Barrett, R.D., Vines, T.H., Bystriansky, J.S. and Schulte, P.M., 2009b. Should I stay or should I go? The Ectodysplasin locus is associated with behavioural differences in three-spine stickleback. *Biology Letters*, *5*(6), pp.788-791.

Bell, M.A. and Foster, S.A. eds., 1994. *The evolutionary biology of the three-spine stickleback*. Oxford University Press.

Buerkle, C.A. and Lexer, C., 2008. Admixture as the basis for genetic mapping. *Trends in ecology & evolution*, *23*(12), pp.686-694.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J. and Myers, R.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *science*, *327*(5963), pp.302-305.

Chappell, L.H., 1995. The biology of diplostomatid eye flukes of fishes. *Journal of Helminthology*, *69*(2), pp.97-101.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. and Kingsley, D.M., 2004. The genetic architecture of parallel armour plate reduction in three spine sticklebacks. *PLoS biology*, *2*(5), p.e109.

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.

De Roij, J. 2011. *PhD Thesis - Spatial variation in host-parasite interactions in the Three-spined stickleback*, University of Nottingham, U.K.

De Roij, J. and MacColl, A.D., 2012. Consistent differences in macroparasite community composition among populations of three-spined sticklebacks, Gasterosteus aculeatus L. *Parasitology*, *139*(11), p.1478.

De Roij, J., Harris, P.D. and MacColl, A.D., 2011. Divergent resistance to a monogenean flatworm among three-spined stickleback populations. *Functional Ecology*, *25*(1), pp.217-226.

Eizaguirre, C. and Lenz, T.L., 2010. Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of fish biology*, *77*(9), pp.2023-2047.

Eizaguirre, C., Lenz, T.L., Kalbe, M. and Milinski, M., 2012. Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology letters*, *15*(7), pp.723-731.

El Nagar, A. and MacColl, A.D., 2016. Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. In *Proc. R. Soc. B*. 283(1836), p. 20160691. The Royal Society.

Folstad, I. and Karter, A.J., 1992. Parasites, bright males, and the immunocompetence handicap. *The American Naturalist*, *139*(3), pp.603-622.

Forrester, G.E., 1991. Social rank, individual size and group composition as determinants of food consumption by humbug damselfish, Dascyllus aruanus. *Animal Behaviour*, *42*(5), pp.701-711.

Frandsen, F., Malmquist, H.J. and Snorrason, S.S., 1989. Ecological parasitology of polymorphic Arctic charr, Salvelinus alpinus (L.), in Thingvallavatn, Iceland. *Journal of Fish Biology*, *34*(2), pp.281-297.

Goldenberger, D., Perschil, I., Ritzler, M. and Altwegg, M. (1995). A simple universal DNA extraction procedure using sds and proteinase-k is compatible with direct pcr amplification. *Pcr-Methods and Applications* **4**(6): 368-370.

Grant, P.R. and Grant, B.R., 2006. Evolution of character displacement in Darwin's finches. *science*, *313*(5784), pp.224-226.

Guttel, Y. and Ben-Ami, F., 2014. The maintenance of hybrids by parasitism in a freshwater snail. *International Journal for Parasitology*, *44*(13), pp.1001-1008.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D. and Berner, D., 2019. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution letters*, *3*(1), pp.28-42.

Hall, T.A., 1999, January. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* (Vol. 41, No. 41, pp. 95-98). [London]: Information Retrieval Ltd., c1979-c2000.

Hamilton, W.D. and Zuk, M., 1982. Heritable true fitness and bright birds: a role for parasites?. *Science*, *218*(4570), pp.384-387.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A., 2010. Population genomics of parallel adaptation in three-spine stickleback using sequenced RAD tags. *PLoS genet*, *6*(2), p.e1000862.

Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T. and Durbin, R., 2002. The Ensembl genome database project. *Nucleic acids research*, *30*(1), pp.38-41.

Janeway, C.A., Travers, P., Walport, M. and Schlomchik, M.J., 2005. Immunobiology: The Immune System in Health and Disease, 6 Edn. Garland Science Publishing, New York.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012a. The genomic basis of adaptive evolution in three-spine sticklebacks. *Nature*, *484*(7392), pp.55-61.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M., Absher, D.M., Myers, R.M., Reimchen, T.E., Deagle, B.E. and Schluter, D., 2012b. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current biology*, *22*(1), pp.83-90.

Kalbe, M. and Kurtz, J., 2006. Local differences in immunocompetence reflect resistance of sticklebacks against the eye fluke Diplostomum pseudospathaceum. *Parasitology*, *132*(1), p.105.

Karvonen, A. and Seehausen, O., 2012. The role of parasitism in adaptive radiations— when might parasites promote and when might they constrain ecological speciation?. *International Journal of Ecology*, *2012*.

Kawakami, T. and Butlin, R.K., 2001. Hybrid zones*In*: eLS. John Wiley, New York. http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.

Kingsley, D.M., Peichel, C.L., Ostlund-Nilsson, S. and Mayer, I., 2007. The molecular genetics of evolutionary change in sticklebacks. *Biology of the Three-spined Stickleback*, pp.41-81.

Knudsen, R., Amundsen, P.A. and Klemetsen, A., 2003. Inter-and intra-morph patterns in helminth communities of sympatric whitefish morphs. *Journal of Fish Biology*, *62*(4), pp.847-859.

Knudsen, R., Kristoffersen, R. and Amundsen, P.A., 1997. Parasite communities in two sympatric morphs of Arctic charr, Salvelinus alpinus (L.), in northern Norway. *Canadian Journal of Zoology*, *75*(12), pp.2003-2009.

Komar, A.A., 2009. Single nucleotide polymorphisms. *Methods in Molecular Biology*, *578*.

Konijnendijk, N., Raeymaekers, J.A., Vandeuren, S., Jacquemin, L. and Volckaert, F.A., 2013. Testing for local adaptation in the Gasterosteus–Gyrodactylus host–parasite system. *Evolutionary Ecology Research*, *15*(4), pp.489-502.

Kurtz, J., Kalbe, M., Aeschlimann, P.B., Häberli, M.A., Wegner, K.M., Reusch, T.B. and Milinski, M., 2004. Major histocompatibility complex diversity influences parasite resistance and innate immunity in sticklebacks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1535), pp.197-204.

Little, T.J., 2002. The evolutionary significance of parasitism: do parasite-driven genetic dynamics occur ex silico?. *Journal of Evolutionary Biology*, *15*(1), pp.1-9.

Maan, M.E., Van Rooijen, A.M., Van Alphen, J.J. and Seehausen, O.L.E., 2008. Parasite-mediated sexual selection and species divergence in Lake Victoria cichlid fish. *Biological Journal of the Linnean Society*, *94*(1), pp.53-60.

MacColl, A.D., 2009. Parasite burdens differ between sympatric three-spined stickleback species. *Ecography*, *32*(1), pp.153-160.

MacColl, A.D., 2011. The ecological causes of evolution. *Trends in Ecology & Evolution*, *26*(10), pp.514-522.

MacColl, A.D. and Chapman, S.M., 2010. Parasites can cause selection against migrants following dispersal between environments. *Functional ecology*, *24*(4), pp.847-856.

Mahmud, M.A., Bradley, J.E. and MacColl, A.D., 2017. Abiotic environmental variation drives virulence evolution in a fish host-parasite geographic mosaic. *Functional Ecology*.pp.1-7.

Malmquist, H.J., 1992. Phenotype-specific feeding behaviour of two arctic charr Salvelinus alpinus morphs. *Oecologia*, *92*(3), pp.354-361.

Marchinko, K.B. and Schluter, D., 2007. Parallel evolution by correlated response: lateral plate reduction in three-spine stickleback. *Evolution: International Journal of Organic Evolution*, *61*(5), pp.1084-1090.*Evolution*, *63*(1), pp.127-138.

Marshall, W. and Grosell, M. 2006. Ion transport, osmoregulation and acid-base balance. *In:* The Physiology of Fishes. 177–230. CRC press.

McCairns, R.S., Bourget, S. and Bernatchez, L., 2011. Putative causes and consequences of MHC variation within and between locally adapted stickleback demes. *Molecular ecology*, *20*(3), pp.486-502.

Miller, C.T., Beleza, S., Pollen, A.A., Schluter, D., Kittles, R.A., Shriver, M.D. and Kingsley, D.M., 2007. cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell*, *131*(6), pp.1179-1189.

Mills, M.G., Greenwood, A.K. and Peichel, C.L., 2014. Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo*, *5*(1), pp.1-10.

Neerincx, A., Lautz, K., Menning, M., Kremmer, E., Zigrino, P., Hösel, M., Büning, H., Schwarzenbacher, R. and Kufer, T.A., 2010. A role for the human nucleotide-binding domain, leucine-rich repeat-containing family member NLRC5 in antiviral responses. *Journal of Biological Chemistry*, *285*(34), pp.26223-26232.

Ostlund-Nilsson, S., Mayer, I. and Huntingford, F.A. eds., 2006. *Biology of the three-spined stickleback*. CRC press.

Owen, S.F., Barber, I. and Hart, P.J.B., 1993. Low level infection by eye fluke, Diplostomum spp., affects the vision of three-spined sticklebacks, Gasterosteus aculeatus. *Journal of Fish Biology*, *42*(5), pp.803-806.

R Core Team, 2020. R: a language and environment for statistical computing. Version 3.6.3 [computer program]. R Foundation for Statistical Computing, Vienna, Austria.

Rauch, G., Kalbe, M. and Reusch, T.B., 2006. Relative importance of MHC and genetic background for parasite load in a field experiment. *Evolutionary Ecology Research*, *8*(2), pp.373-386.

Reimchen, T.E. and Nosil, P., 2001. Ecological causes of sex-biased parasitism in three-spine stickleback. *Biological Journal of the Linnean Society*, *73*(1), pp.51-63.

Robertson, S., 2016. *PhD Thesis - The ecological and evolutionary importance of immune system variation in the three-spined stickleback*, University of Nottingham.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific reports*, *7*, p.42677.

Scharsack, J.P. and Kalbe, M., 2014. Differences in susceptibility and immune responses of three-spined sticklebacks (Gasterosteus aculeatus) from lake and river ecotypes to sequential infections with the eye fluke Diplostomum pseudospathaceum. *Parasites & vectors*, *7*(1), p.109.

Scharsack, J.P., Kalbe, M., Harrod, C. and Rauch, G., 2007. Habitat-specific adaptation of immune responses of stickleback (Gasterosteus aculeatus) lake and river ecotypes. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1617), pp.1523-1532.

Team, R.C., 2020. R: a language and environment for statistical computing. Version 3.6.3 [computer program]. R Foundation for Statistical Computing, Vienna, Austria.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. and Mugue, N.S., 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. *PLoS Genet*, *10*(10), p.e1004696.

Thiel, T., Kota, R., Grosse, I., Stein, N. and Graner, A., 2004. SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Research*, *32*(1), pp. e5-e5.

Tompkins, D.M. and Begon, M., 1999. Parasites can regulate wildlife populations. *Parasitology today*, *15*(8), pp.311-313.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012.Primer3—new capabilities and interfaces. *Nucleic acids research*, *40*(15), pp.e115-e115.

Wafer, R., Tandon, P. and Minchin, J.E., 2017. The role of peroxisome proliferator-activated receptor gamma (PPARG) in adipogenesis: Applying knowledge from the fish aquaculture industry to biomedical research. *Frontiers in endocrinology*, *8*, p.102.

Wegner, K.M., Kalbe, M., Kurtz, J., Reusch, T.B. and Milinski, M., 2003. Parasite selection for immunogenetic optimality. *Science*, *301*(5638), pp.1343-1343.

Wegner, K.M., Kalbe, M., Milinski, M. and Reusch, T.B., 2008. Mortality selection during the 2003 European heat wave in three-spined sticklebacks: effects of parasites and MHC genotype. *BMC Evolutionary Biology*, *8*(1), p.124.

Yu, J., Carroll, T.J., Rajagopal, J., Kobayashi, A., Ren, Q. and McMahon, A.P., 2009. A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. *Development*, *136*(1), pp.161-171.

Chapter 5: Variation in the skin microbiomes of stickleback

Abstract

Skin-associated microbiota and their relationship with environmental gradients are poorly studied for the important ecological and evolutionary model fish, the three-spined stickleback *Gasterosteus aculeatus*. Environmental factors based on water chemistry (e.g. pH, salinity) and genetic influences (fish population) might be responsible for the composition and diversity of microbiomes of this migratory anadromous fish. To address this issue, culture-dependent molecular methods were applied that showed significant effects of environment contrasts such as freshwater-saltwater, high pH-low pH in shaping the skin microbiome communities of stickleback. Overall bacterial load in the skin samples did not reveal remarkable differences. In contrast, qualitative analysis of identified bacterial species through 16sRNA gene sequencing revealed strong variation in their relative frequency among different groups of sticklebacks (freshwater-saltwater, high pH-low pH, resident and anadromous of saltwater and freshwater). The diversity of bacteria was evident also at genus level among different groups based on salinity, pH and ecotypes. Most of the bacterial species (e.g. *Aeromonas* sp., *Pseudomonas* sp., *Yersinia* sp. etc.) belonged to the phylum Proteobacteria which are mostly gram-negative and pathogenic to fish. The gram-negative, pathogenic *Flavobacterium* sp. was also found from the phylum Bacteroidetes. There were some representative genera of gram-positive bacteria identified from other phyla such as *Exiguobacterium* and *Carnobacterium* under class Bacilli of phylum Firmicutes and *Frigoribacterium* under phylum Actinobacteria. The most dominant and common taxa were four overlapping genera (*Aeromonas*, *Pseudomonas*, *Shewanella* and *Psychrobacter*) found in almost all groups of sticklebacks. There were some habitat-specific species identified particularly in freshwater (*Providencia alcalifaciens*, *Stenotrophomonas rhizophila*, *Klebsiella aerogenes, K. pneumoniae, Pseudomonas gessardii, P. poae and P. syringae etc.*) and in saltwater stickleback (*Halomonas* sp., *Marinomonas* sp., *Comamonas jiangduensis, Pseudomonas psychrophile, P. brenneri* etc.*) reflecting the effects of ecological factors. Besides, the relative abundance and diversity of bacterial species composition in the stickleback of alkaline habitats was significantly higher than in acidic habitats indicating the environmental influence on skin microbiome. However, the higher diversity of bacterial communities in the resident fish supported the effect of transition from high salinity sea water to the freshwater may cause physiological changes associated with migration of anadromous fish.

1. Introduction

Generally, microbiota is an assemblage of bacteria, archaea, fungi and viruses present in a defined environment (Marchesi and Ravel, 2015). The term 'microbiome' was first introduced to signify the ecological community of commensal, symbiotic, and pathogenic microorganisms occurring in the human body (Lederberg and McCray, 2001). Animal-associated microbiota form complex communities and have crucial beneficial and pathogenic consequences that affect their host fitness. For instance, fish-associated microbiomes play a pivotal role in the host bodies for their pathogenic (causative agents of fish diseases) and commensal activities (e.g. probiotics). Fish possess bacterial populations on or in their skin, gills, digestive tract, light-emitting organs as well as internal organs (kidney, liver and spleen) which perform important roles in movement, gastrointestinal activities, immune system, nutrition, drag reduction, disease and light emission (Austin, 2006). Probiotic microbes (e.g., *Bacillus* sp., *Lactobacillus* sp., *Enterococcus* sp., *Carnobacterium* sp.) beneficially affect the host by improving its intestinal balance and also increasing the fitness of the host, mainly through the exclusion of opportunistic pathogens (Balcazar et al*.*, 2006; Cruz et al., 2012).

Commensal bacteria associated with fish mucosal surfaces (skin) are beneficial for fish health as they can compete with pathogens for space and nutrients (Larsen et al., 2013). Some microbiota plays a role in the first line of defence (colonization resistance) against opportunist pathogens, relying on the equilibrium between the relative abundance and diversity of endogenous bacteria (Sylvain et al., 2016). In contrast, multiple bacterial genera including members of *Vibrio, Streptococcus, Aeromonas, Flavobacterium, Photobacterium, Pasteurella, Tenacibacterium, Pseudomonas, Lactococcus, Edwarsiella, Yersinia, Renibacterium* and *Mycobacterium* have been reported as pathogens of many teleost fish and may cause disease (Austin and Austin, 2007; Larsen et al., 2015). Fish skin bacteria are also associated with the production of friction-reducing polymers which are important for the movement of fish through the water column by smoothing the surface and decreasing frictional drag (Sar and Rosenberg, 1989). Bacteria of fish also affect the development of certain immune responses through gut-associated lymphoid tissue (GALT) which can differentiate potentially pathogenic and commensal microbiota and determine whether tolerance or an immune response should be induced within the host body (Perez et al., 2010). Despite several studies of the effect of gut microbiome functions on fish health and immunity, very little is known about skin microbiome communities and their role in fish (Sylvain et al., 2016). With increasing outbreaks of fish disease, especially cutaneous disease and high mortality of economically-important fish (e.g. Atlantic salmon *Salmo salar*, trout *S. trutta* and carp *Cyprinus* sp.), studies relating to fish microbiome (bacteria) including

their composition, distribution, commensal and pathogenic roleare inevitably important to address different challenges in wild fisheries, farm management and conservation of fish.

As aquatic poikilothermic animals, fish are continuously exposed to microorganisms present in the surrounding aquatic environment (Kar et al., 2008). Skin is a large fish organ having direct contact with water and acts as a multifunctional organ by contributing to protection, communication, sensory perception, locomotion, respiration, ion regulation, excretion, and thermal regulation (Elliot, 2011). The skin of teleosts secretes mucus which exhibits immune functions and acts as a primary protective barrier against invading microorganisms such as pathogens (bacteria, virus), parasites, etc. (Salinas et al., 2011; Gomez et al., 2013). Skin microbiome diversity is comparatively higher than for other parts of the fish body such as gills and gut (Lowrey et al., 2015; Chiarello et al., 2015). Bacteria associated with fish mucosal surfaces (skin) are found to be species-specific and Proteobacteria, Actinobacteria, Bacteroidetes Cyanobacteria, Verrucomicrobia and Firmicutes have been reported as the most abundant bacterial phyla of fish skin (Larsen et al., 2013; Boutin et al., 2013; Larsen et al., 2015; Lowrey et al., 2015; Lokesh and Kiron, 2016).

To date, skin and gill associated microbiota and their relationship with environmental gradients are poorly studied (Leonard et al., 2014; Lokesh and Kiron, 2016). Some studies revealed that the structure of the bacterial communities of skin and gills in fish are simpler than those of mammals, reflecting the impact of the surrounding water (Horsley, 1977; Sylvain et al., 2016). On the other hand, studies on Atlantic salmon *S. salar* (Cipriano and Dove, 2011; Landeira-Dabarca et al., 2013), Whiting, *Merlangius merlangus* (Smith et al., 2007), Gibel carp, *Carassius auratus gibelio* and Bluntnose black bream, *Megalobrama amblycephala* (Wang et al., 2010) and Lionfish, *Pterois* sp. (Stevens and Olson, 2013) indicated that feeding regime, pathogen invasion, fish culture practices, indigenous bacterial assemblage and host specificity influence microbiome compositions of skin and gills of fishes rather than surrounding water. The differences in microbiome assembly between pre - (pelagic larvae phase) and post - settlement (benthic juvenile and adult) damsel fish (Pomacentridae) and cardinal fish (Apogonidae) families are likely driven by a combination of physiological changes associated with development (Parris et al., 2016). The genetic characterization of the host may also be responsible for the symbiotic microbial community assemblage (Kenny and Valdivia, 2009). Recent studies on two different host types (aquaculture and a wild strain) of Atlantic salmon fry found significant influence of host genotype on both skin and gut microbiota (Øygarden, 2017). In addition, bacterial flora of fish can be greatly influenced by differences across latitude and temperature gradients (Cahill, 1990), salinity (Horsley,

1973; Lozupone and Knight, 2007), pH (Fierer and Jackson, 2006; Chu et al., 2010), seasonality (Gilbert et al., 2009, 2012) and ecological interactions (Steele et al., 2011).

The physiological, immunological and pathological perspectives at the transition phase of anadromous or catadromous fish in marine or freshwater have been well studied (Folmar and Dickhoff, 1980; Barton, 2002; Huang et al., 2015) but comprehensive knowledge is still lacking on the role of the environment in shaping the microbiome communities (e.g. bacterial communities) during the transition. Fishes may encounter a new spectrum of microbes during their movement between marine and freshwater environments and it is likely that this movement or migratory behaviour may significantly affect their microbiome composition. For example, the skin-associated bacterial communities of Atlantic salmon have been reported to reshape (increase in species richness and phylogenetic diversity) under laboratory conditions while transferring from freshwater to seawater (Lokesh and Kiron, 2016). Moreover, some studies indicate that marine–freshwater transitions have been infrequent events during the diversification of microbes and that most of these transitions happened a long time ago in evolutionary terms (Logares et al., 2009). For example, marine and freshwater microbes in the gills of South American wild silversides, *Odontesthes argentinensis* and *O. bonariensis* (Atheriniformes) were substantially different and phylogenetically-distinct from each other (Hughes et al., 2017). In addition to the surrounding environment, the abundance of surface bacterial flora on fish can be affected by pathogen invasion, parasite infestations in relation to the physiological response of host (Cipriano and Dove, 2011). However, factors responsible for the composition and distribution of microbiomes are still addressable issues for migratory fishes like anadromous three-spined stickleback *G. aculeatus*, lamprey *Priscomyzon* sp., shad *Hilsa* sp. etc. Therefore, multiple factors combining surrounding water chemistry (e.g. pH, salinity) and genetic influences (fish population) should be considered to investigate the microbiome composition of fish for eco-evolutionary purposes and the overall management and conservation of fish.

From a multitude of studies and analyses in the past few decades, we have learned that the diversity of microbes is even greater than imagined. Different techniques from conventional culture techniques to advanced molecular methods have been used to characterise fish microbial communities (Austin, 2006). Most of the culture-based studies of fish microbiota were conducted using selective or non-selective isolation media, followed by a series of morphological and biochemical assays (Horsley, 1973; Trust and Sparrow, 1974; Sakata et al., 1981; Kamei et al., 1985; Sugita et al., 1988; Cahill, 1990; Montes, 1999). Culture-dependent techniques are used for the study of numbers and higher-level taxonomic composition of the bacterial populations (Austin, 2006) by which it is difficult to determine the precise relationships between environment and fish

microflora (Cahill, 1990). However, this technique is often time consuming and less accurate (Asfie et al., 2003) and leads to a very uncertain picture of the total microbial community (Spanggaard et al., 2000; Nayak, 2010). Therefore, culture-dependent techniques for bacterial community studies may suffer from inconsistencies, low sensitivity and overall, reflect a biased overview of the bacterial diversity (Sevellec et al., 2014) but it can still be useful in combination with other molecular techniques (Wiik et al., 1995; Clarridge, 2004).

The 16S ribosomal RNA (16S rRNA) gene is the most prominent target gene in many molecular methods to establish the classification of organisms. This gene has been used extensively for phylogenetic studies because of its highly conserved nature (Woese and Fox, 1977; Woese et al., 1990). In addition to highly conserved primer binding sites, sequences of the 16S rRNA gene also possess nine hyper-variable regions (V1-V9) that can provide species-specific signature sequences useful for identification of bacteria (Kolbert and Persing, 1999; Pereira et al., 2010). Chakravorty et al., (2007) characterized V1 - V8 in 110 different bacterial species and found that V3 and V6 regions can identify most bacterial genera because of their high diversity compared to V4 and V5 regions. The major molecular techniques targeting the 16S rRNA gene includes fluorescence in situ hybridization (FISH), Terminal restriction fragment length polymorphism (T-RFLP), Amplified ribosomal DNA restriction analysis (ARDRA), Single strand conformational polymorphism (SSCP) and Denaturing-gradient gel electrophoresis (DGGE) (Fukuda et al., 2016). Among these techniques, DGGE was used in the present study in combination with 16s rRNA sequencing to obtain comparative bacterial diversity among different samples of different locations. This approach provides a rapid, cheaper and convenient molecular method for the detection of bacterial species from mixed microbial populations which allows qualitative analysis of their diversity through direct comparision (Maiwore et al., 2009; 2012).

DGGE is the most common fingerprinting method based on the separation of a specific segment of the rRNA gene amplified on a denaturing polyacrylamide gel. It is an electrophoretic technique to identify single base changes in a segment of DNA. Generally, DGGE combines two stages: firstly, PCR product amplification from extracted DNA using 16S rRNA gene universal primers with an additional Guanine-Cytosine (GC) enriched sequence (GC-clump) and secondly, electrophoresis in a polyacrylamide gel containing a linear gradient of DNA denaturant. Separation of PCR products in DGGE (PCR-DGGE) is based on the decrease of the electrophoretic mobility of partially-melted doubled-stranded DNA molecules in polyacrylamide gels containing a linear gradient like formamide and urea at 60°C. Molecules with different sequences will have a different melting points and will stop migrating at various positions in the gel (Muyzer et al.,

1993; Leesing, 2005; Maiwore et al., 2012). The final result is a gel with a pattern of bands which is a visual profile of the most abundant species in the studied microbial community. In addition, specific bands on the gel can be excised and sequenced for subsequent taxonomic identification.

Many studies have been reported to apply this technique for the investigation of bacterial community composition of fish and surface water (Murray et al., 1996; Øvreas et al., 1997; Moeseneder et al., 1999; Riemann et al., 1999; Le Nguyen et al., 2007; Montet et al., 2008; Maiworé et al., 2009; Tatsadjieu et al., 2010). The major advantage of the PCR-DGGE method is that it can be used to analyse both cultivable and non-cultivable, anaerobic and aerobic bacteria and can rapidly detect the diversity and difference of community compositions of bacteria of each sample prior to sequencing (Yang et al., 2001). Despite its broad application, some disadvantages of this technique are related to the handling of polyacrylamide gels, maintenance of optimal denaturing, confirmation of results by cloning or sequencing of particular bands, comparison of patterns across gels and limited resolution of band detection with complex bacterial communities and microbes of low abundance (Muyzer, 1999; Danilo, 2004) may need careful and specialized application. In the present study, 16SrRNA gene sequence analyses with DGGE was applied to study the skin microbiome in different populations of a model fish, the three-spined stickleback including their characterization and composition in different environmental condition.

Three-spined stickleback (*Gasterosteus aculeatus*) (hereafter 'stickleback') is a perfectly suited model organism in eco-evolutionary research due to its wide distribution across different habitat types such as saltwater and freshwater (Scharsack et al., 2016). Previous work on stickleback from the island of North Uist has shown systematic variation between isolated natural populations in natural parasite burdens (De Roij and MacColl, 2012), resistance to infections (De Roij et al., 2011; El Nagar and MacColl, 2016), immune system variation (Robertson et al., 2017) and the virulence of parasites (Mahmud et al., 2017). Few studies have been reported recently on the gut microbiome of stickleback considering environmental, diet, immunity and host genetic divergence (Bolnick et al., 2014; Smith et al., 2015; Milligan-Myhre et al. 2016; Steury et al., 2019; Small et al., 2019). However, the effect of environmental variation such as saltwater and freshwater on the diversity of the skin microbiome of stickleback fish populations has not been studied yet. In this chapter, by considering the small size and slime-less body surface of stickleback, culture-dependent PCR-DGGE with 16sRNA gene (V1-V9 and V3 region) sequencing techniques were executed to perform an initial descriptive survey of the general skin microbiomes of stickleback on North Uist, including their compositions

and distribution in response to variable water chemistry (salinity, pH) and fish ecotypes (anadromous and resident). Therefore, I collected stickleback samples from seven different lochs of North Uist with different water chemistry (freshwater with acidic or alkaline condition and saltwater) to address the following research questions:

1. Are there any obvious effects of major environmental differences?

To address this question, I investigated the composition and diversity of bacterial taxa (up to genus level) between all fish that live permanently in freshwater and all fish that live permanently in saltwater. I was also interested in differences between high pH (alkaline) and low pH (acidic) habitats within freshwater as studies on the effect of lower pH level have shown more significant changes in cutaneous microbiome than intestinal microbiome (Sylvain et al., 2016). Variation in microbiome composition was expected on stickleback skin samples of freshwater and saltwater or between freshwater high pH and low pH populations.

2. Is microbiome more determined by fish genetics or environment?

To address this question, I investigated differences in the frequency and diversity of bacterial genera between resident and anadromous fish, where both have been sampled in freshwater lochs. I also observed the differences in microbiome composition between anadromous fish sampled in saltwater and anadromous fish sampled in freshwater. It is likely that there may be some distinctiveness of microbiome abundance on specific fish populations due to their genetic or physiological response to microbiomes. The skin microbiota of stickleback may also be affected by the ecotypes (anadromous and resident) of fish populations. I expected the differences of bacterial species composition in different populations or ecotypes of stickleback as the transition of anadromous fish has been known to reshape (increase in species richness and diversity) the microbiota associated with the skin (Lokesh and Kiron, 2016).

2. Materials and Methods

2.1 Study area

The present study was designed to collect samples from seven different lochs on North Uist with different water chemistry (freshwater with low pH or high pH and saltwater) and different ecotypes (anadromous and resident) of stickleback fish (Fig. 1). The evolution of high phenotypic diversity of stickleback populations across the North Uist island in the Western Isles of Scotland (Giles, 1983; Magalhaes et al., 2016) offers a unique opportunity to study the variation of the microbiome in different environmental conditions.

Fig.1 Sites for the collection of samples from seven lochs in North Uist, Scotland (A. Map of United Kingdom B. Loch position in North Uist with road map).

2.2 Schematic flow chart of the study design

Three-spined stickleback fish sample collection from seven lochs on North Uist, Scotland

↓

Collection of swabs from the skin by using culture-swab in 2017, preserved at -20^{0}C.

Collection of swabs from the skin by using culture-independent e-swab in 2019, preserved at -20^{0}C. Collection of loch water by filtration of sample water, filter preserved at -20^{0}C.

↓

Culture of e-swab and filtered water samples in Tryptone Soya agar (TSA) medium following the spread plate method for enumeration and isolation of enough bacteria from the slime-less skin surface of stickleback.

↓

Processing of selected e-swab and culture-swab colonies for molecular analysis (bacterial DNA extraction through culture-dependent method).

↓

Identification of individual bacteria from selected colony types of e-swab and culture-swab samples through 16sRNA gene amplification of V1-V9 region and subsequent sequencing and matching through BLAST search within nucleotide database.

↓

Processing of selected culture-swab samples from three representative lochs (Tros, Hosta and Ardh) collected in 2017 for Denaturing Gradient Gel Electrophoresis (DGGE) of V3 region of 16s RNA gene and subsequent sequencing of specific gel segment from Hosta and Ardh for identification of individual bacteria through BLAST search.

↓

Statistical analysis of data to find out variation in the total bacterial load (Total viable bacterial count) collected from e-swab samples and relative abundance as well as diversity of identified bacterial genus from the sequencing of V1-V9 and V3 regions of 16sRNA gene across different groups of sticklebacks on the basis of salinity (freshwater and saltwater), pH (high pH and low pH) and ecotypes (freshwater-resident, saltwater-resident and anadromous).

2.3 Sample collection

Stickleback fish were collected from North Uist during 06-10 May 2017 and 07-16 May 2019. Fish were caught in Gee's Minnow Traps (Gee traps, Dynamic Aqua, Vancouver, Canada), set overnight (approximately 16h) in seven lochs covering freshwater, saltwater, high pH and low pH habitats (Table 1). The water quality parameters temperature, absolute conductivity and pH were recorded from seven sites using a multi-parameter probe (Multi340/set, WTW, Germany) (Table1). Water samples (2 litres each) from each loch were also filtered through disposable sterile filter units with 25-micron mesh (Thermo Scientific Nalgene, USA) and preserved at -20^0C for the analysis of the microbial composition in loch water. From the seven lochs, 130 live fish samples were haphazardly selected and transported immediately to the laboratory in loch water in darkened boxes with continuous aeration. In the laboratory, skin mucus was collected by gently stroking a sterile cotton mini tip swab along the flank (including fins and tail) of each fish in aseptic conditions. Swabs were preserved in semi-solid and liquid Amies transport medium of both culture (97 samples) and e-swab (Copan Liquid Amies Elution Swab, Italy) (33 samples) respectively, at -20^0C for further processing.

2.4 Fish ecotypes

Fish were euthanized with overdose of MS-222 (400 mgL^{-l}) following Schedule 1 techniques according to UK Home Office regulations and finally death was confirmed by destruction of the brain. For the detection of ecotypes (anadromous or resident) of fishes, the standard length, weight and body texture were recorded.

2.5 Processing of swab samples for molecular analysis

Among the collected samples, a subset was processed for culture-dependent study of stickleback skin microbiome. Firstly, 33 e-swab samples from seven different lochs in 2019 were selected to observe the variation of bacterial load based on environmental variation (freshwater-marine water and low pH-high pH) and fish ecotype (anadromous and resident) (Table 1). Secondly, 18 samples collected from three lochs (Tros, Hosta and Ardh) in 2017 representing three different types of habitat were selected for the isolation of bacteria from culture-swab of stickleback skin samples using the DGGE technique (Table 2). Culture bacteria from pure culture was used for the identification of individual bacteria through sequencing of V1-V9 region and V3 region (DGGE) of 16s RNA.

Table 1. Sampling locations, habitat types and physicochemical parameters of seven lochs on North Uist, Scotland (FW = Freshwater, SW = Saltwater, Anad = Anadromous, Resi = Resident) sampled for the enumeration and isolation of bacteria through culture of 33 e-swabs samples collected in 2019.

Sl.	Loch	Location	Habitat	No. of e-swab samples	pH	Temperature (^{0}C)	Conductivity (μS/cm)	Date of sample collection
1.	Scadavay (Scad)	57°35'6"N; 7°14'10"W	FW low pH	4 FW	6.5	13.2	130.9	15.05.2019
2.	Tormasad (Torm)	57°33'45"N; 7°19'1"W	FW low pH	5 FW	7.0	13.8	162.5	13.05.2019
3.	Trosavat (Tros)	57°35'3"N; 7°24'45"W	FW low pH	1 Anad 2 Resi	6.6	16.4	165.9	14.05.2019
4.	Hosta (Host)	57°37'40"N; 7°29'18"W	FW high pH	6 FW	8.5	12.2	432	13.05.2019
5.	Grogary (Grog)	57°36'54"N; 7°30'40"W	FW high pH	6 FW	8.3	14.6	340	07.05.2019
6.	Ard Heisker (ArdH)	57°34'48"N; 7°24'48"W	SW	3 Anad 2 Resi	8.3	13.2	47,400	11.05.2019
7.	Duin(Duin)	57°38'35"N; 7°12'40"W	SW	2 Anad 2 Resi	8.4	13.3	27,460	15.05.2019

2.5.1 Enumeration and isolation of bacteria through culture of e-swabs

Bacteria of stickleback skin samples were cultured in Tryptone Soya agar (TSA) medium (40g/L, Oxoid) following the spread plate method (Sanders, 2012). All e-swab samples preserved at -20^0C were thawed at 4^0C for 24hrs to maintain consistent condition and the liquid medium mixed (1ml in volume) using a vortex for 5 seconds to release the sample from the swab tip. At first, 100µl of e-swab suspension was transferred into a microcentrifuge tube (MCT) for tenfold serial dilution with 900 µl of sterile 0.9% NaCl solutions and then 100µl aliquots from 10^{-5} dilution was inoculated directly onto duplicate TSA plates. The inoculum was spread homogenously through a sterile plastic spreader and incubated under aerobic conditions at 13°C for 96hrs. Enumeration of bacterial load was determined by direct count of bacterial colonies using fluorescence light and expressed as colony forming units per millilitre (cfu/ml) (Collins and Lyne, 1984). Individual bacterial colonies were selected for pure culture based on observable differences in colony morphology (Koneman et al., 1997) and sub-cultured onto new TSA plates again, following the same incubation method. Culture bacteria were then harvested from the plate with Tryptone Soya Broth (TSB) diluted with sterile glycerol (600 µl TSB + 400 µl glycerol) and used as pure culture for individual bacteria identification through sequencing of V1-V9 region of 16s RNA. Bacterial load of water samples collected from seven lochs were also processed in the same method by direct placing of filter paper on the TSA plate and subsequent incubation under aerobic conditions at 13^0C for 96hrs.

2.5.2 Isolation of bacteria from culture-swabs

Swabs selected from 18 fish samples of Tros, Hosta and Ardh collected in 2017 were inoculated directly onto duplicate TSA plates and incubated under anaerobic or aerobic conditions at 15°C for 48hrs (Table 2). Anaerobic conditions were maintained using the culture anaerobic atmosphere generation bag (AnaeroGen Thermo Scientific) in airtight five litre plastic containers. Bacteria were harvested completely from these culture plates using a sterile spreader and collected in approximately 1.5 ml of liquid media (TSB) to create a liquid bacterial suspension. This solution was then diluted with sterile glycerol to a final concentration of 25% glycerol and stored at -20^0C. This solution was used as mixed cultured bacterial solution for DGGE (V3 region) method from twelve fish samples and further pure culture for molecular analysis of 16s RNA from six fish samples (Table 2).

To identify individual bacterial species through sequencing of V1-V9 of 16s RNA, present in the culture-swab, one representative swab from each location and fish type

(anadromous or resident) (samples 200, 210, 252, 264, 322, and 332, Table 2) were sub-cultured onto new TSA plates. 100µl of mixed culture inoculum was distributed by a spreader and incubated as previously described under aerobic and anaerobic conditions. Single bacterial colonies were selected for pure culture based on observable differences in colony morphology and inoculated onto TSA plate again following same incubation method. Culture bacteria were then harvested from the plate with TSB broth and used as a pure culture for individual bacteria identification through sequencing of V1-V9 of 16s RNA. All procedures were performed under aseptic conditions.

Table 2: Sampling of ecotypes for isolation of bacteria from culture-swabs collected from three representative lochs (Tros, Hosta and Ardh) of North Uist, Scotland.

Location	Habitat	Ecotype	No. of samples	Swab Number
Tros	FW Low pH	Anadromous	3	**252**, 254, 256
		Resident	3	**264**, 266, 268
Hosta	FW High pH	Anadromous	3	**332**, 334, 336
		Resident	3	**322**, 326, 328
Ardh	Saltwater	Anadromous	3	**200**, 202, 204
		Resident	3	**210**, 212, 214

* Samples in bold were pure cultured and sequenced to identify individual bacteria through V1-V9 region of 16s RNA.

2.6 Molecular analysis through 16sRNA gene

For molecular analysis of culture-dependent bacteria, two different techniques were followed: a) study of microbial variation through amplification of V1-V9 region of 16s RNA gene from pure culture of e-swab and culture-swab and b) study of microbial variation through DGGE of amplified V3 region of 16s RNA. As a final step, all the sequences collected from these two methods were characterised for qualitative analysis of microbial variation among different groups of sticklebacks.

2.6.1 Study of microbial variation through amplification of V1-V9 region of 16sRNA gene from pure culture of e-swab and culture-swab

For identification of individual bacteria, 34 colonies of pure culture (aerobic) e-swabs samples collected from 15 fish of seven lochs (Scad, Torm, Tros, Grog, Hosta, Ardh and Duin) and 12 colonies (6 Aerobic and 6 Anaerobic) of culture-swab from 6 representative

fish of three lochs (Tros, Hosta and Ardh) were selected on the basis of colony characteristics to amplify the V1-V9 region of 16sRNA. Pure culture colonies were morphologically distinguished from mixed culture plate with context of size, colour, shape and appearance of each colony both in aerobic (AP) and anaerobic (ANAP) culture (Fig. 2, detail in Appendix Table S4).

| AR 74 | 4 | Aerobic | A. Medium
B. Large
C. Large
D. Large | A. Golden
B. White
C. Creamy
D. Creamy | A. Round and thin
B. Irregular, concave
C. Round and thick
D. Pentagonal | A. Transparent
B. Dull
C. Shiny
D. Shiny |
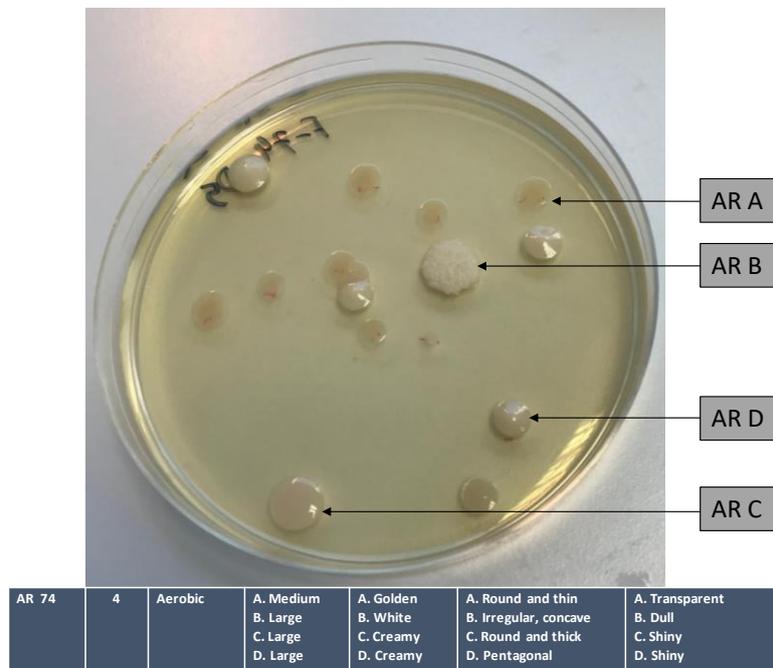
Fig.2 Representative mixed culture plate (TSA) of a saltwater sample (AR 74) from Ardh showing four types of colonies (AR A-D) based on their morphological characteristics like size, colour, shape and appearance of each colony distinguished for pure culture and subsequent sequencing.

DNA was extracted from glycerol stock material of pure culture bacterial colonies by using QIAamp Power Faecal DNA Kit (QIAGEN, USA) with slight modification as described in the manufacturer's protocol, presented in Appendix S2. PCR reactions were generally fixed in 25µL volumes to amplify the V1-V9 region of 16sRNA. Each reaction contained 1µL DNA template, 17.8µl molecular biology grade water, 2.5µl of 10x Thermo pol PCR reaction buffer (New England BioLabs), 2.5µl of 2mM stock dNTP nucleotide mix (200µM each dNTP) (Thermo Scientific), 0.5µl of forward primer (1pmol/µl), 0.5µl of reverse primer (1pmol/µl) (Sigma) and 0.2µl of Taq DNA polymerase (approx. 2 units' /reaction tube) (New England BioLabs) (Table 3). The PCR reaction was performed in a thermocycler (BIO Rad, Mexico) and the following conditions was applied with heated lid at 105°C and initial denaturation for 30s at 95°C, followed by a total of 35 cycles of denaturation at 95°C for 30s, annealing at 56.5°C (depending on primer Tm) for 30s, and extension at 68°C for 1 min, and a final extension at 68°C for 5 min and hold at 12°C.

Table 3: The primers (including their properties) used in the present study

Primer ID (Forward/ Reverse)	Target Region	Primer Sequence (5'-3')	Tm	Approx. Product Size
FW(8f) ReV (1492R)	Bacterial 16sRNA (V1- V9)	5'-AGAGTTTGATCCTGGCTCAG -3' 5'- GGTTACCTTGTTACGACTT-3'	56.5^{0}C	1484bp
FW (P 960) with GC clamp ReV (P 963)	Bacterial 16sRNA (V3)	5'-CGCCCGCCGCGCCCCGCGCCCGGC CCGCCGCCACCGCCGCACTYCTACG GRAGGCWGC-3' 5'- ATTACCGCGGCTGCTGG-3	63^{0}C	196bp

PCR products were separated on a 1% agarose / TAE (2 M tris base,1 M Glacial Acetic Acid, 0.05 M EDTA, pH 8.0) gel. To determine the approximate size of the amplified products, 5 µl of 1Kb Hyper ladder (Bio line) was loaded as a molecular weight marker. About 5 µl of 6x gel-loading dye (0.25% Bromophenol-blue, 40% sucrose, 0.25% Xylene-cyanole and 0.25% Orange G) was added into 5 µl of PCR product, mixed well and loaded into the well with 5 µl volume. The DNA was separated at 90-100Volts, 110mAMP for 25 mins to 1 hr. DNA was visualised by post-staining in 0.5ug/ml ethidium bromide solution for 15-30 min and observed using UV-illumination (Photo Doc-II-imaging system, UVP, UK). For DNA purification, ExoSAP-IT PCR product clean-up kits (Thermo-Fisher Scientific) were used according to the manufacturer's protocol. After purification, quantification was performed by Bio-drop and then DNA was diluted to 15µg/ml where possible and sent to Source Biosciences (UK) for DNA sequence determination.

2.6.2 Study of microbial variation through DGGE of V3 region of 16s RNA

In the present study, culture-swab samples of 6 fish (total 18 fish), with two types of ecotypes (anadromous and resident) from each location (Tros, Hosta and Ardh) were analysed by DGGE to determine qualitative differences in the microbial population of stickleback skin samples (Table 2). Firstly, PCR was performed for 16s rRNA V1-V9 region with 16s RNA universal primer V1(8f) and V9 (1492R) developed by Turner et al. (1999) (Table 4) and the PCR reaction volume and conditions were the same as those described above. PCR product was cleaned by the NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL, Germany) as described in Appendix Fig. S3. Purified V1-V9 DNA samples was used as template for DGGE separation by amplifying the 16s rRNA V3 variable region with degenerate universal primers (Table 4). The following reagents and conditions were used for V3 regionamplification with 50µl reaction volume: 1 µL DNA

template, 32.5µL molecular biology grade water, 10 µl of 1X Phusion GC buffer (New England BioLabs), 1 µl of 200 mM dNTPs (Thermo Scientific), 2.5 µl of 0.50µM forward primer (P960 with GC clamp), 2.5 µl of 0.50 µM reverse primer (P963) and 0.5 µl of 1.0 units of Phusion DNA Polymerase (New England BioLabs). To avoid non-specific hybridizations due to complementary micro-sequences and to improve the specificity of the reaction, a ''touchdown'' PCR was performed for V3 region under the following conditions- initial denaturation at 98°C for 30 sec. and 5 cycles of denaturation at 98°C for 15 sec, then annealing at 68°C for 20 sec and extension at 72°C for 30 sec followed by 30 cycles of 98°C for 15 sec, 63°C for 20 sec and 72°C for 30 sec and a final elongation step for 72^0C for 5 mins and hold 12^0C.

PCR products were then separated by DGGE using the DCode Universal Mutation detection system (Bio-Rad, USA). Two parallel denaturing gels (20-60% and 30-55%) were prepared containing the following compositions:

Table 4: Composition of reagents used for preparation of denaturing gel

Component	20%	30%	55%	60%
Water	To 15ml	To 15ml	To 15ml	To 15ml
40% Acrylamide	3ml	3ml	3ml	3ml
50x TAE	0.3ml	0.3ml	0.3ml	0.3ml
Urea	1.3g	1.9g	3.4g	3.7g
Formamide	1.2ml	1.8ml	3.3 ml	3.6ml
Glycerol	0.3ml	0.3ml	0.3ml	0.3ml

750 µl of sealing solution (1.5ml 0% solution, 13.5 µl 10% Ammonium Persulfate and 1.35 µl TEMED) was added to the vertical casting frame and allowed to set before the addition of the gradient gel. 20 µl of 2x loading dye (2% Bromophenol Blue, 2% Xylene Cyanol, 30%Glycerol and 70% water) were mixed with 45 µl volumes of PCR products and loaded with 20 µl/well. Electrophoresis was performed at 60°C in 1x TAE buffer (2 M tris base,1 M Glacial Acetic Acid, 0.05 M EDTA, pH 8.0) at 60 V for 16 h. After electrophoresis, the gels were stained using GelStar nucleic acid gel stain (Lonza, Switzerland) for 15-20 min and visualized on a UV Trans-illuminator.

The amplified DNA product within the gel was collected using a sterile scalpel blade after visualisation on a UV-transilluminator and was cleaned by using NucleoSpin Gel and PCR Clean-up kit. DNA quantification was performed by Qubit ds DNA BR Assay Kit (Life Technologies) on Qubit Fluorimeter (Invitrogen) and the protocol is described in Appendix S4). After quantification, DNA was diluted to 15µg/ml where possible and sent

to Source Biosciences (UK) for DNA sequence determination. Bacterial identification was performed using BLAST based searches within nucleotide database (National Center for Biotechnology Information, NCBI) to determine the highest homolog to a bacterial group.

2.7 Data analysis

All data were collated in Excel (Microsoft) and statistical analyses were conducted using R, version 3.6.3 (R Core Team, 2020).

Study of microbial abundance through culture of e-swabs

To analyse the microbial variation between freshwater and saltwater, freshwater low pH and high pH habitat and ecotypes (anadromous and resident) of stickleback, e-swab and culture-swab samples were collected from three types of lochs of North Uist on the basis of their water chemistry (salinity and pH). The total viable bacterial counts (expressed as cfu/ml) were estimated for quantitative study of bacterial load variation among e-swab samples of 7 lochs (Scad, Torm, Tros, Grog, Hosta, Ardh and Duin). A one-way analysis of variance (ANOVA) was performed to test for differences in bacterial load between freshwater and saltwater (salinity), freshwater low pH and high pH habitat (pH) and anadromous and resident ecotypes of freshwater stickleback. A two-sample t-test was performed to compare the mean bacterial load of loch water and fish samples of seven lochs. A generalized linear model (GLM) with a Gaussian distribution and identity-link function was fitted to quantify the association among observed variables of stickleback samples. These includes ecotype of fish, salinity, pH of loch water as predictor variables and bacterial load as response variable. Stepwise regression with a combination of forward and backward selection based on likelihood ratio tests was conducted for the GLM model.

Study of microbial variation through amplification of 16sRNA gene

For qualitative analysis, individual bacteria were identified from 84 nucleotide sequences of 16sRNA collected from the 34 colonies of e-swab (V1-V9 region), 12 pure culture colonies of culture-swab (V1-V9 region) and 35 samples of mixed culture colonies of DGGE (V3 region) (Appendix Table. S5). Each individual sequence was matched with a reference database using NCBI-BLAST based searches with the closest cultured strains having percent sequence identities as low as 85.4%. Frequency of bacterial taxa (up to genus) were calculated from the percentage (%) of identified genus observed in total fish of each category to analyse their variation among four groups of stickleback: 1. Freshwater (residents in Scad, Torm, Tros, Hosta and Grog) and saltwater (residents and anadromous in Ardh and Duin) 2. Freshwater low pH (residents in Scad and Torm) and freshwater high pH (residents in Hosta and Grog) 3. Anadromous and resident of

freshwater habitat (Tros and Hosta) and 4. Anadromous of saltwater (Ardh and Duin) and freshwater (Tros, Hosta and Grog) habitat. A chi-square test of independence was performed to observe the association between environmental factors and bacterial composition in each group of stickleback samples.

Simpson's (Simpson, 1949) and Shannon diversity index (Shannon and Wiener, 1963) were calculated using the following formula to observe the variation of bacterial genera with various species identified within the above four groups of sticklebacks:

A. Simpson's index (Index of dominance)

$$D=1-\sum(n/N)^2$$

Where, n= the total number of organisms of a species/genus

N= the total number of organisms of all species/genus

B. Shannon index (Index of information statistics)
$$H=-\sum(n/N) \, Ln \, (n/N)$$

Where, n= the number of individuals of one species/genus

N= the total number of individuals of all species/genus

Ln=Natural logarithm

$\sum$=sum of the calculations

In DGGE analysis, each individual discrete band represented a unique sequence' or phylotype of bacteria in a population (Muyzer et al., 1995; van Hannen et al., 1999). The variation in the composition of bacterial population between two representative lochs of freshwater and saltwater (Hosta and Ardh), ecotypes (anadromous and resident) and culture types (anaerobic and aerobic) were observed as percentage of identified genus from the sequences of V3 region of 16sRNA gene. Diversity of sequenced genus in each fish was calculated using Simpson's and Shannon diversity index as above.

3. Results

3.1 Study of microbial abundance through culture of e-swabs

The viable aerobic bacterial count ranged from 10.5×10^6 to 9.36×10^8 cfu/ml among 33 e-swab samples of stickleback skin collected from seven lochs (Scad, Torm, Tros, Hosta, Grog, Duin and Ardh) of North Uist. The mean bacterial load was highest in Duin (6.33×10^8 cfu/ml) and lowest in Grog (2.70×10^8 cfu/ml) (Fig. 3A) but not significantly different among lochs. The overall mean bacterial load of skin samples of fish was significantly higher than the water samples of seven lochs (Two-sample t-test, t = 8.46, df = 12, p<0.001). However, there were no significant differences in mean bacterial load of skin samples between freshwater and saltwater, freshwater low pH and high pH or anadromous and resident ecotypes including freshwater stickleback (Table 5; Fig. 3B-D).

Table 5. Mean bacterial load from the culture of e-swab samples from 33 individual fish grouped on the basis of salinity (freshwater and saltwater), pH (freshwater low pH and high pH) of habitat (loch) or ecotypes (anadromous and resident including freshwater) of stickleback.

Sl.	Category	Type	N	Average ($\times 10^8$) cfu/ml ± S.E.
1	Habitat (Salinity)	Freshwater	24	3.65 ± 0.49
2.		Saltwater	9	4.82 ± 0.86
1.	Habitat (pH)	FW High pH	12	3.19 ± 0.60
2.		FW Low pH	12	4.10 ± 0.78
1.	Ecotype	Anadromous	6	4.55 ± 0.76
2.		SW resident	4	5.03 ± 1.72
3.		FW resident	23	3.63 ± 0.51
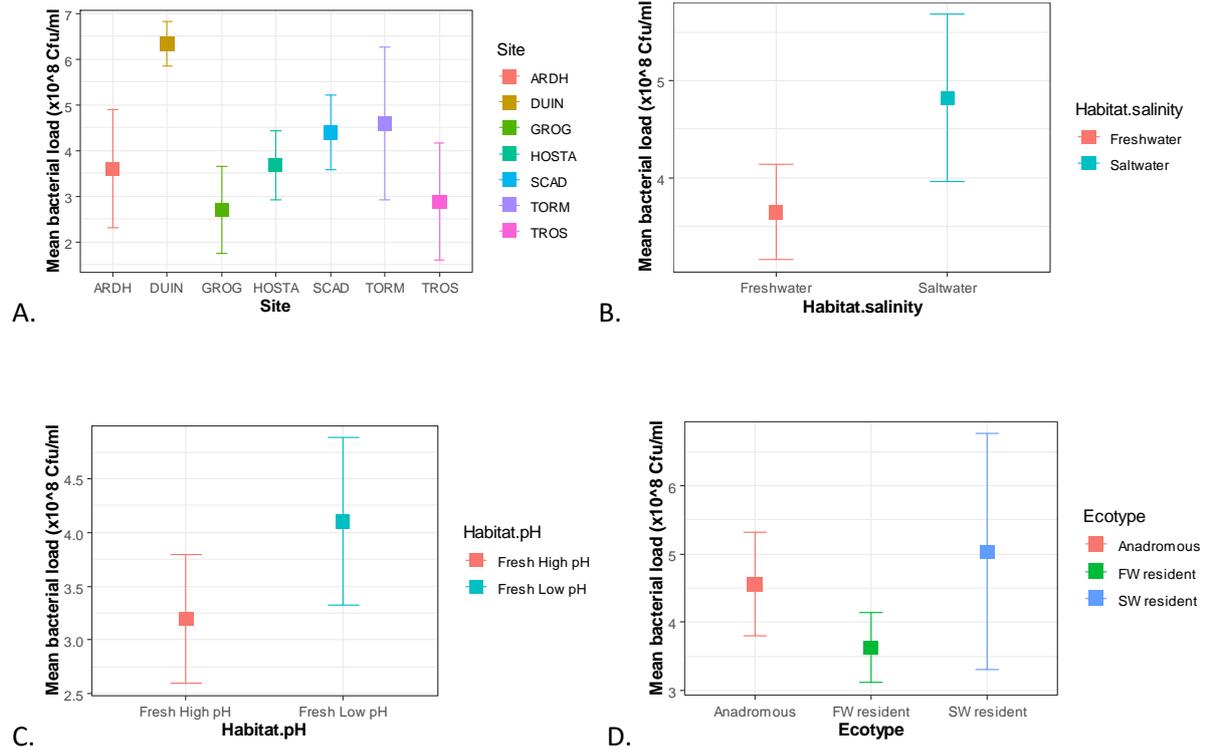
Fig. 3 Average (x10^8 cfu/ml ± S.E.) bacterial load from the culture of 33 e-swab samples collected from North Uist which shows no significant differences among A) seven lochs or between B) freshwater and saltwater habitat C) freshwater low pH and high pH habitat and D) in ecotypes of stickleback (anadromous, saltwater-resident and freshwater-resident).

3.2 Study of microbial variation through sequencing of 16sRNA gene

From 29 sticklebacks collected from seven lochs (Scad, Torm, Tros, Grog, Hosta, Ardh and Duin) on North Uist, 79 sequences of mixed and pure culture colonies were characterised based on their partial 16sRNA gene analysis (V1-V9 and V3 region) (Appendix, Table S5). These comprised a total of 31 bacterial species from 17 genera representing the four phyla: Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes. Among them, most of the identified genera belonged to the phylum Proteobacteria (76%) including various species of *Aeromonas* sp., *Pseudomonas* sp., *Shewanella* sp., *Psychrobacter* sp., *Klebsiella* sp., *Enterobacteriaceae* sp., *Providencia* sp., *Stenotrophomonas* sp., *Comamonas* sp., *Gamma proteobacterium*, *Halomonas* sp., *Marinomonas* sp., *Yersinia* sp. (Fig. 4). Two sequences of isolated colonies did not match with any reference data of NCBI.
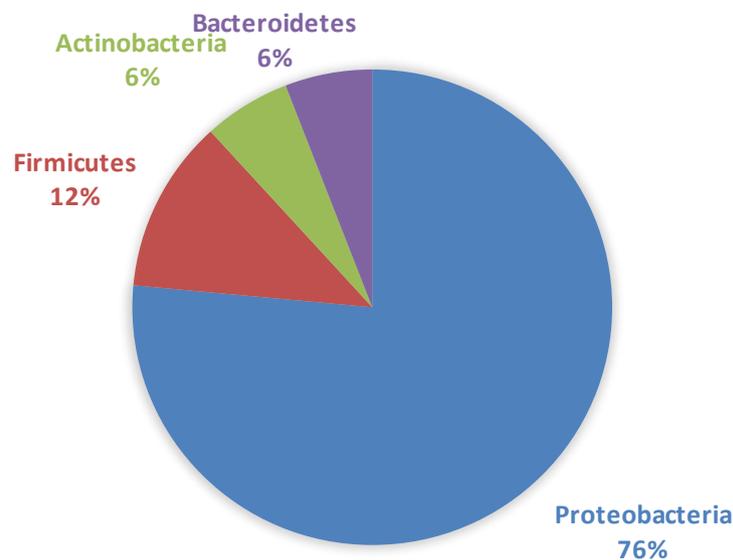
Fig. 4 Diversity of bacterial community identified from the skin swab of stickleback of North Uist shows Proteobacteria as the most dominant phylum with highest frequency of identified genera than the other three phyla.

The overall composition of bacterial strains showed significant variation among stickleback of four different groups: salinity (freshwater and saltwater), pH (freshwater low pH and freshwater high pH) and ecotypes (anadromous in freshwater and freshwater resident) as well as anadromous stickleback of freshwater and saltwater (Table 6). There was little difference in overall diversity, but little overlap in the genera present on fish caught in freshwater and saltwater (Table 7). There was significantly higher diversity of bacterial genera among stickleback in freshwater high pH (Hosta and Grog) than freshwater low pH (Scad, Torm and Tros) lochs (Table 8). *Shewanella* sp. and *Pseudomonas* sp. were the dominant genera found in the samples of freshwater (35.7%)

and low pH fish (50%) whereas *Pseudomonas* sp. was the dominant taxa in saltwater fish (45.5%). *Aeromonas* sp. was the most abundant bacterial genus in high pH stickleback samples (50%) but absent in low pH fish.

Table 6: Results ofSimpson's diversity index (D) and Chi-square test of independence to compare the variation of bacterial genera within different groups of sticklebacks.

Sl.	Group	Simpson's diversity index (D)	Shannon diversity index (H)	Chi-square test of independence
1.	Freshwater	0.840	2.087	χ^2 = 149.24, df =17, p<0.001
	Saltwater	0.837	2.097	
2.	Freshwater high pH	0.746	3.425	χ^2 = 161.42, df = 10, p<0.001
	Freshwater low pH	0.694	1.352	
3.	Freshwater anadromous	0.444	0.636	χ^2 = 168.35, df = 8, p<0.001
	Freshwater resident	0.851	2.046	
4.	Freshwater anadromous	0.520	0.858	χ^2 = 186.15, df = 8, p<0.001
	Saltwater anadromous	0.768	1.756	

Table 7: Frequency of identified bacteria through the amplification of V1-V9 and V3 regions of 16sRNA of 67 nucleotide sequences of the 67 mixed and pure culture samples of e-swab and culture-swab from 25 stickleback collected from five freshwaters (Scad, Torm, Tros, Hosta and Grog) and two saltwater (Duin and Ardh) lochs of North Uist, Scotland.

Bacterial genus (17 genera)	Frequency in freshwater fish	Frequency in saltwater fish
Aeromonas sp.	14.3% (2)	9.1% (1)
***Pseudomonas* spp.**	**35.7% (5)**	**45.5%** (5)
Psychrobacter spp.	7.1% (1)	27.3% (3)
***Shewanella* spp.**	**35.7% (5)**	36.4% (4)
Carnobacterium sp.	14.3% (2)	0.0% (0)
Enterobacteriaceae sp.	7.1% (1)	0.0% (0)
Frigoribacterium sp.	7.1% (1)	0.0% (0)
Klebsiella spp.	14.3% (2)	0.0% (0)
Providencia sp.	7.1% (1)	0.0% (0)
Stenotrophomonas sp.	7.1% (1)	0.0% (0)
Comamonas sp.	0.0% (0)	9.1% (1)
Exiguobacterium sp.	0.0% (0)	9.1% (1)
Flavobacterium sp.	0.0% (0)	27.3% (3)
Gamma proteobacterium	0.0% (0)	9.1% (1)
Halomonas sp.	0.0% (0)	9.1% (1)
Marinomonas sp.	0.0% (0)	9.1% (1)
Yersinia sp.	0.0% (0)	9.1% (1)
Uncultured bacterium	7.1% (1)	9.1% (1)
Total fish	**14**	**11**

Table 8: Frequency of bacterial populations identified through the amplification of V1-V9 and V3 regions of 16sRNA of 41 nucleotide sequences of the 41 mixed and pure culture colony samples of e-swab and culture-swab from 18 stickleback collected from three freshwater low pH (Scad, Torm and Tros) and two freshwater high pH (Hosta and Grog) lochs of North Uist, Scotland.

Bacterial genus (10 genera)	Frequency in FW high pH fish	Frequency in FW low pH fish
***Aeromonas* sp.**	**50% (5)**	0**%** (0)
Carnobacterium sp.	10**%** (1)	12.5**%** (1)
***Pseudomonas* sp.**	10**%** (1)	**50% (4)**
***Shewanella* sp.**	40**%** (4)	**50% (4)**
Klebsiella sp.	20**%** (2)	0**%** (0)
Providencia sp.	10**%** (1)	0**%** (0)
Psychrobacter sp.	10**%** (1)	0**%** (0)
Stenotrophomonas sp.	10**%** (1)	0**%** (0)
Uncultured bacterium	10**%** (1)	0**%** (0)
Enterobacteriaceae bacterium	0**%** (0)	12.5**%** (1)
Frigoribacterium sp.	0**%** (0)	12.5**%** (1)
Total fish	**10**	**8**

There was higher diversity of identified bacterial genera among samples of resident fish than anadromous fish collected from 31 nucleotide sequences of 4 anadromous and 7 resident sticklebacks of two freshwater lochs (Hosta and Tros) (Table 6). In freshwater anadromous and resident group, only two genera of various species (*Shewanella* sp. and *Aeromonas* sp.) were present in the samples of anadromous fish (Fig. 5). Similarly, frequency of identified bacterial genera showed significant variation between freshwater and saltwater anadromous fish collected from 30 nucleotide sequences of 10 sticklebacks (Table 6). Anadromous fish collected from freshwater (Hosta and Tros) showed less diversity in the composition of bacteria than the anadromous fish collected from saltwater (Duin and Ardh) representing only two genera (*Shewanella* sp. and *Aeromonas* sp.) (Fig. 6).
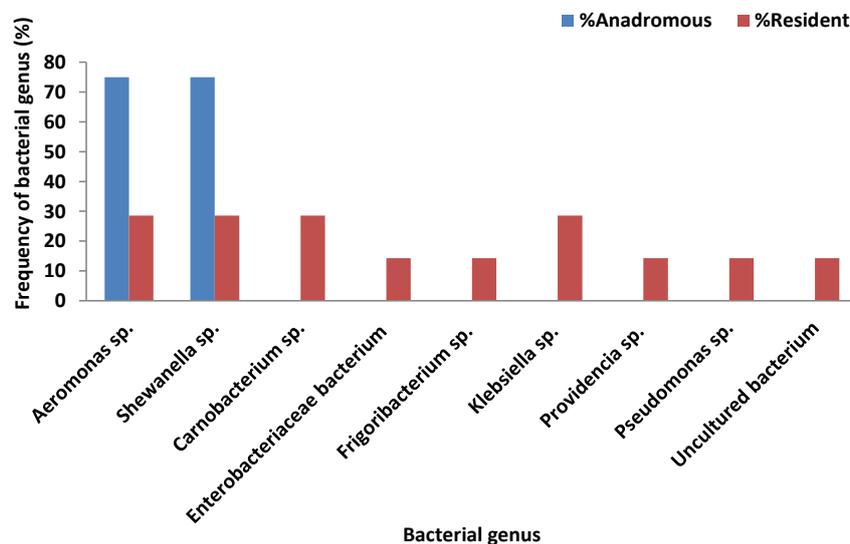
Fig. 5: Frequency (%) of identified bacteria from the skin samples of stickleback shows more diversity in resident fish than anadromous fish collected from two freshwater (Hosta and Tros) lochs of North Uist, Scotland.
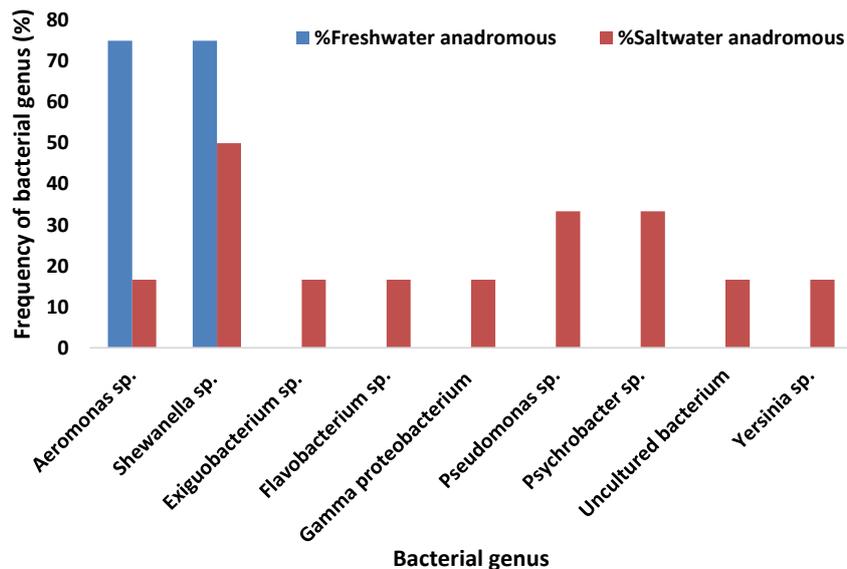
Fig. 6: Frequency (%) of identified bacteria shows high diversity in saltwater-anadromous than freshwater-anadromous collected from two freshwater (Hosta and Tros) and two saltwater lochs (Duin and Ardh) of North Uist, Scotland.

Qualitative analysis of DGGE gel (with concentration of 20-60%) showed comparatively different banding patterns for the V3 region of 16s RNA in the samples collected from two types of fishes (anadromous and resident) from a saltwater loch (Ardh) and a freshwater loch (Hosta) (Fig. 7A and C). One band in the DGGE profile indicated amplification of a sequence from only one species of bacteria. There was a significantly different bacterial composition based on the selected 35 sequenced individual bands of Ardh (19) and Hosta (16) samples as a representative of saltwater and freshwater environment (χ^2 = 137.17, df = 9, p<0.001). In both cases, five common previously identified genera (*Shewanella* sp., *Marinomonas* sp., *Pseudomonas* sp., *Psychrobacter sp.* and *Aeromonas* sp.) were added as a ladder to compare the similar length of bands among the samples. *A. rivuli* was present in all samples of both lochs represented by the ladder no. 5. Freshwater fish collected from Hosta (Diversity index, D = 2.03, H = 1.55) showed less diversity in the composition of identified bacteria than the saltwater fish collected from Ardh (Diversity index, D = 4.0, H = 0.83; Fig. 5). In addition, *Aeromonas* sp. with various species was the most abundant taxa (62.5%) along with two other genera (*Shewanella* sp. and *Klebsiella* sp.) among sequenced data of Hosta whereas Ardh comprised six different genera with *Shewanella* sp. as a dominant taxon (31.6%) (Fig. 7B and D). Among three identified genera of Hosta, there was a significantly different pattern between aerobic and anaerobic culture type (χ^2 = 10.42, df = 2, p = 0.005) as well as anadromous and resident fish (χ^2 = 16.19, df = 2, p < 0.001). Similarly, there was significant variation of bacterial composition (χ^2 = 14.32, df = 7, p = 0.045) between aerobic and anaerobic culture type as well as anadromous and resident fish (χ^2 = 32.78, df = 7, p< 0.001) of Ardh.

190

Ardh

A.

Band id	Bacterial species in Ardh
1BA	*Shewanella* sp.
1EA	*Flavobacterium* sp.
1FA	*S. baltica*
3BA	*Pseudomonas* sp.
5GA	Uncultured bacteria
6BA	*Flavobacterium* sp.
7AR	*S. baltica*
7FA	*S. baltica*
8AA	*Shewanella* sp.
8BA	*Flavobacterium* sp.
8EA	No significant similarity found
8FA	*Pseudomonas* sp.
8CA	*Pseudomonas* sp.
10AA	*Psychrobacter aquimaris*
11AA	*Halomonas* sp.
12BA	*S. baltica*
1DA	*Flavobacterium* sp.
13AA	*Pseudomonas* sp.
9AA	Gamma proteobacterium

B.

Hosta

C.

Band id	Bacterial species in Hosta
1BH	*S. putrefaciens*
3GH	*S. putrefaciens*
4BH	*Aeromonas* sp.
5HH	*A. sobria*
6AH	*S. putrefaciens*
7BH	*A. salmonicida*
8AH	*S. seohaensis*
9AH	*K. pneumoniae*
10AH	*A. sobria*
10BH	*A. salmonicida*
10CH	*Aeromonas* sp.
11AH	*A. rivuli*
12HH	*A. rivuli*
13HH	*Aeromonas* sp.
14HH	*A. Sobria*
15HH	*Shewanella* sp.

D.

Fig. 7 Variation in the banding pattern of DGGE resulted from the PCR amplification and separation of amplified V3 region (196 bp) of 16sRNA in a denaturing gel with a concentration of 20-60% for stickleback samples of two lochs: (A) Ardh and (C) Hosta. One band in the DGGE profile indicate amplification of a sequence from only one species of bacteria. Figure B and D presented the list of bacterial genera from the sequencing of specific band collected from 12 culture-swab samples (6 aerobic and 6 anaerobic) from two types of sticklebacks (6 anadromous and 6 resident) from a representative saltwater loch (Ardh) and a freshwater loch (Hosta). Ardh showing more variation with six different genera than the Hosta with mostly *Aeromonas* spp.

4.Discussion

The parallel evolution of three-spined stickleback populations to a diverse range of new freshwater environments with conspicuous divergence in morphology, physiology and behaviour provides a unique opportunity to study the effect of environmental changes on the skin microbiome of fish (Bell and Foster, 1994; Ostlund-Nilsson et al., 2006). In the present study, I present a preliminary description of the microbial variation between freshwater and saltwater, freshwater low pH and high pH habitat and ecotypes (saltwater resident, freshwater resident and anadromous) of stickleback collected from seven lochs of North Uist with different water chemistry (salinity and pH).The hypothesis of this study was that bacterial load in culture-dependent method would be same on the skin of freshwater and saltwater or other groups of sticklebacks, but that the composition and diversity of the bacterial community might be varied by the effect of environmental factors. Therefore, a traditional culture-dependent technique was combined with a molecular method (16s RNA gene sequence analyses with DGGE) to study the skin microbiome in different populations of stickleback including their characterization and composition in different environmental conditions. To identify the diversity and richness of bacterial taxa, Simpson's and Shannon diversity index were calculated for each sample fish from the identified genera of sequenced data.

Quantitative analysis of bacterial load through traditional culture-dependent methods from the skin samples of stickleback showed high densities (10^6-10^8 cfu/ml) but no significant differences among different groups of sticklebacks on the basis of salinity (freshwater and saltwater), pH (high pH and low pH) and ecotypes (freshwater resident, saltwater resident and anadromous). Although the viable bacterial count was comparatively higher than the standard bacterial load (10^2-10^4/cm^2 in the fish skin sample), this is supported by previous finding of culture-dependent studies on rainbow trout, *O. mykiss* (Diler et al., 2000; Austin 2006) and Atlantic salmon (Landeira-Dabarca et al., 2013). This result was expected for cultured bacterial count because of the direct and continuous exposure of stickleback skin surface to the surrounding environment which serve as a common source of bacterial colonization. Another reason might be the use of inoculated swabs with nutrient-enriched culture medium and the growth of culture-dominant species in the culture-dependent method. However, the bacterial load of skin samples of stickleback was higher than the bacterial densities of surrounding loch water which may be due to the presence of cutaneous mucus as a nutrient source for colonizing bacteria of fish. This observation was consistent with the similar findings in zebra fish (*Danio rerio*), whiting (*Merlangius merlangus*) and Atlantic salmon (*S. salar*) (Rawls et al., 2004; Smith et al., 2007; Landeira-Dabarca et al., 2013).

Traditional culture-dependent methods have a limited capacity (small set of organisms) to characterize microbiota of water in comparison to advanced culture-independent approaches such as whole genome sequencing on Illumina sequencing platform (Miseq or HiSeq). However, it can be used to identify bacteria to the species level in contrast to other modern molecular technique that rarely allowed identification below the order (Vaz-Moreira et al., 2011). In the present study, culture-dependent 16sRNA gene sequencing data from e-swab and culture-swab samples discovered a group of bacterial species that showed significant differences in their relative frequency among sticklebacks of different groups (freshwater-saltwater, high pH-low pH, freshwater anadromous-resident and anadromous of salt-freshwater). The phylum Proteobacteria was the most abundant group among skin microbiome constituted the majority of identified taxa and was present in all of the study groups of sticklebacks. This result corresponds to several previous studies of skin microbiome diversity in different species of fish such as Atlantic salmon (Cipriano and Dove, 2011; Lokesh and Kiron, 2016), red snapper, *Lutjanus campechanus* (Arias et al., 2013), Atlantic cod, *Gadus morhua* (Wilson et al., 2008), Gulf killifish, *Fundulus grandis* (Larsen et al., 2015), large-mouth bass, *Micropterus salmoides*, bluegill, *Lepomis macrochirus* and spotted gar, *Lepisosteus oculatus* (Ray, 2016). The presence of bacterial genera from three other phyla: Firmicutes, Actinobacteria and Bacteroidetes in the skin samples of stickleback is also supported by previous microbiome studies of fish (Sullam et al., 2012; Larsen et al., 2013; Eichmiller et al., 2016) including the gut microbiome of stickleback (Milligan-Myhre et al., 2016).

Most of the bacterial species identified from the sequenced data belonged to gram-negative, pathogenic bacterial genera of fish (*Aeromonas* sp., *Pseudomonas* sp., *Shewanella* sp., *Psychrobacter* sp., *Klebsiella* sp., *Enterobacteriaceae* sp., *Providencia* sp., *Stenotrophomonas* sp., *Comamonas* sp., *Halomonas* sp., *Yersinia* sp. and *Flavobacterium* sp.). This finding mirrored previous studies on skin microbiome which confirmed these species as a common colonizer of fish reflecting the effect of environmental and ecological factors (Larsen et al., 2013; Benhamed et al., 2014; Smith et al., 2015; Lokesh and Kiron, 2016). Among the common bacterial genera in all groups of stickleback, the pathogenic properties of *Aeromonas* spp. and *Pseudomonas* spp. are well documented (Austin and Austin, 2007). *Aeromonas* sp. is regarded as an important disease-causing pathogen of fish and other cold-blooded species; for example, *Aeromonas salmonicida* was reported for depletion of sea trout (*S. trutta*) populations by causing furunculosis, as one of the most common diseases of fish (Janda and Abbott, 2010). However, some species of *Pseudomonas* sp. and *Psychrobacter* sp. have been reported as commensal bacteria in cold-water wild Atlantic cod (Lazado et al., 2010).

In this study, the relative abundance of bacterial genera identified in sample fish was significantly different between freshwater and saltwater stickleback. There was very little variation in the diversity of bacterial species between saltwater and freshwater fish including four overlapping genera (*Aeromonas* sp., *Pseudomonas* sp., *Shewanella* sp. and *Psychrobacter* sp.) which appear as a core set of bacterial genera with the highest abundance. *Pseudomonas* sp. was the most common bacteria identified in almost all populations except Hosta. Some species of *Pseudomonas* sp. were found distinctively in freshwater (*P. gessardii, P. poae and P. syringae*) and saltwater (*P. psychrophile, P. brenneri*) stickleback. A recent study on the gut microbiota of stickleback based on the culture-dependent 16sRNA gene sequencing has shown *Pseudomonas* sp. and *Shewanella* sp. as a shared bacterial genus between oceanic and freshwater fish populations (Milligan-Myhre et al., 2016). In the present study, there were some bacterial taxa present only in freshwater (*Carnobacterium* sp., *Enterobacteriaceae* sp., *Frigoribacterium* sp., *Klebsiella* spp., *Providencia* sp. and *Stenotrophomonas* sp.) and in saltwater (*Comamonas* sp., *Exiguobacterium* sp., *Flavobacterium* sp., Gamma proteobacterium, *Halomonas* sp., *Marinomonas* sp. and *Yersinia* sp.). Among them, *Marinomonas* was considered as one of the biomarkers of the saltwater bacterial groups (Lokesh and Kiron, 2016). Besides, banding profiles of DGGE also showed qualitative differences between freshwater and saltwater skin microbial abundance of stickleback along with less diversity in the composition of sequenced data of freshwater samples. The result of DGGE provided the semi-quantitative analysis of the major bacterial species present in the samples of representative freshwater and saltwater loch as only the most abundant species could be detected by this method (Le Nguyen et al., 2007; Cénit et al., 2014). Previous studies showed that salinity of water predominately influences the microbial community in fish, especially skin and gut microbiome which is also evident in the present study (Sullam et al., 2012; Wong and Rawls, 2012; Schmidt et al., 2015; Lokesh and Kiron, 2016).

The pH of loch water is another environmental factor that influenced the composition of bacterial community of stickleback skin samples. There was significantly higher relative abundance and diversity of bacterial species composition in the stickleback of alkaline habitat than the acidic habitat of North Uist. It has been known that skin mucosal surface of fish acts as an active site for colonization of microbial community and is in direct contact with the surrounding environment (Austin, 2006). Generally, stickleback in acidic waters have a comparatively smaller body size and fewer lateral plates than stickleback in alkaline waters which may be responsible for the reduced diversity in the associated bacterial community in the low pH samples. On the other hand, alkaline pH

produces high volume of fish epidermal mucus by reducing the bioactive components (such as immunoglobulins, lysozymes, protease inhibitors, complement proteins, antibacterial peptides etc.) and antimicrobial activity which may create the appropriate condition for highest adhesion affinity of skin microbiome (Balebona et al., 1995; Al-Arifa, et al., 2011). This contrasts somewhat with the findings of Sylvain et al., (2016) who demonstrated a significantly disturbed composition of skin microbiota in an Amazonian fish, Tambaqui (*Colossoma macropomum*) with increased relative abundance of the phylum Proteobacteria but decreased abundance of the genus Flavobacterium in response to a reduction in pH. Here, the samples were collected from the natural source with constant pH of loch water rather than an experimental exposure which may influence the composition of bacterial community.

The bacterial composition and diversity of the skin bacterial community was determined in ecotypes of stickleback to address the distinctiveness of microbiome abundance on specific fish populations due to their genetic or physiological response to microbiomes. There were significant differences in the relative abundance of bacterial genera between anadromous and resident stickleback of freshwater environment with higher diversity in the resident fish. This could be the effect of transition from high salinity sea water to the freshwater environment along with a possible role of physiological change associated with migration of anadromous fish. In contrast, the high diversity of freshwater-resident fish microbiome is likely the regular bacterial species that colonize in constant environmental conditions (Milligan-Myhre et al., 2016). Besides, the properties of the external environment can strongly determine the composition of microbial community and any change in the rearing environmental condition might affect the skin-associated microbial balance (Yoshimizu and Kimura, 1976; Fierer and Jackson, 2006; Schmidt et al., 2015; Mohammed and Arias, 2015). A similar trend was also observed in the bacterial composition of freshwater and saltwater anadromous groups where freshwater anadromous stickleback showed less diversity and relative abundance with only two isolated shared genera (*Aeromonas* sp. and *Shewanella* sp.). This result is consistent with a study on skin microbiota of Atlantic salmon where transition from freshwater to seawater reshapes the bacterial community with increased diversity in the new environment (Lokesh and Kiron, 2016). However, the destabilised skin bacterial community of anadromous fish in freshwater might be controlled by fish genetics similar to gut microbiota of stickleback which reported higher divergent microbiota in highly genetically-divergent populations (Kenny and Valdivia, 2009; Smith et al., 2015).

In conclusion, the present work demonstrated that environmental factors such as salinity and pH had significant impact on the composition and diversity of skin microbiome of

stickleback. This was a small-scale preliminary study combining traditional culture-dependent and modern molecular method which precisely recorded some important pathogenic bacterial genera such as *Aeromonas* sp., *Pseudomonas* sp., *Shewanella* sp., *Klebsiella* sp., *Yersinia* sp. and *Flavobacterium* sp. etc. from the skin of stickleback. This data can be utilized for species-specific disease study or eco-evolutionary purposes and the overall management and conservation of fish. There were several limitations experienced in the present study in term of sampling period and location, use of swab, limited sample and sequencing of large volume of data etc. However, the present findings will provide a baseline data for the further large-scale research based on 16S rRNA gene sequencing and future metagenomic study with whole genome sequencing of stickleback microbiomes.

References

Al-Arifa, N., Mughal, M.S., Hanif, A. and Batool, A., 2011. Effect of alkaline pH on bioactive molecules of epidermal mucus from Labeo rohita (Rahu). *Turkish Journal of Biochemistry/Turk Biyokimya Dergisi*, *36*(1).

Arias, C.R., Koenders, K. and Larsen, A.M., 2013. Predominant bacteria associated with red snapper from the northern Gulf of Mexico. *Journal of aquatic animal health*, *25*(4), pp.281-289.

Asfie, M., Yoshijima, T. and Sugita, H., 2003. Characterization of the goldfish fecal microflora by the fluorescent in situ hybridization method. *Fisheries science*, *69*(1), pp.21-26.

Austin, B. and Austin, D.A., 2007. Characteristics of the pathogens: Gram-negative bacteria. *Bacterial Fish Pathogens: Diseases of Farmed and Wild Fish*, Springer, Cham pp.81-150.

Austin, B., 2006. The bacterial microflora of fish, revised. *The Scientific World Journal*, *6*, pp.931-945.

Balcázar, J.L., De Blas, I., Ruiz-Zarzuela, I., Cunningham, D., Vendrell, D. and Muzquiz, J.L., 2006. The role of probiotics in aquaculture. *Veterinary microbiology*, *114*(3), pp.173-186.

Balebona, M.C., Morinigo, M.A., Faris, A., Krovacek, K., Månsson, I., Bordas, M.A. and Borrego, J.J., 1995. Influence of salinity and pH on the adhesion of pathogenic Vibrio strains to Sparus aurata skin mucus. *Aquaculture*, *132*(1-2), pp.113-120.

Barton, B.A., 2002. Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. *Integrative and comparative biology*, *42*(3), pp.517-525.

Bell, M.A. and Foster, S.A. eds., 1994. *The evolutionary biology of the three-spine stickleback*. Oxford University Press.

Benhamed, S., Guardiola, F.A., Mars, M. and Esteban, M.Á., 2014. Pathogen bacteria adhesion to skin mucus of fishes. *Veterinary microbiology*, *171*(1-2), pp.1-12.

Bolnick, D.I., Snowberg, L.K., Hirsch, P.E., Lauber, C.L., Knight, R., Caporaso, J.G. and Svanbäck, R., 2014. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology letters*, *17*(8), pp.979-987.

Boutin, S., Bernatchez, L., Audet, C. and Derôme, N., 2013. Network analysis highlights complex interactions between pathogen, host and commensal microbiota. *PLoS One*, *8*(12), p.e84772.

Cahill, M.M., 1990. Bacterial flora of fishes: a review. *Microbial ecology*, *19*(1), pp.21-41.

Cénit, M.C., Matzaraki, V., Tigchelaar, E.F. and Zhernakova, A., 2014. Rapidly expanding knowledge on the role of the gut microbiome in health and disease. *BiochimicaetBiophysicaActa (BBA)-Molecular Basis of Disease*, *1842*(10), pp.1981-1992.

Chakravorty, S., Helb, D., Burday, M., Connell, N. and Alland, D., 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of microbiological methods*, *69*(2), pp.330-339.

Chiarello, M., Villéger, S., Bouvier, C., Bettarel, Y. and Bouvier, T., 2015. High diversity of skin-associated bacterial communities of marine fishes is promoted by their high variability among body parts, individuals and species. *FEMS microbiology ecology*, *91*(7).

Chu, H., Fierer, N., Lauber, C.L., Caporaso, J.G., Knight, R. and Grogan, P., 2010. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology*, *12*(11), pp.2998-3006.

Cipriano, R.C. and Dove, A., 2011. Far From Superficial: Microbial Diversity Associated With the Skin and Mucus of Fish.*In*: Cipriano R, Schelkunov I (eds) Health and diseases of aquatic organisms: bilateral perspectives. MSU Press, East Lansing, pp 156–167.

Clarridge, J.E., 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical microbiology reviews*, *17*(4), pp.840-862.

Collins, C.H. and Lyne, P.M., 1984. Microbiological methods. Fifth edition continued. *Microbiological methods.*

Cruz, M.P., Ibáñez, A.L., Monroy Hermosillo, O.A. and RamírezSaad, H.C., 2012. Use of probiotics in aquaculture. *ISRN microbiology*, pp. 1-13.

Danilo, E. 2004. PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. J. Microbiol. Methods 56: 297–314.

De Roij, J. and MacColl, A.D., 2012. Consistent differences in macroparasite community composition among populations of three-spined sticklebacks, *Gasterosteus aculeatus* L. *Parasitology*, *139*(11), pp.1478-1491.

De Roij, J., Harris, P.D. and MacColl, A.D., 2011. Divergent resistance to a monogenean flatworm among three-spined stickleback populations. *Functional ecology*, *25*(1), pp.217-226.

Diler, Ö.Z.N.U.R., ALTUN, S., ÇALIKUŞU, F. and DİLER, A., 2000. A study on qualitative and quantitative bacterial flora of the rainbow trout (Oncorhynchus mykiss) living in different fish farms. *Turkish Journal of Veterinary and Animal Sciences*, *24*(3), pp.251-260.

Eichmiller, J.J., Hamilton, M.J., Staley, C., Sadowsky, M.J. and Sorensen, P.W., 2016. Environment shapes the fecal microbiome of invasive carp species. *Microbiome*, *4*(1), pp.1-13.

El Nagar, A. and MacColl, A.D., 2016. Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1836), p.20160691.

Elliott, D.G., 2011. The many functions of fish integument. *In*: Farrell, A.P.(Ed.), Encyclopedia of Fish Physiology: From Genome to Environment. Academic Press, San Diego, pp. 471–475.

Fierer, N. and Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *103*(3), pp.626-631.

Folmar, L.C. and Dickhoff, W.W., 1980. The parr—Smolt transformation (smoltification) and seawater adaptation in salmonids: A review of selected literature. *Aquaculture*, *21*(1), pp.1-37.

Fukuda, K., Ogawa, M., Taniguchi, H., and Saito, M., 2016. Molecular approaches to studying microbial communities: targeting the 16S ribosomal RNA gene. *J. UOEH* 38: 223–232.

Gilbert, J.A., Field, D., Swift, P., Newbold, L., Oliver, A., Smyth, T., Somerfield, P.J., Huse, S. and Joint, I., 2009. The seasonal structure of microbial communities in the Western English Channel. *Environmental microbiology*, *11*(12), pp.3132-3139.

Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbrück, L., Reeder, J., Temperton, B., Huse, S., McHardy, A.C., Knight, R., Joint, I. and Somerfield, P., 2012. Defining seasonal marine microbial community dynamics. *The ISME journal*, *6*(2), p.298.

Giles, N., 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *Journal of zoology*, *199*(4), pp.535-544.

Gomez, D., Sunyer, J.O. and Salinas, I., 2013. The mucosal immune system of fish: the evolution of tolerating commensals while fighting pathogens. *Fish & shellfish immunology*, *35*(6), pp.1729-1739.

Horsley, R.W., 1973. The bacterial flora of the Atlantic salmon (*Salmo salar* L.) in relation to its environment. *Journal of Applied Microbiology*, *36*(3), pp.377-386.

Horsley, R.W., 1977. A review of the bacterial flora of teleosts and elasmobranchs, including methods for its analysis. *Journal of Fish biology*, *10*(6), pp.529-553.

Huang, Z., Ma, A., Wang, X.A., Lei, J., Li, W., Wang, T., Yang, Z. and Qu, J., 2015. Interaction of temperature and salinity on the expression of immunity factors in different tissues of juvenile turbot Scophthalmus maximus based on response surface methodology. *Chinese journal of oceanology and limnology*, *33*(1), pp.28-36.

Hughes, L.C., Somoza, G.M., Nguyen, B.N., Bernot, J.P., González-Castro, M., Díaz de Astarloa, J.M. and Ortí, G., 2017. Transcriptomic differentiation underlying marine-to-freshwater transitions in the South American silversides *Odontesthesargentinensis* and *O. bonariens*is (Atheriniformes). *Ecology and Evolution*. 7:5258–5268.

Janda, J.M. and Abbott, S.L., 2010. The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clinical microbiology reviews*, *23*(1), pp.35-73.

Kamei, Y., Sakata, T. and Kakimoto, D., 1985. Microflora in the alimentary tract of Tilapia: characterization and distribution of anaerobic bacteria. *The Journal of General and Applied Microbiology*, *31*(2), pp.115-124.

Kar, N., Roy, R.N., Sen, S.K. and Ghosh, K., 2008. Isolation and characterization of extracellular enzyme producing bacilli in the digestive tracts of rohu, Labeorohita

(Hamilton) and Murrel, Channa punctatus (Bloch). *Asian Fisheries Science*, *21*(4), pp.421-434.

Kenny, B. and Valdivia, R., 2009. Host–microbe interactions: bacteria. *Current Opinion in Microbiology*, 12, pp. 1-3.

Kolbert, C.P. and Persing, D.H., 1999. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Current opinion in microbiology*, *2*(3), pp.299-305.

Koneman E.W., Allen S. D., Janda W. M., Schreckenberger P. C., Winn W. J., editors. 1997. Color Atlas and Textbook of Diagnostic Microbiology. 5th ed. Philadelphia: Lippincott Williams & Wilkins; p. 98-102.

Landeira-Dabarca, A., Sieiro, C. and Alvarez, M., 2013. Change in food ingestion induces rapid shifts in the diversity of microbiota associated with cutaneous mucus of Atlantic salmon Salmo salar. *Journal of fish biology*, *82*(3), pp.893-906.

Larsen, A., Tao, Z., Bullard, S.A. and Arias, C.R., 2013. Diversity of the skin microbiota of fishes: evidence for host species specificity. *FEMS microbiology ecology*, *85*(3), pp.483-494.

Larsen, A.M., Bullard, S.A., Womble, M. and Arias, C.R., 2015. Community Structure of Skin Microbiome of Gulf Killifish, Fundulusgrandis. *Microbial ecology*, *70*(2), pp.534-544.

Lazado, C.C., Caipang, C.M.A., Rajan, B., Brinchmann, M.F. and Kiron, V., 2010. Characterization of GP21 and GP12: two potential probiotic bacteria isolated from the gastrointestinal tract of Atlantic cod. *Probiotics and antimicrobial proteins*, *2*(2), pp.126-134.

Le Nguyen, D.D., Ngoc, H.H., Dijoux, D., Loiseau, G. and Montet, D., 2008. Determination of fish origin by using 16S rDNA fingerprinting of bacterial communities by PCR-DGGE: An application on Pangasius fish from Viet Nam. *Food control*, *19*(5), pp.454-460.

Lederberg, J. and McCray, A.T., 2001. OmeSweetOmics--A Genealogical Treasury of Words. *The Scientist*, *15*(7), pp.8-8.

Leesing, R., 2005. *Identification and validation of specific markers for traceability of aquaculture fish for import/export* (Doctoral dissertation, Ph. D. dissertation. University of Montpellier 2, France).

Logares, R., Bråte, J., Bertilsson, S., Clasen, J.L., Shalchian-Tabrizi, K. and Rengefors, K., 2009. Infrequent marine–freshwater transitions in the microbial world. *Trends in microbiology*, *17*(9), pp.414-422. Salmo salar. *Journal of fish biology*, *82*(3), pp.893-906.

Lokesh, J. and Kiron, V., 2016. Transition from freshwater to seawater reshapes the skin-associated microbiota of Atlantic salmon. *Scientific reports*, *6*, p. 19707–19716.

Lowrey, L., Woodhams, D.C., Tacchi, L. and Salinas, I., 2015. Topographical mapping of the rainbow trout (Oncorhynchus mykiss) microbiome reveals a diverse bacterial community with antifungal properties in the skin. *Applied and environmental microbiology*, *81*(19), pp.6915-6925.

Lozupone, C.A. and Knight, R., 2007. Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, *104*(27), pp.11436-11440.

Magalhaes, I.S., D'Agostino, D., Hohenlohe, P.A. and MacColl, A.D., 2016. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular ecology*, *25*(17), pp.4319-4336.

Mahmud, M.A., Bradley, J.E. and MacColl, A.D., 2017. Abiotic environmental variation drives virulence evolution in a fish host-parasite geographic mosaic. *Functional Ecology*. pp.1-7.

Maïworé, J., Tatsadjieu, N.L., Montet, D., Loiseau, G. and Mbofung, C.M.F., 2009. Comparison of bacterial communities of tilapia fish from Cameroon and Vietnam using PCR-DGGE (polymerase chain reaction-denaturing gradient gel electrophoresis). *African journal of Biotechnology*, *8*(24).

Maiwore, J., TatsadjieuNoune, L., Goli, O. T., Montet, D. and Mbofung, C. M. F. 2012. Influence of technological treatments on bacterial communities in tilapia *(Oreochromisniloticus)* as determined by 16S rDNA fingerprinting using polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE), *African Journal of Biotechnology* 11(34), pp.8586-8593.

Marchesi, J.R. and Ravel, J., 2015. The vocabulary of microbiome research: a proposal. *Microbiome*, *3*(1), p.31.

Milligan-Myhre, K., Small, C.M., Mittge, E.K., Agarwal, M., Currey, M., Cresko, W.A. and Guillemin, K., 2016. Innate immune responses to gut microbiota differ between oceanic and freshwater threespine stickleback populations. *Disease models & mechanisms*, *9*(2), pp.187-198.

Moeseneder, M.M., Arrieta, J.M., Muyzer, G., Winter, C. and Herndl, G.J., 1999. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. *Applied and environmental microbiology*, *65*(8), pp.3518-3525.

Mohammed, H.H. and Arias, C.R., 2015. Potassium permanganate elicits a shift of the external fish microbiome and increases host susceptibility to columnaris disease. *Veterinary Research*, *46*(1), pp.1-13.

Montes, M., Pérez, M.J. and Nieto, T.P., 1999. Numerical taxonomy of gram-negative, facultative anaerobic bacteria isolated from skin of turbot (Scophthalmus maximus) and surrounding water. *Systematic and applied microbiology*, *22*(4), pp.604-618.

Montet, D., Le Nguyen, D.D., El Sheikha, A., Condur, A., Metayer, I. and Loiseau, G., 2008. Application of PCR-DGGE in determining food origin: Cases studies of fish and fruits. *In*: Greening the Food Chain 3 and 4; 3: Traceability: Tracking and tracing in the food chain. *Aspects of Applied Biol*., 87: 11-22.

Murray, A.E., Hollibaugh, J.T. and Orrego, C., 1996. Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Applied and Environmental Microbiology*, *62*(7), pp.2676-2680.

Muyzer, G., 1999. DGGE/TGGE a method for identifying genes from natural ecosystems. *Current opinion in microbiology*, *2*(3), pp.317-322.

Muyzer, G., De Waal, E.C. and Uitterlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and environmental microbiology*, *59*(3), pp.695-700.

Muyzer, G., Teske, A., Wirsen, C.O. and Jannasch, H.W., 1995. Phylogenetic relationships of Thiomicrospira species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Archives of Microbiology*, *164*(3), pp.165-172.

Nayak, S.K., 2010. Role of gastrointestinal microbiota in fish. *Aquaculture Research*, *41*(11), pp.1553-1573.

Ostlund-Nilsson, S., Mayer, I. and Huntingford, F.A. eds., 2006. *Biology of the three-spined stickleback*. CRC press.

Ovreås, L., Forney, L., Daae, F.L. and Torsvik, V., 1997. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Applied and environmental microbiology*, *63*(9), pp.3367-3373.

Øygarden, E.T., 2017. *Influence of Genetic Background and Environmental Factors on the Skin and Gut Microbiota of Atlantic salmon (Salmo salar) Fry* (Master's thesis), Norwegian University of Science and Technology (NTNU).

Parris, D.J., Brooker, R.M., Morgan, M.A., Dixson, D.L. and Stewart, F.J., 2016. Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement. *Peer J*, *4*, pp. 2412.

Pereira, F., Carneiro, J., Matthiesen, R., van Asch, B., Pinto, N., Gusmão, L. and Amorim, A., 2010. Identification of species by multiplex analysis of variable-length sequences. *Nucleic acids research*, *38*(22):203-203.

Pérez, T., Balcázar, J.L., Ruiz-Zarzuela, I., Halaihel, N., Vendrell, D., De Blas, I. and Múzquiz, J.L., 2010. Host-microbiota interactions within the fish intestinal ecosystem. *Mucosal immunology*, *3*(4), pp.355.

R. Core Team, 2020. R: a language and environment for statistical computing. Version 3.6.3 [computer program].R Foundation for Statistical Computing, Vienna, Austria.

Rawls, J.F., Samuel, B.S. and Gordon, J.I., 2004. Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proceedings of the National Academy of Sciences*, *101*(13), pp.4596-4601.

Ray, C.L., 2016.*Master's thesis -Characterization of the gut and skin microbiomes of wild-caught fishes from Lake Guntersville*.Auburn University, Alabama, USA.pp.93.

Riemann, L., Steward, G.F., Fandino, L.B., Campbell, L., Landry, M.R. and Azam, F., 1999. Bacterial community composition during two consecutive NE Monsoon periods in the Arabian Sea studied by denaturing gradient gel electrophoresis (DGGE) of rRNA genes. *Deep Sea Research Part II: Topical Studies in Oceanography*, *46*(8), pp.1791-1811.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. No evidence of local adaptation of immune responses to Gyrodactylus in three-spined stickleback (Gasterosteus aculeatus). *Fish & shellfish immunology*, *60*, pp.275-281.

Sakata, T., Sugita, H., Mitsuoka, T., Kakimoto, D. and Kadota, H., 1981. Characteristics of Obligate Anaerobic Bacteria in the Intestines of Freshwater Fish. *Bulletin of the Japanese Society of Scientific Fisheries*, *47*(3), pp.421-427.

Salinas, I., Zhang, Y.A. and Sunyer, J.O., 2011. Mucosal immunoglobulins and B cells of teleost fish. *Developmental & Comparative Immunology*, *35*(12), pp.1346-1365.

Sanders, E.R.,2012. Aseptic laboratory techniques: plating methods. *JoVE (Journal of Visualized Experiments)*, (63), p.e3064.

Sar, N. and Rosenberg, E., 1989. Fish skin bacteria: production of friction-reducing polymers. *Microbial ecology*, *17*(1), pp.27-38.

Scharsack, J.P., Franke, F., Erin, N.I., Kuske, A., Büscher, J., Stolz, H., Samonte, I.E., Kurtz, J. and Kalbe, M., 2016. Effects of environmental variation on host–parasite interaction in three-spined sticklebacks (Gasterosteus aculeatus). *Zoology*, *119*(4), pp.375-383.

Schmidt, V.T., Smith, K.F., Melvin, D.W. and Amaral-Zettler, L.A., 2015. Community assembly of a euryhaline fish microbiome during salinity acclimation. *Molecular ecology*, *24*(10), pp.2537-2550.

Sevellec, M., Pavey, S.A., Boutin, S., Filteau, M., Derome, N. and Bernatchez, L., 2014. Microbiome investigation in the ecological speciation context of Lake Whitefish (Coregonusclupeaformis) using next-generation sequencing. *Journal of evolutionary biology*, *27*(6), pp.1029-1046.

Shannon, C.E. and Wiener, W., 1963. The mathematical theory of communities. *Urbana: University of Illinois press*.

Simpson, E.H., 1949. Measurement of diversity. *Nature*, *163*(4148), pp.688-688.

Small, C.M., Currey, M., Beck, E.A., Bassham, S. and Cresko, W.A., 2019. Highly reproducible 16S sequencing facilitates measurement of host genetic influences on the stickleback gut microbiome. *Msystems*, *4*(4).

Smith, C.C., Snowberg, L.K., Caporaso, J.G., Knight, R. and Bolnick, D.I., 2015. Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *The ISME journal*, *9*(11), pp.2515-2526.

Smith, C.J., Danilowicz, B.S. and Meijer, W.G., 2007. Characterization of the bacterial community associated with the surface and mucus layer of whiting (Merlangius merlangus). *FEMS microbiology ecology*, *62*(1), pp.90-97.

Spanggaard, B., Huber, I., Nielsen, J., Nielsen, T., Appel, K.F. and Gram, L. 2000. The microflora of rainbow trout intestine: a comparison of traditional and molecular identification. *Aquaculture*, 182, pp. 1–15.

Steele, J.A., Countway P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y., Chow, C.E., Sachdeva, R., Jones, A.C., Schwalbach, M.S., Rose, J.M., Hewson, I., Patel, A., Sun, F., Caron, D.A., Fuhrman, J.A.. 2011. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J.,* 5, pp.1414–1425.

Steury, R.A., Currey, M.C., Cresko, W.A. and Bohannan, B.J., 2019. Population genetic divergence and environment influence the gut microbiome in Oregon threespine stickleback. *Genes*, *10*(7), p.484.

Stevens, J.L. and Olson, J.B., 2013. Invasive lionfish harbor a different external bacterial community than native Bahamian fishes. *Coral reefs*, *32*(4), pp.1113-1121.

Sugita, H., Tsunohara, M., Ohkoshi, T. and Deguchi, Y., 1988. The establishment of an intestinal microflora in developing goldfish (Carassiusauratus) of culture ponds. *Microbial ecology*, *15*(3), pp.333-344.

Sullam, K.E., Essinger, S.D., Lozupone, C.A., O'CONNOR, M.P., Rosen, G.L., Knight, R.O.B., Kilham, S.S. and Russell, J.A., 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular ecology*, *21*(13), pp.3363-3378.

Sylvain, F.É., Cheaib, B., Llewellyn, M., Correia, T.G., Fagundes, D.B., Val, A.L. and Derome, N., 2016. pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (Colossomamacropomum). *Scientific reports*, 6: 32032.

Tatsadjieu, N.L., Maïworé, J., Hadjia, M.B., Loiseau, G., Montet, D. and Mbofung, C.M.F., 2010. Study of the microbial diversity of Oreochromisniloticus of three lakes of Cameroon by PCR-DGGE: Application to the determination of the geographical origin. *Food control*, *21*(5), pp.673-678.

Trust, T.J. and Sparrow, R.A.H., 1974. The bacterial flora in the alimentary tract of freshwater salmonid fishes. *Canadian Journal of Microbiology*, *20*(9), pp.1219-1228.

Turner, S., Pryer, K.M., Miao, V.P. and Palmer, J.D., 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *Journal of Eukaryotic Microbiology*, *46*(4), pp.327-338.

Van Hannen, E.J., Zwart, G., van Agterveld, M.P., Gons, H.J., Ebert, J. and Laanbroek, H.J., 1999. Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. *Applied and Environmental Microbiology*, *65*(2), pp.795-801.

Vaz-Moreira, I., Egas, C., Nunes, O.C. and Manaia, C.M., 2011. Culture-dependent and culture-independent diversity surveys target different bacteria: a case study in a freshwater sample. *Antonie Van Leeuwenhoek*, *100*(2), pp.245-257.

Wang, W., Zhou, Z., He, S., Liu, Y., Cao, Y., Shi, P., Yao, B. and Ringø, E., 2010. Identification of the adherent microbiota on the gills and skin of poly-cultured gibel carp (Carassius auratus gibelio) and bluntnose black bream (Megalobrama amblycephala Yih). *Aquaculture Research*, *41*(9), pp.e72-e83.

Wiik, R., Stackebrandt, E., Valle, O., Daae, F.L., RØDSETH, O.M. and Andersen, K., 1995. Classification of fish-pathogenic vibrios based on comparative 16S rRNA analysis. *International Journal of Systematic and Evolutionary Microbiology*, *45*(3), pp.421-428.

Wilson, B., Danilowicz, B.S. and Meijer, W.G., 2008. The diversity of bacterial communities associated with Atlantic cod Gadus morhua. *Microbial ecology*, *55*(3), pp.425-434.

Woese, C.R. and Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences*, *74*(11), pp.5088-5090.

Woese, C.R., Kandler, O. and Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, *87*(12), pp.4576-4579.

Wong, S. and Rawls, J.F., 2012. Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Molecular Ecology*, 21(13), pp.3100–3102.

Yang, C.H., Crowley, D.E. and Menge, J.A., 2001. 16S rDNA fingerprinting of rhizosphere bacterial communities associated with healthy and Phytophthora infected avocado roots. *FEMS Microbiology Ecology*, *35*(2), pp.129-136.

Yoshimizu, M. and Kimura, T., 1976. Study on the intestinal microflora of salmonids. *Fish pathology*, *10*(2), pp.243-259.

Chapter 6: General Discussion

Phenotypic and genotypic variation caused by different agents of natural selection are extensively considered in order to understand the environmental adaptation of a species (Endler, 1986; Kingsolver et al., 2011; MacColl, 2011; Wadgymar et al., 2017). In this thesis, I identify several abiotic and biotic features of the environment that may shape adaptive evolution of the three-spined stickleback (*G. aculeatus*) in relation to their underlying phenotypic and genotypic traits. In this chapter, I review the main findings of this thesis, with the aim of highlighting the most interesting observations, discussing possible explanations and limitations, and providing suggestions for future research.

At first, I examined the morphological diversity of different stickleback populations collected from North and South Uist and found significant variation in their patterns among resident sticklebacks in relation to the pH of the loch water. This suggests that the divergence of armour traits of stickleback populations collected from two neighbouring islands could be influenced by an abiotic factor of the environment. I then observed the structure of a putative hybrid zone between a freshwater loch and the sea to examine the phenotypic and genotypic variation among hybrids of ancient marine anadromous and newly-adapted freshwater stickleback populations. Morphological traits such as standard length of fish, lateral plate count and parasite abundance showed mosaic patterns whereas body shape, 1st and 2nd dorsal spines, pelvic spines and pelvic structure exhibited clinal patterns across the hybrid zone. This suggests that spatial variation or salinity might work as an environmental agent of natural selection on local adaptation between parental populations. My development of PCR-RFLP based SNPs assays allowed me to study allele frequency variation, highlighting the strength of selection across the geographical gradient based on salinity. I used naturally-recombinant generations of stickleback from a natural hybrid zone to observe the variation in selection due to environmental differences. There was a significant correlation between phenotype and genotype of hybrid zone stickleback which suggests the probability of mutations in the selected loci being associated with phenotypic traits. The strong genetic differentiation and linkage disequilibrium among selected loci suggests strong selection leading to prezygotic reproductive isolation among loch, hybrid zone and sea stickleback populations.

I conducted three laboratory infection experiments to examine the parasitic abundance of stickleback under experimental conditions, in comparison to natural observations of wild populations. I used the previously developed candidate gene SNPs assays to observe the genetic links to parasite resistance of stickleback. The abundance of the

common ectoparasite (*Gyrodactylus* sp.) and freshwater endoparasite (*D. pseudospathecum*) of stickleback showed different patterns in relation to the genotypes of selected adaptive loci (Eda, PPARA, WNT7B and NLRC5), which emphasises the role of parasites in ecological selection for environmental adaptation of fish. To control the genetic background of each individual fish, I used F2 generation hybrids of susceptible freshwater and resistant marine parent populations of stickleback for two *Gyrodactylus* experiments. Lab-raised F1 hybrids from the already admixed Hosta hybrid zone on North Uist were used for the *D. pseudospathecum* experiment. This approach helped with the screening of genotypes in relation to parasite burdens in a controlled environment in comparison to natural observation of the same parasites with different genetic response. Lastly, I examined the relative abundance and diversity of bacterial species in the stickleback skin samples and identified environmental and genetic influences on their composition in a small-scale observation. The variation in the composition of stickleback skin microbiomes across different populations was shown to relation to salinity and pH of water indicating the potential role of these factors in local adaptation.

6.1 Phenotypic adaptation influenced by environmental agents of natural selection

Phenotypic adaptation e.g. through morphological changes represents the most important adaptive response to divergent natural selection, interacting with the biotic and abiotic factors of the environment. How do abiotic agents of natural selection influence adaptive divergence of closely-related lineages? As the environment varies, organisms exhibit diverse adaptation trajectories, which often include changes in appearance, accompanied by genetic changes (Xue et al., 2019). Previous studies on North Uist stickleback have shown that diverse ecological factors such as water chemistry, nutrient content, competitors, predators and parasites might be responsible for their adaptive divergence among different loch populations (Giles, 1983; De Roij et al., 2011; MacColl et al., 2013; MacColl and Aucott, 2014; Magalhaes et al., 2016; Haenel et al., 2019). Here, I have shown that the differences in morphological diversity of North and South Uist sticklebacks are influenced by the pH of loch water (chapter 2). My main observation was that migratory anadromous stickleback collected from both islands did not show any remarkable variation in length or armour traits, but resident salt and freshwater fish in North Uist exhibited much more variation in all such traits than South Uist. This may indicate that the restricted environmental variance associated with abiotic factors of the loch water on South Uist may be less favourable to the diversification underpinning the local adaptation of stickleback lineages. For example, stickleback possessing remarkable reduction of armour plates and spines in acidic lochs

of North Uist were absent in the neutral to slightly alkaline lochs of South Uist. The pH of water is correlated with various properties of a water body, such as salinity, dissolved calcium ions and nutrient content, which may affect the presence or absence of predators and competitors as well as the whole ecosystem (MacColl et al., 2013; Magalhaes et al., 2016; Smith et al., 2020). However, the influence of other factors such as temperature, water depth, resource competitors or predators could be examined in South Uist lochs to assess the morphological and genetical diversification of closely-related lineages.

Environmental divergence of a natural hybrid zone also highlighted the phenotypic variation between freshwater and marine stickleback and revealed hybrids with intermediate characteristics. Through a classical cline-fit model (Derryberry et al., 2014), I observed a general decreasing trend in the majority of morphological traits in relation to distance from the sea due to environmental selection regimes (chapter 3). Specifically, I found steep transitions near hybrid zone 1 of the Hosta stream with a narrow width and centre for fish length and total plate count, which suggests their bimodal distribution is due to low dispersal and strong disruptive selection (Jiggins and Mallet, 2000; Pedersen et al., 2017). It also suggests the existence of a premating barrier for reproductive isolation associated with body size and plate number due to strong assortative mating between two divergent forms of stickleback (Jiggins and Mallet, 2000; Jones et al., 2006). In addition, clines for armour traits including spines, body shape and parasite abundance were broader with a slight transition and unimodal distribution across the small-scale geographical range of the hybrid zone, suggesting limited reproductive isolation among the hybrids due to weak selection on such traits (Gay et al., 2008). The pattern of phenotypic clines with significantly noncoincident and discordant centres and widths might be the action of divergent selection regimes for each type of adaptive trait with various functions in fitness and survival in nature (Parsons et al., 1993; Vines et al., 2016; Morales-Rozo et al., 2017). However, a preliminary study on putative hybrid zone based on only morphological variation of armour traits also supported the consistant hybrid generation under selection over the 100 – 200 m distance of the Hosta stream (Gostick, 2008). Further studies on different natural hybrid zones could examine the effect of spatial variation with a smooth transition of salinity, or by targeting a single factor of selection for examining the changes in the patterns of each phenotypic trait.

Equivalent environmental selection pressure may be responsible for the evolution of similar phenotypic traits as adaptations to changing ecological conditions (Bolnick et al., 2018). Divergence in the number of lateral plates of stickleback between derived

freshwater populations and ancestral marine population is considered as the most common example of a parallel evolutionary trait (Barrett, 2010). A recent study demonstrated an association between body length and lateral armour plates which indicates the adaptation of the low plated morph as a consequence of the small size of stickleback in response to environmental selection (e.g., pH and temperature) (Smith et al., 2020). In this hybrid zone study, I found significant associations between stickleback body size (length) and shape with each other among different populations of Hosta stream and loch, which suggests their similar adaptation in equivalent environmental selection. Similarly, armour traits including pelvic girdle and spines were associated to body shape, which may indicate the adaptive advantage for protection against different types of predators in the Hosta steam and loch (Reimchen, 1994). Sex of fish also showed a significant relationship with almost all morphological traits such as body size, shape, armour girdle and spines, suggesting that sexual dimorphism may also play a vital role in phenotypic divergence in environmental adaptation (Kitano et al., 2007; Albert et al., 2008; Aguirre et al., 2008; Ouyang et al., 2018). However, observation of specific phenotypic traits related to environmental agents of selection still requires more attention to understand the adaptive evolution of stickleback.

6.2 Understanding the genetic basis of adaptation driven by natural selection

To characterize the genetic variation in the environmental adaptation of stickleback, I have applied a candidate gene approach for some previously identified adaptive loci that appear to be under selection, and which are involved in parallel divergence into different habitats. For this, the most convenient and cost-effective PCR-RFLP method was chosen to detect specific SNPs from the selected segment of a panel of 10 candidate genes related to salinity and immunity. These genes were: 1. PPARA (Peroxisome proliferator-activated receptor alpha a), 2. WNT7B (Wingless-type MMTV integration site family, member 7Ba), 3. ATP1A1 (ATPase Na+/K+ Transporting Subunit Alpha 1), 4. CSMD1 (CUB and Sushi multiple domains 1), 5. MFAP1(Microfibrillar-associated protein 1), 6. ABCD3 (ATP-binding cassette, sub-family A (ABC1), member 3) for salinity and 7. LEMD3 (LEM domain containing 3), 8. NLRC5 (NLR family CARD domain containing 5), 9. MUC2 (Mucin 2), 10. MUC5B (Mucin 5B) for immunity. The main limitation of this method is the application of an appropriate restriction endonuclease enzyme discriminating target SNPs within the alignment of amplified PCR product. For example, I could not find any fixed SNP differences between freshwater and saltwater stickleback of North Uist after scanning different segments of the ATP1a1 gene, which is important for osmoregulation and might be due to the inversion of three large regions between the two ecotypes (Hohenlohe et al., 2010; Jones, et al., 2012; Terekhanova et al., 2014; Vines et al., 2016). It is also time-consuming to develop a complete assay for a specific

gene and I was successful in developing only three assays from a panel of ten genes in this study. Here, I identified specific SNP differences in PPARA, WNT7B and NLRC5 loci, which proved to be fixed in ancestral marine and diverged freshwater sticklebacks collected from North Uist. I validated those using assays of the genotypes of hybrid populations of freshwater and saltwater stickleback in a natural hybrid zone, as well as F1 and F2 generations of laboratory-crossed populations (chapter 3 and 4). The major objective was to establish these SNP assays as 'non-neutral nuclear markers' for adaptive loci to examine the role of natural selection in driving adaptive evolutionary change in stickleback.

In chapter 3, I observed significant genetic differentiation across the Hosta hybrid zone, with results echoing that of several previous studies on stickleback hybrid zones between estuary and freshwater environments that represent a gradient for many environmental factors such as salinity, vegetation and predators (Jones, et al., 2006; Taugbøl et al., 2014; Ravinet et al., 2015; Vines et al., 2016; Pedersen et al., 2017). Here, selection varies across the hybrid zone, shown through the divergence in allele frequencies of four loci (Eda, PPARA, WNT7B and NLRC5) among different populations of three groups: ancestral anadromous fish (GROG), hybrid zone fish and freshwater loch fish. The most common Eda locus corresponding to plate morph of stickleback exhibited a bimodal distribution and step cline with a narrow centre and width which indicates strong selection for a sharp transition in the lower part of the hybrid zone, close to the sea. In contrast, PPARA and WNT7B loci showed shallow transitions in the upper part of the hybrid zone with a narrow width, indicating strong direct selection and the existence of reproductive isolation among populations of the Hosta stream and loch fish. The cline for the NLRC5 locus was broader in width, corresponding to a unimodal distribution across the hybrid zone that indicates comparatively weaker selection than other loci (Szymura and Barton, 1991; Gay et al., 2008). Generally, genetic differences may cluster together as concordance clines to move as one unit which allows the occurrence of sub-species (Barton and Hewitt, 1985; Barton, 1993). The apparent shift of the clines in this study might be the genetic additivity which showed some tendency to dominance by the marine allele at the Eda locus and freshwater allele at PPARA and WNT7B loci (Rohwer and Wood, 1998; Brumfield et al., 2001). Other reasons for discordant and noncoincidence clines of such loci might be due to their particular effects on fitness as a selective advantage, or to reduce the frequency of deleterious gene combinations by asymmetrical gene flow or introgression (Barton, 1993; Jaarola, et al., 1997).

Another explanation for the strong genetic diversity in the Hosta hybrid zone can be observed from the average Fst value (0.261) for all loci, indicative of an overall

intermediate level of inbreeding among populations. The highest genetic differentiation was seen in pairwise comparisons of pure freshwater (loch) and anadromous (GROG) stickleback, indicating the existence of reproductive isolation consistent with other studies on stickleback (Jones, et al., 2006; Taugbøl et al., 2014). The significant deviation from HWE and high positive Fst value for most of the populations, especially inside the loch, indicates that a heterozygote deficit and strong selection, rather than drift, contributes to the divergence of hybrid zone stickleback. This may result from strong linkage disequilibrium (LD) among Eda-PPARA and PPARA-WNT7B loci located in the same chromosome (IV), which suggests that strong endogenous selection (independently of the environment) is influencing the genetic differentiation of the hybrid zone, especially in the population of downstream and loch fish. The association of such genomic loci can extend the effects of selection at one locus to others, which may create a genome-wide barrier to introgression (Barton and Hewitt, 1989). Furthermore, I investigated the frequency of trans-Atlantic (At) and European (EU) lineages based on SNPs differences in the mitochondrial Cyt b gene and demonstrated that there was an absence of the At lineage across the Hosta stream, but that it was found in both GROG and loch fish. This interesting genetic differentiation might be the effect of selection on the mitochondrial divergence due to physiological adaptation for freshwater colonization of stickleback (Orti, et al., 1994; Mäkinen and Merilä, 2008; Ishikawa et al., 2019). However, it has been recommended to examine as many genetic markers of adaptive and non-adaptive loci as possible to describe the genetic architecture of a natural hybrid zone (Brumfield et al., 2001). If the clines of a hybrid zone are maintained by a balance between random dispersal and selection against hybrids, it can be considered as a 'tension zone' (Barton and Hewitt, 1989). I did not attempt to estimate the strength of selection and rate of recombination related dispersal, but this could be addressed at a fine scale in the Hosta hybrid zone.

Genotype-phenotype associations provide an important evolutionary biological insight into the role of genetic changes that underlie phenotypic variation and adaptation to divergent environments (Peichel and Marques, 2017). Here, Eda, PPARA and WNT7B were found to be important pleiotropic genes that had large effects underlying major adaptive phenotypic traits of stickleback. Different populations of Hosta stream and loch fish showed significant relationships with these loci in relation to divergence in morphological traits including body size, shape, plate morph, pelvic girdle and all spines. The association of the Eda locus with lateral plate count has long been the focus of research into the genetic basis of adaptation under selection: this locus tends to account for around 80% of the variation in plate number in stickleback (Colosimo et al., 2004, 2005). This study reported an association between PPARA and plate count, which

showed a similar trend to Eda and plate morph, indicating the similar effects of selection on such loci. PPARA locus also exhibited a significant association with divergent body shape and armour traits, which might be due to its consistent physiological function influenced by strong environmental selection. The WNT7B locus related to the Wnt signalling pathway involved in the development of the fish body also showed a strong association with body size and armour traits including pelvic girdle and spines.

The Eda gene has already been suggested to have pleiotropic effects (simultaneous effects of a single gene on multiple phenotypic traits) for a diverse range of phenotypic variations related to growth, pigmentation, behaviour, development of sensory system and immune response (Barrett et al., 2009a; Barrett et al., 2009b; Mills et al., 2014; Robertson et al., 2017). Previously, PPARA and WNT7B loci have been identified as genes in a genomic region under selection for the adaptive divergence of salt and freshwater stickleback (Jones, et al., 2012). Nevertheless, this study reported association of additional multiple phenotypic traits, which to my knowledge, is the first evidence that relates body size, shape, defence and anti-predators features to the pleiotropic effects of these loci. Another reason might be the tight genetic linkage of Eda, PPARA and WNT7B loci, that control such an extensive phenotypic association due to strong selection under divergent environmental conditions (Slatkin, 2008). The occurrence of phenotypic co-variation due to allelic divergence at a single locus (pleiotropy) is also evident in some plants and animals such as the flowering plant *Petunia* (Hermann et al., 2013), *Boechera stricta* (Prasad et al., 2012), fruit fly *Drosophilamelanogaster* (Carbone et al., 2006; Wittkopp and Beldade, 2009), deer mice *Peromyscus maniculatus* (Linnen et al., 2013) and humans (Flint and Mackay, 2009). Further investigation into the phenotype, genotype and related fitness effects of these alleles could be examined at a broader scale in stickleback or another model organism, to understand the potential effects of mutations on shaping divergent phenotypes in environmental adaption.

6.3 Phenotypic and genotypic basis of parasite-mediated selection

Selection imposed by parasites as a biotic agent can influence the physical, behavioural, or defence system of host organisms, or may reduce viability or fecundity of immigrants and hybrids to facilitate the process of speciation (Karvonen and Seehausen, 2012). Previous work on North Uist stickleback demonstrated parasite-mediated divergent selection by describing persistent heterogeneity in infections among different populations inhabiting different geographical areas (De Roij et al., 2011). In chapter 4, I investigated relationships between infection status of ectoparasite and endoparasite abundance, and armour plate phenotypes of stickleback in natural and laboratory environments. In a

natural hybrid zone, fish exhibiting the partially plated morph showed higher resistance to *Gyrodactylus,* while low plated fish morphs were more resistant to *Diplostomum*. This might be due to the influence of the host condition in relation to different environmental factors such as salinity, or the presence of other parasites in maintaining local adaptation of common saltwater ectoparasite and freshwater endoparasite of stickleback. A similar response to partially plated fish was also observed in the experimental conditions with this ectoparasite, which may indicate a genetic control of the immune system for exhibiting the resistance to this group of parasites. This type of co-variation between host population phenotype and parasite was evident in previous studies of stickleback in North Uist (MacColl and Chapman, 2010; De Roij, 2011; El Nagar and MacColl, 2016; Robertson et al., 2017).

In contrast, the high abundance of *Gyrodactylus* in the completely plated morph might be due to local selection for easy attachment of this leech-like monogenean ectoparasite in the rough surface of stickleback skin. I also found a significant association between a higher growth rate and the partially plated morph stickleback in the experimental environment, which might be related to the low infection levels of this parasite as a fitness consequence. Sex of fish was associated with only ectoparasite burden in wild and laboratory experiments which indicates a potentially complex relationship between this parasite and host phenotypic traits for sexual selection and cost of mating behaviour. The mode of infection influenced by the complex life cycle of the digenean *Diplostomum, which* depends on multiple hosts (Birds, snails and fish), might cause differences in response to infection with monogenean, single host, hermaphroditic *Gyrodactylus* (Barrett et al., 2008)*.* However, the interaction of other morphological and physiological features related to adaptive and innate immune responses of fish could be associated with parasite-mediated selection of stickleback.

The role of host genetics in the occurrence of parasite-mediated selection is an interesting evolutionary ecology topic of host-parasite interactions (Little and Ebert, 2000). Genetic variability of stickleback in relation to parasite susceptibility, both in natural and experimental conditions, may contribute to local adaptation between marine and freshwater populations (MacColl and Chapman, 2010; El Nagar and MacColl, 2016; Robertson et al., 2017). I found a strong association between Eda, PPARA and WNT7B loci with both species of parasite, which suggests that parasites may act as a selective agent on these loci. Previous studies on the Eda locus of stickleback already demonstrated its association with immune gene expression and *G. arcuatus* susceptibility in a semi-natural saltwater environment: higher parasitic abundance was seen in low plated fish corresponding to Eda L allele (Robertson et al., 2017). In chapter 4, I

conducted an infection experiment in freshwater with F2 hybrids with *G. gasterostei* collected from freshwater stickleback and found a higher abundance in the homozygous marine genotype corresponding to completely plated fish. This opposite pattern suggests a complex relationship, with possibly additional environmental factors (such as stressdue to salinity change) determining the parasite susceptibility in relation to host genotype. The Eda genotype also showed a significant association with abundance of the endoparasite *Diplostomum* in an artificial infection experiment using laboratory-crossed F1 hybrids of Hosta population. The fish carrying the homozygous marine allele (CC) experienced a higher infection burden corresponding to plate morph which may be due to susceptibility in the freshwater environment.

The PPARA locus was also associated with the growth of infected stickleback under experimental conditions, as well as *Gyrodactylus* abundance under natural conditions. This indicates a genetic effect on the ectoparasite burden of stickleback. The strong association of the WNT7B locus with only *Diplostomum* under both natural and experimental conditions indicates a significant genetic effect on this endoparasite. The pattern of *Diplostomum* abundance was contrasting in the two groups: fish carrying the homozygous marine allele exhibited the lowest burden in the wild and highest burden in the control experiment. PPARA and WNT7B genes are known to be associated with physiological aspects of fish and other organisms, such as the metabolism of fatty acids and developmental signalling of kidney function. They have also been identified as major loci in saltwater-freshwater as well as basic-acidic divergence of stickleback (Yu et al., 2009; Jones et al., 2012; Wafer et al., 2017; Haenel et al., 2019). This study demonstrated for the first time an association of parasite resistance with PPARA and WNT7B loci, which could explain the genetic basis of parasite-mediated selection. This might be the pleiotropic effect of each of these adaptive loci for multiple phenotypic traits in response to environmental selection. The variation in response of Eda, PPARA and WNT7B loci to different types of parasites (ectoparasite and endoparasite) in different environmental conditions, suggests that their genetic effects for parasite susceptibly are individual, not a result of linkage as a haplotype block. Future study could be based on other factors related to immunity of fish, including the fitness effect, which are required to better elucidate the actual reason for variation in genetic response of these loci to parasitic infection.

6.4 The effects of environmental differences in shaping stickleback microbiome

The role of the environment in shaping the skin-associated bacterial communities can provide information to understand the effects of selection on transition phase of migratory anadromous stickleback between marine and freshwater. In chapter 5, I

applied culture-dependent PCR-DGGE method with 16sRNA gene (V1-V9 and V3 region) sequencing techniques which was able to distinguish only culturable bacteria in a selected nutrient medium representing around 10% of total microflora (Amann et al., 1995; Hugenholtz et al., 1998). I observed a significant variation in the diversity of microbiomes in response to variable water chemistry (salinity, pH) and ecotypes (anadromous and resident) of stickleback on North Uist. This suggests that there might be some genetic or physiological responses of stickleback to microbiomes which provides some distinctiveness of bacterial abundance on specific fish populations (Eichmiller et al., 2016).

Qualitative analysis of the DGGE technique showed comparatively different banding patterns for the V3 region of 16s RNA in anadromous and resident fish collected from a saltwater (Ardh) and a freshwater (Hosta) loch. Simpson's and Shannon diversity indices also showed significant differences among the identified genera of different ecological groups, suggesting an influence of the environment in shaping the diversity of stickleback microbiome. For example, there was clear distinctiveness for some bacterial taxa which are found specifically in saltwater (*Comamonas* sp., *Exiguobacterium* sp., *Flavobacterium* sp., Gamma proteobacterium, *Halomonas* sp., *Marinomonas* sp. and *Yersinia* sp.) and in freshwater (*Carnobacterium* sp., *Enterobacteriaceae* sp., *Frigoribacterium* sp., *Klebsiella* spp., *Providencia* sp. and *Stenotrophomonas* sp.). The influence of salinity (Horsley, 1973; Lozupone and Knight 2007; Sullam et al., 2012; Wong and Rawls, 2012; Schmidt et al., 2015; Lokesh and Kiron, 2016) and pH (Chu et al., 2010; Sylvain et al., 2016) are evident in shaping bacterial flora of other fish species which supports this finding for skin microbiota of stickleback. However, the preliminary study based on quantitative analysis of bacterial load did not show any significant differences in the skin samples of contrasting environments, including freshwater-saltwater, high pH-low pH and anadromous-resident stickleback. This pattern was expected considering the continuous exposure of the skin surface to the surrounding water affecting similar colonization method of bacteria to fish. The present data might be insufficient to provide concrete evidence of skin microbiome divergence between freshwater and saltwater environments due to the comparatively low sample size, or culture-dependent approaches for a vast diverse microbiome study. Further studies based on culture-independent method such as environmental metagenomics are likely to be required, as well as more extensive sampling at different times of the year, in order to provide conclusive results on microbiome divergence.

6.5 Concluding remarks

The main goal of the research in this thesis was to investigate morphological and molecular variations in the evolutionary 'supermodel' stickleback fish in response to various selection pressures between freshwater and saltwater ecotypes and their role in environmental adaptation and speciation. I have shown that water pH, salinity, macro-parasite abundance and the skin microbiome all act as integral agents of selection and influence the morphological and genotypic diversity of stickleback for environmental adaptation. I pinpointed the occurrence of divergence due to precise mutations in the adaptive loci driving several adaptive phenotypic traits, in order to understand the evolutionary mechanism of stickleback from the ancestral marine to freshwater form. I found that single genetic changes in Eda, PPARA and WNT7B loci have pleiotropic effects on multiple phenotypic traits of stickleback, which might facilitate the adaptation process in natural populations. This high-resolution candidate gene mapping will provide the base line study for identifying specific genes in a particular genomic region in association with adaptive phenotypic traits of fish under the influence of abiotic and biotic agents of natural selection. However, several issues remain to be addressed including experimental manipulation for a comprehensive understanding of evolutionary diversification and the potential adaptive consequences of these genes or more genes underlying phenotypic variation in divergent environment.

References

Aguirre, W.E., Ellis, K.E., Kusenda, M. and Bell, M.A., 2008. Phenotypic variation and sexual dimorphism in anadromous threespine stickleback: implications for postglacial adaptive radiation. *Biological Journal of the Linnean Society*, *95*(3), pp.465-478.

Albert, A.Y., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R., Balabhadra, S., Kingsley, D.M. and Schluter, D., 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution: International Journal of Organic Evolution*, *62*(1), pp.76-85.

Amann, R.I., Ludwig, W. and Schleifer, K.H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, *59*(1), pp.143-169.

Barrett, R.D., Rogers, S.M. and Schluter, D., 2009a. Environment specific pleiotropy facilitates divergence at the Ectodysplasin locus in threespine stickleback. *Evolution: International Journal of Organic Evolution*, *63*(11), pp.2831-2837.

Barrett, R.D., Vines, T.H., Bystriansky, J.S. and Schulte, P.M., 2009b. Should I stay or should I go? The Ectodysplasin locus is associated with behavioural differences in threespine stickleback. *Biology Letters*, *5*(6), pp.788-791.

Barrett, L.G., Thrall, P.H., Burdon, J.J. and Linde, C.C., 2008. Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in ecology & evolution*, *23*(12), pp.678-685.

Barrett, R.D.H., 2010. Adaptive evolution of lateral plates in three-spined stickleback Gasterosteus aculeatus: A case study in functional analysis of natural variation. *Journal of Fish Biology*, *77*(2), pp.311-328.

Barton, N.H. and Hewitt, G.M., 1985. Analysis of hybrid zones. *Annual review of Ecology and Systematics*, *16*(1), pp.113-148.

Barton, N.H. and Hewitt, G.M., 1989. Adaptation, speciation and hybrid zones. *Nature*, *341*(6242), pp.497-503.

Barton, N.H., 1993. Why species and subspecies?. *Current Biology*, *3*(11), pp.797-799.

Bolnick, D.I., Barrett, R.D., 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.

Brumfield, R.T., Jernigan, R.W., McDonald, D.B. and Braun, M.J., 2001. Evolutionary implications of divergent clines in an avian (Manacus: Aves) hybrid zone. *Evolution*, *55*(10), pp.2070-2087.

Carbone, M.A., Jordan, K.W., Lyman, R.F., Harbison, S.T., Leips, J., Morgan, T.J., DeLuca, M., Awadalla, P. and Mackay, T.F., 2006. Phenotypic variation and natural selection at Catsup, a pleiotropic quantitative trait gene in Drosophila. *Current Biology*, *16*(9), pp.912-919.

Chu, H., Fierer, N., Lauber, C.L., Caporaso, J.G., Knight, R. and Grogan, P., 2010. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology*, *12*(11), pp.2998-3006.

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.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. and Kingsley, D.M., 2004. The genetic architecture of parallel armor plate reduction in three spine sticklebacks. *PLoS biology*, *2*(5), p.e109.

De Roij, J. 2011. *PhD Thesis - Spatial variation in host-parasite interactions in the Three-spined stickleback*, University of Nottingham, U.K.

De Roij, J., Harris, P.D. and MacColl, A.D., 2011. Divergent resistance to a monogenean flatworm among three-spined stickleback populations. *Functional ecology*, *25*(1), pp.217-226.

Derryberry, E.P., Derryberry, G.E., Maley, J.M. and Brumfield, R.T., 2014. HZAR: hybrid zone analysis using an R software package. *Molecular ecology resources*, *14*(3), pp.652-663.

Eichmiller, J.J., Hamilton, M.J., Staley, C., Sadowsky, M.J. and Sorensen, P.W., 2016. Environment shapes the fecal microbiome of invasive carp species. *Microbiome*, *4*(1), pp.1-13.

El Nagar, A. and MacColl, A.D., 2016, August. Parasites contribute to ecologically dependent postmating isolation in the adaptive radiation of three-spined stickleback. In *Proc. R. Soc. B*. 283(1836), p. 20160691. The Royal Society.

Endler, J.A., 1986. *Natural selection in the wild*. Princeton University Press.

Flint, J. and Mackay, T.F., 2009. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome research*, *19*(5), pp.723-733.

Gay, L., Crochet, P.A., Bell, D.A. and Lenormand, T., 2008. Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution: International Journal of Organic Evolution*, *62*(11), pp.2789-2806.

Giles, N., 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *Journal of zoology*, *199*(4), pp.535-544.

Gostick D., 2008. *MS thesis*: Morphological variation and selection within a putative hybrid zone of the three-spined stickleback *Gasterosteus aculeatus*, University of Nottingam, UK.

Haenel, Q., Roesti, M., Moser, D., MacColl, A.D. and Berner, D., 2019. Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution letters*, *3*(1), pp.28-42.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T. and Kuhlemeier, C., 2013. Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in Petunia. *Current Biology*, *23*(10), pp.873-877.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A., 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genet*, *6*(2), p.e1000862.

Horsley, R.W., 1973. The bacterial flora of the Atlantic salmon (*Salmo salar* L.) in relation to its environment. *Journal of Applied Microbiology*, *36*(3), pp.377-386.

Hugenholtz, P., Goebel, B.M. and Pace, N.R., 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of bacteriology*, *180*(18), pp.4765-4774.

Ishikawa, A., Kabeya, N., Ikeya, K., Kakioka, R., Cech, J.N., Osada, N., Leal, M.C., Inoue, J., Kume, M., Toyoda, A. and Tezuka, A., 2019. A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science*, *364*(6443), pp.886-889.

Jaarola, M., Tegelström, H. and Fredga, K., 1997. A contact zone with noncoincident clines for sex-specific markers in the field vole (Microtus agrestis). *Evolution*, *51*(1), pp.241-249.

Jiggins, C.D. and Mallet, J., 2000. Bimodal hybrid zones and speciation. *Trends in ecology & evolution*, *15*(6), pp.250-255.

Jones, F.C., Brown, C., Pemberton, J.M. and Braithwaite, V.A., 2006. Reproductive isolation in a three-spine stickleback hybrid zone. *Journal of evolutionary biology*, *19*(5), pp.1531-1544.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S. and Birney, E., 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), p.55.

Karvonen, A. and Seehausen, O., 2012. The role of parasitism in adaptive radiations—when might parasites promote and when might they constrain ecological speciation?. *International Journal of Ecology*, *2012*.

Kingsolver, J.G. and Diamond, S.E., 2011. Phenotypic selection in natural populations: what limits directional selection?. *The American Naturalist*, *177*(3), pp.346-357.

Kitano, J., Mori, S. and Peichel, C.L., 2007. Sexual dimorphism in the external morphology of the threespine stickleback (Gasterosteus aculeatus). *Copeia*, *2007*(2), pp.336-349.

Linnen, C.R., Poh, Y.P., Peterson, B.K., Barrett, R.D., Larson, J.G., Jensen, J.D. and Hoekstra, H.E., 2013. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, *339*(6125), pp.1312-1316.

Little, T.J. and Ebert, D., 2000. The cause of parasitic infection in natural populations of Daphnia (Crustacea: Cladocera): the role of host genetics. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1457), pp.2037-2042.

Lokesh, J. and Kiron, V., 2016. Transition from freshwater to seawater reshapes the skin-associated microbiota of Atlantic salmon. *Scientific reports*, *6*, p. 19707–19716.

Lozupone, C.A. and Knight, R., 2007. Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, *104*(27), pp.11436-11440.

MacColl, A.D. and Aucott, B., 2014. Inappropriate analysis does not reveal the ecological causes of evolution of stickleback armour: a critique of Spence et al. 2013. *Ecology and Evolution*, *4*(18), pp.3509-3513.

MacColl, A.D. and Chapman, S.M., 2010. Parasites can cause selection against migrants following dispersal between environments. *Functional ecology*, *24*(4), pp.847-856.

MacColl, A.D., 2011. The ecological causes of evolution. *Trends in Ecology & Evolution*, *26*(10), pp.514-522.

MacColl, A.D., Nagar, A.E. and de Roij, J., 2013. The evolutionary ecology of dwarfism in three-spined sticklebacks. *Journal of Animal Ecology*, *82*(3), pp.642-652.

Magalhaes, I.S., D'Agostino, D., Hohenlohe, P.A. and MacColl, A.D., 2016. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular ecology*, *25*(17), pp.4319-4336.

Mäkinen, H.S. and Merilä, J., 2008. Mitochondrial DNA phylogeography of the three-spined stickleback (Gasterosteus aculeatus) in Europe—evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, *46*(1), pp.167-182.

Mills, M.G., Greenwood, A.K. and Peichel, C.L., 2014. Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo*, *5*(1), pp.1-10.

Morales-Rozo, A., Tenorio, E.A., Carling, M.D. and Cadena, C.D., 2017. Origin and cross-century dynamics of an avian hybrid zone. *BMC evolutionary biology*, *17*(1), pp.1-18.

Orti, G., Bell, M.A., Reimchen, T.E. and Meyer, A., 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution*, *48*(3), pp.608-622.

Ouyang, X., Gao, J., Xie, M., Liu, B., Zhou, L., Chen, B., Jourdan, J., Riesch, R. and Plath, M., 2018. Natural and sexual selection drive multivariate phenotypic divergence along climatic gradients in an invasive fish. *Scientific reports*, *8*(1), pp.1-22.

Parsons, T.J., Olson, S.L. and Braun, M.J., 1993. Unidirectional spread of secondary sexual plumage traits across an avian hybrid zone. *Science*, *260*(5114), pp.1643-1646.

Pedersen, S.H., Ferchaud, A.L., Bertelsen, M.S., Bekkevold, D. and Hansen, M.M., 2017. Low genetic and phenotypic divergence in a contact zone between freshwater and marine sticklebacks: gene flow constrains adaptation. *BMC evolutionary biology*, *17*(1), pp.1-16.

Peichel, C.L. and Marques, D.A., 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1713), p.20150486.

Prasad, K.V., Song, B.H., Olson-Manning, C., Anderson, J.T., Lee, C.R., Schranz, M.E., Windsor, A.J., Clauss, M.J., Manzaneda, A.J., Naqvi, I. and Reichelt, M., 2012. A

gain-of-function polymorphism controlling complex traits and fitness in nature. *science*, *337*(6098), pp.1081-1084.

Ravinet, M., Hynes, R., Poole, R., Cross, T.F., McGinnity, P., Harrod, C. and Prodöhl, P.A., 2015. Where the lake meets the sea: strong reproductive isolation is associated with adaptive divergence between lake resident and anadromous three-spined sticklebacks. *PloS one*, *10*(4), p.e0122825.

Reimchen, T.E., 1994. Predators and morphological evolution in threespine stickleback. *The evolutionary biology of the threespine stickleback*, pp.240-276.

Robertson, S., Bradley, J.E. and MacColl, A.D., 2017. Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific reports*, *7*, p.42677.

Rohwer, S. and Wood, C., 1998. Three hybrid zones between Hermit and Townsend's warblers in Washington and Oregon. *The Auk*, *115*(2), pp.284-310.

Schmidt, V.T., Smith, K.F., Melvin, D.W. and Amaral-Zettler, L.A., 2015. Community assembly of a euryhaline fish microbiome during salinity acclimation. *Molecular ecology*, *24*(10), pp.2537-2550.

Slatkin, M., 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*(6), pp.477-485.

Smith, C., Zięba, G., Spence, R., Klepaker, T. and Przybylski, M., 2020. Three-spined stickleback armour predicted by body size, minimum winter temperature and pH. *Journal of Zoology*, *311*(1), pp.13-22.

Sullam, K.E., Essinger, S.D., Lozupone, C.A., O'CONNOR, M.P., Rosen, G.L., Knight, R.O.B., Kilham, S.S. and Russell, J.A., 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular ecology*, *21*(13), pp.3363-3378.

Sylvain, F.É., Cheaib, B., Llewellyn, M., Correia, T.G., Fagundes, D.B., Val, A.L. and Derome, N., 2016. pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (Colossomamacropomum). *Scientific reports*, 6: 32032.

Szymura, J.M. and Barton, N.H., 1991. The genetic structure of the hybrid zone between the fire-bellied toads Bombina bombina and B. variegata: comparisons between transects and between loci. *Evolution*, *45*(2), pp.237-261.

Taugbøl, A., Junge, C., Quinn, T.P., Herland, A. and Vøllestad, L.A., 2014. Genetic and morphometric divergence in threespine stickleback in the Chignik catchment, Alaska. *Ecology and Evolution*, *4*(2), pp.144-156.

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. and Mugue, N.S., 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. *PLoS Genet*, *10*(10), p.e1004696.

Vines, T.H., Dalziel, A.C., Albert, A.Y., Veen, T., Schulte, P.M. and Schluter, D., 2016. Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution*, *70*(5), pp.1023-1038.

Wadgymar, S.M., Lowry, D.B., Gould, B.A., Byron, C.N., Mactavish, R.M. and Anderson, J.T., 2017. Identifying targets and agents of selection: innovative methods to

evaluate the processes that contribute to local adaptation. *Methods in Ecology and Evolution*, *8*(6), pp.738-749.

Wafer, R., Tandon, P. and Minchin, J.E., 2017. The role of peroxisome proliferator-activated receptor gamma (PPARG) in adipogenesis: Applying knowledge from the fish aquaculture industry to biomedical research. *Frontiers in endocrinology*, *8*, p.102.

Wittkopp, P.J. and Beldade, P., 2009, February. Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. In *Seminars in cell & developmental biology* (Vol. 20, No. 1, pp. 65-71). Academic Press.

Wong, S. and Rawls, J.F., 2012. Intestinal microbiota composition in fishes is influenced by host ecology and environment.*Molecular Ecology*, 21(13), pp.3100–3102.

Xue, B., Sartori, P. and Leibler, S., 2019. Environment-to-phenotype mapping and adaptation strategies in varying environments. *Proceedings of the National Academy of Sciences*, *116*(28), pp.13847-13855.

Yin, A., Korzh, S., Winata, C.L., Korzh, V. and Gong, Z., 2011. Wnt signaling is required for early development of zebrafish swimbladder. *PloS one*, *6*(3), p.e18431.

Yu, J., Carroll, T.J., Rajagopal, J., Kobayashi, A., Ren, Q. and McMahon, A.P., 2009. A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. *Development*, *136*(1), pp.161-171.

Appendix

Supplementary materials of chapter 3-5

S1. Method of SNPs assay development for PPARA, WNT7B and NLRC gene

For PPARA gene, we have selected 610bp from ChrIV: 19851440-19852050 region of groupIV:19849404-19859507 (contig 7763, Ensembl, BROAD S1) targeting 17 SNPs filtered from the whole genome sequencing data of Scotland stickleback. From 3 pairs of primers designed for BstUI restriction enzyme, PCR amplification of 331bp was successful by primer pair 2 (chapter 3, Table 2) with very specific band. After treatment with BstUI enzyme (according to manufacturer protocol, NEB, UK), three types of bands were observed in -a single uncut band for marine PP allele (331bp), two cut bands (273 and 58 bp) for freshwater RR allele and three bands (331, 273 and 58 bp) for heterozygous allele (PR) having one copy of each allele (chapter 3, Fig. 11A). To confirm this results, 3 samples were sequenced from each group of samples and found 4 SNPs differences between freshwater and marine water fish.

For WNT7B gene, we have selected 1000bp from ChrIV:19900521:19901520 region of groupIV:19899773-19905382 (contig 7763, Ensembl, BROAD S1) targeting 15 SNPs filtered from the previous whole genome sequencing data of the laboratory. From 3 pairs of primers designed for SacII restriction enzyme, PCR amplification of 429bp was successful by primer pair 1 (chapter 3, Table 2) with very specific band but SacII enzyme did not cut the PCR product. Then 2 samples were sequenced from each group of samples to get the sequence and found a deletion with three SNPs differences between freshwater and marine water fish to design for another enzyme. After treated with Msal enzyme, three types of bands were observed in -a single uncut band for freshwater TT allele (429bp), two cut bands (333 and 96 bp) for marine WW allele and three bands for heterozygous (WT) form (chapter 3, Fig 11B).

For NLRC5 gene, we have selected 100bp from ChrXIX: 2453701:2454700 region of groupXIX: 2446925-2473806 (contig 8695, Ensembl, BROAD S1) targeting 32 SNPs filtered from the reference whole genome sequencing data. From 2 primer pair designed for EcoRI restriction enzyme, PCR amplification of 460bp was successful by primer pair 1(chapter 3, Table 2) with very specific band. After treatment with EcoRI enzyme, it did not cut the amplified PCR product. Then we sent two samples for sanger sequencing to find out any differences between freshwater and marine fish and found 3 SNPs differences between them. From the sequence, we designed new enzyme AluI for RFLP

treatment and found three types of bands: a single uncut band for Freshwater LL allele (460bp), two cut bands (355 and 125 bp) for marine NN allele and three bands (460, 355 and 125 bp) for heterozygous allele (NL) sharing from the parent allele (chapter 3, Fig. 11C).
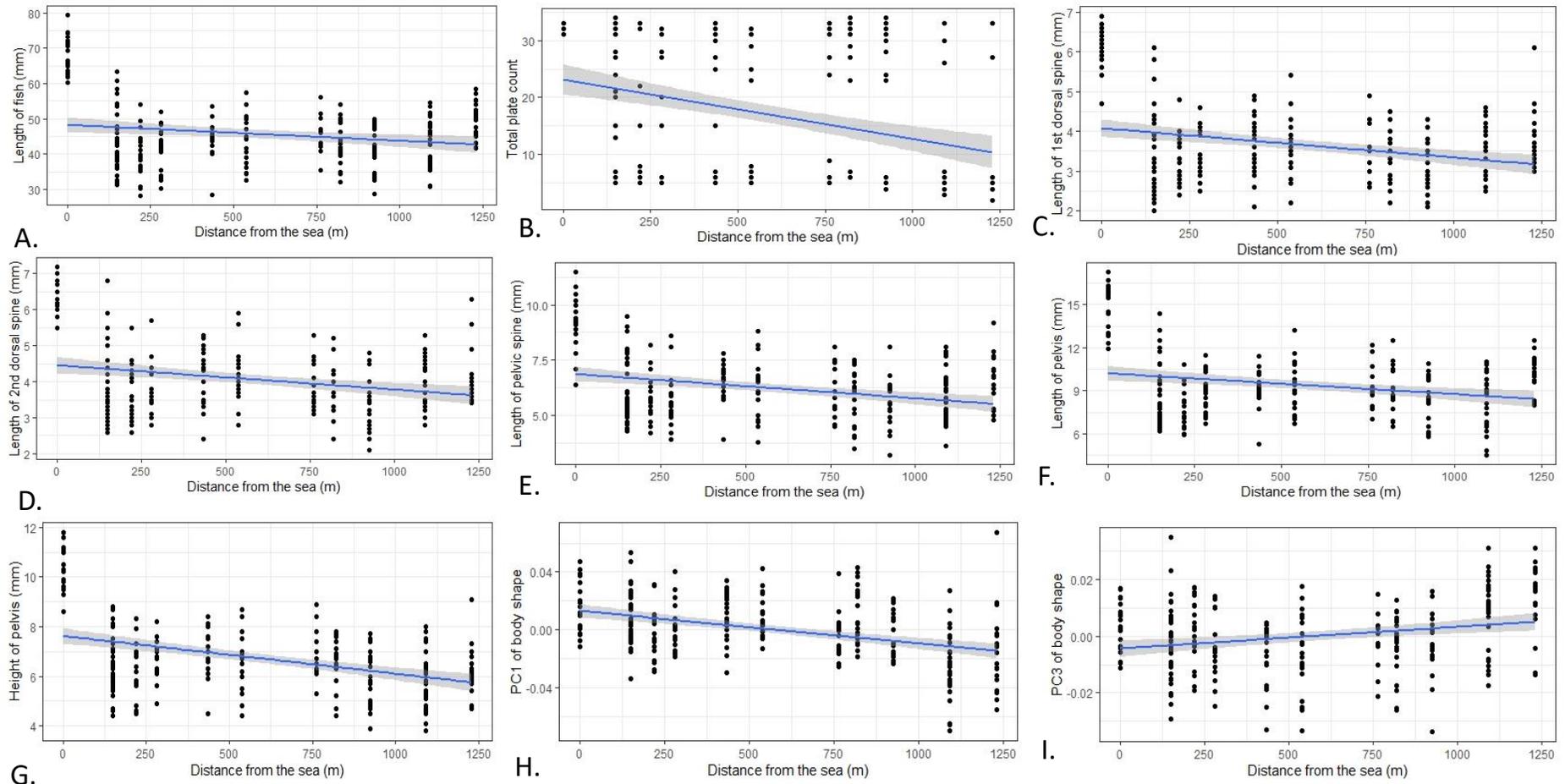
Fig. S1 A-I Morphological traits: standard length, total plate count, length of 1st dorsal spine, 2nd dorsal spine, pelvic spine, length and height of pelvis, PC1 and PC3 of body shape for individual fish showing variation with a gradual decreasing trend (except shape PC3) relative to distance from the sea of Hosta hybrid zone.

Table S1: Summary of the genotypes of four genes (Eda, PPARA, WNT7B and NLRC5) typed collected from the 10 sites of hybrid zone including loch and 1 site of GROG, North Uist, Scotland. The genotype CC/ PP/ WW/ NN = homozygous marine allele, CL/ PR/ WT/ NL = heterozygote and LL/ RR/ TT/ LL = homozygous freshwater allele.

Genotype / Site	Distance (m)	No. of fish	Eda (n = 236)			PPARA (n = 236)			WNT7B (n = 231)			NLRC5 (n = 233)		
			CC	CL	LL	PP	PR	RR	WW	WT	TT	NN	NL	LL
GROG	0	20	19	1	0	20	0	0	20	0	0	17	2	1
HB Zone 1	150	34	5	11	18	21	13	0	17	4	12	16	4	14
HB Zone 2	220	20	1	7	12	14	6	0	12	3	5	4	3	13
HB Zone 3	280	19	2	10	7	15	4	0	13	4	2	10	4	3
HB Zone 4	434	20	3	9	8	14	4	2	13	4	3	15	1	4
HB Zone 5	537	20	0	9	11	12	8	0	14	4	1	6	2	12
HB Zone 6	762	13	1	5	7	9	3	1	6	5	2	5	6	2
HB Zone 7	820	20	2	8	10	16	3	1	9	4	6	1	5	13
HB Zone 8	925	20	2	11	7	14	6	0	8	6	6	3	6	11
Loch Zone 9	1090	30	0	5	25	5	9	16	4	2	23	3	8	19
Loch Zone 10	1230	20	1	1	18	2	2	16	1	2	16	4	4	12

Table S2: Summary of the phenotype (Plate morph) and genotypes of four genes typed for 100 samples of the *Gyrodactylus* sp. infected fish in experiment 2.

Sl	Genetic factors	Homozygous marine	Heterozygous	Homozygous freshwater
1.	Eda gene	17	60	23
2.	PPARA	24	56	20
3.	WNT7B	47	39	14
4.	NLRC5	60	34	6
	Phenotype	Completely-plated	Partially-plated	Low-plated
5.	Plate morph	61	21	18

Table S3: Summary of the phenotype (Plate morph) and genotypes of four genes typed for 90 samples of the *Diplostomum* sp. infected stickleback in experiment 3.

Sl	Genetic factors	Homozygous marine	Heterozygous	Homozygous freshwater
1.	Eda gene	21	31	38
2.	PPARA	50	38	2
3.	WNT7B	56	25	9
4	NLRC5	54	17	19
	Phenotype	Completely-plated	Partially-plated	Low-plated
5.	Plate morph	32	29	29

S2. Protocol for isolation of DNA using QIAamp Power Faecal DNA kit (QIAGEN, USA)

100 µl of preserved bacteria culture sample was added into Dry Bead tube with 750 µl of Power Bead Solution along with 60 µl of C1 solution and vortexed briefly. Samples were then heated at 65^0C for 10 mins in a heat block incubator and the tissue was disrupted in a mini-Bead beater (Bio Spec) at maximum speed (1800 rpm) at 25^0C for 10minutes.The tubes were then centrifuged (Thermo Scientific, UK) (all centrifugation steps were conducted at 13,000×g) for 1 minute and 500 µl of supernatant transferred into a new 2ml collection tube (DNase and RNase free). 250 µl of C2 solution was added and vortex for 1 minute and incubated at 8^0C for 5 minutes. After centrifugation as above, 600 µl of supernatant was collected into a new 2ml collection tube and 200 µl C3 solution was added along with subsequent vortex (1 min), incubation (5 mins at 8^0C) and centrifugation. After that 750 µl of supernatant was transferred into a new 2ml collection tube by avoiding the pellet and added 1200 µl C4 solution and vortex for 5sec.

Then about 650 μl of supernatant was loaded into a MB spin column and centrifuged. Then discarded the flow through and repeated until all the supernatant had been processed (nearly 3 times load). After that about 500 μl C5 solution was added for washing and centrifuged. Discarded the flow through and centrifuged the tubes as above again to remove any residual and placed the MB spin column in a new 2ml collection tube. DNA was eluted using 100 ul of C6 solution in the final step of extraction and preserved at -20^{0}C before further analysis.

S3. Protocol for DNA purification using NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL, Germany)

Firstly, the collected gel was weighted and added double volume of Buffer NT1 (e.g. for 100 mg gel, 200 μl of NT1 was added) incubated for 5–10 min at 50 °C until the gel dissolved. Then 700 μl of sample was loaded into NucleoSpin Clean-up Column and centrifuged at 11,000 x g for 30 sec. To wash, 700μL Buffer NT3 was added to the Column and repeated the centrifugation step as above. The flow-through was discarded and centrifuged again for 1 min to remove the NT3 buffer completely. Finally, 15 μl of pre-heated (70^{0}C) NE buffer was added into NucleoSpin Column set in a new 1.5 mL micro-centrifuge tube. The DNA was collected after 1 min incubation and centrifugation as above and stored at -20^{0}C for further analysis.

S4. Protocol for quantification of DNA using Qubit ds DNA BR Assay Kit (Life Technologies)

For working solution, 2000 μl of Qubit dsDNA BR Buffer was mixed with 1 μL of Qubit dsDNA BR Reagent. The aliquots of 198 μl of working solution were distributed into 0.5 ml tubes. Then 2 μl of sample was added and vortexed (3-5 seconds), incubated for 2 minutes at room temperature and read on Qubit Fluorometer (Invitrogen).

Table S4: Mixed culture bacterial colony characteristics of e-swab and culture swab of stickleback skin samples cultured in aerobic and anaerobic conditions, collected from seven lochs of North Uist (SC = Scad, TM = Torm, TR = Tros, GR = Grog, HS = Hosta, AR = Ardh and DN = Duin).

Swab type	Fish/ swab Id.	No. of selected colonies	Type of culture	Size of colony	Colour of colony	Shape of colony	Appearance of colony
E-swab	SC 97	2	Aerobic	A. Medium B. Large	A. Creamy B. Whitish	A. Round & convex B. Oval, deep in the middle	A. Shiny B. Shiny
	SC 98	1	Aerobic	Medium	Creamy	Round and convex in the middle	Dull
	TM 48	1	Aerobic	Medium	Light pink	Oval, thick in the middle	Dull
	TM 50	2	Aerobic	A. Large B. Small	A. Creamy B. Whitish	A. Round, convex & thick B. Round & smooth surface	A. Shiny B. Shiny
	TM 51	1	Aerobic	Small	White	Round	Shiny
	TR 81	4	Aerobic	A. Small B. Large C. Large D. Small	A. White B. Yellow C. White D. Dark yellow	A. Round B. Round C. Irregular D. Round	A. Shiny B. Shiny C. Shiny D. Shiny
	GR 5	2	Aerobic	A. Medium B. Large	A. Dark yellow B. Creamy	A. Round and convex B. Oval and convex in the middle	A. Shiny B. Dull
	GR 6	1	Aerobic	Large	Whitish	Round and convex in the middle	Shiny
	HS 39	1	Aerobic	Large	Yellow	Round, convex	Shiny
	HS 40	3	Aerobic	A. Large B. Large C. Small	A. Pinkish B. Creamy C. White	A. Round, convex B. Round, convex in the middle C. Round	A. Dull B. Shiny C. Shiny

Table S4: (Cont.)

Swab type	Fish/ swab Id.	No. of selected colonies	Type of culture	Size of colony	Colour of colony	Shape of colony	Appearance of colony
E-swab	AR 74	4	Aerobic	A. Medium B. Large C. Large D. Large	A. Golden B. White C. Creamy D. Creamy	A. Round and thin B. Irregular, concave C. Round and thick D. Pentagonal	A. Transparent B. Dull C. Shiny D. Shiny
	AR 19	3	Aerobic	A. Large B. Very large C. Medium	A. Creamy B. Off-white C. Whitish	A. Round B. Irregular, concave C. Round, concave	A. Shiny B. Dull C. Shiny
	AR 20	2	Aerobic	A. Large B. Medium	A. Whitish B. Golden	A. Round B. Round	A. Shiny B. Shiny
	DN 110	4	Aerobic	A. Large B. Medium C. Medium D. Medium	A. Creamy B. White C. White D. Whitish	A. Fluffy, irregular round B. Round, concave C. Round, convex D. Star-shape, concave	A. Shiny B. Dull C. Shiny D. Dull
	DN 108	3	Aerobic	A. Small B. Large C. Medium	A. White B. Creamy C. Orange	A. Flower-shape, concave B. Round C. Round	A. Dull B. Shiny C. Shiny
Culture-swab	TS 252	1	Anaerobic	Small	White	Irregular	Shiny
	TS 264	1	Aerobic	Medium	Whitish	Circular	Slightly shiny
	HS 322	4	A. Aerobic B. Aerobic C. Anaerobic D. Anaerobic	A. Medium B. Large C.Small C. Large	A. Deep brown B. Light brown C. White D. Cream	A. Irregular, Slightly Oval B. Circular, convex C. Circular D. Irregular	A. Dull B. Dull C. Dull D. Shiny
	HS 332	2	A. Aerobic B. Anaerobic	A. Large B. Large	A. Brownish B. Cream	A. Irregular, convex B. Irregular	A. Dull & opaque B. Shiny

Table S4: (Cont.)

Swab type	Fish/ swab Id.	No. of selected colonies	Type of culture	Size of colony	Colour of colony	Shape of colony	Appearance of colony
Culture-swab	AR 200	2	A. Aerobic B. Anaerobic	A. Large B. Tiny	A. Cream B. White	A. Circular B. Circular	A. Dull B. Shiny & transparent
	AR 210	2	A. Aerobic B. Anaerobic	A. Medium B. Tiny	A. Cream B. Transparent	A. Regular & flat B. Not Clear	A. Shiny B. Not clear

Table S5: Identified bacteria through the amplification of V1-V9 and V3 region (DGGE) of 16sRNA from the mixed and pure culture of e-swab (ES) and culture-swab (CS) colonies in aerobic and anaerobic conditions in two ecotypes of stickleback: resident (Resi) and anadromous (Anad) collected from seven lochs [three freshwater low pH, two freshwater(FW) high pH and two saltwater(SW) loch] of North Uist, Scotland.

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
1	V1-V9	CS	AR 210	210 AP2	Aerobic	Ardh	Resi	SW	SW	*Psychrobacter sp.*	100%	328
2	V1-V9	CS	AR 210	210ANAp1	Anaerobic	Ardh	Resi	SW	SW	*Marinomonas sp.*	99%	172
3	V1-V9	CS	AR 200	200 AP2	Aerobic	Ardh	Anad	SW	SW	*Shewanella baltica*	100%	85
4	V1-V9	CS	AR 200	200ANAP1	Anaerobic	Ardh	Anad	SW	SW	*Shewanella* sp.	99%	102
5	V1-V9	CS	TR 252	252ANAP2	Anaerobic	Tros	Anad	FW	FW low pH	*Shewanella* sp.	99%	149
6	V1-V9	CS	TR 264	264 AP1	Aerobic	Tros	Resi	FW	FW low pH	*Pseudomonas sp.*	99%	228
7	V1-V9	CS	HS 322	322 AP3	Aerobic	Hosta	Resi	FW	FW high pH	*Shewanella glacialipiscicola*	99%	228
8	V1-V9	CS	HS 322	322 AP2	Aerobic	Hosta	Resi	FW	FW high pH	*Shewanella baltica*	94%	112
9	V1-V9	CS	HS 322	322ANAP1	Anaerobic	Hosta	Resi	FW	FW high pH	*Shewanella putrefaciens*	100%	98
10	V1-V9	CS	HS 322	322ANAP2	Anaerobic	Hosta	Resi	FW	FW high pH	*Aeromonas sp.*	100%	84
11	V1-V9	CS	HS 332	332 AP2	Aerobic	Hosta	Anad	FW	FW high pH	*Shewanella putrefaciens*	100%	128

Table S5: (Cont.)

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
12	V1-V9	CS	HS 332	332ANAP3	Anaerobic	Hosta	Anad	FW	FW high pH	*Aeromonas* sp.	100%	132
13	V1-V9	E-swab	SC 97	SC1_A	Aerobic	Scad	Resi	FW	FW low pH	*Pseudomonas syringae*	98%	589
14	V1-V9	E-swab	SC 97	SC1_B	Aerobic	Scad	Resi	FW	FW low pH	*Pseudomonas poae*	98%	808
15	V1-V9	E-swab	SC 98	SC_C	Aerobic	Scad	Resi	FW	FW low pH	*Pseudomonas baetica*	98%	1274
16	V1-V9	E-swab	TM 48	TM_A	Aerobic	Torm	Resi	FW	FW low pH	*Shewanella sp.*	87%	248
17	V1-V9	E-swab	TM 50	TM_B	Aerobic	Torm	Resi	FW	FW low pH	*Shewanella morhuae*	90%	344
18	V1-V9	E-swab	TM 50	TM_C	Aerobic	Torm	Resi	FW	FW low pH	*Pseudomonas fluorescens*	99%	605
19	V1-V9	E-swab	TM 51	TM_D	Aerobic	Torm	Resi	FW	FW low pH	*Shewanella morhuae*	87%	742
20	V1-V9	E-swab	TR 81	TR_A	Aerobic	Tros	Resi	FW	FW low pH	*Carnobacterium maltaromaticum*	99%	656
21	V1-V9	E-swab	TR 81	TR_B	Aerobic	Tros	Resi	FW	FW low pH	*Enterobacteriaceae bacterium*	98%	619
22	V1-V9	E-swab	TR 81	TR_C	Aerobic	Tros	Resi	FW	FW low pH	No significant similarity found		
23	V1-V9	E-swab	TR 81	TR_D	Aerobic	Tros	Resi	FW	FW low pH	*Frigoribacterium sp.*	98%	1199
24	V1-V9	E-swab	GR 5	GR_A	Aerobic	Gros	Resi	FW	FW high pH	*Stenotrophomonas rhizophila*	97%	1020
25	V1-V9	E-swab	GR 5	GR_B	Aerobic	Gros	Resi	FW	FW high pH	*Pseudomonas gessardii*	98%	908
26	V1-V9	E-swab	GR 6	GR_C	Aerobic	Gros	Resi	FW	FW high pH	*Psychrobacter cibarius*	98%	1189

Table S5: (Cont.)

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
27	V1-V9	E-swab	HS 40	HS_A	Aerobic	Hosta	Resi	FW	FW high pH	*Shewanella baltica*	85%	915
28	V1-V9	E-swab	HS 40	HS_B	Aerobic	Hosta	Resi	FW	FW high pH	*Providencia alcalifaciens*	97%	1092
29	V1-V9	E-swab	HS 39	HS_C	Aerobic	Hosta	Resi	FW	FW high pH	*Klebsiella aerogenes*	96%	1119
30	V1-V9	E-swab	HS 40	HS_D	Aerobic	Hosta	Resi	FW	FW high pH	*Carnobacterium maltaromaticum*	98%	918
31	V1-V9	E-swab	AR 74	AR_A	Aerobic	Ardh	Resi	SW	SW	*Comamonas jiangduensis*	98%	719
32	V1-V9	E-swab	AR 74	AR_B	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas psychrophila*	96%	971
33	V1-V9	E-swab	AR 74	AR_C	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas fluorescens*	97%	684
34	V1-V9	E-swab	AR 74	AR_D	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas fluorescens*	99%	650
35	V1-V9	E-swab	AR 19	AR_E	Aerobic	Ardh	Anad	SW	SW	*Aeromonas hydrophila*	99%	1116
36	V1-V9	E-swab	AR 19	AR_F	Aerobic	Ardh	Anad	SW	SW	*Psychrobacter cibarius*	98%	388
37	V1-V9	E-swab	AR 19	AR_G	Aerobic	Ardh	Anad	SW	SW	*Pseudomonas baetica*	96%	947
38	V1-V9	E-swab	AR 20	AR_H	Aerobic	Ardh	Anad	SW	SW	*Psychrobacter cibarius*	99%	700
39	V1-V9	E-swab	AR 20	AR_I	Aerobic	Ardh	Anad	SW	SW	*Yersinia kristensenii*	98%	1171
40	V1-V9	E-swab	DN 110	DN_A	Aerobic	Duin	Resi	SW	SW	*Shewanella putrefaciens*	95%	366
41	V1-V9	E-swab	DN 110	DN_B	Aerobic	Duin	Resi	SW	SW	*Pseudomonas psychrophila*	99%	607

Table S5: (Cont.)

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
42	V1-V9	E-swab	DN 110	DN_C	Aerobic	Duin	Resi	SW	SW	Uncultured *Pseudomonas sp.*	89%	175
43	V1-V9	E-swab	DN 110	DN_D	Aerobic	Duin	Resi	SW	SW	*Pseudomonas psychrophila*	98%	1119
44	V1-V9	E-swab	DN 108	DN_E	Aerobic	Duin	Anad	SW	SW	*Pseudomonas psychrophila*	99%	594
45	V1-V9	E-swab	DN 108	DN_F	Aerobic	Duin	Anad	SW	SW	*Pseudomonas brenneri*	99%	614
46	V1-V9	E-swab	DN 108	DN_G	Aerobic	Duin	Anad	SW	SW	*Exiguobacterium undae*	99%	892
47	V3	CS	HS 336	11AH	Aerobic	Hosta	Anad	FW	FW high pH	*Aeromonas rivuli*	97%	159
48	V3	CS	HS 332	7BH	Aerobic	Hosta	Anad	FW	FW high pH	Uncultured *Aeromonas* sp.	96%	160
49	V3	CS	HS 332	5GH	Aerobic	Hosta	Anad	FW	FW high pH	*Aeromonas sobria*	96%	133
50	V3	CS	HS 334	12HH	Aerobic	Hosta	Anad	FW	FW high pH	*Aeromonas rivuli*	98%	158
51	V3	CS	HS 336	15HH	Aerobic	Hosta	Anad	FW	FW high pH	*Shewanella sp.*	97%	159
52	V3	CS	HS 332	14HH	Aerobic	Hosta	Anad	FW	FW high pH	*Aeromonas sobria*	96%	161
53	V3	CS	HS 332	3FH	Anaerobic	Hosta	Anad	FW	FW high pH	*Shewanella putrefaciens*	94%	168
54	V3	CS	HS 336	10AH	Anaerobic	Hosta	Anad	FW	FW high pH	*Aeromonas sobria*	97%	138
55	V3	CS	HS 332	10BH	Anaerobic	Hosta	Anad	FW	FW high pH	*Aeromonas salmonicida*	97%	158

Table S5: (Cont.)

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
56	V3	CS	HS 322	1BH	Aerobic	Hosta	Resi	FW	FW high pH	*Shewanella putrefaciens*	94%	161
57	V3	CS	HS 322	6AH	Aerobic	Hosta	Resi	FW	FW high pH	*Shewanella putrefaciens*	98%	160
58	V3	CS	HS 328	10CH	Aerobic	Hosta	Resi	FW	FW high pH	*Aeromonas* sp.	99%	159
59	V3	CS	HS 328	13HH	Aerobic	Hosta	Resi	FW	FW high pH	*Aeromonas* sp.	99%	159
60	V3	CS	HS 326	8AH	Anaerobic	Hosta	Resi	FW	FW high pH	*Uncultured bacterium*	91%	131
61	V3	CS	HS 328	9AH	Anaerobic	Hosta	Resi	FW	FW high pH	*Klebsiella pneumoniae*	93%	100
62	V3	CS	HS 322	4BH	Anaerobic	Hosta	Resi	FW	FW high pH	Uncultured *Aeromonas* sp.	99%	156
63	V3	CS	AR 202	7a	Aerobic	Ardh	Anad	SW	SW	*Shewanella baltica*	99%	148
64	V3	CS	AR 202	1BA	Aerobic	Ardh	Anad	SW	SW	Uncultured *Shewanella* sp	98%	120
65	V3	CS	AR 202	8aA	Aerobic	Ardh	Anad	SW	SW	Uncultured *Shewanella* sp.	99%	143
66	V3	CS	AR 204	12BR	Aerobic	Ardh	Anad	SW	SW	*Shewanella baltica*	94%	164
67	V3	CS	AR 200	1FA	Anaerobic	Ardh	Anad	SW	SW	Uncultured bacterium	97%	135
68	V3	CS	AR 202	1EA	Anaerobic	Ardh	Anad	SW	SW	*Flavobacterium sp.*	97%	151
69	V3	CS	AR 202	9AA	Anaerobic	Ardh	Anad	SW	SW	*Gamma proteobacterium*	96%	164

Table S5: (Cont.)

Sl no.	Sequence region	Swab type	Fish id	Colony id	Culture type	Location	Ecotype	Habitat (salinity)	Habitat (pH)	Bacterial species	NCBI identity matching	Amplified base pair
70	V3	CS	AR 200	7FA	Anaerobic	Ardh	Anad	SW	SW	*Shewanella baltica*	97%	158
71	V3	CS	AR 214	6bR	Anaerobic	Ardh	Resi	SW	SW	*Flavobacterium sp.*	93%	148
72	V3	CS	AR 214	8FA	Anaerobic	Ardh	Resi	SW	SW	*Pseudomonas sp.*	98%	150
73	V3	CS	AR 210	5gA	Aerobic	Ardh	Resi	SW	SW	Uncultured bacteria	92%	166
74	V3	CS	AR 210	10aR	Aerobic	Ardh	Resi	SW	SW	*Psychrobacter aquimaris*	96%	162
75	V3	CS	AR 212	11aR	Aerobic	Ardh	Resi	SW	SW	*Halomonas sp.*	98%	158
76	V3	CS	AR 214	3BA	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas sp.*	96%	103
77	V3	CS	AR 212	8BR	Aerobic	Ardh	Resi	SW	SW	*Flavobacterium sp.*	97%	156
78	V3	CS	AR 212	1DA	Aerobic	Ardh	Resi	SW	SW	*Flavobacterium sp.*	97%	152
79	V3	CS	AR 214	13AR	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas sp.*	93%	163
80	V3	CS	AR 214	8CA	Aerobic	Ardh	Resi	SW	SW	*Pseudomonas sp.*	98%	144
81	V3	CS	AR 212	8EA	Aerobic	Ardh	Resi	SW	SW	No significant similarity found		

*13 species matched with below 95%