Evolutionary genomics of pelvic spine reduction in *Gasterosteus aculeatus*

Dhurba Adhikari

A thesis for the degree of
Philosophiae Doctor (PhD)

PhD in Biosciences no. 61 (2024)
Faculty of Biosciences and Aquaculture

PhD in Biosciences no. 61 (2024)
Dhurba Adhikari
Evolutionary genomics of pelvic spine reduction in *Gasterosteus aculeatus*

© **Dhurba Adhikari**
ISBN: 978-82-93165-60-6

Print: Trykkeriet Nord University

Nord University
N-8049 Bodø
Tel: +47 75 51 72 00
www.nord.no

Preface

The thesis is submitted in fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Faculty of Biosciences and Aquaculture (FBA), Nord University, Bodø, Norway. The studies included in this dissertation represent original research as a part of the Stipendiat program.

The PhD project team consisted of the following members:

Dhurba Adhikari, M.Sc., FBA, Nord University: PhD candidate

Jarle Tryti Nordeide, Professor, FBA, Nord University : Main supervisor

Steinar Daae Johansen, Professor, FBA, Nord University: Co-supervisor

Truls Borg Moum, Professor, FBA, Nord University: Co-supervisor

Dhurba Adhikari,

Bodø, Norway, January 2024

Acknowledgement

I would like to express my sincere gratitude to my main supervisor, Prof. Jarle Tryti Nordeide, for his invaluable guidance, support, and suggestions throughout my PhD journey. His generosity with time, whether for problem-solving, paperwork, or writing, has been truly appreciated. His efforts have significantly eased my life at Nord University during my PhD project.

I am equally grateful to my co-supervisors, Prof. Steinar Daae Johansen and Prof. Truls Borg Moum, for their immense support, help, and insights. Their expertise in genomics and evolutionary biology has been the linchpin of this project, which would not have been completed without their masterful contributions.

Special thanks to Dr. Bård Ove Karlsen from Nordland Hospital for his support in bioinformatics analysis, innovative ideas, and assistance in paper writing, which have been invaluable. I extend my appreciation to Prof. Joost Andre M Raeymaekers, head of the Genomics Department, for his support during my PhD.

My gratitude also goes to Dr. Lars Martin Jakt, Dr. Tor Erik Jørgensen, Dr. Joanna Babiak, Dr. Irina Vladimirovna Smolina, Martina Elisabeth Luise Kopp, Jeanett Stegen, Guro Gill Prestegård Aars, and Irene Stork for their support during my PhD. A heartfelt thank you to all staff members at Nord University who have assisted me in various ways throughout my PhD.

I am thankful to my colleagues Dr. Deepak Pandey, Arif Khan, Dr. Prabhugouda, Aruna Mattanavile Shankregowda, and Chinmayi for their support during my PhD. I also would like to thank my Nepalese friends Dr. Sudip Tiwari, Dr. Ranjan Parajuli, Manoj Rimal, Samjhana Timalsina, Dr. Gagan Poudel, Bharati Mainali, Sushil Banjade, Rupa Thapa, Surya Prasad Dahal, Alina Thapa, and Sujan Khadka for their support during my PhD. I would like to extend my gratitude to my friend Dr. Isha Gharti for her suggestion, motivation and unconditional support during my PhD. I also would like to

extend my thanks to my friends Saroj Bajracharya, Dr. Shiva Bhandari, Dr. Namrata Basnet, Pratik Bhattarai, Sanjib Acharya, Ambika Aryal, and Nitu Basnet for their encouragement.

I owe a debt of gratitude to my parents for their unconditional support, love, motivation, and suggestions they provided during my life and PhD. Without their support, I would not have been able to get to where I am. I love you, Dad and Mum. I would also like to give special mention to my sister Radhika, my brother Bijaya, my brother-in law Bir, and my niece Rabina for their support and motivation. I would also like to thank to extended family members, including my uncles, aunts, cousins, and others (Anju, Upendra, Sabitri, Umakanta, Prince, Sudip, Surakshya, Priyanka, and so on) for their unwavering support and inspiration.

I would like to thank Nord University for providing me with the opportunity to pursue my PhD. And, finally, I am grateful to God, the divine, an unseen power, for blessings that have guided me through this journey of life.

Contents

List of figures and tables

List of abbreviations

AER	Apical ectodermal ridge
AF	Allele frequency
BMP	Bone morphogenic protein
CGV	Cryptic genetic variation
COI	Cytochrome C oxidase subunit I
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
EDA	Ectodysplasin
FGF	Fibroblast growth factor
GWAS	Genome-wide association studies
Hand2	Heart and neural crest derivatives expressed 2
Hox	Homeobox
JNK	Jun-N-terminal-kinase
mtDNA	Mitochondrial DNA
NGS	Next generation sequencing
nDNA	nuclear DNA
NAD1	NADH dehydrogenase subunit 1
PCGs	Protein-coding genes
Pitx1	Pituitary homeobox transcription factor 1
QTL	Quantitative trait locus
RAD	Restriction site associated DNA
RA	Retinoic acid
rRNA	Ribosomal RNA
SGV	Standing genetic variation
Shh	Sonic hedgehog
Tbx4/5	T-box transcription factor 4/5
TADs	Topologically associating domains

tRNA	Transfer RNA
WNT	Wingless/Integrated
ZPA	Zone of polarizing activity

List of papers

Paper I: Adhikari, D.; Hanssen, I.K.; Johansen, S.D.; Moum, T.B. & Nordeide, J.T. (2023). Pitx1 Enhancer Variants in Spined and Spine-Reduced Subarctic European Sticklebacks. *Fishes*, 8: 164.

Paper II: Karlsen, B.O.; **Adhikari, D.**; Jørgensen, T.E,; Hanssen, I.K.; Moum, T.B.; Nordeide J.T. & Johansen S.D. (2023). Two distinct maternal lineages of threespine stickleback (*Gasterosteus aculeatus*) in a small Norwegian subarctic lake. (Manuscript)

Paper III: Adhikari, D.; Karlsen, B.O.; Johansen S.D.; Nordeide J.T. & Moum, T.B. (2023). The genomics of postglacial vicariance and freshwater adaptations in European subarctic threespine sticklebacks. (Manuscript)

Abstract

Marine threespine sticklebacks (*Gasterosteus aculeatus*) have colonised numerous freshwater lakes since the last ice age, exhibiting a notable loss of body armour, such as reduced lateral plate numbers and reduced pelvic structures. Especially, the loss of lateral plates is a classic example of parallel evolution. The reduction of pelvic structures in sticklebacks is a rarer phenomenon and has previously been associated with the recurrent yet varying deletions within the pelvic enhancer regions *Pel*A and *Pel*B. The two enhancers regulate the expression of the paired-like homeodomain transcription factor gene *Pitx*1. Comprehensive genome-wide studies have further corroborated the adaptive nature of pelvic reduction in freshwater sticklebacks. An upper lake in a subarctic Norwegian watercourse, known for its unusual pelvic spine morphology in threespine sticklebacks, contains a mix of completely spined, asymmetrically spined, and spineless specimens. Contrasting this, downstream lakes and a nearby marine site contain only the completely spined morph. The main aims of this study are *(i)* to examine the molecular variations between the pelvic enhancers of spined and spine-reduced sticklebacks, and *(ii)* to explore the genome-wide diversity and differentiation among sticklebacks from the three populations.

Sticklebacks were collected from the two small freshwater lakes and a nearby marine site in subarctic Norway. Sanger sequencing was used to examine the variation in nucleotide sequences at the pelvic enhancers *Pel*A and *Pel*B from the three sites. Moreover, next-generation sequencing methods were used to sequence the mitogenomes and the nuclear genomes from sticklebacks, after pooling DNA from group of 40 specimens from each of the three sites. In addition, the mitogenome of two single specimens from the upper lake was sequenced individually. The sequences from the pooled samples were analysed by the PoPoolation tool, to calculate the population genetics parameter: nucleotide diversity (π), Tajima's neutrality statistic (T_D), and the fixation index (F_{ST}).

The Sanger sequencing showed that the variation at *Pel*A between the three sites was mainly due to variable numbers of repeats at three fragile TG-repeat loci at the *Pel*A enhancer region. A unique 58 bp deletion upstream of and adjacent to TG-repeats III of the *Pel*A enhancer was found in all examined specimens from the upper lake but not in the two other sites. No polymorphism was found at *Pel*B among the examined specimens. Next-generation sequencing of the mitogenomes of both two single specimens and the 40 pooled groups detected two distinct Euro-North-American (ENA) mitogenome haplotypes in the upper lake. Only one of them was present in the downstream lake, whereas the marine sample contained more mitogenome diversity. Next-generation sequencing of the 40-pooled groups indicate that the overall nucleotide diversity (π) is generally higher at the marine site than in the two freshwater populations. Nuclear genetic differentiation between the two freshwater populations appeared greater than between the marine population and each of the freshwater populations.

No association was found between the presence of pelvic spine polymorphism in the upper lake and variations in (*i*) the *Pitx*1 enhancers, (*ii*) the 58 bp region upstream of and adjacent to TG-repeats III, or (*iii*) the two mitogenome haplotypes. However, the nuclear genomes of spine reduced and fully spined specimens from the upper lake differed at a specific region at chromosome 9, where a gene *Hand*2, previously implicated in limb development, was identified. We speculate that the polymorphism in pelvic spine length of sticklebacks in the upper lake is due to either (*i*) the short *Pel*A which leads to a tipping point for transcription to occur or not, leaving spined and spine-reduced individuals to develop based on additional genetic factors, or (*ii*) other genomic regions, for example at *Hand*2 at chromosome 9. The first and second scenarios suggest phenotypic parallel evolution caused by the same or different loci, respectively, in this North European lake. Especially, the second scenario is different from previous studies of threespine stickleback populations.

Sammendrag

Marine trepigga stingsild (*Gasterosteus aculeatus*) har kolonisert talrike innsjøar sidan siste istid. Marine stingsild har laterale og dorsale piggar og beinplater på sidene av kroppen for å beskytta dei mot fiendar, og etter ei tid i ferskvatn mistar dei ofte beinplatene og i nokre populasjonar også dei laterale piggane («sidepiggar»). Dette fenomenet er eit døme på parallell evolusjon. Tidlegare studiar, i hovudsak frå Nord-Amerika, har vist at tilbakedanning av sidepiggar skyldast tillegg eller tap av DNA («indels») på kontroll-områda *Pel*A og *Pel*B. Dette er to loci som kontrollerer om genet som kodar for piggane «homeodomain transcription factor gene» (*Pitx*1) vert uttrykt eller ikkje. Tidlegare studie av genomet har vist at heilt eller delvis tap av sidepiggar er ei adaptiv tilpassing i mange stingsild populasjonar i ferskvatn. Ein innsjø øvst i eit vassdrag i Vesterålen i Nord-Norge har ei uvanleg blanding av stingsild med sidepiggar med full lengde, asymmetrisk piggar, og stingsild utan sidepiggar. Stingsild i innsjøen like nedanfor i same vassdrag, og i sjøen like i nærheita, har derimot alle normal lengde på sidepiggane. Formålet med dette studiet er å (*i*) undersøke om skilnader i nukleotid-sekvensane i *Pel*A og *Pel*B er årsak til at nokre stingsild utviklar sidepiggar og andre ikkje også i denne populasjonen, og (*ii*) å studere genetisk diversitet innan i kvar populasjon, og genetisk forskjell mellom dei tre populasjonane.

Stingsild vart fanga i kvar av dei to ferskvatna øvst i Froskelandsvassdraget og i sjøen like ved, i Vesterålen, Norge. Nukleotid sekvensen i kontrollområda *Pel*A og *Pel*B vart sekvensert med Sanger sekvensering. Mitokondrie- og kjerne-DNA vart sekvensert med Neste-generasjon-sekvensering, etter å ha slått saman DNA frå 40 stingsild-individ frå kvar av populasjonane. I tillegg vart mitogenomet til to enkeltfisk frå det øvre vatnet sekvensert individuelt. Sekvensane frå dei samanslåtte gruppene vart analysert med programvara PoPoolation, for å estimere populasjon-genetiske parametrar som nukleotid diversitet (π), Tajimas nøytralitet parameter (T_D), og fikserings-indeks (F_{ST}).

Sanger-sekvenseringa viste at variasjonen i kontroll-området *Pel*A mellom individ frå dei to vatna og sjøen, i hovudsak skyldast ulik tal på repeterte thymine-guanine-sekvensar (TG-sekvensar) i tre ustabile område på kontrollområdet *Pel*A. Lengda på *Pel*A, særleg for ein av TG-sekvensane (TG- sekvens III), var kortare i alle undersøkte stingsild frå det øvre vatnet samanlikna med individ frå dei to andre populasjonane. I tillegg var 58 basepar, oppstraums for denne korte TG-sekvensen, borte i det øvre vatnet samanlikn med dei to andre populasjonane. Alle undersøkte stingsild frå dei tre populasjonane hadde identiske sekvensar på det andre kontrollområdet *Pel*B. Neste-generasjon sekvensering av mitogenomet, både frå dei to enkelt-individa og frå den samanslått gruppa av 40 individ frå det øvre vatnet, viste to ulike Euro-Nord-amerikanske (ENA) mitogenom haplotypar. Berre ein av desse ENA mitogenom haplotypane vart funne i det nedre vatnet. Mitogenomet til dei marine stingsilda innehald meir genetisk variasjon. Neste-generasjon sekvensering av kjerne-DNA i dei samanslåtte gruppene av 40 individ viste generelt høgare nukleotid diversiteten (π) blant stingsild i sjøen enn i dei to ferskvass-innsjøane. Genetisk forskjell var større mellom fisk i dei to ferskvass-innsjøane enn mellom marine stingsild og individ i kvar av dei to ferskvass- innsjøane. Ingen samanheng vart påvist mellom sidepigg-variant (normal, asymmetrisk og ingen sidepigg) og verken (*i*) variasjon i *Pitx*1 kontrollområda, eller (*ii*) dei 58 bp oppstrøms for TG-III, eller (*iii*) dei to mitogenom haplotypane. Kjerne-DNA til individ med reduserte og med normale sidepiggar frå det øvre vatnet var ulike på kromosom 9. *Hand*2 vart identifisert på kromosom 9. Tidlegare studie på andre artar har knytt *Hand*2 til utvikling av lemer.

Det blir spekulert om side-pigg polymorfisme i det øvre vatnet skyldast (*i*) at den korte *Pel*A med ustabil Z-DNA medfører eit vippepunkt for om transkripsjon skjer eller ikkje. Slik at om individ får sidepiggar eller ikkje også vert påverka av andre genetiske faktorar. Alternativt, (*ii*) utvikling av side-piggar vert kontrollert på andre deler av genomet, som t.d. *Hand*2 på kromosom 9. Det siste scenarioet vil i så fall vera døme på parallell evolusjon koda for på ulike stader på genomet. Dette vil i så fall vera ulikt tidlegare studie av stingsild populasjonar med tilbakedanna sidepiggar.

1 Introduction

Evolution is the phenomenon where species undergo changes over generations, accumulating differences from their ancestors as they gradually adapt to varying environments throughout time (Campbell *et al.*, 2020). Evolutionary biology aims to elucidate the diversity among organisms, explore the origins and historical development of this diversity, and understand the natural processes that drive and maintain this diversity (Skelton *et al.*, 1993).

1.1 Local adaptation

Local adaptation is an evolutionary mechanism in which species develop traits that confer an advantage in particular local environments, without considering how those traits affect their fitness in other habitats (Nignan *et al.*, 2022). Natural selection is the primary driving force behind adaptive evolution, elevating the frequency of genetic variants associated with improved reproductive fitness. This facilitates better adaptation to the specific environment in which a population resides (Frankham *et al.*, 2017). Populations adapt to their local environment either through directional, balancing, or diversifying selection. Directional selection increases the occurrence of beneficial alleles in a population and drives local adaptation. Balancing selection maintains genetic diversity by favouring heterozygotes or rare alleles, thereby maintaining polymorphism within a population. This retention of genetic diversity enables the population to better adapt to the changing environment (Brandt *et al.*, 2018, Frankham *et al.*, 2017). Diversifying selection occurs when individuals at both extremes of a phenotype are favoured over those with an intermediate phenotype (Skelton *et al.*, 1993). Over time, this type of selection can lead to distinct subpopulations within the same species, each adapted to different environmental conditions. This form of selection enhances the fitness of a population by responding to its local environmental pressure and contributes to local adaptation (Frankham *et al.*, 2017, White and Butlin, 2021).

1.2 Parallel evolution

Parallel evolution occupies a special position in evolutionary biology as it offers compelling evidence for local adaptation (Bolnick *et al.*, 2018). The definition of parallel evolution that I use in this study is the "independent evolution of the same trait in closely related lineages" (Schluter *et al.*, 2004). However, several different definitions of parallel evolution have been suggested including those by Arendt and Reznick (2008) and others.

Previously, studies have demonstrated adaptations to similar phenotypes either in different populations of the same species, or in different subspecies (Hoekstra *et al.*, 2006, reviewed by Arendt and Reznick, 2008). Arendt and Reznick (2008) discriminated between studies that documented similar phenotypic changes due to (*i*) "different" genetic changes (changes at different loci or different locations within the same locus; 17 studies), and (*ii*) "similar" genetic changes (11 studies). For example, in Florida, several populations of a rodent species (*Peromyscus poliotiotus*), which display variations in fur colour, have been observed (Hoekstra *et al.*, 2006). The mainland populations have relatively dark-coloured fur compared to the light-coloured fur of isolated populations living in sand dunes along the Atlantic coast and the Gulf coast of Florida. The light-coloured fur is presumed to be an adaptation for improved camouflage in the sand dunes. The light-coloured fur of the Gulf coast population is due to a single nucleotide substitution causing a charge-changing amino acid variant in the melanocortin 1 receptor gene (*Mc1r*), whereas this variant is absent in the mainland and the three examined Atlantic coast populations. The authors suggest that the light-coloured fur in the Atlantic coast population is due to mutation(s) in genes other than the *MC1r*. This indicates that the evolution of similar phenotypes in different populations (or subspecies) can result from mutations at various genomic locations, as well as from identical mutations at the same locus. Moreover, several phenotypic traits are coded for at a number of loci in a genetic pathway (Arendt and Reznick, 2008, Le Rouzic *et al.*, 2011, White and Rabago-Smith, 2011). Mutations at

different locations in the genome may affect various genes that are part of the same genetic pathway, which in turn influences the phenotype (Arendt and Reznick, 2008, Sharma *et al.*, 2011).

In isolated populations, parallel evolution could be governed either by standing genetic variation (SGV), which seems to be more common, or more rarely, the occurrence of novel mutations (Schluter and Conte, 2009).

Novel mutation: When the specimens enter into a novel environment, they encounter new selection pressures. In response to these novel challenges, the specimens may adapt through novel (*de novo*) mutations that arise in the population (Barrett and Schluter, 2008). The majority of new mutations are likely to be detrimental, while some are neutral, and only a few are advantageous. Detrimental mutations negatively impact an organism's fitness and are important for understanding how inbreeding can lead to reduced reproductive fitness in a population (Frankham *et al.*, 2017). On the other hand, beneficial mutations contribute to adaptive evolutionary processes (Bisschop *et al.*, 2020, Frankham *et al.*, 2017). Further, neutral mutations do not have a significant impact on an organism's fitness; these mutations are useful to assess demographic processes, such as genetic drift, rates of gene flow, and to construct evolutionary relationships (Frankham *et al.*, 2017).

Standing genetic variation (SGV): SGV refers to the pre-existing genetic variation, or the presence of multiple alleles at particular loci, within a population. When populations already harbour advantageous genetic variants capable of thriving under similar selective pressures, parallel evolution can occur through the independent selection of these pre-existing variants. In fact, adaptation is generally expected to occur more rapidly from SGV (Barrett and Schluter, 2008) compared to novel mutations due to three main reasons: (a) Firstly, the readily available beneficial alleles can immediately act at higher frequencies, eliminating the need to wait for new mutations to arise (Barrett and Schluter, 2008). (b) Secondly, these beneficial alleles are typically older than new mutations and may have undergone previous selection in

different environments (Barrett and Schluter, 2008). Such SGV has already passed through a "selective filter", increasing the likelihood that large-effect alleles are advantageous and enabling a higher probability of parallel evolution (Schluter *et al.*, 2004). (c) Finally, novel mutations usually appear as single copies in a population and are more likely to be lost through genetic drift. However, beneficial alleles from SGV persist in multiple copies at the onset of selection, increasing their chances of becoming fixed in the population (Prezeworski *et al.*, 2005). The probability of fixation increases with the magnitude of the beneficial effect and with a larger effective population size. Furthermore, over a wide range of selection pressure, the probability of fixation is high for SGV while negligible for new mutations (Barrett and Schluter, 2008, Schluter *et al.*, 2004). When adaptation arises from a new mutation (or an allele), the chance of a new allele becoming fixed in a population is shaped by various factors, including the allele's dominance coefficient. Such an allele usually appears at low frequency and is often in heterozygous form. As a result, dominant beneficial alleles tend to rise in frequency more rapidly than their recessive counterparts, boosting their likelihood of becoming permanently established. This concept is commonly referred to as 'Haldane's Sieve' (Barrett and Schluter, 2008, Charlesworth, 1992, Turner, 1981).

Alleles derived from SGV are expected to predominate in most cases when adaptation occurs within short timescales (Hermisson and Pennings, 2005). There are additional factors that could further increase the frequency of alleles present as SGV beyond what is expected. For example, gene flow from populations experiencing different environmental conditions, or even hybridisation with other species, could preserve relatively high amounts of SGV despite negative selection. Alternatively, alleles that are deleterious under specific environmental conditions might be hidden from selection because they do not have any effects on the phenotype in the ancestral environment, and may be referred to as "cryptic genetic variation (CGV)". The CGV, a type of SGV, is the dark matter of biology that does not typically contribute to produce the usual range of observable phenotypes within a population. However, it remains available to modify a phenotype that arises after environmental change or the

introduction of novel alleles, representing an essential source of physiological and evolutionary potential (Gibson and Dworkin, 2004).

1.3 Study of the genomic basis of local adaptation

Identifying the genetic basis of local adaptation is crucial in addressing fundamental questions in evolutionary biology and natural selection. It allows us to investigate whether natural selection predominantly operates on SGV or novel mutations. Additionally, it provides insights into the intriguing question of whether common genetic solutions evolve independently across various populations. Moreover, it allows us to unravel the distribution and nature of adaptive loci, whether they involve regulatory elements, coding sequences, non-coding sequences, or structural components of the genome (Hoban *et al.*, 2016). One of the structural components of a genome involves chromosomal inversions—reversed DNA regions linked to adaptation and speciation. These inversions can shield large chromosomal segments containing numerous genes from recombination and facilitate local adaptation even in the presence of high gene flow (Wellenreuther and Bernatchez, 2018). As a result, they can contribute to genetic differentiation among populations. Inversions are increasingly being investigated for their possible significance in evolutionary processes, spanning from mating systems to environmental adaptation and speciation (Wellenreuther and Bernatchez, 2018).

In recent years, a variety of methods such as genetic differentiation outlier tests, genetic environment association, quantitative trait locus (QTL) mapping (Savolainen *et al.*, 2013), genome-wide association studies (GWAS) (Yang *et al.*, 2011), and population-specific selective sweeps (Fariello *et al.*, 2013) have been used to study the genomic basis of adaptation of an organism (Hoban *et al.*, 2016). Generally, the choice of a particular method depends on the existing knowledge of the biological system under investigation. When traits driving local adaptation are known, the associated genetic regions can be identified with linkage mapping or genome-wide association approaches (Robinson *et al.*, 2014, Savolainen *et al.*, 2013). For rapid selection on novel

mutations or alleles, the genomic signatures of selective sweeps can spotlight population-specific signatures of selection (Hohenlohe *et al.*, 2010b). And, when traits driving local adaptation are unknown, researchers can explore loci with notable genetic differentiation among populations (Beaumont, 2005). Genetic differentiation, usually calculated as F_{ST}, measures the genetic difference between two or more populations (Gregorius, 1987). The value of F_{ST} ranges from 0 to 1; the lower and the higher F_{ST} values indicate lesser and higher difference in the allele frequency, respectively, among the populations (Holsinger and Weir, 2009).

Further, other population genetic parameters based on allele frequency distributions, such as nucleotide (or genetic) diversity (π), and Tajima's D (T_D), allow us to trace population-specific processes as well as signatures of selection at particular loci. Nucleotide diversity is defined as the average number of nucleotide differences per site between DNA sequences in a population sample (Nei and Li, 1979). Gene flow can introduce new alleles into a population, potentially providing genetic material that can be beneficial in the local environment. By increasing genetic diversity, gene flow can enhance the adaptive potential of a population, giving it a broader genetic toolkit to respond to environmental changes.

In addition, random changes in allele frequencies within populations take place due to random sampling effects in finite populations rather than natural selection (Gandon and Nuismer, 2009, Frankham *et al.*, 2017). Genetic drift also arises from both bottleneck and founder events. A bottleneck—characterized by a sharp reduction in population size followed by a subsequent expansion—can dramatically reduce the genetic diversity within a population. In addition, the founder effect occurs when a small group of individuals separates from a larger population to establish a new population in a different habitat, leading to reduced genetic diversity (Daco *et al.*, 2022, Frankham *et al.*, 2017).

Tajima's D is a population genetic test statistic based on the difference between two measures of genetic diversity, the mean pairwise difference (π) between

sequences in a population sample and the number of polymorphic sites (s) (Tajima, 1989). A T_D of approximately zero implies a neutrally evolving sequence, while a negative T_D suggests purifying selection (removing deleterious alleles while preserving beneficial ones), and a positive T_D suggests balancing selection (crucial for maintaining polymorphism) (Biswas and Akey, 2006). T_D is also influenced by demographic processes, however. Specifically, a recent population bottleneck may have little effect on π while s can be substantially reduced giving the locus a transient positive T_D value, whereas a recent population expansion would tend to preserve novel mutations, increasing s more than π and producing a negative T_D value (Tajima, 1989).

Further, if we identify genetic regions associated with adaptations, we can also explore them with the basic molecular biology tools such as PCR, gel electrophoresis, Sanger Sequencing etc. If we need to examine the genome-wide regions, next-generation sequencing has made it possible in recent years.

1.3.1 Next-generation sequencing (NGS)

NGS, also known as high-throughput sequencing, is a powerful technique that enables the generation of vast amounts of DNA sequence data in a single operation. It allows for the sequencing of various targets, including genomes (DNAs), transcriptomes (RNAs), small RNAs, amplicons or large genomic fragments such as BAC clones (Karlsen *et al.*, 2013). This technology enables researchers to efficiently and rapidly analyse genetic material on a large scale, opening up new possibilities for genomic and transcriptomic studies. One of the methods of NGS is pooled DNA sequencing.

Pooled DNA sequencing involves combining equal amounts of DNA from several individuals to create a single DNA pool, which is a cost-effective option for obtaining population genomic data (Anand *et al.*, 2016). However, this pooling approach leads to the loss of information on haplotypes, heterozygosity and linkage disequilibrium (Cutler and Jensen, 2010). Additionally, DNA pooling can pose challenges in accurately calling SNPs and estimating allele frequency (AF), particularly at low coverage and

depth and with a lower number of individuals (Cutler and Jensen, 2010), which can potentially result in false positive variants (Anand *et al.*, 2016). Including a larger number of individuals in a pool (Gautier *et al.*, 2013, Rode *et al.*, 2018) can significantly reduce sequencing errors and improve the accuracy of AF estimation, but achieving accurate AF estimation requires each individual in the pool to contribute an equal amount of DNA (Anand *et al.*, 2016). Furthermore, it is worth noting that most analysis tools were designed initially for individual genome sequencing, and until approximately 20 years ago, no tools were available for analysing pooled sequence data. The PoPoolation tools (Kofler *et al.*, 2011a, Kofler *et al.*, 2011b) were developed to address this gap, which is now commonly used for analysing pooled DNA sequencing data (e.g. Karlsen *et al.*, 2013, Love *et al.*, 2016).

1.3.2 Molecular genetic markers to evaluate local adaptation

Molecular genetic markers in vertebrates can be derived from either nuclear DNA (nDNA) or mitochondrial DNA (mtDNA) (Chan *et al.*, 2021). These markers are employed to detect polymorphisms, which can be in the form of base deletions, insertions, or substitutions between different alleles within a single population or among multiple populations (Guo-Liang, 2013). Analysis of nDNA and mtDNA markers provides estimates of genetic diversity and differentiation. Commonly used DNA markers are Restriction fragment length polymorphism, Simple sequence length polymorphism, Short tandem repeat, and Single nucleotide polymorphism (SNP) (Amiteye, 2021). Among these markers, SNPs are considered the most useful for assessing genetic diversity (Nadeem *et al.*, 2018), due to their biallelic nature, widespread occurrence, dense distribution throughout the genome, low mutation rate, and lack of homoplasy (Brumfield *et al.*, 2003, Zimmerman *et al.*, 2020). SNPs can be found in coding sequences, non-coding regions of genes, or in intergenic regions between genes, with varying frequencies across different chromosomal regions (Guo-Liang, 2013). Advances in molecular markers and genome sequencing have paved the way for a deeper understanding of the extent and patterns of genetic diversity in vertebrates (Nadeem *et al.*, 2018).

Over the past three decades, mtDNA has gained popularity as a marker in vertebrate studies due to its particular characteristics. First, mtDNA can be easily amplified in laboratory settings as it exists in multiple copies within cells (Galtier *et al.*, 2009). Moreover, mitochondrial gene content and gene order are highly conserved across vertebrates, featuring minimal gene duplications and molecular recombination, a complete absence of introns, and short intergenic regions (Gissi *et al.*, 2008). Second, mtDNA exhibits a higher mutation rate compared to nDNA, resulting in greater variation within natural populations and providing insights into population history over shorter timeframes. Third, mtDNA analysis is a convenient and cost-effective method for studying base sequences in natural environments. Assessing the complete mtDNA sequences (mitogenomics) has greatly increased the resolution of mtDNA studies. This makes mtDNA a valuable genetic marker for studying population genetic patterns and variation (Chan *et al.*, 2021). In general, paternal mtDNA is actively eliminated before, during, and after fertilization, and since only the mother's egg contributes mtDNA to the offspring, there is little opportunity for recombination between different individuals' mtDNA. Thus, mtDNA is exclusively inherited from the maternal line, acting as a single, non-recombining locus. This makes phylogenetic reconstruction based on mtDNA relatively straight forward. The complete mitogenome has emerged as a valuable maternal-line marker with significantly higher resolution compared to that of single mitochondrial genes. Relationships among teleost fishes have been inferred based on mitogenomes at various taxonomic levels, from division to species (Miya *et al.*, 2003, Miya and Nishida, 2015). The Atlantic cod mitogenome is one of the best studied among the teleosts, featuring 277 complete individual-level mitogenomes (Jørgensen *et al.*, 2018) and pooled mitogenome sequencing of 44 specimen from each of two cod ecotypes (Karlsen *et al.*, 2014). There are some precautions that should be noted, however, with respect to the suitability of mitogenomes for assessing vertebrate population history, due to increasing evidence of rare recombination events, positive selection, erratic evolutionary rates, and the presence of heteroplasmy (reviewed by Ameur *et al.*, 2011, Galtier *et al.*, 2009).

1.4 Threespine stickleback

The threespine stickleback (*Gasterosteus aculeatus*) is a widely distributed fish species of the teleost group that initially originated in marine environments and is now found in various aquatic habitats across the northern hemisphere. It exhibits significant phenotypic diversity and inhabits marine, brackish, and freshwater environments (Bell, 1994, Wootton, 1976, Wootton, 1984). It has a spindle-shaped and laterally compressed body, often 4-6 cm in size, with a short caudal fin (Wootton, 1976, Wootton, 1984). Instead of scales, its body is protected by rows of bony plates on each lateral side, as well as dorsal and pelvic spines (Figure 1).

Over the past fifty years, the threespine stickleback has deepened our understanding in various fields such as ecology, behaviour, toxicology, vertebrate evolution, and developmental biology. This knowledge has paved the way for significant discoveries across numerous biological domains. The threespine stickleback, with its small body size, widespread distribution, high fecundity, and short generation time, has been a suitable model organism. In laboratory settings, it can be studied through either natural mating or artificial fertilization (Bell, 1994, Kingsley *et al.*, 2004, Kingsley and Peichel, 2006). The rapid colonisation and parallel adaptation across the northern hemisphere make threespine stickleback an invaluable model species for studying population histories and the molecular basis of parallel evolution and local adaptation in the 21st century (Colosimo *et al.*, 2005, DeFaveri *et al.*, 2011, Fang *et al.*, 2020, Hohenlohe *et al.*, 2010a, Jones *et al.*, 2012).

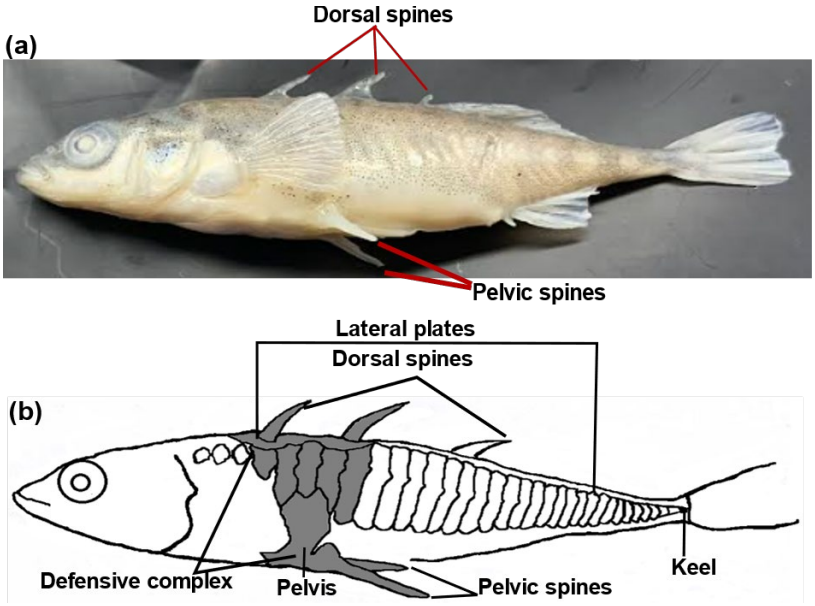


Figure 1. *Morphology of threespine stickleback: (a) Threespine stickleback from Lake Storvatnet; (b) schematic diagram of threespine stickleback showing lateral plates, dorsal spines and pelvic spines. Two dorsal spines connecting to the lateral plates and pelvic spines make a defensive complex.*

1.5 Origin, distribution and dispersal of the threespine stickleback

The contemporary populations of threespine sticklebacks originated from the Pacific Ocean (Fang *et al.*, 2018). A phylogenetic study (Fang *et al.*, 2018) based on restriction site associated DNA (RAD) sequencing data hypothesised three distinct clades in the Pacific Ocean (Figure 2). The first clade consists of freshwater populations in the Eastern Pacific Basin, specifically in Vancouver and the Queen Charlotte Islands of British Columbia. The second clade includes a marine population around Vancouver Island, while the third clade comprises populations from the Western Pacific Basin, such as a freshwater population in Alaskan Kodiak and marine populations in Russia's Anadyr Bay of the Bering Sea, as well as populations from the East Coast of Japan. The Eastern Pacific lineage further consists of two branches that diverged from the Pacific lineage within ca. 60 thousand years. One is the "Southern European Clade", distributed in the Mediterranean area, from the Iberian Peninsula to the Black Sea. The

other, the Trans-Atlantic Clade, is younger than the other major phylogeographic lineages and is estimated to have originated after the last glaciation period (< 20 thousand years ago). This Trans-Atlantic clade is subdivided into the "Trans-Atlantic subclade" (inhabiting the Western Atlantic region), and the "North Sea-Baltic Sea subclade", which includes populations from the Norwegian Sea, the North Sea, the Baltic Sea, the White sea, and the Barents Sea (Fang *et al.*, 2018).

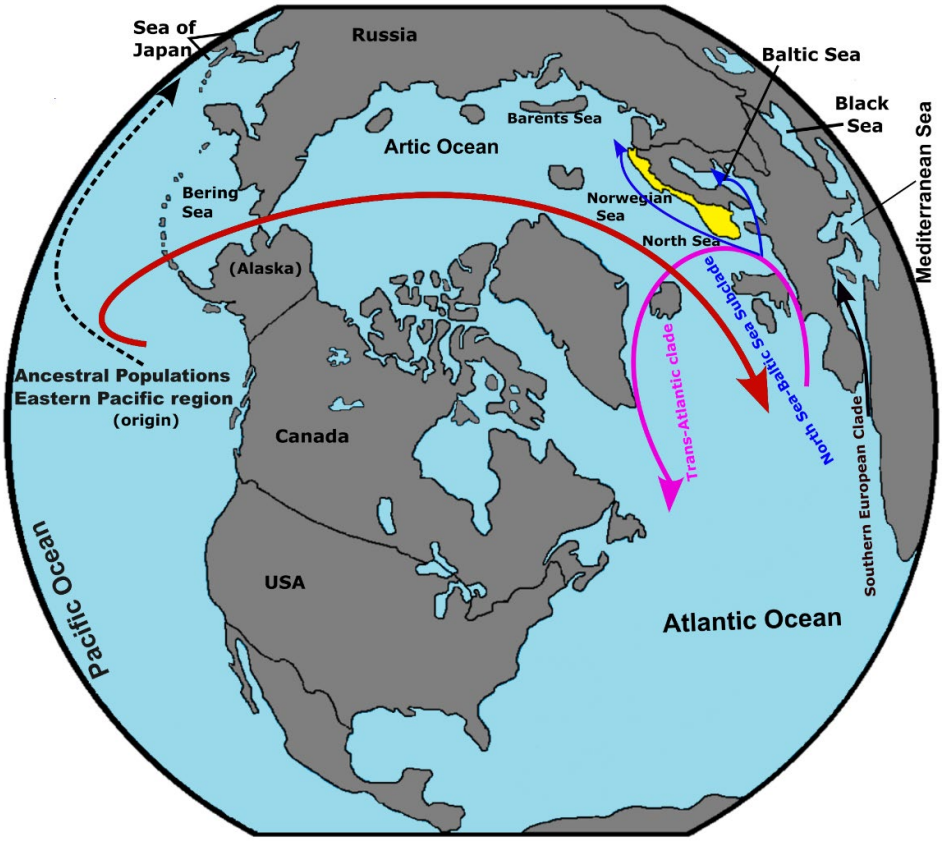


Figure 2. *Phylogenetic clades of threespine sticklebacks. The threespine stickleback originated in the Eastern Pacific region and colonised the Atlantic ocean via the Arctic Ocean (red arrow). Here, the two branches Southern European Clade (black) and Trans-Atlantic clade diverged from the Eastern pacific lineage. The trans-Atlantic clade has two subclades: Trans-Atlantic subclade (pink) and North-sea-Baltic Sea Subclade (blue). Yellow colour highlights the location of the sampling country, Norway, in our study. The figure is based on Fang et al. (2018).*

Marine sticklebacks are anadromous and exhibit remarkable adaptability to changes in salinity, enabling them to thrive in both marine and freshwater environments. After the last ice age, when marine sticklebacks colonised and adapted to freshwater systems, the remarkable and widespread morphological changes observed were the repeated reduction in the number of lateral plates and pelvic spines (Barrett *et al.*, 2011, Baumgartner and Bell, 1984, Bell, 1994, Bell, 2001, Cresko *et al.*, 2004, Fang *et al.*, 2020, Giles, 1983, Klepaker *et al.*, 2012, Laurentino *et al.*, 2022, O'Brown *et al.*, 2015, Reimchen, 1980, Shapiro *et al.*, 2004). The repeated emergence of similar freshwater ecotypes across wide geographical areas, presents an invaluable opportunity to explore the mechanisms of adaptive change and parallel evolution.

1.6 Lateral plate development and their variation in sticklebacks

Stickleback can be categorised into four types based on the presence of lateral plates: *i)* "completely-plated" individuals have an entire row of lateral plates extending from the head to the base of the caudal fin, typically numbering up to 36 (O'Brown *et al.*, 2015, Schröder *et al.*, 2023, Wootton, 1976); *ii)* "partially-plated" individuals have fewer lateral plates, but possess either a fully or partially developed keel on the tail; *iii)* "low-plated" individuals have only a few anterior lateral plates and lack a keel; and *iv)* "no-plated" individuals lack lateral plates altogether (Klepaker *et al.*, 2012, Wootton, 1976). The pelvic spines are linked to lateral plates via the ascending branch, and dorsal spines are connected to the lateral plates through a basal plate (Klepaker *et al.*, 2012, Wootton, 1976). Together, these lateral plates, dorsal spines, and pelvic structures form a defensive complex (Figure 1), making sticklebacks more difficult for gape-limited predators to swallow (O'Brown *et al.*, 2015, Schröder *et al.*, 2023). The lateral plates involved in this defensive complex are referred to as structural plates, while the remaining plates are called non-structural plates (Schröder *et al.*, 2023). In "completely-plated" morphs, these non-structural plates play a vital role in evading predatory fish. Predators take longer to process and swallow "completely-plated" morphs compared to the "low-plated" one (Bell, 2001, Hagen and Gilbertson, 1973,

Reimchen, 2000). When confronted by gape-limited predators, sticklebacks can lock both their pelvic and dorsal spines in an erect position, making it more difficult for these predators to swallow them (Wootton 1976). These defensive mechanisms of sticklebacks have been the focus of much interest among researchers because of their obvious and readily observable co-variability with environmental conditions, and thus, that they appear to be of crucial importance in stickleback evolution.

Marine sticklebacks typically possess a complete set of 36 bony lateral plates extending from the head to the tail. In contrast, freshwater fish often exhibit a reduced number of plates, usually around 7, located in the anterior region of the body (O'Brown *et al.*, 2015). The distribution of lateral plate morphs in freshwaters varies both regionally and among different habitats within regions. For instance, the low-plated morphs predominate in ponds and lakes, while the complete-plated morphs are more prevalent in high-gradient streams (Baumgartner and Bell, 1984). Hagen and Gilbertson (1973) claim that local natural selection supports the evolution of low-plated morphs in freshwater populations. Evidently, predation is a prominent selective pressure that plays a significant role in the evolution of low-plated morphs in freshwater systems (Hagen and Gilbertson, 1973).

The reduction in the number of lateral bony plates in freshwater populations is commonly attributed to selection acting on SGV (Colosimo *et al.*, 2005, Schluter and Conte, 2009). This facilitates the relatively rapid evolution of low-plated morphs in freshwater systems. One of the intriguing questions is how SGV associated with freshwater adaptation has been maintained within marine populations following the post-glaciation period (Reid *et al.*, 2021). Schluter and Conte (2009), proposed a “transporter” hypothesis to explain the maintenance of freshwater-adaptive SGV (Schluter and Conte, 2009). This hypothesis proposes a series of steps. Initially, alleles from a freshwater-adapted population are transferred to the sea through a hybridisation event between individuals from the marine and freshwater populations. As subsequent generations undergo recombination, the freshwater-adapted genotype

breaks down, leading to marine individuals carrying varying numbers of freshwater-adapted alleles, ranging from none to a small amount. In the next step, a new freshwater is formed elsewhere where a glacier is retreating, and is colonised by marine individuals, introducing SGV from the sea. Finally, through selection and recombination, the frequency of freshwater-adapted alleles gradually increases in the new location, leading to reassembly of the freshwater-adapted alleles. This model advocates the role of hybridisation as a creative force in evolution and diversification. Besides, the model suggests that gene flow facilitates parallel evolution of multiple traits and contributes to speciation on a large geographic scale. However, the "transporter" hypothesis does not address the origin or timing of advantageous mutations. Instead, it proposes that once a mutation reaches appreciable frequency in a freshwater population, it could participate in the transporter process (Reid *et al.*, 2021, Schluter and Conte, 2009).

1.7 The genetics of lateral plate development

The ectodysplasin (*Eda*) gene on chromosome 4, the expression of which is controlled by an enhancer (cis-regulatory sequences of DNA that interact with corresponding transcription factors to enhance the rate of transcription of the specific gene), accounts for the majority of variation in the number of bony plates (Colosimo *et al.*, 2004, Cresko *et al.*, 2004, O'Brown *et al.*, 2015). The *Eda* encodes a signalling protein essential for the development of the skeleton, skin, and other tissues, and it accounts for over 75% of the observed variation in lateral plate number (Colosimo *et al.*, 2004, Cresko *et al.*, 2004). O'Brown and colleagues (2015) observed that a T→G base pair change in the *Eda* is responsible for lateral plate reduction in freshwater sticklebacks (O'Brown *et al.*, 2015). However, several unrelated modifier genes or loci of minor importance have been identified, each responsible for controlling 5-10% of the variance in plate numbers (Colosimo *et al.*, 2004, Cresko *et al.*, 2004, O'Brown *et al.*, 2015, Peichel *et al.*, 2001). For instance, O'Brown and colleagues observed that Wingless/Integrated (WNT) signalling pathway can alter lateral plate development and

modify the expression of the *Eda* (O'Brown *et al.*, 2015). In addition, multiple signalling pathways have been implicated in lateral plate development, including c-Jun N-terminal kinase (JNK) (Sinha *et al.*, 2002), bone morphogenic protein (BMP) (Han *et al.*, 2018), and fibroblast growth factor (FGF) signalling pathways (Häärä *et al.*, 2012), alongside WNT signalling pathway (Cai *et al.*, 2021). It is important to note that several genes and proteins regulate these signalling pathways, making it challenging to have a complete overview of all factors involved in lateral plate development.

1.8 Mechanisms of pelvis development in sticklebacks

Sticklebacks display a diverse array of fins and spines: *i)* dorsal fins, usually accompanied by three dorsal spines; *ii)* pelvic fins, generally consisting of a pelvic spine on each side of the body and a single soft fin ray; *iii)* pectoral fins; *iv)* a caudal fin and *v)* anal fins. Pectoral fins are located on the sides of the body and aid in swimming (Wootton, 1976). Each pelvic spine is accompanied by an anterior process, an ascending branch, and a posterior process, forming a complete pelvic structure (pelvis) on both sides of the body. The pelvic spines are linked to lateral plates via the ascending branch, and dorsal spines connected to the lateral plates through a basal plate (Klepaker *et al.*, 2012, Wootton, 1976). The pelvic fins of sticklebacks can be considered homologous to vertebrate hind limbs, while their pectoral fins are homologous to fore limbs (Don *et al.*, 2013, Zhu *et al.*, 2012).

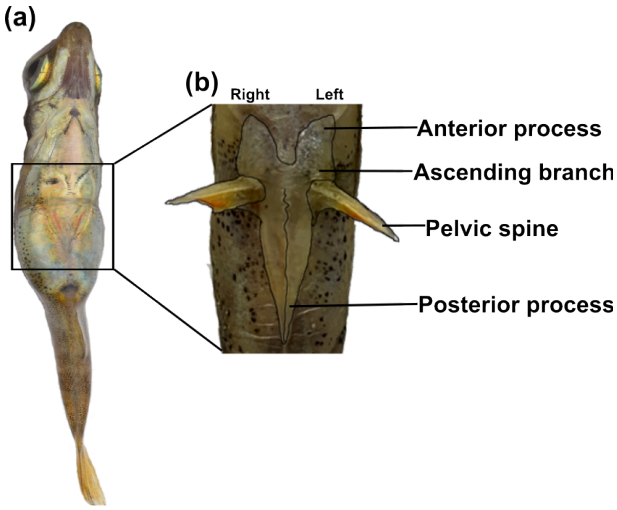


***Figure 3**. Pelvic structure of threespine stickleback. (a) Ventral view of threespine stickleback; (b) the pelvis structure (also termed "pelvic girdle") consisting of pairs of "anterior process", "ascending branch", "pelvic spines", and "posterior process".*

During development, the lateral plate mesoderm, situated at the periphery of the embryo, gives rise to both fore and hind limbs. These limbs develop along three axes: the proximal-distal (P-D) axis, the anterior-posterior (A-P) axis, and the dorsal-ventral (D-V) axis (Petit *et al.*, 2017). Proper limb development requires coordination among these axes, regulated by genes and their associated regulatory elements such as enhancers, promoters, activators, and repressors (Petit *et al.*, 2017).

The P-D axis refers to the differentiation of limb structures from the proximal (e.g. shoulders) to the distal regions (e.g. fingers). This differentiation is regulated by signals emanating from the apical ectodermal ridge (AER), an epithelial region at the tip of the limb-bud. The AER maintains the limb bud mesenchyme in a state conducive to proliferation and development (Petit *et al.*, 2017, Zeller *et al.*, 2009). Although forelimbs and hindlimbs initially share morphological similarities, their specific identities are determined by two paralogous transcription factors: T-box transcription factor 4 (*Tbx*4) for hindlimbs, and *Tbx*5 for forelimbs. These genes are regulated by the homeobox (*Hox*) gene families, which are organised in clusters, namely *Hox*A, *Hox*B, *Hox*C and *Hox*D. *Hox*A and *Hox*D are particularly crucial during limb development and

are regulated by elements within two flanking topologically associating domains (TADs : genomic regions where the frequency of chromatin interactions is high) (Petit *et al.*, 2017). In tetrapods, expression of *Hoxd*9 may diminish at later stages of limb bud development, whereas in threespine sticklebacks, *Hoxd*9 expression is maintained during pelvic fin bud development (Don *et al.*, 2013). Clustered *Hox* genes are expressed along two axes: rostral axis and caudal axis. Along the rostral axis, Hox proteins activate *Tbx*5 expression by binding to its enhancer in the prospective forelimb bud mesenchyme. In contrast, along the caudal axis, Hox proteins activate Pituitary homeobox transcription factor 1 (Pitx1) to bind to a hindlimb specific enhancer that subsequently drives the expression of *Tbx*4 in the proliferative hind-limb bud mesenchyme. In the case of threespine sticklebacks, these enhancers for *Pitx1* are called *Pel* enhancers (Petit *et al.*, 2017).

Furthermore, the *Pel*A enhancer, which is not well conserved outside teleosts, is located upstream of the *Pitx*1 gene and contains several transcription factor binding sites, of which a crucial feature is multiple TG-repeats (Chan *et al.*, 2010). In addition, the downstream enhancer of the *Pitx1*, *Pel*B, which is conserved between mammals and teleosts—particularly sticklebacks—may also play a role in the development of pelvic spines (Thompson *et al.*, 2018). Both Tbx4 and Tbx5 ultimately trigger the expression of the fibroblast growth factor 10 (Fgf10) in the mesenchyme, which in turn induces the expression of Fgf8 in the AER. Fgf10 and Fgf8 interact through a positive feedback loop mechanism that is crucial for the development of limb buds into forelimbs and hindlimbs (Minguillon *et al.*, 2009, Nishimoto *et al.*, 2014, Petit *et al.*, 2017). Moreover, retinoic acid (RA) influences the development of P-D axis by promoting the activity of *Tbx*4, *Tbx*5 and sonic hedgehog (*Shh*) genes (Figure 4) (Feneck and Logan, 2020, Lewandoski and Mackem, 2009).

The A-P axis, that determines the positioning of structures along the limb from front to back, is controlled by signals like SHH, originating from the zone of polarizing activity (ZPA) signalling centre (Petit *et al.*, 2017). Mutual antagonism between the

transcription factors Heart and Neural Crest Derivatives Expressed 2 (Hand2) and Gli3 regulates of Shh expression in the limb bud mesenchyme (Charité *et al.*, 2000, Galli *et al.*, 2010, te Welscher *et al.*, 2002, reviewed by Zuniga, 2015). The *Hand*2 collaborates with *Hoxd*13 to activate *Shh* expression (Galli *et al.*, 2010). This SHH signalling pathway prevents the processing of the *Gli*3 repressor (Gli3R), which would otherwise suppress the *Hand*2 regulation (Figure 4). A complex array of genes is subsequently expressed during limb bud outgrowth, playing a pivotal role in patterning and specifying the correct number and identity of digits. Further, the D-V axis, running from the back of the hand to the palm (or differentiating structures from the back to belly region), is regulated by WNT signalling molecules (Petit *et al.*, 2017).

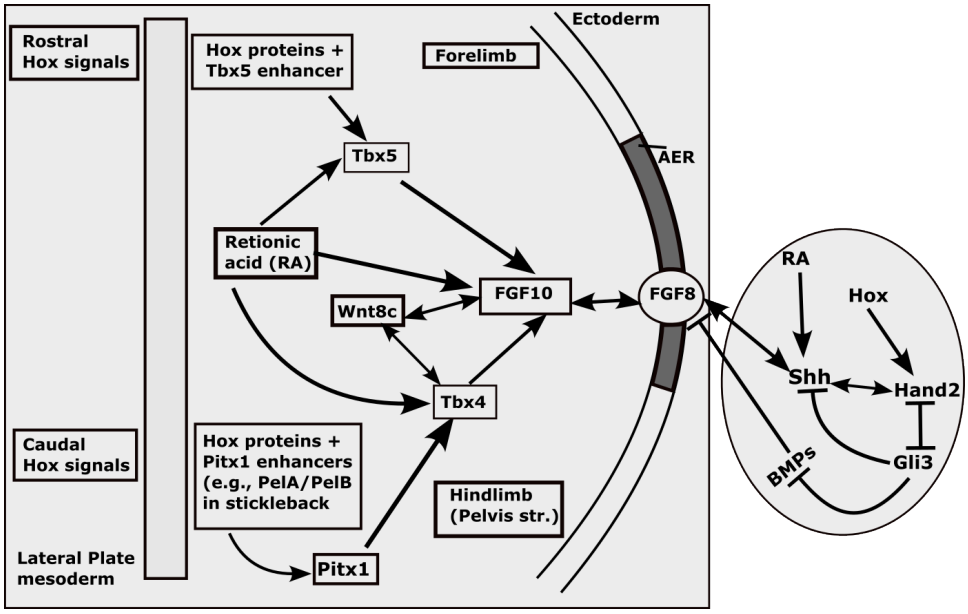


Figure 4. *Mechanism of forelimb and hindlimb development in vertebrates (adapted from Reid et al., 2021, Petit et al., 2017).*

1.9 Reduction of pelvic structures in sticklebacks

1.9.1 Evolutionary hypotheses behind pelvic reduction

Repeated evolution of reduced pelvic structures in threespine sticklebacks is backed by three hypotheses. One widely discussed hypothesis supporting the reduction of pelvis in freshwater environment is the “predation hypothesis”. It suggests that the presence of pelvic spines is positively correlated with predation pressure from gape-limited predators (Hagen and Gilbertson, 1972, Moodie, 1972, Reimchen, 1994). Several instances of reduced pelvic spine development in freshwater populations have been documented in Alaska, Western Canada, Iceland, Scotland, and Norway (Bell *et al.*, 1993, Bell and Ortí, 1994, Chan *et al.*, 2010, Klepaker and Ostbye, 2008, Peichel *et al.*, 2001, Shapiro *et al.*, 2004). This reduction in pelvic spines could be a result of selection driven by invertebrate larvae, which can latch onto and grip the spines of juvenile sticklebacks (Reimchen, 1980). However, there are differing perspectives on this matter (see Marchinko, 2009, Zeller *et al.*, 2012). Thus, the absence of or reduced length of spines in sticklebacks may be favoured in environments with low fish and bird predators, but a high abundance of insect predators, and vice versa. An alternative hypothesis is the "calcium hypothesis", which suggests that low calcium ion concentration in freshwater may favour pelvic spine reduction (Giles, 1983). Ultimately, the "predation-calcium hypothesis" proposes that a combination of predator presence and low calcium ion concentration could account for the evolution of pelvic reduction in sticklebacks (Bell *et al.*, 1993).

1.9.2 Molecular genetic mechanism behind pelvic reduction

As mentioned in the context of reduced plate morphs in freshwater sticklebacks, the repeated evolution of sticklebacks with reduced pelvic structures is likely facilitated by SGV. Here, the primary focus is on a few widely reported genes responsible for pelvic reduction.

The *Pitx*1 is considered an important factor in hindlimb development and the specification of hindlimb identity. In mice, the absence of *Pitx*1 expression leads to the

loss of hindlimb-like characteristics (Don *et al.*, 2013). Whereas, in fugu (*Takifugu rubripes*), a lack of *Hoxd9* expression may result in the loss of pelvic fins (Don *et al.*, 2013). Furthermore, in sticklebacks, the absence of *Pitx*1 expression is associated with the reduction or complete loss of pelvic spines. Genome mapping analyses have identified a chromosomal region containing the *Pitx*1 that is essential for hindlimb development in vertebrates, including the development of pelvic spines in sticklebacks (Chan *et al.*, 2010, Coyle *et al.*, 2007, Cresko *et al.*, 2004, Shapiro *et al.*, 2004).

In most sticklebacks with reduced or absent pelvic spines, the protein-coding region of the *Pitx*1 remains conserved and apparently functional, but its expression is absent in the pelvic region (Cole *et al.*, 2003, Shapiro *et al.*, 2004). Studies in mammals have concluded that the *Pel*B enhancer is essential for hind-limb development, but its role in sticklebacks is still a subject of interest. However, insertions and deletions (indels) in the *Pel*B enhancer have been observed in the pelvic-reduced benthic sticklebacks from Paxton Lake in British Columbia (Thomson *et al.*, 2018). The study reported partial deletion in *Pel*B and a full deletion of the *Pel*A enhancer region in pelvic reduced fish, highlighting the essential role of *Pel*A in pelvic reduction, while the role of *Pel*B remains unclear (Thompson *et al.*, 2018). Besides these enhancers, there may be other, currently uncharacterized regulatory regions affecting *Pitx*1 expression.

Chromosome mapping analyses have shown that the *Pitx*1 is located at the telomeric region of the linkage group 7, representing a DNA fragile site with a high probability of intrachromosomal rearrangements, such as deletions. These deletions might explain the observed reduction in pelvic spines in sticklebacks (Chan *et al.*, 2010, Xie *et al.*, 2019). In addition to the *Pitx*1, another closely related gene, *Pitx*2, has also been reported in vertebrates (Cole *et al.*, 2003, Tickle and Cole, 2004). The product of *Pitx*2 likely contributes to the asymmetry of the pelvis, potentially leading to the loss of pelvic spines on one side and their reduction on the other (Cole *et al.*, 2003). However, the role of *Pitx*2 in pelvic spine reduction or loss is not fully understood. Additionally, loci on chromosomes 2, 4, and 8 have been proposed to play a role in

fine-tuning the length of pelvic spines (Peichel *et al.*, 2001, Roberts Kingman *et al.*, 2021, Shapiro *et al.*, 2004).

1.10 Molecular tools available for threespine sticklebacks

Several genomic resources and tools are now available for studying threespine stickleback (Chan *et al.*, 2010, Kingsley *et al.*, 2004, Kingsley and Peichel, 2006, Wucherpfennig *et al.*, 2019). In 2003, David Kingsley highlighted the importance of sticklebacks in enhancing our understanding of vertebrate adaptation and forwarded the proposal to sequence the threespine stickleback genome to the National Human Genome Research Institute (reviewed by Reid *et al.*, 2021). The first draft of the threespine stickleback genome, *gasAcu-1*, derived from paired-end Sanger sequencing of an inbred female from Bear Paw Lake, Alaska, was released in 2006 and published in 2012 (Jones *et al.*, 2012, reviewed by Reid *et al.*, 2021). This freshwater ecotype genome spanned about 450 Mb. Subsequent refinements in the genome assembly is based on the Pacific Biosciences long read sequencing technology derived from benthic male strain of Paxton Lake, British Columbia (Nath *et al.*, 2021). This has furnished a high-quality reference genome for sticklebacks.

Threespine sticklebacks display diverse behaviours that vary both within and among populations (Bell, 1994). As a result, this species has been a focal point for behavioural studies for an extended period (Bell, 1994, Huntingford and Ruiz-Gomez, 2009). With the recent availability of a high-quality reference genome and the declining cost of generating genomic and transcriptomic data, opportunities to investigate the links between genes and behavioural traits are emerging. Employing microarrays and RNA-seq transcriptome techniques, researchers can detect alterations in distinct brain areas and tissues, elucidating transcriptional pathways triggered by various stimuli (reviewed by Reid *et al.*, 2021).

Many studies on host-microbe interactions have focused on traditional model organisms, such as mice (Douglas, 2019). However, the threespine stickleback, with its

extensive ecological and genomic resources, offers an exceptional opportunity to study host-microbiome interactions in natural environments. Simultaneously, the interaction dynamics between sticklebacks and their parasites give insights into how animals develop defences against parasites (Barber, 2013). An interesting example is the tapeworm, *Schistocephalus solidus*, which is commonly found in freshwater sticklebacks. This tapeworm is particularly useful as a model due to its detailed life cycle, helping researchers gain deeper understanding of host-parasite interactions (Barber and Scharsack, 2010).

With advancements in NGS methods, genome-wide studies that identify candidate regions associated with phenotypic traits have become increasingly popular in evolutionary genomics. However, these regions often encompass numerous genes, making it challenging to pinpoint the specific genes responsible for the observed traits (Itan *et al.*, 2015, Miller *et al.*, 2007). In such cases, targeted gene editing becomes an invaluable tool for elucidating the role of specific genes in trait development. The CRISPR-Cas9 system has emerged as a streamlined and efficient method for genome editing. Successful applications of CRISPR-Cas9 have been documented across various model systems, from yeast (DiCarlo *et al.*, 2013) to mice (Yang *et al.*, 2013), including sticklebacks (Hart and Miller, 2017, Wucherpfennig *et al.*, 2019). Wucherpfennig *et al.* (2019) applied the CRISPR-Cas9 technique to produce *Eda* mutant sticklebacks by editing the *Eda* loci. Consequently, mutations in *Eda* led to noticeable phenotypic changes, including the loss of lateral plates and soft fin rays, as well as substantial reductions in tooth and gill raker numbers. Thus, the ability to precisely edit genomes introduces a powerful dimension to the stickleback model system, enabling detailed investigations into the roles of individual genes and regulatory elements (Wucherpfennig *et al.*, 2019).

1.11 Molecular genetic markers of sticklebacks

Molecular genetic markers from both the mitogenome (mtDNA) and the nuclear genome (nDNA) can be used to study the evolutionary genomics of threespine sticklebacks.

nuclear genome (nDNA)*:* The nDNA of a threespine stickleback contains genetic information inherited from both parents and is transmitted in a biparental manner. According to the latest genome assembly (GAculeatus_UGA_version5), the total size of the stickleback's nuclear genome is approx. 472 Mb, estimated to comprise a total of 28,393 genes and 103 pseudogenes. These genes encode a total of approximately 45,382 proteins, with 317 rRNA sequences, 2,151 tRNA sequences, and 4,303 other RNA sequences. The GC content of the genome is approx. 44.6947 % (Nath *et al.*, 2021).

mitogenome (mtDNA)*:* The closed circular mitochondrial genome of threespine stickleback (GAculeatus_UGA_version5) is about 16,545 base pairs (bp) long (Figure 5). It contains the 37 genes commonly found in vertebrate mtDNA, which includes 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, and two ribosomal RNA genes. Of these, two rRNAs, 12 PCGs, and 14 tRNAs correspond to the heavy (H) strand, while ND6 and the remaining eight tRNAs correspond to the light (L) strand. Commonly used genetic markers within teleost mtDNA include protein-coding genes such as cytochrome c oxidase subunit I (COI) and NADH dehydrogenase subunit 1 (NAD1) (Chan *et al.*, 2021, Nath *et al.*, 2021), but cytochrome b (CytB) and the control region (CR) are most frequently used in sticklebacks (DeFaveri *et al.*, 2011, Mäkinen and Merilä, 2008).

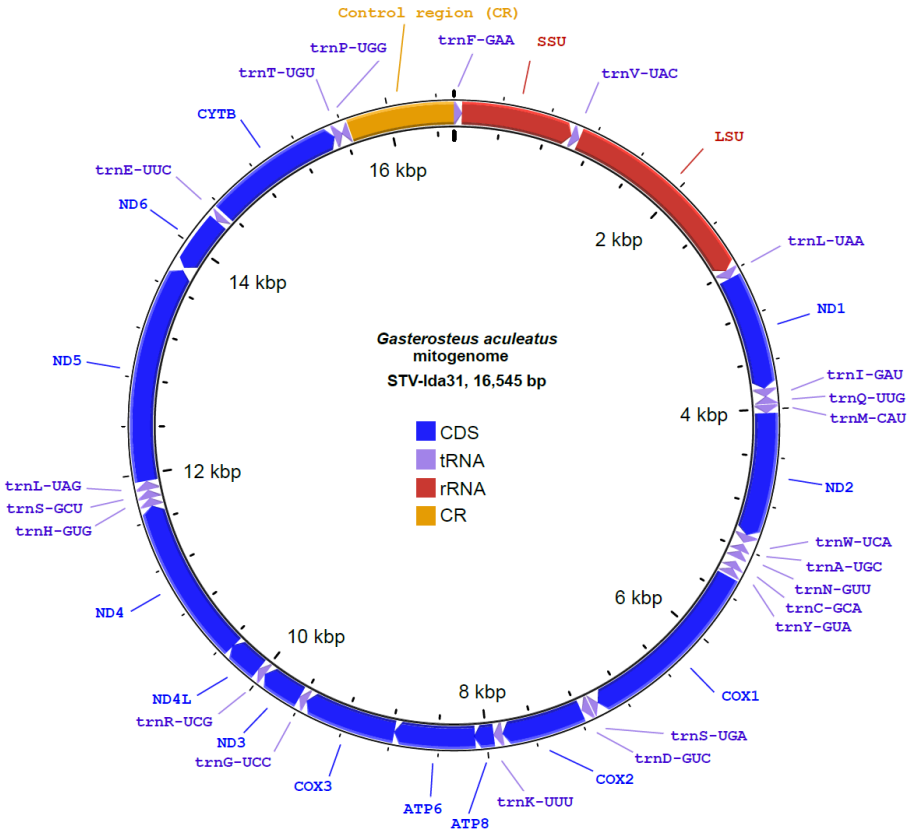


Figure 5. *Schematic representation of the circular mitogenome of threespine stickleback. Genes encoded by the heavy strand are indicated by arrows in the clock wise direction; genes encoded by the light strand are indicated by arrows in the anti clock wise directions.*

2 Aims of the study

Marine threespine sticklebacks colonised and adapted to freshwater habitats after the last ice age, followed by reduction in anti-predator armour such as pelvic structures. In most cases, reduction in pelvic structures fits with the predation hypothesis, the calcium hypothesis or the predation-calcium hypothesis. However, a few isolated freshwater pelvic-spine reduced stickleback populations, including Lake Storvatnet in Northern Norway (Figure 6), do not fit with the pelvic-reduction hypotheses (Klepaker and Ostbye, 2008, Klepaker *et al.*, 2012). Building on this intriguing finding, this study was designed to examine whether similar genetic mechanisms underlie stickleback adaptations in geographically distant regions, specifically in northern America and northern Europe.

In this study, the initial hypothesis posits that pelvic reduction in threespine sticklebacks correlates with variations in the *Pel* enhancers of the *Pitx*1. The main objectives are *(i)* to examine the molecular variations between the pelvic enhancers of spined and spine-reduced sticklebacks, and *(ii)* to explore the genome-wide diversity and differentiation among sticklebacks from three populations, aiming to identify genomic regions underlying the selection for spineless variants in a freshwater system. Overall, this study is divided into three papers, each with distinct specific objectives, as outlined below:

1. To investigate the *Pitx*1 enhancer variation in spined and spine-reduced stickleback populations **(Paper I)**
2. To study whole mtDNA (mitogenome) sequencing in assessing the maternal lineages of sticklebacks in a freshwater population **(Paper II)**
3. To provide an initial assessment of the genome-wide diversity and differentiation among a marine and two vicariant freshwater stickleback populations **(Paper III)**

4. To conduct a bulk segregant analysis of spined and spineless sticklebacks from a freshwater population **(Paper III)**

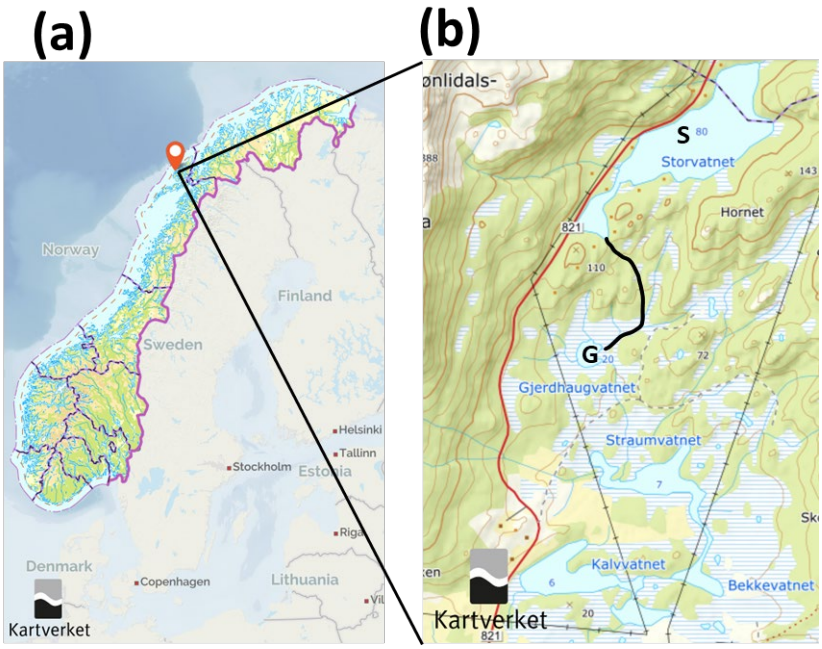


Figure 6. *Sampling sites* ***(a)*** *Norway with the study site.* ***(b)*** *Map sourced from www.norgeskart.no, accessed on 17th October, 2023) showing the upper Lake Storvatnet S (68°46'49"N, 15°9'36"E; 80 m altitude) and the second Lake Gjerdhaugvatnet G (20 m altitude) in the water course. The marine sampling site, approximately 8 km from the two lakes, at the tidal mouth of a small river at Sandstrand (68°44'45"N, 15°20'42"E), is not shown. The black curve line connecting the two lakes indicates the presence of a brook between them.*

3 Summary of methods

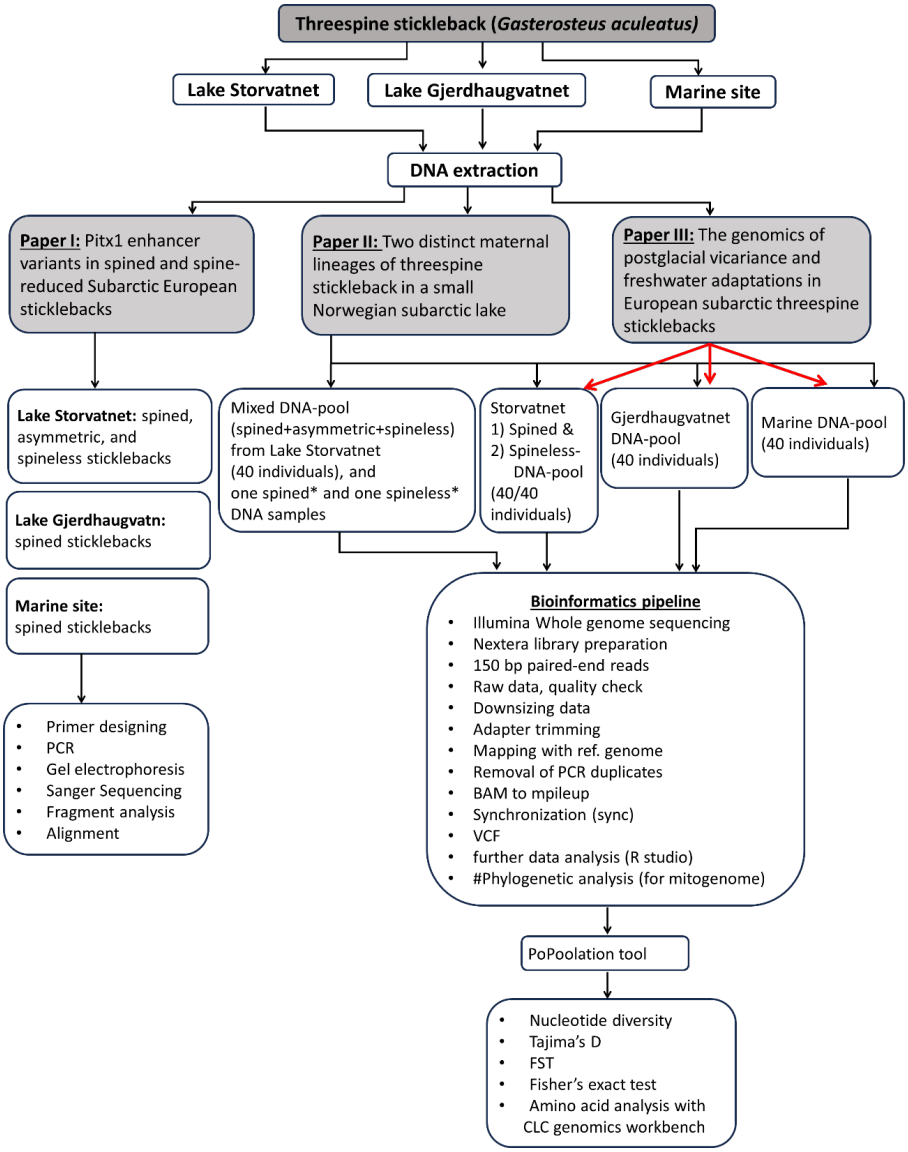


Figure 7. *Flow-chart of the overall methods applied in this study. (Note: red arrows indicate the samples used for Paper III). Note: The two individual mitogenomes* were sequenced by Ion Torrent technology.*

4 Major findings

Paper I: *Pitx*1 enhancer variants in spined and spine-reduced subarctic European sticklebacks

Out of the 304 specimens from Lake Storvatnet, 113 (37%) were completely-spined (symmetric), 99 (33%) were asymmetrically spined, and 92 (30%) were spineless. Notably, none of the specimens were missing the entire pelvic girdle. Among the sticklebacks from Lake Storvatnet with asymmetric spines, those with a left bias were more numerous than their right-biased counterparts. All specimens collected from Lake Gjerdhaugvatnet (N = 73) and from the marine site (N = 50) were completely-spined (CPS = 8) and symmetrical.

The allelic variation in the *Pel*A enhancer of *Pitx*1 among specimens from all three locations was identified. This allelic variation primarily resulted from variable counts of TG dinucleotides across three TG-repeats (TG-repeats I, II, and III) arrays (Figure 8). TG-repeats III especially displayed a diverse range of length variants. In addition, another polymorphism was identified upstream of, and adjacent to, TG-repeats III in specimens from Lake Storvatnet. At this upstream flanking region, all specimens from Lake Storvatnet had at least one allele with a 58 bp deletion compared to the reference (GenBank accession no. GU130435) (Figure 8). Meanwhile, all examined specimens from Lake Gjerdhaugvatnet and the marine site showed only variation at TG-repeats III, but not the adjacent flanking regions. Further, no indels were revealed by DNA sequencing of the *Pel*B region among the examined specimens from all three sites.

The stickleback population in Lake Storvatnet is unique due to the presence of a 58 bp polymorphism adjacent to TG-repeats III. However, a clear correlation between this 58 bp polymorphism and pelvic reduction in sticklebacks from Lake Storvatnet was not observed.

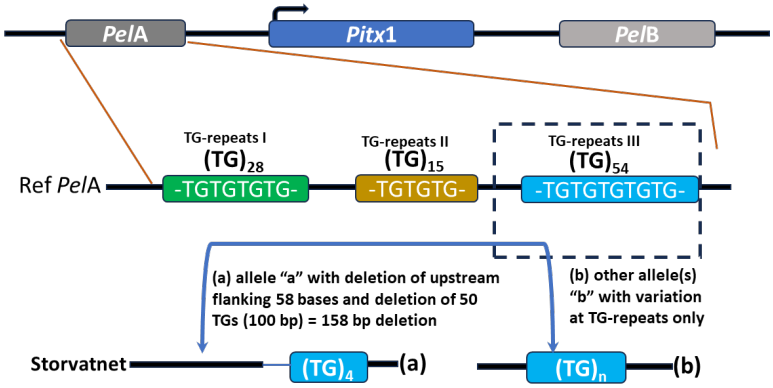


Figure 8. *The relative position of Pitx1 and its enhancers PelA and PelB. The PelA enhancer further consists of three TG-repeats: TG-repeats I, II, and III consisting of* $(TG)_{28}$*,* $(TG)_{15}$ *and* $(TG)_{54}$*, respectively (based on reference sequence GU130435 from Salmon River, British Columbia). The two blue arrows emerging from the dashed-rectangular area indicate the presence of multiple alleles at or adjacent to TG-repeats III. Allele "a" is a variant in which 58 bases upstream of TG-III are deleted (in addition to deletion of 50 TGs) compared to the reference, and allele "b" represents another variation where only the count of the TGs varies.*

Paper II: Two distinct maternal lineages of threespine stickleback (*Gasterosteus aculeatus*) in a small Norwegian subarctic lake

Mitogenomic analysis of the threespine stickleback has revealed two main clades: the Trans-North-Pacific (TNP) haplotype group, and the Euro-North-American (ENA) haplotype group. Complete mitogenome sequences of a completely spined stickleback (STV-Ida31) and a spineless stickleback (STV-Ida61) from Lake Storvatnet differed at 88 nucleotide positions (excluding the HTR array). The mitogenomes STV-Ida31 and STV-Ida61 represent two distinct branches of the ENA haplotype group. STV-Ida31 clustered with mitogenomes from nearby Lake Gjerdhaugvatnet, while STV-Ida61 was more closely related to a marine mitogenome from Altafjord, Northern Norway. Thus, it is speculated that Lake Storvatnet could have experienced two colonisation events of sticklebacks, with the second colonisation likely involving spineless sticklebacks carrying the STV-Ida61 mitogenome haplotype. It is notable that mitogenomes STV-Ida31 and STV-Ida61 both are interspersed with US mitogenome haplotypes from Lake Superior specimens (Wisconsin, USA).

Only one major mitochondrial haplotype (GHV-Maj1) was found in Lake Gjerdhaugvatnet. Most SNPs observed in Lake Storvatnet were also found in the marine sample. In contrast, Lake Gjerdhaugvatnet harbours some unique and fixed SNPs (e.g., F_{ST} = 1.0 in ND4) not observed in Lake Storvatnet or the marine sample, underscoring the genetic differentiation between specimens from lake Storvatnet and Lake Gjerdhaugvatnet.

Further assessment of potential mitogenomic differentiation among various pelvic morph specimens in Lake Storvatnet did not reveal any genetic differentiation of mitogenomes, showing that mitogenome haplotypes are not associated with the present pelvic morphs in Lake Storvatnet.

Paper III: The genomics of postglacial vicariance and freshwater adaptations in European subarctic threespine sticklebacks

In this initial assessment of genome-wide diversity and differentiation among sticklebacks in a marine population and two freshwater lakes, the most prominent genetic signatures of selection, found in 13 out of 21 autosomes, were pointed out. The nucleotide diversity $\boldsymbol{\pi}$ values, and neutrality index Tajima's D **(T_D)** were not uniformly distributed and displayed extensive variations along the genome among the three populations. Genomic regions suggestive of balancing and purifying selection were identified, contributing to the understanding of the selective forces shaping these populations.

The average genome-wide F_{ST} value of 0.189 revealed high genetic differentiation between the two freshwaters Lake Storvatnet and Lake Gjerdhaugvatnet despite the fact that both lakes belong to the same water course, which suggests selection, restricted gene flow between the populations, and genetic drift (especially in the downstream lake).

Several candidate genes for bone/organ development were pinpointed, including known and novel loci. Aligning with prior research, the presence of ancient

chromosomal inversions was inferred, indicating their persistent role in SGV and parallel adaptation of sticklebacks in freshwaters.

Bulk segregant analysis revealed, as expected, a low overall genetic differentiation between spined and spineless sticklebacks from Lake Storvatnet. However, significant genetic differentiation between these two groups was identified in a few genomic regions. The most prominent differentiation was found at a specific region (F_{ST}-island) of chromosome 9, which includes the gene *Hand*2 among several others, potentially playing a role in pelvic spine development.

5 General Discussion

This research focuses on threespine sticklebacks collected from two freshwater lakes within the same watercourse, upper Lake Storvatnet and neighbouring Lake Gjerdhaugvatnet, as well as a marine site. Only completely (pelvic) spined sticklebacks were found in Lake Gjerdhaugvatnet and the marine site, whereas Lake Storvatnet hosted polymorphic sticklebacks, comprising completely spined, asymmetrically spined, and spineless forms. No clear correlation was observed between sequence variations in the *Pel*A enhancer and pelvic-reduced morphs in Lake Storvatnet. Further assessment of mitogenomes among various pelvic morph specimens in Lake Storvatnet suggests that this population comprises two distinct mitogenome clades, each belonging to separate branches of the ENA group. These two mitogenome haplotypes are not associated with the contemporary pelvic morphs found in Lake Storvatnet. In contrast, only one of the main clades found in Lake Storvatnet population is present in Lake Gjerdhaugvatnet. Additionally, Lake Gjerdhaugvatnet harbours some unique and fixed SNPs not observed in Lake Storvatnet or the marine sample, underscoring the genetic differentiation between mitogenomes from Lake Storvatnet and Lake Gjerdhaugvatnet.

The analysis continued with an assessment of genome-wide diversity and differentiation across the three stickleback populations. Utilizing the nucleotide diversity ($\boldsymbol{\pi}$), neutrality index Tajima's D (T_D) and F_{ST}, several genomic regions and genes potentially involved in adaptation of sticklebacks to freshwater were identified. The presence of ancient chromosomal inversions was inferred based on signatures of diversity and differentiation. Thes inversions may harbour adaptive variation, facilitating the recurrent evolution of freshwater adaptations. Additionally, a greater genetic differentiation between the pair of freshwater populations (Lake Storvatnet and Lake Gjerdhaugvatnet) was detected compared to pairs of freshwater and marine stickleback populations (Storvatnet-marine and Gjerdhaugvatnet-marine pairs), in line with the mitogenome analysis. Further, the examination of genome-wide

differentiation between spined and spineless stickleback specimens from Lake Storvatnet, revealed an "F_{ST}-island" on chromosome 9. This "F_{ST}-island" encompasses several genes, including the *Hand*2, which could be involved in pelvic spine reduction.

5.1 No direct association between variations in *Pel*A enhancer sequences and pelvic-reduced sticklebacks in Lake Storvatnet

The consistent presence of 58-bp-depleted-variant alleles upstream to TG-repeats III of *Pel*A across completely spined, asymmetrically spined and spineless specimens from Lake Storvatnet suggests that there is no direct association between this specific *Pel*A sequence variation and the phenomenon of pelvic reduction (Paper I). In contrast, earlier research identified a larger deletion associated with pelvic reduction in the *Pel*A enhancer: 1868-bp in sticklebacks from Lake Paxton, 973-bp from HUMP Lake, and 757-bp from Beer Paw Lake in North America (Chan *et al.*, 2010). Nevertheless, the 58 bp deletion identified in this study does align, at least in part, with the deletions previously reported in North American populations (Paper I).

In Lake Paxton, approximately 80% of the adult specimens (benthic morph) completely lack a pelvic girdle (McPhail, 1992). This rate contrasts sharply with Hump Lake and Bear Paw Lake, where only 12.7% and 7.6% of specimens, respectively, display a complete absence of the pelvic girdle (Bell and Ortí, 1994). In Lake Storvatnet, there are no reported instances of specimens completely lacking pelvic girdles. Inferring a connection between the size of *Pel*A deletions and the absence of pelvic spines (and girdle) based on limited samples from three North American and one North European stickleback populations could be inconclusive. Interestingly, while variations in the *Pel*B enhancer region were detected in spineless specimens from Lake Paxton (Thompson *et al.*, 2018), such indels were absent in the Lake Storvatnet specimens.

While the findings from this study did not identify an association between the *Pel*A enhancer variation and pelvic reduction, previous studies have reported involvement of *Pel* enhancers in pelvic reduction (Chan *et al.*, 2010, Thompson *et al.*,

2018). The dinucleotide TG-repeats are known to form left-handed fragile Z-DNA, which is more susceptible to deletions (Liu *et al.*, 2001, Xie *et al.*, 2019). This Z-DNA configuration opens up the chromatin structure, facilitating the binding of transcription factors to the enhancer (Liu *et al.*, 2001). Such binding leads to chromatin-dependent activation of promoters, thereby initiating transcription (Liu *et al.*, 2001). Notably, the *Pel*A variants observed among sticklebacks in Lake Storvatnet are comparatively short. Given this, it is reasonable to hypothesise that these *Pel*A variants are at a pivotal juncture. This might dictate whether Z-DNA formation and subsequent transcription takes place, influencing the emergence of both spined and spine-reduced individuals. In Lake Storvatnet, developmental trajectories for the pelvic apparatus could be further impacted by other genomic regions, genetic components, epistatic and epigenetic interactions, and environmental factors.

The absence of a direct association between the *Pel*A variants and pelvic spine status in sticklebacks from Lake Storvatnet introduces nuanced layers of complexity to understanding of the mechanisms underpinning pelvic reduction. Previous studies using linkage mapping and QTL analysis have identified additional loci that might fine-tune pelvic structures (Peichel *et al.*, 2001, Shapiro *et al.*, 2004). To further explore the potential mechanisms and population histories possibly associated with the pelvic-reduced phenotype, maternal lineages of Lake Storvatnet were studied in Paper II.

5.2 Mitogenome analysis could mean that Lake Storvatnet has experienced a second invasion

Phylogenetic analyses of mitogenomes done in Paper II identified two major haplotypes STV-Ida 31 and STV-Ida61 from Lake Storvatnet belonging to the two distinct clusters of the ENA mitogenome clades. Among two major haplotypes found in Lake Storvatnet, STV-Ida31 closely aligns with the mitochondrial haplotype GHV-Maj1 found in the downstream Lake Gjerdhaugvatnet. This suggests that STV-Ida31 and GHV-Maj1 represent the original maternal lineages of sticklebacks in the watercourse. The other main haplotype, STV-Ida61, bears a close relationship to a

marine haplotype from Altafjord in Northern Norway, which is roughly 400 km away in a straight aerial line. The presence of STV-Ida61 in Lake Storvatnet and Altafjord, and absence in Lake Gjerdhaugvatnet bring about the intriguing question: what is the origin of STV-Ida61 in Lake Storvatnet? It is speculated that STV-Ida61 was introduced during a second colonisation by marine sticklebacks carrying the genetic disposition for spinelessness. However, an association between mitogenome haplotypes and pelvic morphs among the present-day sticklebacks was not found in Lake Storvatnet.

To investigate the genetics of spinelessness in Lake Storvatnet, genome-wide diversity and differentiation were examined in Paper III. The analysis focused on comparison of *(i)* sticklebacks from two freshwater populations and a marine population, and *(ii)* spined and spineless groups within Lake Storvatnet population.

5.3 To what extent are sticklebacks from Lake Storvatnet genetically different from those in Lake Gjerdhaugvatnet and a nearby marine site?

One of the important findings presented in the Paper III is that sticklebacks in Lake Storvatnet exhibit pronounced genetic differentiation from those in downstream Lake Gjerdhaugvatnet, despite inhabiting the same watercourse. Additionally, the genetic differentiation between sticklebacks in Lake Storvatnet and the marine site is slightly higher when compared to the genetic differentiation between sticklebacks in Lake Gjerdhaugvatnet and the marine site.

Several factors likely contribute to the genetic differentiation between the sticklebacks of Lake Storvatnet and those in the downstream lake and the nearby marine site. *(i)* Restricted current gene flow between two freshwaters: as posited by Klepaker *et al.* (2012), a lack of recent gene flow between the two lakes could be contributing to the observed genetic differentiation. The presence of small waterfalls between the lakes, in the present time, acts as physical barriers (see Figure 6), making upstream migration for sticklebacks of Lake Gjerdhaugvatnet practically unfeasible.

Downstream migration from Lake Storvatnet is also considered unlikely, but cannot be entirely ruled out. The recolonisation of threespine sticklebacks in Northern Europe, particularly in marine waters, began between 17.1 – 37.3 thousand years ago (kya), according to Fang *et al.* (2018). This implies that marine waters in Northern Norway were colonised by sticklebacks before freshwater lakes such as Lake Storvatnet and Lake Gjerdhaugvatnet. Historical records on how sticklebacks colonised these lakes are absent. Possibly, the ancestors of Lake Storvatnet's sticklebacks, located at an altitude of 80 meters, well above the "marine limit" (the maximum altitude of the sea surface relative to today's sea level since the last ice-age) of 35 m, have been transported by humans or birds, or migrated via a historical stream before the formation of the current waterfalls. In contrast, Lake Gjerdhaugvatnet, located at 20 m altitude, below the marine limit, suggests that gene flow between its stickleback population and others, including the nearby marine site, could have been more significant and prolonged, thus influencing their genetic composition. This might result in the sticklebacks from Lake Gjerdhaugvatnet being genetically less distinct from marine sticklebacks, yet showing greater genetic differentiation when compared to sticklebacks in Lake Storvatnet. *(ii)* Genetic drift: Lake Storvatnet, being larger, likely supports a greater number (N) of sticklebacks and has a larger effective population size (Ne). In contrast, Lake Gjerdhaugvatnet, which is 8 to 10 times smaller than Lake Storvatnet, likely hosts fewer sticklebacks and has a smaller effective population size. Consequently, genetic drift might be more pronounced in Lake Gjerdhaugvatnet, increasing the genetic differentiation from Lake Storvatnet. The long-term effective population size is also negatively affected by population fluctuations. It is likely that population numbers fluctuate more in Gjerdhaugvatnet, as conditions in such a small and shallow lake could be more easily affected by environmental factors, such as weather. *(iii)* Natural selection: In Lake Storvatnet, genetic drift might play a less significant role due to its larger effective population size. This allows natural selection to have a more pronounced impact on this population. This dynamic could explain the observed higher diversity in nuclear-DNA, and phenotypic traits (such as pelvic morphs) in Lake

Storvatnet compared to Lake Gjerdhaugvatnet. Additionally, the presence of a single dominant mitogenome haplotype in Lake Gjerdhaugvatnet, contrasted with two dominant mitogenome haplotypes in lake Storvatnet, further underscores the genetic differentiation between the populations of these two lakes.

Furthermore, the specific regions (genes) along the genome where the sticklebacks from Lake Storvatnet exhibit high genetic differentiation, as identified in this study, are listed in Table 1 below. Some of the genes listed below have been previously reported to be associated with bone formation and development in sticklebacks, while others are known to be associated with bone development in other vertebrates. Their potential role in bone development in sticklebacks should not be ignored. This assessment serves as an initial exploration into the genes that putatively underpin local adaptation in Lake Storvatnet sticklebacks. However, individual assessments of candidate genes and further molecular studies are required to understand their involvement in phenotypic traits, which is beyond the scope of the current study.

Table 1. *list of genes related to bone and organ development that have higher genetic differentiation in sticklebacks from Lake Storvatnet compared to the other two populations*

Chrom	Genes	Description	Function
1	*Tbx*4	T-box transcription factor 4	developing the hindlimbs into pelvic appendages (Petit *et al.*, 2017)
4	*Wnt*7ba	Wingless-Type MMTV Integration Site Family, Member 7B	Bone morphogenesis (Bollaert *et al.*, 2019, Zhang *et al.*, 2009)
	*Hbp*1 (LOC120817816)	HMG box-containing protein 1-like	Regulates Wnt signalling pathway (Bollaert *et al.*, 2019)
	*Sox*5	SRY-box transcription factor 5	Regulates Wnt signalling pathway (Bollaert *et al.*, 2019)
5	*Gdf*10	growth/differentiation factor 10-like gene	interdigital webbing (Cheng *et al.*, 2016)
	*Fhl*2	four and a half LIM domain protein 2 like gene	limb morphogenesis, dorsoventral patterning, inhibition of osteoblast differentiation, and cartilage formation (Lai *et al.*, 2006)
	*Bmp*2	bone morphogenic protein 2-like gene	Bone morphogenesis (Kishimoto *et al.*, 1997)
	*Sox*9a	transcription factor Sox-9-A-like gene	Cartilage formation (Cresko *et al.*, 2003)
6	*Fgf*8a	fibroblast growth factor 8a	developing the hindlimbs into pelvic appendages (Cole *et al.*, 2003)
7	*H2az*	histone H2A.Z implications in chromatin structure	implications in chromatin structure (Giaimo *et al.*, 2019)
	*Mllt*3	MLLT3 super elongation complex subunit	hematopoietic stem cell maintenance (Germano *et al.*, 2022)
	Bbx	HMG box transcription factor BBX	central nervous system development (Chen *et al.*, 2014)
	Bsx	brain-specific homeobox	brain-specific functions (Cremona *et al.*, 2004)
	H2a		implications in chromatin structure (Giaimo *et al.*, 2019)
14	*Brinp*1	bone morphogenetic protein/retinoic acid inducible neural-specific 1	Brain development (Berkowicz *et al.*, 2016)
	*Lhx*3	LIM homeobox 3	embryonic development, neuronal identity (Srivastava *et al.*, 2010)
16	*Pou*1*f*1	POU class 1 homeobox 1	regulator of growth hormone, pituitary organogenesis (Işık and Bilgen, 2019)
19	*Alx* homeobox genes	ALX homeobox 1	neural tube closure, limb development, and craniofacial development (McGonnell *et al.*, 2011)

Furthermore, this study has identified putative inverted regions on three chromosomes, which could contribute significantly to the observed genetic differences between the stickleback populations.

5.4 Presence of chromosomal inversions along the genome indicates balancing selection

Three prominent regions of putative chromosomal inversions (reversed DNA regions) with sizes of 0.5 Mb, 0.4 Mb, and 2 Mb were detected on chromosomes 1, 11, and 21, respectively. These putative inversions were based solely on signatures of population genetic parameters, including nucleotide diversity (π), Tajima's D (T_D), and F_{ST}, but the findings are consistent with previous studies on North American stickleback populations (Jones *et al.* 2012) and Russian stickleback populations (Terekhanova *et al.* 2014). These putative inversions of ancient origin have been suggested as to play a role in divergent selection in marine and freshwater environments. Hence, these inversions seem to be examples of the reuse of shared genetic variation at a global scale (Jones *et al.*, 2012). Notably, there was a large chromosomal inversion on chromosome 21. In the T_D analysis, this inversion region in Lake Storvatnet stickleback genome displayed higher T_D values compared to the genomes of sticklebacks from the other two populations. These elevated T_D values suggest that multiple alleles in this region are under balancing selection in sticklebacks from Lake Storvatnet.

Despite a century of research, many questions about the adaptive role of chromosomal inversions remain inadequately understood (Hoffmann and Rieseberg, 2008, Kapun and Flatt, 2019, Kapun *et al.*, 2023, Kirkpatrick and Barton, 2006). In recent years, inversions have been recognized as *(i)* unlikely to be selectively neutral, *(ii)* having a role in protecting extensive chromosomal regions, which may contain hundreds of genes, from recombination, thereby facilitating local adaptation, *(iii)* decreasing the effective population size, causing increased genetic drift, *(iv)* causing increased genetic hitchhiking and *(v)* accelerating the accumulation of weakly deleterious mutations (Berdan *et al.*, 2023). A challenge in studying inversions is that

their evolution is often affected by multiple interacting processes, impeding efforts to distinguish between the different mechanisms (Frank, 2014, Nosil *et al.*, 2023). Nonetheless, more comprehensive studies are necessary to understand the detailed molecular mechanism behind chromosomal inversions.

5.5 Why is there pelvic polymorphism in Lake Storvatnet sticklebacks?

In Lake Storvatnet, the presence of polymorphic sticklebacks—both spined and spine-reduced morphs—prompted this study. Existing evolutionary hypotheses, particularly those concerning predation and calcium-based theories, do not appear to provide a comprehensive explanation for the pelvic reduction observed in this lake (Klepaker *et al.*, 2012). The present study did not establish any association between variations in pelvic enhancers and the phenomenon of pelvic spine reduction (Paper I). Furthermore, Paper III failed to uncover any variations in the *Pitx*1 gene and nearby loci from genome-wide analysis, despite its acknowledged role in the development of pelvic spines in sticklebacks. This lack of information is due to the unavailability of annotations for *Pitx*1 and adjacent loci in both the current and the previous versions of threespine stickleback genome assemblies. The *Pitx*1, located at the sub-telomeric region, harbours repetitive sequences such as TG dinucleotide repeats, making it susceptible to deletion and posing challenges in its incorporation in genome assemblies. Further, the presence of two mitogenome haplotypes in Lake Storvatnet did not establish any association between pelvic morphs and mitogenome haplotypes (Paper II).

While examining genome-wide diversity and differentiation between the spined and the spineless groups within Lake Storvatnet, similar values of nucleotide diversity (π), neutrality index (T_D), and F_{ST} across the genomes of both groups were observed, testifying the consistency of the approach as both groups belong to the same population. Intriguingly, a cluster of higher F_{ST} points (for simplicity hereafter referred to as a "F_{ST}-island") spanning from 4.0 to 4.4 Mb region at the chromosome 9 was

observed. Within this F_{ST}-island, a gene *Hand*2 was identified, which participates in the genetic pathway influential in pelvic development (Figure 4). The *Hand*2 is involved in diverse cellular processes related to heart and fore/hindlimb development in zebrafish and mouse (Osterwalder *et al.*, 2014, Yelon *et al.*, 2000), and along with its upstream enhancer collaborates with *Hoxd*13 to activate the Sonic Hedgehog (SHH) pathway during limb bud development (Galli *et al.*, 2010). Notably, the concurrent absence of both *Hand*2 and *Gli3R* results in a complete loss of anteroposterior (AP) polarity along the proximo-distal axis (Galli *et al.*, 2010). Especially since there are currently no reported studies demonstrating the involvement of the *Hand*2 in pelvic spine development, this study opens up avenues for further research into the role of the *Hand*2 in the development of pelvic spines in threespine sticklebacks. Furthermore, multiple candidate genes potentially related to limb development, were also identified. These findings imply that the traits in questions could sometimes be the result of complex genetic networks, and mutational events in any part of these networks can have an impact on the resulting traits.

The pelvic spine reduction in sticklebacks in Lake Storvatnet might be claimed to be an example of parallel evolution of the same trait, which is caused by *de novo* mutation and not by standing genetic variation. Except in a few well-studied systems, it is not known how common genetic parallelism – repeated use of the same genetic changes underlying phenotypic parallelism - is in sticklebacks (Poore *et al.*, 2022). Studies from several North-East American populations of pelvic spine-reduced sticklebacks have documented indels, at the same genomic region at and near *Pel*A, although the size of the indels vary between the populations (Chan *et al.* 2010). This is to a large extent contrary to the investigation of *Pitx*1, done in this study, based on PCR, fragment analysis and DNA sequencing, which suggested that there is no clear association between variation of the *Pel* enhancer of *Pitx*1 and pelvic reduced morphotypes in Lake Storvatnet. However, it is speculated that the sizes of *Pel*A variants could be at a tipping point for transcription to occur *or not* in sticklebacks from Lake Storvatnet (Paper I). Thus, North-East American stickleback populations on one

hand and the population in Lake Storvatnet (North Europe) on the other hand, seem to be examples of phenotypic parallel evolution by changes at different loci (or regions at the genome, see the paragraph “Parallel evolution” in the “Introduction”). This result concurs with suggestions that parallelism at the genetic level is greater between populations at regional compared to global scales (Paccard *et al.*, 2019). Parallel evolution of phenotypic traits with a non-parallel genetic basis has previously been demonstrated in threespine stickleback gill rakers length (Glazer *et al.*, 2015), teeth development (Ellis *et al.*, 2015) and many morphological traits in QTL crosses of four different lake-stream pairs (Poore *et al.*, 2022).

Even though this study did not conclusively answer the question of the mechanism behind the presence of polymorphic sticklebacks in Lake Storvatnet, or identify the specific genomic region responsible for spinelessness in this lake, there were several unique findings. These include: *(i)* a unique 58 bp variation adjacent to TG repeat III, and this variation might act as a critical threshold that influences the transcriptional activity *(ii)* “F_{ST}-island” region along the genome, *(iii)* several candidate genes underlying limb development, and *(iv)* two distinct mitogenome haplotypes in sticklebacks from Lake Storvatnet. These features make the Lake Storvatnet population appear as unique among stickleback populations.

6 Conclusion and future perspectives

In the upper lake, varied pelvic spine morphs and two distinct mitogenome haplotypes have been identified. However, no link was found between *Pel*A enhancer variations and the reduced pelvic spines. It is proposed that the reduction in pelvic spines might be linked to the *Hand*2 in these sticklebacks. To further confirm the role of *Hand*2 in spine reduction, assessing this gene through PCR and sequencing is essential. Similarly, identifying the enhancers regulating the *Hand*2 is crucial, akin to the *Pel* enhancers' role in regulating the *Pitx*1, which contributes to pelvic spine reduction in North American stickleback populations. Hence, it would be interesting to figure out if the parallel evolution of pelvic-reduced sticklebacks in Lake Storvatnet is an example of phenotypic parallel evolution or phenotypic-genotypic parallel evolution. Another challenge encountered in this study was obtaining all dinucleotide TG-repeats from the *Pel*A enhancer using either Sanger Sequencing or short-read NGS. Sequencing genomic regions that contain repeats are particularly challenging due to difficulties in accurately capturing these repetitive sequences. To address this issue, more recent methods of long-read sequencing could prove beneficial. These techniques excel in sequencing DNA with repeats because they can read much longer DNA fragments, typically >10 kb, which allow for a more accurate representation of the DNA sequence (Amarasinghe *et al.*, 2020).

Two distinct mitogenome clades were detected in the upper lake's stickleback population, indicating these represent different mitogenome clusters of the Euro-North-American lineage. However, no correlation was observed between these mitogenome haplotypes and the reduction of pelvic spines in this lake's population. Interestingly, of these two haplotypes, only one was identified in a marine source located 400 km away. To further understand the distribution of these haplotypes in the marine population, it's essential to analyse mitogenome sequences from additional marine sites from subarctic Norway. A Master-thesis aimed at identifying mitogenome sequences from subarctic marine populations is currently underway.

This study employed pooled-DNA sequencing for genome-wide diversity and differentiation analysis. While cost-effective for population genetics studies, this approach does not capture individual genome and haplotype information. To address this limitation, individual DNA sequencing is recommended, with long-read sequencing offering additional benefits due to its higher accuracy in resolving complex genomic regions. To pinpoint the genomic regions responsible for pelvic spine reduction in the upper lake's sticklebacks, a transcriptomic study could provide valuable gene expression analysis. Additionally, conducting breeding experiments between spined and spineless specimens from the upper lake and specimens from a second lake would offer further insights into the genetic mechanisms underlying this trait. In addition, closing in on the genomic location of the locus/loci for spinelessness could be achieved by carrying out a QTL analysis with fish from Lake Storvatnet. Further, genome editing techniques could be useful to examine the role of the particular gene in pelvic spine reduction.

References

Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E. & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol,* 21**:** 30.

Ameur, A.; Stewart, J.B.; Freyer, C.; Hagström, E.; Ingman, M.; Larsson, N.-G. & Gyllensten, U. (2011). Ultra-Deep Sequencing of Mouse Mitochondrial DNA: Mutational Patterns and Their Origins. *PLoS Genet,* 7**:** e1002028.

Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon,* 7**:** e08093.

Anand, S.; Mangano, E.; Barizzone, N.; Bordoni, R.; Sorosina, M.; Clarelli, F.; Corrado, L.; Martinelli Boneschi, F.; D'alfonso, S. & De Bellis, G. (2016). Next Generation Sequencing of Pooled Samples: Guideline for Variants' Filtering. *Sci Rep,* 6**:** 33735.

Arendt, J. & Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol,* 23**:** 26-32.

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

Barber, I. & Scharsack, J.P. (2010). The three-spined stickleback-*Schistocephalus solidus* system: an experimental model for investigating host-parasite interactions in fish. *Parasitology,* 137**:** 411-424.

Barrett, R.D.; Paccard, A.; Healy, T.M.; Bergek, S.; Schulte, P.M.; Schluter, D. & Rogers, S.M. (2011). Rapid evolution of cold tolerance in stickleback. *Proc Biol Sci,* 278**:** 233-238.

Barrett, R.D.H. & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol Evol,* 23**:** 38-44.

Baumgartner, J.V. & Bell, M.A. (1984). Lateral Plate Morph Variation in California Populations of the Threespine Stickleback, *Gasterosteus aculeatus*. *Evolution,* 38**:** 665-674.

Beaumont, M.A. (2005). Adaptation and speciation: what can Fst tell us? *Trends Ecol Evol,* 20**:** 435-440.

Bell, M. A. (1994). Palaeobiology and evolution of threespine stickleback. In M. A. Bell & S. A. Foster (ed.), *The evolutionary biology of the threespine stickleback*. Oxford University Press, Oxford, pp 438-471.

Bell, M.A. (2001). Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetica,* 112**:** 445-461.

Bell, M.A. & Ortí, G. (1994). Pelvic Reduction in Threespine Stickleback from Cook Inlet Lakes: Geographical Distribution and Intrapopulation Variation. *Copeia,* 1994**:** 314-325.

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

Berdan, E.L.; Barton, N.H.; Butlin, R.; Charlesworth, B.; Faria, R.; Fragata, I.; Gilbert, K.J.; Jay, P.; Kapun, M.; Lotterhos, K.E.; Mérot, C.; Durmaz mitchell, E.; Pascual, M.; Peichel, C.L.; Rafajlović, M.; Westram, A.M.; Schaeffer, S.W.; Johannesson, K. & Flatt, T. (2023). How chromosomal inversions reorient the evolutionary process. *J Evol Biol,* n/a**:** 1-22.

Berkowicz, S.R.; Featherby, T.J.; Qu, Z.; Giousoh, A.; Borg, N.A.; Heng, J.I.; Whisstock, J.C. & Bird, P.I. (2016). Brinp1−/−mice exhibit autism-like behaviour, altered memory, hyperactivity and increased parvalbumin-positive cortical interneuron density. *Mol Autism,* 7**:** 22.

Bisschop, G.; Setter, D.; Rafajlović, M.; Baird, S.J.E. & Lohse, K. (2020). The impact of global selection on local adaptation and reproductive isolation. *Philos Trans R Soc Lond B Biol Sci,* 375**:** 20190531.

Biswas, S. & Akey, J.M. (2006). Genomic insights into positive selection. *Trends Genet,* 22**:** 437-446.

Bollaert, E.; De Rocca Serra, A. & Demoulin, J.-B. (2019). The HMG box transcription factor HBP1: a cell cycle inhibitor at the crossroads of cancer signaling pathways. *Cell Mol Life Sci,* 76.

Bolnick, D.I.; Barrett, R.D.H.; Oke, K.B.; Rennison, D.J. & Stuart, Y.E. (2018). (Non)Parallel Evolution. *Annu Rev Ecol Evol Syst,* 49**:** 303-330.

Brandt, D.Y.C.; César, J.; Goudet, J. & Meyer, D. (2018). The Effect of Balancing Selection on Population Differentiation: A Study with HLA Genes. *Genes Genom Genet,* 8**:** 2805-2815.

Brumfield, R.T.; Beerli, P.; Nickerson, D.A. & Edwards, S.V. (2003). The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol Evol,* 18**:** 249-256.

Cai, Z.; Deng, X.; Jia, J.; Wang, D. & Yuan, G. (2021). Ectodysplasin A/Ectodysplasin A Receptor System and Their Roles in Multiple Diseases. *Front Physiol,* 12.

Campbell, N.A.; Urry, L.A.; Cain, M.L.; Wasserman, S.A. & Minorsky, P.V. (2020). *Biology: A Global Approach, Global Edition*: Pearson Education.

Chan, A.H.E.; Chaisiri, K.; Saralamba, S.; Morand, S. & Thaenkham, U. (2021). Assessing the suitability of mitochondrial and nuclear DNA genetic markers for molecular systematics and species identification of helminths. *Parasit Vectors,* 14**:** 233.

Chan, Y.F.; Marks, M.E.; Jones, F.C.; Villarreal, G., Jr.; Shapiro, M.D.; Brady, S.D.; Southwick, A.M.; Absher, D.M.; Grimwood, J.; Schmutz, J.; Myers, R.M.; Petrov, D.; Jonsson, B.; Schluter, D.; Bell, M.A. & Kingsley, D.M. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science,* 327**:** 302-305.

Charité, J.; Mcfadden, D.G. & Olson, E.N. (2000). The bHLH transcription factor dHAND controls Sonic hedgehog expression and establishment of the zone of polarizing activity during limb development. *Development,* 127**:** 2461-2470.

Charlesworth, B. (1992). Evolutionary Rates in Partially Self-Fertilizing Species. *Am Nat,* 140**:** 126-148.

Chen, T.; Zhou, L.; Yuan, Y.; Fang, Y.; Guo, Y.; Huang, H.; Zhou, Q. & Lv, X. (2014). Characterization of Bbx, a member of a novel subfamily of the HMG-box superfamily together with Cic. *Dev Genes Evol,* 224**:** 261-268.

Cheng, C.-W.; Hsiao, J.-R.; Fan, C.-C.; Lo, Y.-K.; Tzen, C.-Y.; Wu, L.-W.; Fang, W.-Y.; Cheng, A.-J.; Chen, C.-H.; Chang, I.-S.; Jiang, S.S.; Chang, J.-Y. & Lee, A.Y.-L. (2016). Loss of GDF10/BMP3b as a prognostic marker collaborates with TGFBR3 to enhance chemotherapy resistance and epithelial-mesenchymal transition in oral squamous cell carcinoma. *Mol Carcinog,* 55**:** 499-513.

Cole, N.J.; Tanaka, M.; Prescott, A. & Tickle, C. (2003). Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. *Curr Biol,* 13**:** R951-R952.

Colosimo, P.F.; Hosemann, K.E.; Balabhadra, S.; Villarreal, G., Jr.; Dickson, M.; Grimwood, J.; Schmutz, J.; Myers, R.M.; Schluter, D. & Kingsley, D.M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science,* 307**:** 1928-1933.

Colosimo, P.F.; Peichel, C.L.; Nereng, K.; Blackman, B.K.; Shapiro, M.D.; Schluter, D. & Kingsley, D.M. (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol,* 2**:** E109.

Coyle, S.M.; Huntingford, F.A. & Peichel, C.L. (2007). Parallel evolution of Pitx1 underlies pelvic reduction in Scottish threespine stickleback (*Gasterosteus aculeatus*). *J Hered,* 98**:** 581-586.

Cremona, M.; Colombo, E.; Andreazzoli, M.; Cossu, G. & Broccoli, V. (2004). Bsx, an evolutionary conserved Brain Specific homeoboX gene expressed in the septum, epiphysis, mammillary bodies and arcuate nucleus. *Gene Expr Patterns,* 4**:** 47-51.

Cresko, W.; Yan, Y.L.; Baltrus, D.; Amores, A.; Singer, A.; Rodríguez-Marí, A. & Postlethwait, J. (2003). Genome Duplication, Subfunction Partitioning, and Lineage Divergence: Sox9 in Stickleback and Zebrafish. *Dev Dyn:,* 228**:** 480-489.

Cresko, W.A.; Amores, A.; Wilson, C.; Murphy, J.; Currey, M.; Phillips, P.; Bell, M.A.; Kimmel, C.B. & Postlethwait, J.H. (2004). Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci U.S.A.,* 101**:** 6050-6055.

Cutler, D.J. & Jensen, J.D. (2010). To pool, or not to pool? *Genetics,* 186**:** 41-43.

Daco, L.; Matthies, D.; Hermant, S. & Colling, G. (2022). Genetic diversity and differentiation of populations of Anthyllis vulneraria along elevational and latitudinal gradients. *Ecol Evol,* 12**:** e9167.

Defaveri, J.; Shikano, T.; Shimada, Y.; Goto, A. & Merilä, J. (2011). Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution,* 65**:** 1800-1807.

Dicarlo, J.E.; Norville, J.E.; Mali, P.; Rios, X.; Aach, J. & Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res,* 41**:** 4336-4343.

Don, E.K.; Currie, P.D. & Cole, N.J. (2013). The evolutionary history of the development of the pelvic fin/hindlimb. *J Anat,* 222**:** 114-133.

Douglas, A.E. (2019). Simple animal models for microbiome research. *Nat Rev Microbiol,* 17**:** 764-775.

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

Fang, B.; Kemppainen, P.; Momigliano, P.; Feng, X. & Merilä, J. (2020). On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nat Ecol Evol,* 4**:** 1105-1115.

Fang, B.; Merilä, J.; Ribeiro, F.; Alexandre, C.M. & Momigliano, P. (2018). Worldwide phylogeny of three-spined sticklebacks. *Mol Phylogenet Evol,* 127**:** 613-625.

Fariello, M.I.; Boitard, S.; Naya, H.; Sancristobal, M. & Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics,* 193**:** 929-941.

Feneck, E. & Logan, M. (2020). The Role of Retinoic Acid in Establishing the Early Limb Bud. *Biomolecules,* 10**:** 312.

Frank, S.A. (2014). Generative models versus underlying symmetries to explain biological pattern. *J Evol Biol,* 27**:** 1172-1178.

Frankham, R.; Ballou, J.D.; Ralls, K.; Eldridge, M.; Dudash, M.R.; Fenster, C.B.; Lacy, R.C. & Sunnucks, P. (2017). Genetic Management of Fragmented Animal and Plant Populations. Oxford University Press, Oxford, pp 15-40.

Galli, A.; Robay, D.; Osterwalder, M.; Bao, X.; Bénazet, J.D.; Tariq, M.; Paro, R.; Mackem, S. & Zeller, R. (2010). Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. *PLoS Genet,* 6**:** e1000901.

Galtier, N.; Nabholz, B.; Glémin, S. & Hurst, G.D.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol,* 18**:** 4541-4550.

Gandon, S. & Nuismer, S.L. (2009). Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *Am Nat,* 173**:** 212-224.

Gautier, M.; Foucaud, J.; Gharbi, K.; Cézard, T.; Galan, M.; Loiseau, A.; Thomson, M.; Pudlo, P.; Kerdelhué, C. & Estoup, A. (2013). Estimation of population allele frequencies from next-generation sequencing data: pool-versus individual-based genotyping. *Mol Ecol,* 22**:** 3766-3779.

Germano, G.; Porazzi, P. & Felix, C.A. (2022). Leukemia-associated transcription factor mllt3 is important for primitive erythroid development in zebrafish embryogenesis. *Dev Dyn,* 251**:** 1728-1740.

Giaimo, B.D.; Ferrante, F.; Herchenröther, A.; Hake, S.B. & Borggrefe, T. (2019). The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin,* 12**:** 37.

Gibson, G. & Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nat Rev Genet,* 5**:** 681-690.

Giles, N. (1983). The possible role of environmental calcum levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback, *Gasterosteus aculeatus*. *J Zool,* 199**:** 535-544.

Gissi, C.; Iannelli, F. & Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity,* 101**:** 301-320.

Glazer, A.M.; Killingbeck, E.E.; Mitros, T.; Rokhsar, D.S. & Miller, C.T. (2015). Genome Assembly Improvement and Mapping Convergently Evolved Skeletal Traits in Sticklebacks with Genotyping-by-Sequencing. *Genes Genom Genet,* 5**:** 1463-1472.

Gregorius, H.R. (1987). The relationship between the concepts of genetic diversity and differentiation. *Theor Appl Genet,* 74**:** 397-401.

Guo-Liang, J. (2013). Molecular Markers and Marker-Assisted Breeding in Plants. In S.V. Andersen (ed), *Plant Breeding from Laboratories to Fields.* Rijeka: IntechOpen, pp 45-83.

Häärä, O.; Harjunmaa, E.; Lindfors, P.H.; Huh, S.H.; Fliniaux, I.; Åberg, T.; Jernvall, J.; Ornitz, D.M.; Mikkola, M.L. & Thesleff, I. (2012). Ectodysplasin regulates activator-inhibitor balance in murine tooth development through Fgf20 signaling. *Development,* 139**:** 3189-3199.

Hagen, D. & Gilbertson, L. (1972). Geographic variation and environmental selection in *Gasterosteus aculeatus L.* in the Pacific Northwest, America. *Evolution***:** 32-51.

Hagen, D.W. & Gilbertson, L.G. (1973). Selective predation and the intensity of selection acting upon the lateral plates of threespine sticklebacks. *Heredity,* 30**:** 273-287.

Han, X.; Yoshizaki, K.; Miyazaki, K.; Arai, C.; Funada, K.; Yuta, T.; Tian, T.; Chiba, Y.; Saito, K.; Iwamoto, T.; Yamada, A.; Takahashi, I. & Fukumoto, S. (2018). The transcription factor NKX2-3 mediates p21 expression and ectodysplasin-A signaling in the enamel knot for cusp formation in tooth development. *J Biol Chem,* 293**:** 14572-14584.

Hart, J.C. & Miller, C.T. (2017). Sequence-Based Mapping and Genome Editing Reveal Mutations in Stickleback Hps5 Cause Oculocutaneous Albinism and the casper Phenotype. *Genes Genom Genet,* 7**:** 3123-3131.

Hermisson, J. & Pennings, P.S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics,* 169**:** 2335-2352.

Hoban, S.; Kelley, J.L.; Lotterhos, K.E.; Antolin, M.F.; Bradburd, G.; Lowry, D.B.; Poss, M.L.; Reed, L.K.; Storfer, A. & Whitlock, M.C. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am Nat,* 188**:** 379-397.

Hoekstra, H.E.; Hirschmann, R.J.; Bundey, R.A.; Insel, P.A. & Crossland, J.P. (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science,* 313**:** 101-104.

Hoffmann, A.A. & Rieseberg, L.H. (2008). Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu Rev Ecol Evol Syst,* 39**:** 21-42.

Hohenlohe, P.A.; Bassham, S.; Etter, P.D.; Stiffler, N.; Johnson, E.A. & Cresko, W.A. (2010a). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet,* 6**:** e1000862.

Hohenlohe, P.A.; Phillips, P.C. & Cresko, W.A. (2010b). Using population genomics to detect selection in natural populations: key concepts and methodological considerations *Int J Plant Sci,* 171**:** 1059-1071.

Holsinger, K.E. & Weir, B.S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nat Rev Genet,* 10**:** 639-650.

Huntingford, F.A. & Ruiz-Gomez, M.L. (2009). Three-spined sticklebacks *Gasterosteus aculeatus* as a model for exploring behavioural biology. *J Fish Biol,* 75**:** 1943-1976.

Işık, R. & Bilgen, G. (2019). Associations between genetic variants of the POU1F1 gene and production traits in Saanen goats. *Arch Anim Breed,* 62**:** 249-255.

Itan, Y.; Gerbault, P. & Pines, G. (2015). Evolutionary Genomics. *Evol Bioinform Online,* 11**:** 53-55.

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

Jørgensen, T.E.; Karlsen, B.O.; Emblem, Å.; Breines, R.; Andreassen, M.; Rounge, T.B.; Nederbragt, A.J.; Jakobsen, K.S.; Nymark, M.; Ursvik, A.; Coucheron, D.H.; Jakt, L.M.; Nordeide, J.T.; Moum, T. & Johansen, S.D. (2018). Mitochondrial genome variation of Atlantic cod. *BMC Res Notes,* 11**:** 397.

Kapun, M. & Flatt, T. (2019). The adaptive significance of chromosomal inversion polymorphisms in *Drosophila melanogaster*. *Mol Ecol,* 28**:** 1263-1282.

Kapun, M.; Mitchell, E.D.; Kawecki, T.J.; Schmidt, P. & Flatt, T. (2023). An Ancestral Balanced Inversion Polymorphism Confers Global Adaptation. *Mol Biol Evol,* 40.

Karlsen, B.O.; Emblem, Å.; Jørgensen, T.E.; Klingan, K.A.; Nordeide, J.T.; Moum, T. & Johansen, S.D. (2014). Mitogenome sequence variation in migratory and stationary ecotypes of North-east Atlantic cod. *Mar Genomics,* 15**:** 103-108.

Karlsen, B.O.; Klingan, K.; Emblem, Å.; Jørgensen, T.E.; Jueterbock, A.; Furmanek, T.; Hoarau, G.; Johansen, S.D.; Nordeide, J.T. & Moum, T. (2013). Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Mol Ecol,* 22**:** 5098-5111.

Kingsley, D.M. & Peichel, C. (2006). The molecular genetics of evolutionary change in sticklebacks. In S. Östlund-Nilsson, I. Mayer, & F. A. Huntingford (ed), *Biology of the three-spined stickleback*. CRC press, Taylor & Francis Group, Boca Raton, pp 41-81.

Kingsley, D.M.; Zhu, B.; Osoegawa, K.; Jong, P.J.; Schein, J.E.; Marra, M.A.; Peichel, C.L.; Amemiya, C.T.; Schluter, D.; Balabhadra, S.; Friedlander, B.; Cha, Y.M.; Dickson, M.C.; Grimwood, J.; Schmutz, J.; Talbot, W.S. & Myers, R.M. (2004). New genomic tools for molecular studies of evolutionary change in threespine sticklebacks. *Behaviour,* 141**:** 1331-1344.

Kirkpatrick, M. & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics,* 173**:** 419-434.

Kishimoto, Y.; Lee, K.-H.; Zon, L.; Hammerschmidt, M. & Schulte-Merker, S. (1997). The molecular nature of zebrafish swirl: BMP2 function is essential during early dorsoventral patterning. *Development,* 124**:** 4457-4466.

Klepaker, T.; Ostbye, K.; Bernatchez, L. & Vollestad, L.A. (2012). Spatio-temporal patterns in pelvic reduction in threespine stickleback (*Gasterosteus aculeatus L*.) in Lake Storvatnet. *Evol Ecol Res,* 14**:** 169-191.

Klepaker, T.O. & Ostbye, K. (2008). Pelvic anti-predator armour reduction in Norwegian populations of the threespine stickleback: a rare phenomenon with adaptive implications? *J Zool,* 276**:** 81-88.

Kofler, R.; Orozco-Terwengel, P.; De Maio, N.; Pandey, R.; Nolte, V.; Futschik, A.; Kosiol, C. & Schlötterer, C. (2011a). PoPoolation: A Toolbox for Population Genetic Analysis of Next Generation Sequencing Data from Pooled Individuals. *PloS One,* 6**:** e15925.

Kofler, R.; Pandey, R. & Schlötterer, C. (2011b). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics,* 27**:** 3435-3436.

Lai, C.F.; Bai, S.; Uthgenannt, B.A.; Halstead, L.R.; Mcloughlin, P.; Schafer, B.W.; Chu, P.H.; Chen, J.; Otey, C.A.; Cao, X. & Cheng, S.L. (2006). Four and half lim protein 2 (FHL2) stimulates osteoblast differentiation. *J Bone Miner Res,* 21**:** 17-28.

Laurentino, T.G.; Boileau, N.; Ronco, F. & Berner, D. (2022). The ectodysplasin-A receptor is a candidate gene for lateral plate number variation in stickleback fish. *Genes Genom Genet,* 12.

Le Rouzic, A.; Ostbye, K.; Klepaker, T.O.; Hansen, T.F.; Bernatchez, L.; Schluter, D. & Vollestad, L.A. (2011). Strong and consistent natural selection associated with armour reduction in sticklebacks. *Mol Ecol,* 20**:** 2483-2493.

Lewandoski, M. & Mackem, S. (2009). Limb Development: The Rise and Fall of Retinoic Acid. *Curr Biol,* 19**:** R558-R561.

Liu, R.; Liu, H.; Chen, X.; Kirby, M.; Brown, P.O. & Zhao, K. (2001). Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell,* 106**:** 309-318.

Love, R.R.; Steele, A.M.; Coulibaly, M.B.; Traore, S.F.; Emrich, S.J.; Fontaine, M.C. & Besansky, N.J. (2016). Chromosomal inversions and ecotypic differentiation in Anopheles gambiae: the perspective from whole-genome sequencing. *Mol Ecol,* 25**:** 5889-5906.

Mäkinen, H. & Merilä, J. (2008). Mitochondrial DNA phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe - Evidence for multiple glacial refugia. *Mol Phylogenet Evol,* 46**:** 167-182.

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

Mcgonnell, I.M.; Graham, A.; Richardson, J.; Fish, J.L.; Depew, M.J.; Dee, C.T.; Holland, P.W.H. & Takahashi, T. (2011). Evolution of the Alx homeobox gene family: parallel retention and independent loss of the vertebrate Alx3 gene. *Evol Dev,* 13**:** 343-351.

Mcphail, J.D. (1992). Ecology and evolution of sympatric sticklebacks (Gasterosteus): evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. *Can J Zool,* 70**:** 361-369.

Miller, C.T.; Beleza, S.; Pollen, A.A.; Schluter, D.; Kittles, R.A.; Shriver, M.D. & Kingsley, David m. (2007). cis-Regulatory Changes in Kit Ligand Expression and Parallel Evolution of Pigmentation in Sticklebacks and Humans. *Cell,* 131**:** 1179-1189.

Minguillon, C.; Gibson-Brown, J.J. & Logan, M.P. (2009). Tbx4/5 gene duplication and the origin of vertebrate paired appendages. *Proc Natl Acad Sci U.S.A.,* 106**:** 21726-21730.

Miya, M. & Nishida, M. (2015). The mitogenomic contributions to molecular phylogenetics and evolution of fishes: a 15-year retrospect. *Ichthyol Res* 62**:** 29-71.

Miya, M.; Takeshima, H.; Endo, H.; Ishiguro, N.B.; Inoue, J.G.; Mukai, T.; Satoh, T.P.; Yamaguchi, M.; Kawaguchi, A.; Mabuchi, K.; Shirai, S.M. & Nishida, M. (2003). Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol Phyl Evol,* 26**:** 121-138.

Moodie, G. (1972). Predation, natural selection and adaptation in an unusual threespine stickleback. *Heredity,* 28**:** 155-167.

Nadeem, M.A.; Nawaz, M.A.; Shahid, M.Q.; Doğan, Y.; Comertpay, G.; Yıldız, M.; Hatipoğlu, R.; Ahmad, F.; Alsaleh, A.; Labhane, N.; Özkan, H.; Chung, G. & Baloch, F.S. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol Biotechnol Equip,* 32**:** 261-285.

Nath, S.; Shaw, D.E. & White, M.A. (2021). Improved contiguity of the threespine stickleback genome using long-read sequencing. *Genes Genom Genet,* 11.

Nei, M. & Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U.S.A.,* 76**:** 5269-5273.

Nignan, C.; Poda, B.S.; Sawadogo, S.P.; Maïga, H.; Dabiré, K.R.; Gnankine, O.; Tripet, F.; Roux, O. & Diabaté, A. (2022). Local adaptation and colonization are potential factors affecting sexual competitiveness and mating choice in *Anopheles coluzzii* populations. *Sci Rep,* 12**:** 636.

Nishimoto, S.; Minguillon, C.; Wood, S. & Logan, M.P. (2014). A combination of activation and repression by a colinear Hox code controls forelimb-restricted expression of Tbx5 and reveals Hox protein specificity. *PLoS Genet,* 10**:** e1004245.

Nosil, P.; Soria-Carrasco, V.; Villoutreix, R.; De-La-Mora, M.; De Carvalho, C.F.; Parchman, T.; Feder, J.L. & Gompert, Z. (2023). Complex evolutionary processes maintain an ancient chromosomal inversion. *Proc Natl Acad Sci U.S.A.,* 120**:** e2300673120.

O'brown, N.M.; Summers, B.R.; Jones, F.C.; Brady, S.D. & Kingsley, D.M. (2015). A recurrent regulatory change underlying altered expression and Wnt response of the stickleback armor plates gene EDA. *eLife,* 4**:** e05290.

Osterwalder, M.; Speziale, D.; Shoukry, M.; Mohan, R.; Ivanek, R.; Kohler, M.; Beisel, C.; Wen, X.; Scales, S.J.; Christoffels, V.M.; Visel, A.; Lopez-Rios, J. & Zeller, R. (2014). HAND2 targets define a network of transcriptional regulators that compartmentalize the early limb bud mesenchyme. *Dev Cell,* 31**:** 345-357.

Paccard, A.; Hanson, D.; Stuart, Y.E.; Von Hippel, F.A.; Kalbe, M.; Klepaker, T.; Skúlason, S.; Kristjánsson, B.K.; Bolnick, D.I.; Hendry, A.P. & Barrett, R.D.H. (2019). Repeatability of Adaptive Radiation Depends on Spatial Scale: Regional Versus Global Replicates of Stickleback in Lake Versus Stream Habitats. *J Hered,* 111**:** 43-56.

Peichel, C.L.; Nereng, K.S.; Ohgi, K.A.; Cole, B.L.; Colosimo, P.F.; Buerkle, C.A.; Schluter, D. & Kingsley, D.M. (2001). The genetic architecture of divergence between threespine stickleback species. *Nature,* 414**:** 901-905.

Petit, F.; Sears, K.E. & Ahituv, N. (2017). Limb development: a paradigm of gene regulation. *Nat Rev Genet,* 18**:** 245-258.

Poore, H.A.; Stuart, Y.E.; Rennison, D.J.; Roesti, M.; Hendry, A.P.; Bolnick, D.I. & Peichel, C.L. (2022). Repeated genetic divergence plays a minor role in repeated phenotypic divergence of lake-stream stickleback. *Evolution,* 77**:** 110-122.

Prezeworski, M.; Coop, G. & Wall, J.D. (2005). The signature of positive selection on standing genetic variation. *Evolution,* 59**:** 2312-2323.

Reid, K.; Bell, M.A. & Veeramah, K.R. (2021). Threespine Stickleback: A Model System For Evolutionary Genomics. *Annu Rev Genomics Hum Genet,* 22**:** 357-383.

Reimchen, T.E. (1980). Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*: an adaptation to predators? *Can J Zool,* 58**:** 1232-1244.

Reimchen, T.E. (1994). Predators and morphological evolution in threespine stickleback. *Evol Biol threespine stickleback***:** 240-276.

Reimchen, T.E. (2000). Predator Handling Failures of Lateral Plate Morphs in *Gasterosteus aculeatus*: Functional Implications for the Ancestral Plate Condition. *Behaviour,* 137**:** 1081-1096.

Roberts Kingman, G.A.; Vyas, D.N.; Jones, F.C.; Brady, S.D.; Chen, H.I.; Reid, K.; Milhaven, M.; Bertino, T.S.; Aguirre, W.E.; Heins, D.C.; Von Hippel, F.A.; Park, P.J.; Kirch, M.; Absher, D.M.; Myers, R.M.; Di Palma, F.; Bell, M.A.; Kingsley, D.M. & Veeramah, K.R. (2021). Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution. *Sc Adv,* 7**:** eabg5285.

Robinson, M.R.; Wray, N.R. & Visscher, P.M. (2014). Explaining additional genetic variation in complex traits. *Trends Genet,* 30**:** 124-132.

Rode, N.O.; Holtz, Y.; Loridon, K.; Santoni, S.; Ronfort, J. & Gay, L. (2018). How to optimize the precision of allele and haplotype frequency estimates using pooled-sequencing data. *Mol Ecol Resour,* 18**:** 194-203.

Savolainen, O.; Lascoux, M. & Merilä, J. (2013). Ecological genomics of local adaptation. *Nat Rev Genet,* 14**:** 807-820.

Schluter, D.; Clifford, E.; Nemethy, M. & Mckinnon, J. (2004). Parallel Evolution and Inheritance of Quantitative Traits. *Am Nat,* 163**:** 809-822.

Schluter, D. & Conte, G.L. (2009). Genetics and ecological speciation. *Proc Natl Acad Sci U.S.A.,* 106**:** 9955-9962.

Schröder, M.; Windhager, S.; Schaefer, K. & Ahnelt, H. (2023). Adaptability of Bony Armor Elements of the Threespine Stickleback *Gasterosteus aculeatus* (Teleostei: Gasterosteidae): Ecological and Evolutionary Insights from Symmetry Analyses. *Symmetry,* 15**:** 811.

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

Sharma, M.; Cortes-Cruz, M.; Ahern, K.R.; Mcmullen, M.; Brutnell, T.P. & Chopra, S. (2011). Identification of the pr1 gene product completes the anthocyanin biosynthesis pathway of maize. *Genetics,* 188**:** 69-79.

Sinha, S.K.; Zachariah, S.; Quiñones, H.I.; Shindo, M. & Chaudhary, P.M. (2002). Role of TRAF3 and -6 in the activation of the NF-kappa B and JNK pathways by X-linked ectodermal dysplasia receptor. *J Biol Chem,* 277**:** 44953-44961.

Skelton, P.W.; Iain Gilmour & Open University S365 Course Tream (eds) (1993). Evolution : a biological and palaeontological approach. Addison-Wesley: Milton Keynes : in associaiton with the Open University.

Srivastava, M.; Larroux, C.; Lu, D.R.; Mohanty, K.; Chapman, J.; Degnan, B.M. & Rokhsar, D.S. (2010). Early evolution of the LIM homeobox gene family. *BMC Biol,* 8**:** 4.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* 123**:** 585-595.

Terekhanova, N.V.; Logacheva, M.D.; Penin, A.A.; Neretina, T.V.; Barmintseva, A.E.; Bazykin, G.A.; Kondrashov, A.S. & Mugue, N.S. (2014). Fast Evolution from Precast Bricks: Genomics of Young Freshwater Populations of Threespine Stickleback *Gasterosteus aculeatus*. *PLoS Genet,* 10**:** e1004696.

Te Welscher, P.; Fernandez-Teran, M.; Ros, M.A. & Zeller, R. (2002). Mutual genetic antagonism involving GLI3 and dHAND prepatterns the vertebrate limb bud mesenchyme prior to SHH signaling. *Genes Dev,* 16**:** 421-426.

Thompson, A.C.; Capellini, T.D.; Guenther, C.A.; Chan, Y.F.; Infante, C.R.; Menke, D.B. & Kingsley, D.M. (2018). A novel enhancer near the Pitx1 gene influences development and evolution of pelvic appendages in vertebrates. *eLife,* 7**:** e38555.

Tickle, C. & Cole, N.J. (2004). Morphological diversity: taking the spine out of three-spine stickleback. *Curr Biol,* 14**:** R422-424.

Turner, J.R.G. (1981). Adaptation and Evolution in Heliconius: A Defense of NeoDarwinism. *Annu Rev Ecol Syst,* 12**:** 99-121.

Wellenreuther, M. & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol Evol,* 33**:** 427-440.

White, D. & Rabago-Smith, M. (2011). Genotype-phenotype associations and human eye color. *J Hum Genet,* 56**:** 5-7.

White, N.J. & Butlin, R.K. (2021). Multidimensional divergent selection, local adaptation, and speciation. *Evolution,* 75**:** 2167-2178.

Wootton, R.J. (1976). The biology of the sticklebacks. London; New York: Academic Press.

Wootton, R.J. (1984). A functional biology of sticklebacks. Croom Helm, London.

Wucherpfennig, J.I.; Miller, C.T. & Kingsley, D.M. (2019). Efficient CRISPR-Cas9 editing of major evolutionary loci in sticklebacks. *Evol Ecol Res,* 20**:** 107-132.

Xie, K.T.; Wang, G.; Thompson, A.C.; Wucherpfennig, J.I.; Reimchen, T.E.; Maccoll, A.D.C.; Schluter, D.; Bell, M.A.; Vasquez, K.M. & Kingsley, D.M. (2019). DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science,* 363**:** 81-84.

Yang, H.; Wang, H.; Shivalila, C.S.; Cheng, A.W.; Shi, L. & Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell,* 154**:** 1370-1379.

Yang, J.; Lee, S.H.; Goddard, M.E. & Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet,* 88**:** 76-82.

Yelon, D.; Ticho, B.; Halpern, M.E.; Ruvinsky, I.; Ho, R.K.; Silver, L.M. & Stainier, D.Y. (2000). The bHLH transcription factor hand2 plays parallel roles in zebrafish heart and pectoral fin development. *Development,* 127**:** 2573-2582.

Zeller, M.; Lucek, K.; Haesler, M.; Seehausen, O. & Sivasundar, A. (2012). Little evidence for a selective advantage of armour-reduced threespined stickleback individuals in an invertebrate predation experiment. *Evol Ecol,* 26**:** 1293-1309.

Zeller, R.; Lopez-Rios, J. & Zuniga, A. (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nat Rev Genet,* 10**:** 845-858.

Zhang, Y.; Tomann, P.; Andl, T.; Gallant, N.M.; Huelsken, J.; Jerchow, B.; Birchmeier, W.; Paus, R.; Piccolo, S.; Mikkola, M.L.; Morrisey, E.E.; Overbeek, P.A.; Scheidereit, C.; Millar, S.E. & Schmidt-Ullrich, R. (2009). Reciprocal requirements for EDA/EDAR/NF-kappaB and Wnt/beta-catenin signaling pathways in hair follicle induction. *Dev Cell,* 17**:** 49-61.

Zhu, M.; Yu, X.; Choo, B.; Wang, J. & Jia, L. (2012). An antiarch placoderm shows that pelvic girdles arose at the root of jawed vertebrates. *Biol Lett,* 8**:** 453-456.

Zimmerman, S.J.; Aldridge, C.L. & Oyler-Mccance, S.J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genom,* 21**:** 382.

Zuniga, A. (2015). Next generation limb development and evolution: old questions, new perspectives. *Development,* 142**:** 3810-3820.

https://geo.ngu.no/kart/losmasse_mobil/?lang=nor&map=9

Paper I

fishes

MDPI

Article

*Pitx*1 Enhancer Variants in Spined and Spine-Reduced Subarctic European Sticklebacks

Dhurba Adhikari, Ida K. Hanssen, Steinar D. Johansen, Truls B. Moum and Jarle T. Nordeide *

Genomic Division, Faculty of Biosciences and Aquaculture, Nord University, NO-8049 Bodø, Norway
* Correspondence: jarle.t.nordeide@nord.no; Tel.: +47-7551-7331

Abstract: Loss of body armour, sometimes including a reduction in or loss of pelvic spines, is an adaptation observed in many isolated freshwater populations. Pelvic reduction in sticklebacks has previously been associated with recurrent, but variant, deletions within pelvic enhancer regions *Pel*A and *Pel*B, which regulate expression of the homeodomain transcription factor gene *Pitx*1. We investigated variation in nucleotide sequences of pelvic enhancers in sticklebacks collected from two small freshwater lakes in the same watercourse and a nearby marine site in subarctic Norway. Spineless, as well as asymmetrically spined and completely spined sticklebacks are present in the upper lake, while only specimens with complete spines are found at the other lake and the marine site. Observed variation at *Pel*A between the three sites was mainly due to variable numbers of repeats at three fragile TG-repeat loci. The length of *Pel*A, mainly at one of the TG-repeat loci, was consistently shorter among individuals in the upper lake compared with specimens from the two other sites. However, no obvious association was revealed between enhancer variants and pelvic status. No polymorphism was found at *Pel*B. Thus, additional genetic factors and/or environmental cues need to be identified to fully explain the occurrence of pelvic reduction in sticklebacks in this lake.

Keywords: *Gasterosteus aculeatus*; stickleback; pelvic reduction; pelvic spines; TG-repeat; *Pitx*1; *Pel*A; *Pel*B; parallel evolution

Key Contribution: DNA sequence variation among sticklebacks suggests that additional genetic or environmental factors are involved in pelvic reduction than those shown by previous studies.

check for updates

Citation: Adhikari, D.; Hanssen, I.K.; Johansen, S.D.; Moum, T.B.; Nordeide, J.T. *Pitx*1 Enhancer Variants in Spined and Spine-Reduced Subarctic European Sticklebacks. *Fishes* **2023**, *8*, 164. https://doi.org/10.3390/fishes8030164

Academic Editor: Albert Kjartan Dagbjartarson Imsland

Received: 8 January 2023
Revised: 8 March 2023
Accepted: 9 March 2023
Published: 15 March 2023

1. Introduction

Parallel phenotypic evolution has been defined as the independent evolution of the same trait in closely related lineages [1]. Parallel phenotypic evolution in organisms colonising new habitats may be due to either de novo mutations or standing genetic variation in the ancestral population (reviewed by [2]). Authors have advocated in favour of standing genetic variation as the most plausible mechanism due to its likely presence at higher frequencies, immediate availability in the new habitat, and because it has already been tested in similar environments [3–5]. Yet other studies support de novo mutations (reviewed by [6]).

Countless freshwater populations founded by marine ancestors after the last glacial period were trapped and isolated as land uplifted due to the deglaciation [7,8]. In addition, modern times human activity and perhaps birds might have transported species from saline to some freshwater habitats [8]. The threespine stickleback (*Gasterosteus aculeatus*) is one of the species with a marine origin that has colonized freshwater habitats, followed by physiological, behavioural, and morphological adaptations [9–17]. Marine threespine sticklebacks in general are protected against numerous piscivorous predators by strong external bony structures such as rows of lateral bony plates, pelvis structure (also termed pelvic girdle, which includes the pelvic spines), and dorsal spines [18,19]. A reduction in anti-predator armour, such as the lateral bony plates, may occur within a couple of decades of isolation in freshwater [20,21].

Dorsal and pelvic spines are assumed to give efficient protection against gape-limited predators such as fishes and birds, especially since the spines can be locked in the erect position [19]. Hence, the presence and length of the spines have been reported as positively associated with predation pressure from vertebrates [22–24]. However, in Cook Inlet, Alaska, freshwater populations of threespine sticklebacks with complete or partial loss of pelvic spines seem to be relatively abundant [12,18,25] (reviewed by [26]). A few pelvic reduced, freshwater populations have been reported elsewhere as well, e.g., from Western Canada [15,27–30], Iceland [28], Scotland [31], and Norway [16], reviewed by [26]. Such pelvic spine reduction may be selected for by invertebrate larvae, which are able to grab and hold on to the spines of juvenile spined sticklebacks [9] (but see [32,33]). Thus, low abundance of fish and bird predators and high abundance of insect predators could select for absence of spines or reduced spine length in sticklebacks, and vice versa. An alternative hypothesis to this "predation hypothesis" is the "calcium hypothesis", which advocates that low calcium ion concentration in freshwater could favour pelvic spine reduction [10]. Finally, the "predation-calcium hypothesis" argues that the combined effect of predators and low calcium ion concentration would be required to explain the evolution of pelvic reduction in sticklebacks [12].

A major determinant of pelvic development in threespine sticklebacks is the pituitary homeobox transcription factor gene *Pitx*1, located at chromosome 7 [15,28,31,34]. In addition, loci located at chromosome 2 [15], chromosome 4 [15,35], and chromosome 8 [29] have been suggested to play a role in fine-tuning of pelvic spine length. An enhancer element termed *Pel*A located upstream of *Pitx*1 (Figure 1) is reported as essential for the development of pelvic spines, and deletions at this locus have been shown to be associated with pelvic loss and reduction [28]. This *Pel*A pelvic limb enhancer is a cis-regulatory sequence, which contains multiple transcription factor binding sites, interacts with corresponding transcription factors, and enhances the transcription rate of *Pitx*1 [28,36]. Another enhancer element designated *Pel*B that maps downstream of *Pitx*1 (Figure 1) has been suggested to play a role in pelvic spine modification [37].

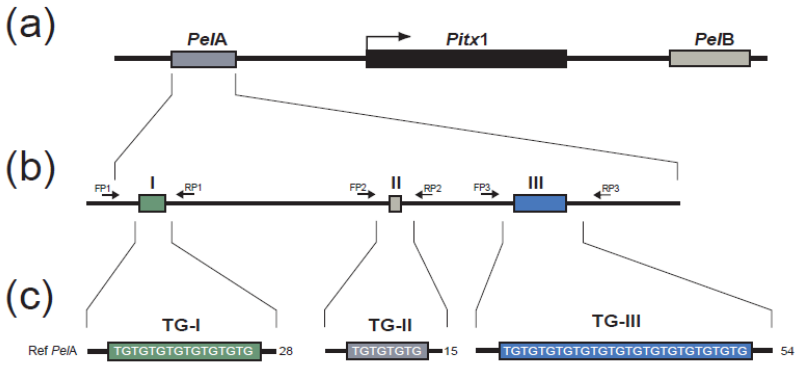


Figure 1. (**a**) *Pitx*1 with the upstream and downstream enhancers *Pel*A and *Pel*B, respectively. (**b**) *Pel*A with TG-repeats I, II, and III, and their relative location. FP1-3 and RP1-3 show the location of the forward and reverse primers used to sequence TG-repeat I, II, and III, respectively. (**c**) The reference sequence (SALR GU130435; 377,852 nt) is from a stickleback collected in Salmon River, British Columbia, and harbours 28, 15, and 54 TG-repeats at TG-repeat I, II, and III, respectively.

The *Pel*B enhancer was identified initially in mammals and is conserved between mice and fish including sticklebacks, in contrast to *Pel*A, which seems less conserved outside teleosts [37].

There are three TG-repeat arrays within the DNA-fragile region of *Pel*A, denoted TG-repeats I–III in the present paper (Figure 1), which likely contribute to deletion mutations

that are functionally related to pelvic reduction [28]. TG-repeats in the *Pel*A region may elicit a left-handed DNA helical structure, called Z-DNA [38–40]. This structure might affect the binding of transcription factors to the corresponding binding sites, causing an increase in the transcription rate [39]. Chan and colleagues [28] reported 9 different deletion patterns from 9 different spineless stickleback populations within the 2.5 kb *Pel*A region. These deletions are partially overlapping in a 488 bp region located at or near TG-repeats I–III [28]. The enhancer region's fragility and capability of forming a secondary DNA structure may explain the deletions of TG repeats I–III within the *Pel*A enhancer and the concomitant loss of pelvic spines in some threespine stickleback populations [39,40].

Studies in mammals have established the role of the *Pel*B enhancer as essential for pelvic hind limb development [37], but the corresponding biological role in sticklebacks is still a subject of interest and discussion. Spineless benthic sticklebacks from Paxton Lake in British Columbia have both a deletion of 125 bp and an insertion of 341 bp at *Pel*B, in addition to the mutations at *Pel*A (as discussed above) [37]. In addition, there might be other regulatory regions affecting pelvic development. For example, another transcriptional regulator, *Pitx*2, which is closely related to *Pitx*1, has been reported in vertebrates [41,42]. *Pitx*2 probably affects pelvis symmetry so that pelvic spines could be completely or partially lost at one side and less reduced at the other [41]. However, the role of the *Pitx*2 in pelvic spine reduction is not fully understood.

Pelvic reduction is reported from only 8 out of more than 200 Norwegian, mainly freshwater, populations examined [16,26]. In one of these lakes, Lake Storvatnet located in subarctic Northern Norway, 60% of the population lack one or both pelvic spines [43]. No pelvic reduction is observed in the downstream Lake Gjerdhaugvatnet in the same watercourse, or from a nearby marine site [43]. Specimens from both of the two freshwater populations have been categorized as "low plated" and marine specimens in this region have been categorized as "partially" and "completely" plated, based on the number of lateral bone plates [43]. Interestingly, Lake Storvatnet also contains an abundant population of brown trout (*Salmo trutta*) and identifiable stickleback parts were found in 19 per cent (N = 86) of the trout stomachs [43]. The abundance of insects, which may potentially prey on juvenile sticklebacks in Lake Storvatnet, was categorized as low [43]. A relatively large part of the population ($\geq$30%) in Lake Storvatnet has grown 2 normal pelvic spines [43].

We studied the phenotypic variation of pelvic spines and the molecular variation at *Pel*A and *Pel*B in a comparison between (i) spined and spine-reduced sticklebacks within Lake Storvatnet, and (ii) Lake Storvatnet sticklebacks and (spined) conspecifics from the downstream Lake Gjerdhaugvatnet and marine specimens. Our hypothesis was that spineless sticklebacks from Lake Storvatnet have large parts of the *Pel*A enhancer deleted, similar to their North American spineless conspecifics [28]. We also hypothesised that more of *Pel*A was deleted in spineless compared with spined specimens in Lake Storvatnet and spined fish in the two nearby sites.

2. Materials and Methods

2.1. Sample Collection

A total of 427 sticklebacks were collected at three locations, 2 freshwater and 1 marine location, at Langøya island in Northern Norway in 2017, 2019, and 2020. The position (EU89 Lat/Lon), altitude, and size of the upper Lake Storvatnet are 68°46′49″ N, 15°9′36″ E, 80 m, and 0.2 km^2, respectively (Figure 2). Lake Gjerdhaugvatnet is a small lake (0.01 km^2) located downstream in the same watercourse at 20 m altitude. The two lakes are connected by an approximately 500 m brook with several waterfalls, which most likely prevent any gene flow between stickleback populations inhabiting the two lakes. Sticklebacks were sampled from marine or brackish water in the tidal mouth of a small river at Sandstrand (68°44′45″ N, 15°20′42″ E), here referred to as the marine site. The marine site is located about 8 km (direct distance) from the two other sampling sites (Figure 2). Sampling was carried out in June 2017, 2019, and 2020 in Lake Storvatnet, in June 2019 and 2020 in Lake Gjerdhaugvatnet, and in June 2020 at the marine site. Traps were deployed at 0.3–1.0 m

depth along the shore and retrieved about 24 h later. The sticklebacks were euthanised and sacrificed by an overdose of tricaine methanesulfonate (MS222) and then rinsed with water. The total length (from head to the posterior part of the caudal fin) of the body was measured by a ruler to the nearest mm. The caudal fin was cut off and discarded. Samples of the posterior fin muscle (about 5 mm in size) were homogenized immediately by bead beating using a Dremel 8220 rotary tool (MP Biomedicals, Solon, OH, USA) and 0.5 mL DNA/RNA Shield solution (Zymo Research, Irvine, CA, USA) and kept at low temperatures before further analyses at the laboratory. Specimens with body size less than 30 mm were discarded.

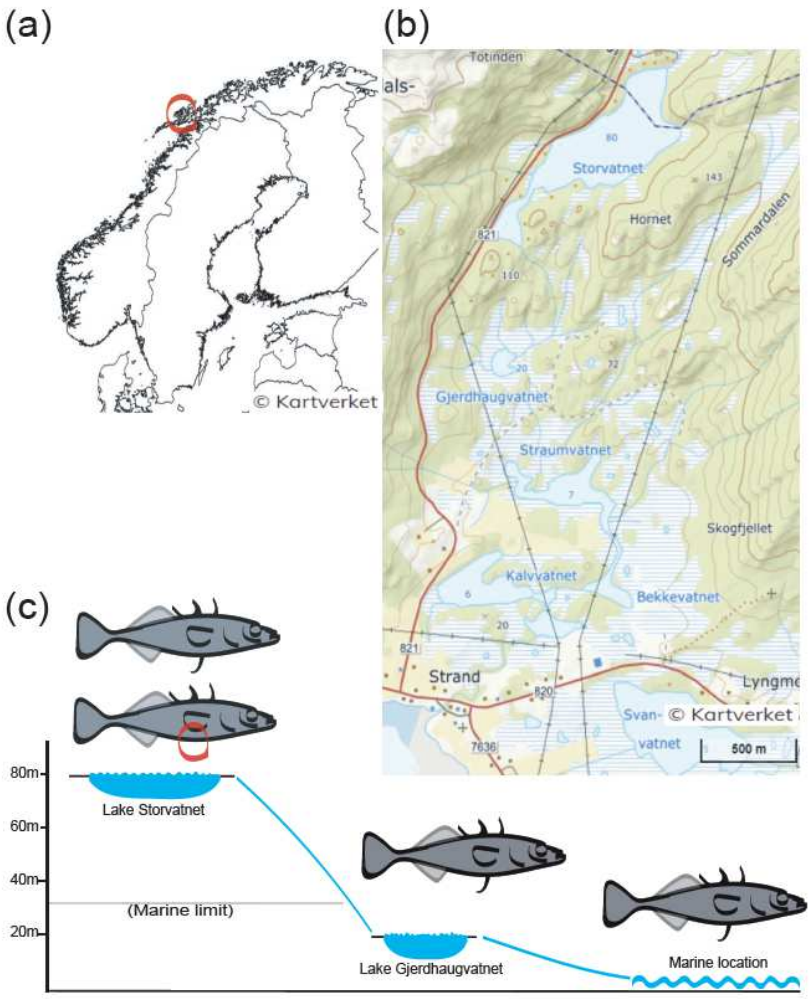


Figure 2. (**a**) Scandinavia with the study site encircled in red. (**b**) Map (www.norgeskart.no accessed on 28 March 2023) of the upper Lake Storvatnet (68°46′49″ N, 15°9′36″ E) and the lower Lake Gjerdhaugvatnet in the watercourse, whereas the marine sampling site (not shown) is located about 8 km from the two lakes. (**c**) Schematic drawing (out of scale) of the three sites showing the symmetric-spined, asymmetric-spined, and the spineless (encircled in red) sticklebacks in Lake Storvatnet. All specimens are symmetric spined at the two other sites.

To measure the Ca^{2+} concentration in Lake Storvatnet and Lake Gjerdhaugvatnet, water samples were collected on 26 June 2021 from about 10 cm depth and about 1 m from land. The water samples were analysed by Labora AS (Bodø, Norway).

2.2. Morphology and Computation of Pelvic Scores

The specimens and their spines, and the tube where each specimen was stored individually, were examined for potentially broken spines. No broken spines were detected. The lengths of the right and left pelvic spines were measured by a digital calliper and a pelvic score (PS) of 0–4 was given to each side of the pelvis. Each side of a complete pelvis consists of an anterior process (ap), an ascending branch (ab), a posterior process (pp), and a pelvic spine (ps). PS 0 is for fish with no pelvic structure at all, PS 1 is for fish with ap only, PS 2 is for fish with ap + ab, PS 3 is for fish with ap + ab + pp, and PS 4 is for fish with a complete pelvis structure (ap + ab + pp + ps). A combined pelvic score (CPS) of 0–8 was assigned to each specimen by adding up the PS of both sides of the pelvis [12] (Supplementary Figure S1). CPS 0 is for fish with no pelvic structure at all, and CPS 8 is for fish with complete pelvic structure including pelvic spines [12]. The individuals were observed under a stereomicroscope (10× or 20× magnifying lens) with gentle pressure on the pelvis by forceps to categorize PS. Samples were divided into three classes: (a) spineless, (b) symmetric spined, and (c) asymmetric-spined specimens (see details in the Supplementary Figure S1). The pelvis was defined as asymmetric if the difference between the length of the 2 pelvic spines was > 0.2 mm.

2.3. DNA Sequencing and Fragment Analysis

Muscle tissue for DNA analyses was taken from a total of 19 specimens from Lake Storvatnet. Seven symmetric-spined, six asymmetric-spined, and six spineless specimens were picked at random after categorising the sticklebacks into the three groups. Moreover, muscle tissue was sampled from 12 random specimens from Lake Gjerdhaugvatnet and 7 from the marine site. Genomic DNA was extracted from these 38 muscle tissue samples using the Monarch genomic DNA purification kit (New England Biolabs). The quality and concentration of DNA were checked with NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA) spectrophotometry. DNA samples included in the study had a concentration of $\geq$ 20 ng/µL and absorbance ratios A260/A280 = 1.80–1.90 and A260/A230 = 1.80–2.50. All amplicons used for DNA sequencing and fragment analyses were produced with LongAmp Taq polymerase (Thermo Fisher Scientific, Waltham, MA, USA).

2.3.1. *Pel*A Sequence Analyses

The genomic sequence of a *Pitx*1 allele from a marine pelvic-complete stickleback from Salmon River, British Columbia, was adopted as the reference sequence for the present study (Genbank accession no. GU130435 (377,852 nt)). Primers were designed based on the reference sequence using the "primer design tools" and "oligo analysis tools" of Eurofins Genomics (https://eurofinsgenomics.eu/) and named according to the position of their 3′ nucleotide in the reference sequence. DNA samples were amplified with forward and reverse primers: 5′-GCC CAA AAC TGA CAA AGC A-3′ (F128812) and 5′-AGC AGC AAA AGC AAA ATG AGA-3′ (R131624) targeting a 2813 bp region containing *Pel*A (*Pel*A amplicon) according to the reference. PCR was conducted with an initial denaturation at 94 °C for 90 s; 30 cycles of denaturation at 94 °C for 30 s; annealing at 59 °C for 20 s; extension at 65 °C for 150 s; and final extension at 65 °C for 10 min. PCR products were inspected by agarose gel electrophoresis, cleaned by ExoSap IT (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol, and subjected to direct sequencing with BigDye v3.1 (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. Each of the segments containing TG-repeat I, II, and III, were sequenced in the forward and reverse directions using the following sequencing primers: 5′-AGG TCC ACA GTA CAG TGC AG-3′ (F128968) (FP1 in Figure 1) and 5′-TGG GAC GAG AAG ATG CCT TCA G-3′ (R129360) (RP1 in Figure 1),

5′-GTC GAA GCA AAG AGG CGA GAC ATC-3′ (F129687) (FP2 in Figure 1) and 5′-TTC TAA AGT GGT CGC TCG GC-3′ (R129962) (RP2 in Figure 1), and 5′-GTT ATG AAG GGC CGA GCG AC-3′ (F129933) (FP3 in Figure 1) and 5′-GCG TGA CCA CAA CAA TCC G-3′ (R130252) (RP3 in Figure 1) (Supplementary Figure S2). Sequencing reactions were treated with magnetic bind and ethyl alcohol, eluted with elution buffer, and run on a 3500xL Genetic Analyzer (Applied Biosystems) following the manufacturer's instructions. Sequencing results were analysed with the help of Finch TV version 1.4.0.

2.3.2. *Pel*A Fragment Analyses

The allelic length variation of the TG-repeats in the *Pel*A region was determined using fragment analyses. All three TG-repeats were amplified from DNA samples with primers corresponding to those applied in sequencing, and the forward primers to identify TG-repeat I, TG-repeat II, and TG-repeat III, were fluorescently labelled with FAM, FAM, and ATT056, respectively. Then, amplicons were diluted to 1:200 and treated with HiDi formamide (Thermo Fisher Scientific, Waltham, MA, USA) and run with standard ladder GeneScan 500 LIZ (Thermo Fisher Scientific, Waltham, MA, USA) on the 3500xL Genetic Analyzer. Fragment data were analysed with the software GeneMarker version 2.6.3. The exact number of TG-repeats for each allele was inferred, based on DNA sequencing and fragment analysis in combination. No conflict was revealed between the Sanger sequencing and the fragment analysis data.

2.3.3. *Pel*B Sequence Analyses

The *Pel*B region was studied for 10, 6, and 5 specimens from Lake Storvatnet, Lake Gjerdhaugvatnet, and from the marine site, respectively. DNA samples were amplified with a forward and a reverse primer: 5′-CAC GGA TTA CTG AGC AGC AA-3′ (F176680) and 5′-AGC TCA AGA CCT CTG GAT GG-3′ (R177688), targeting a 1009 bp region that harbours *Pel*B. PCR conditions consisted of an initial denaturation at 94 °C for 90 s; 25 cycles of denaturation at 94 °C for 30 s; annealing at 59 °C for 20 s; extension at 65 °C for 90 s; and final extension at 65 °C for 10 min. A 671 bp segment of the amplicon where length polymorphism was previously reported by [37] was sequenced in both directions as detailed above, using primers 5′-ACA GAC AGA CAG ACA GAC AG-3′ (F176836) and 5′-TAT ATC AAT CGA GAG AGG AAG AGG-3′ (R177550).

2.3.4. Identification of Single Nucleotide Polymorphisms (SNPs)

Successfully retrieved sequences from all specimens, including the upstream and downstream flanking sequences of TG-repeats I–III of the *Pel*A region, as well as *Pel*B sequences, were aligned to the reference sequence (GU130435) and their SNPs identified. Alignment was conducted with the help of Clustal Omega tools (https://www.ebi.ac.uk/Tools/msa/clustalo/).

The study was carried out according to ethical guidelines stated by the Norwegian Ministry of Agriculture and Food through the Animal Welfare Act. According to these guidelines, we were not required to—and therefore do not—have a specific approval or approval number.

3. Results

3.1. Morphology, Pelvic Scores, and Ca^{2+} Concentration

Descriptive statistics of body length and length of the pelvic spines of sticklebacks from Lake Storvatnet, Lake Gjerdhaugvatnet, and the marine site, are presented in Table 1. Of the 304 specimens from Lake Storvatnet, 113 (37%) were symmetric spined, 99 (33%) were asymmetric spined, and 92 (30%) were spineless (Table 1). The polymorphic sticklebacks in Lake Storvatnet were classified into eleven groups based on their PS and CPS scores (Table 2). Note that none of these specimens had a CPS of 0 which means that none lacked the entire pelvic girdle (Table 2, Supplementary Figure S1). Among the asymmetric-spined sticklebacks from Lake Storvatnet, 29 had right-biased asymmetric

pelvic spines, and 70 had left-biased asymmetric pelvic spines (Table 1). All specimens collected from Lake Gjerdhaugvatnet (N = 73) and from the marine site (N = 50) were fully spined (CPS = 8) and symmetric (Table 1). Of the asymmetric-spined fish, 29 and 70 were right- and left-biased, respectively (Table 1), which is significantly different from unity ($\chi^2 = 16.9$, $p < 0.001$, d.f. = 1, chi-square test). Moreover, after including the one right-biased asymmetric spineless individual (Table 2) "P.v. j"), the difference is significant ($\chi^2 = 16.0$, $p < 0.001$, d.f. = 1).

Table 1. Morphological measurements of threespine sticklebacks from the two freshwater lakes, Lake Storvatnet and Lake Gjerdhaugvatnet, and a marine site. The mean of pelvic spine lengths from Lake Storvatnet was calculated based on specimens with spines and asymmetric-spined specimens only. Pelvic scores (PS) were calculated for both the left and right side of the specimens and vary from 0–4. The combined pelvic score (CPS) is the sum of the PS from both sides and varies from 0–8. "N" is the total number of specimens from each location.

Site	N	Spineless	Symmetric Spined	Asymmetric		Spine Length (cm) (Mean ± Sd)	Body Length (cm) (Mean ± Sd)
				Right-Biased	Left-Biased		
Storvatnet	304	92 (30%)	113 (37%)	29 (10%)	70 (23%)	0.26 ± 0.100	4.7 ± 0.60
Gjerdhaugvatnet	73	0	73	0	0	0.37 ± 0.070	4.1 ± 0.60
Marine	50	0	50	0	0	0.55 ± 0.100	4.8 ± 0.70

Table 2. Number of morphological variants from the three examined populations, Lake Storvatnet, Lake Gjerdhaugvatnet and the marine site. See also Supplementary Figure S1. "n" is the number of each morphological variant from each location. [1] The left column (''P.v.") refers to the pelvic spine morphs as shown in the Supplementary Figure S1.

P.v. [1]	Pelvic Scores		Combined Pelvic Scores (CPS)	Locations (n)			Remarks
	Left PS	Right PS		Storvatn	Gjerdhaugvatn	Marine	
b	4	4	8	113	73	50	Symmetric spined
c	4	4 short	8	35	0	0	Left-biased asymmetry
d	4 short	4	8	22	0	0	Right-biased asymmetry
f	3	4	7	7	0	0	Right-biased asymmetry
e	4	3	7	29	0	0	Left-biased asymmetry
h	4	1	5	5	0	0	Left-biased asymmetry
g	4	2	6	1	0	0	Left-biased asymmetry
i	3	3	6	32	0	0	Spineless
j	1	3	4	1	0	0	Spineless
l	1	1	2	51	0	0	Spineless
k	2	2	4	8	0	0	Spineless

The measured Ca^{2+} concentration in Lake Storvatnet and Lake Gjerdhaugvatnet was 0.90 mg/L and 0.84 mg/L, respectively.

3.2. Allelic Variation of *PelA*

Allelic variation of *PelA* was primarily caused by variable numbers of TG dinucleotides at three TG-repeat arrays. The allelic variation found in Lake Storvatnet differed strikingly from that of the two other sampling sites. TG-repeats I and III in particular, showed a wide range of length variants, with generally shorter arrays found in individuals from Lake Storvatnet. At TG-repeat III, the predominant allele in Lake Storvatnet was $(TG)_4$, compared with $(TG)_{27}$ in Lake Gjerdhaugvatnet and $(TG)_{31}$ at the marine site. The repeat numbers of the most common allelic variants for each of the five groups of sticklebacks examined are shown in Figure 3. Below is a more detailed assessment of the TG-repeat arrays for all specimens analysed.

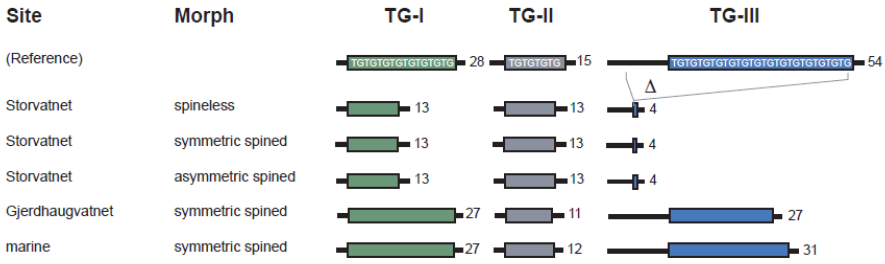


Figure 3. Representative examples of allelic variants of *Pel*A caused by variable numbers of TG dinucleotides at TG-repeats I–III. The repeat number of the most common allelic variants is shown for each of the five groups of sticklebacks examined. For more details see the Supplementary Figures S2 and S3.

3.2.1. TG-Repeat Array I

At TG-repeat I, the repeat length varied from $(TG)_{13}$ to $(TG)_{32}$ among sticklebacks from the three sites sampled (Figure 4a, Supplementary Figures S2 and S3). In Lake Storvatnet, repeat lengths ranged from $(TG)_{13}$ to $(TG)_{26}$ (Figure 4a, Supplementary Figure S2). $(TG)_{13}$ dominated in this lake with a frequency of 0.87, the presence of at least one copy in all specimens except in one spineless individual, and low heterozygosity (0.21). The array lengths of spineless and symmetrical spined specimens in Lake Storvatnet were within the same interval, from $(TG)_{13}$ to $(TG)_{32}$ (Figure 4a, Supplementary Figure S2a,b), and all asymmetrical specimens had $(TG)_{13}$ (Figure 4a, Supplementary Figure S2c). TG-repeat I among specimens from Lake Gjerdhaugvatnet varied from $(TG)_{26}$ to $(TG)_{30}$ and was relatively uniform although all specimens were heterozygous (Figure 4a, Supplementary Figure S3a). Their conspecifics at the marine site varied more at TG-repeat I, from $(TG)_{13}$ to $(TG)_{32}$ with a heterozygosity of 0.86 (Figure 4a, Supplementary Figure S3b).

3.2.2. TG-Repeat Array II

At TG-repeat II, the repeat length varied from $(TG)_9$ to $(TG)_{16}$ among all sticklebacks sampled (Figure 4b, Supplementary Figures S2 and S3). In Lake Storvatnet, repeat lengths ranged from $(TG)_{11}$ to $(TG)_{16}$, and $(TG)_{13}$ was present in all the examined specimens, with an allele frequency of 0.92 (Figure 4b, Supplementary Figure S2). The TG-repeats II of spineless and asymmetric-spined specimens in Lake Storvatnet were within the same interval, from $(TG)_{13}$ to $(TG)_{16}$ repeats, whereas two symmetric individuals were heterozygous $(TG)_{11/13}$ (Figure 4b and Supplementary Figure S2). In Lake Gjerdhaugvatnet, the number of $(TG)_n$ was uniform with all specimens being homozygous for $(TG)_{11}$ (Figure 4b, Supplementary Figure S3), whereas their marine conspecifics varied from $(TG)_9$ to $(TG)_{16}$, with a single heterozygous individual (Figure 4b, Supplementary Figure S3a).

3.2.3. TG-Repeat Array III

At TG-repeat III, the number of repeats varied from $(TG)_4$ to $(TG)_{47}$ among sticklebacks from the three sites sampled (Figure 4c). In Lake Storvatnet, $(TG)_n$ varied from $(TG)_4$ to $(TG)_{43}$ (Figure 4c, Supplementary Figure S2). The short $(TG)_4$ dominated with an allele frequency of 0.76 and the presence of at least one copy in each of the spineless (Supplementary Figure S2a), symmetric-spined (Supplementary Figure S2b), and asymmetric-spined (Supplementary Figure S2c) specimens. Spineless, symmetric-spined and asymmetric-spined specimens in Lake Storvatnet had TG-III repeats within the same interval from $(TG)_4$ to approximately $(TG)_{43}$ (Figure 4c, Supplementary Figure S2). In Lake Gjerdhaugvatnet, the number of repeats at TG-repeat III varied from $(TG)_{25}$ to $(TG)_{30}$, and $(TG)_{27}$ was present with at least 1 copy in all but 1 of the 12 examined specimens and with an allele frequency of 0.79 (Figure 4c, Supplementary Figure S3a). Specimens from the

marine site varied from $(TG)_{24}$ to $(TG)_{47}$ and were heterozygous throughout (Figure 4c, Supplementary Figure S3b).

An additional polymorphism was found upstream of and flanking TG-repeat III in specimens from Lake Storvatnet. At this upstream flanking region, all specimens from Lake Storvatnet had a 58 bp deletion compared with the reference (Figure 5; Supplementary Figure S6-3). In contrast, all of the examined specimens from Lake Gjerdhaugvatnet (N = 12) and the marine site (N = 7) conformed to the reference in this respect.

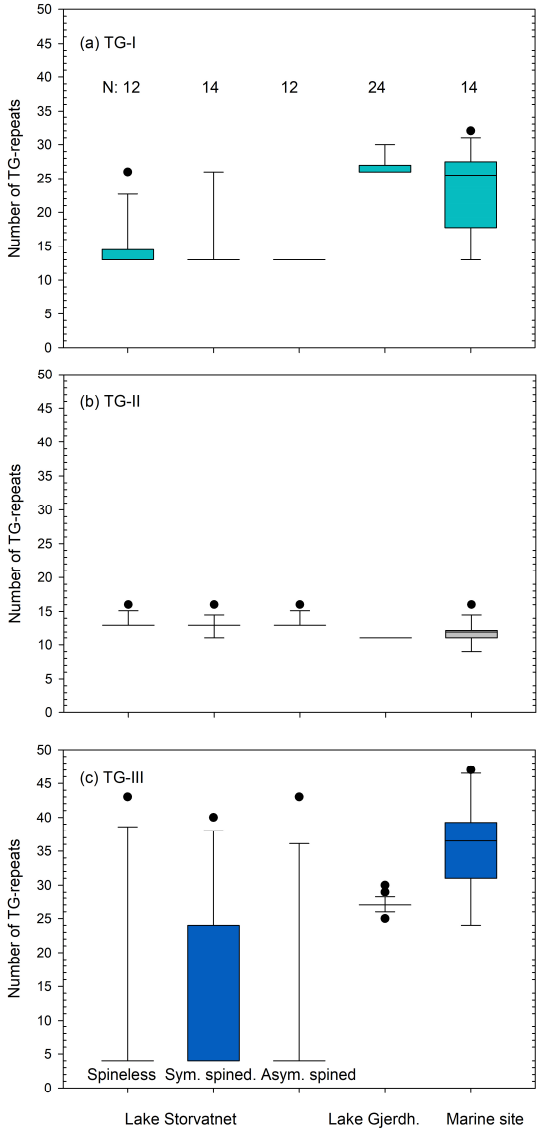


Figure 4. Box–whisker plots showing the number of repeats of (**a**) TG-I, (**b**) TG-II, and (**c**) TG-III, for spineless, symmetric, and asymmetric specimens from Lake Storvatnet, and spined specimens from Lake Gjerdhaugvatnet and the marine site. See also Supplementary Figures S2 and S3. Numbers ("N") in the figure show the number of alleles examined.

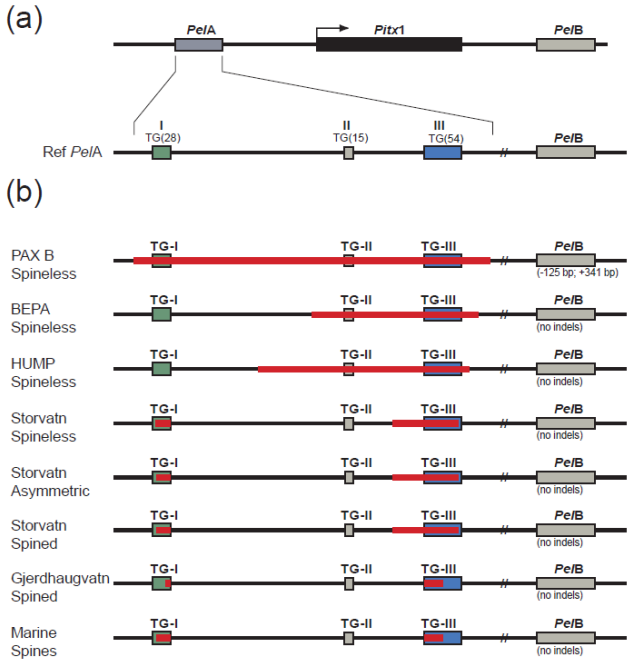


Figure 5. (**a**) The relative positions of the *Pitx*1 gene and its upstream and downstream enhancers, *Pel*A and *Pel*B. In addition, the relative positions within *Pel*A of TG-repeat I (green), TG-repeat II (grey), and TG-repeat III (blue) are also shown. The approximate downstream position (out of scale) of *Pel*B (grey) is indicated as well, as are the number of TG-repeats at TG-repeat I, II, and III for the reference sequence (GU130435) from Salmon River, British Columbia. (**b**) Polymorphism at TG-repeat I, II, and III located at *Pel*A, and at *Pel*B in spineless sticklebacks from three previously published studies from Paxton Benthic Lake (PAXB), Bear Paw Lake (BEPA), and Hump Lake (HUMP) from the west coast of North America [28]. Spineless, asymmetric-spined, and symmetric-spined sticklebacks from the present study sites of Lake Storvatnet (Storvatn), Lake Gjerdhaugvatnet (Gjerdhaugvatn), and the marine sample (Marine) in sub-Arctic Norway at the north-east coast of Europe are also shown. Missing regions at *Pel*A compared with the reference sequence are indicated in red.

3.2.4. Comparing Haplotypes of Spineless and Spined Sticklebacks from Lake Storvatnet

The combination of Sanger sequencing and fragment analyses enabled the haplotypes of *Pel*A to be inferred for specimens that were either homozygous throughout or heterozygous at one of the TG-repeats only. Three spineless (S30, S31, and S42, Supplementary Figure S2a), two symmetric-spined (S03, S07, Supplementary Figure S2b), and four asymmetric-spined specimens (S23, S33, S34 and S43, Supplementary Figure S2c) were all homozygous with haplotypes $(TG)_{13}$, $(TG)_{13}$, and $(TG)_4$ at TG-repeat I, TG-repeat II, and TG-repeat III, respectively. The same haplotype is also the most frequent in each of the 3 groups making up at least 58%, 36%, and 75% of the haplotypes among spineless, symmetric-spined, and asymmetric-spined individuals, respectively (Supplementary Figure S2a–c).

3.3. Allelic Variation of PelB

No indels were revealed by DNA sequencing of the *Pel*B region among the 10, 6, and 5 examined specimens from Lake Storvatnet, Lake Gjerdhaugvatnet, and the marine site, respectively (Figure 5b, Supplementary Figure S6-4). Sequence reads were ambiguous in between two variable poly-G runs (reference 176,958–177,294), but the sequence analyses and gel-based sizing of *Pel*B amplicons from all individuals both supported a lack of sizable

indels within the *Pel*B region (Supplementary Figures S5 and S6-4c). There was no association between pelvic morphs and SNPs upstream of the first poly-G tract (Supplementary Figure S6-4a) and downstream of the last poly-G tract (Supplementary Figure S6-4b).

3.4. Sequence Alignments of PelA and PelB

Sequence analyses revealed a number of SNPs in the pelvic enhancer regions. However, there was no apparent association between nucleotide polymorphism and pelvic status among the sticklebacks (Supplementary Figure S6).

4. Discussion

A causal connection between deletion mutations affecting the enhancer elements of the *Pitx*1 gene and loss of pelvic spines has been established in several independent stickleback populations in North America, making a strong case for parallel evolution by common molecular pathways [28]. As far as we know, the stickleback population in Lake Storvatnet is unique by its lack of any such obvious relationship between indels mapping to *Pitx*1 enhancer regions and pelvic status, ranging from fully spined via asymmetric spined to spineless.

The present study suggests that presence or absence of pelvic spines in Lake Storvatnet sticklebacks is not explained by the TG-repeat regions TG-I, TG-II, and TG-III only. Firstly, some individuals with and without spines have exactly the same haplotypes at these TG-repeats. Secondly, TG-repeats TG-I, TG-II, and TG-III at the enhancer *Pel*A are within the same length range regardless of spine phenotype. Thirdly, nothing indicates that *Pel*B, or the flanking regions of the TG-repeats I–III at *Pel*A, explains the presence or absence of pelvic spines in sticklebacks from Lake Storvatnet. TG-II and TG-III are located within the 488 bp region of *Pel*A which has previously been reported as lacking in several North American spineless sticklebacks [28]. Spineless fish from Lake Paxton (benthic morph) lack large segments of *Pel*A, including TG-repeats I, II, and III, in addition to indels at *Pel*B (Figure 5). Spine-reduced specimens from Bear Paw Lake and Hump Lake also lack relatively large segments of *Pel*A, which include TG-repeats II and III. Bear Paw Lake and Hump Lake sticklebacks have larger TG-repeat I compared with their conspecifics in Lake Storvatnet. In contrast, TG-repeat II is absent in these two North American lakes and present in Lake Storvatnet (Figure 5).

Paxton sticklebacks, which exhibit the most extensive deletions at *Pel*A, also seem to have the least developed anti-predator defence with respect to the pelvis structure (or pelvic girdle). Approximately 80 per cent of the adult specimens (benthic morph) in Lake Paxton lack the entire pelvic girdle according to [27]. This is high compared with 12.7 and 7.6 per cent that lack the entire pelvic girdle in Hump Lake and Bear Paw Lake, respectively [18], and especially compared with the complete absence of such individuals in Lake Storvatnet. Moreover, the percentage of sticklebacks lacking both pelvic spines (which are part of the pelvic structure/girdle) regardless of the rest of the pelvic girdle is $\geq$ 80, 77, 92, and 30 for Paxton Lake, Hump Lake, Bear Paw Lake [18,27], and Lake Storvatnet, respectively. However, it is premature to draw conclusions about any association between the size of *Pel*A and lack of pelvic spines (and pelvic girdle) based on a few individuals from three North American and one North European stickleback population.

The relationship between the *Pel*A enhancer and presence of pelvic spines was nicely demonstrated by Chan and colleagues [28]. Quantitative traits loci analyses and DNA sequencing studies have also pointed at chromosome 7 close to where *Pitx*1 and *Pel*A are located (see Introduction), as a position of loci coding for pelvic spines. Thus, the lack of any association between the *Pel*A variants and pelvic spine status in Lake Storvatnet is challenging to explain, but other genetic loci have been suggested to be involved in the development of pelvic spines as well. Based on linkage mapping and QTL analysis, additional loci suggested to play a role in the fine-tuning of the length of the pelvic spines (not to be confused with loci coding for presence or absence of pelvic spines) seem to be located at chromosome 2 and 4 [15] and chromosome 8 [29].

The percentage of specimens from Lake Storvatnet with asymmetrical pelvic spines is similar to a previous report from the same lake [43]. The significantly higher number of left- compared with right-biased individuals among these asymmetrical fish concurs with previous reports from a majority of populations of pelvic-reduced sticklebacks in North America (see [44]). Bell and collaborators [44] gave an overview of potential reasons for the asymmetrical pelvic spines and suggested that (i) asymmetry is associated with lack of *Pitx*1 expression, and (ii) *Pitx*2 and some other loci or genetic mechanisms may play a role in the asymmetry as well. Sticklebacks in Lake Storvatnet also seem to have a genetic component in the asymmetry of their pelvic spines. This is suggested by the significantly higher abundance of left- compared with right-biased asymmetric specimens compared with the expected abundance with random asymmetry (50% of each). However, such a genetic component does not exclude random phenotypic variation in symmetry due to developmental instability.

At this point, we can only speculate about the reason for the lack of association between *Pel*A and pelvic spine status among sticklebacks in Lake Storvatnet. Firstly, TG-repeats are known to form left-handed, fragile Z-DNA, which is prone to deletions [38,39]. Z-DNA opens up the chromatin structure which allows transcription factors to bind to the enhancer [38]. Thus, TG-repeats of certain lengths creating left-handed Z-DNA sequences may be required for chromatin-dependent activation of promoters and for transcription to occur [38]. The pelvic enhancers might not function effectively in specimens with large TG-repeat regions deleted, such as in 9 different pelvic-reduced stickleback populations with deleted sequences of from 757 to approximately 5000 bp [28]. The *Pel*A variants among sticklebacks in Lake Storvatnet are also relatively short. Thus, one might speculate that the size of *Pel*A variants in this population are at a tipping point for Z-DNA formation and transcription to occur or not, leaving spined and spine-reduced individuals to develop based on additional genetic factors, epistatic and epigenetic effects, and/or environmental cues.

The Ca^{2+} concentration in Lake Storvatnet (0.9 mg/L) is well within the range of 0.07–13 mg/L from 1000 Norwegian freshwater lakes reported by [45]. On the other hand, the Ca^{2+} concentration in Lake Storvatnet is relatively low compared with three Norwegian freshwater lakes inhabited by spineless sticklebacks with 5.5 mg/L and 3.0 mg/L [16] and 7.9 mg/L (unpubl. data, J.T. Nordeide). Thus, the low Ca^{2+} concentration in Lake Storvatnet may be interpreted as strengthening the tipping point hypothesis (above).

An alternative, though not mutually exclusive, explanation for the lack of association between the different variants of *Pel*A and pelvic spine status in Lake Storvatnet has to do with reduced standing genetic variation and subsequent alternative genetic pathways to adapt to freshwater environments. Fang and collaborators [46] suggested that contemporary threespine stickleback populations originated in the Eastern Pacific Ocean and North America, while some sticklebacks subsequently migrated to colonize other regions including the Atlantic Ocean and Northern Europe. Thus, the ancestral populations from the Eastern Pacific region have a higher standing genetic diversity than stickleback populations from other geographical regions [17,46,47]. Such inter-regional differences in standing genetic variation have been suggested to give striking differences in the proportion of loci involved in freshwater adaptations along the west coast of North America and Northern Europe [17,47]. Moreover, Kemppainen and colleagues [48] advocated that *Pitx*1′s role in coding for pelvic spines of pelvic-reduced nine-spined sticklebacks (*Pungitius pungitius*) has been replaced by alternative loci in some North European populations. Pelvic spines in these populations were suggested to be a polygenic trait coded for by loci located near 10 novel QTLs [48]. At the moment we can only speculate whether other loci than *Pel*A and *Pel*B take part in controlling the expression of pelvic spines in some threespine stickleback populations as well, such as the one in Lake Storvatnet. Future whole-genome sequencing of the different polymorphic forms of sticklebacks in Lake Storvatnet and examination of population genetic parameters for genetic diversity and differentiation might contribute to locate alternative loci controlling the expression of pelvic spines (see [49,50]).

The presence of spined, spineless, and asymmetric specimens from Lake Storvatnet, and the lack of spineless fish from the downstream Lake Gjerdhaugvatnet and from the marine site, concurs with results from previous studies of spine morphology from the same sites [16,43]. Comparison of *Pel*A variants between the three sites in the present study revealed a few trends (Figure 3, Figure 4, Supplementary Figures S2 and S3). First, the diversity at *Pel*A of the relatively few specimens examined seems high in the marine threespine sticklebacks compared with those from Storvatnet, and those from Lake Gjerdhaugvatnet in particular. This is as expected according to the founder effect and the putatively larger effective population size of sticklebacks in the sea. Second, *Pel*A variants were in general shorter among Lake Storvatnet sticklebacks than in the two downstream populations, especially due to TG-repeat III. TG-repeat II was of approximately the same length in all three populations, whereas at TG-repeat I, the specimens in Lake Gjerdhaugvatnet have relatively uniform and long TG-repeat sequences.

5. Conclusions

Lake Storvatnet sticklebacks carry unique variants of the *Pel*A enhancer region. No simple association was detected between the pelvic spine status and *Pel*A among sticklebacks from Lake Storvatnet. The *Pel*A enhancers of sticklebacks from Lake Storvatnet were short compared with their spined conspecifics in the downstream Lake Gjerdhaugvatnet and the nearby marine site, yet they were relatively long compared with those of pelvic-spine-reduced threespine sticklebacks from three North American populations. No polymorphism was found at *Pel*B. These results clearly indicate that there are alternative molecular pathways to parallel evolution of pelvic reduction in threespine sticklebacks, which could include epistatic and epigenetic effects, and/or environmental cues.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/fishes8030164/s1, Supplementary Figure S1: Drawings of the ventral part of fully spined and spine-reduced threespine sticklebacks; Supplementary Figure S2: Number of thymine-guanine repeats [$(TG)_n$] at the enhancer *Pel*A of spined and spine-reduced specimens from Lake Storvatnet; Supplementary Figure S3: Number of thymine-guanine repeats [$(TG)_n$] at the enhancer *Pel*A of fully spined specimens from Lake Gjerdhaugvatnet and a nearby marine site; Supplementary Figure S4: DNA sequencing of *Pel*A enhancers from threespine sticklebacks and a reference sequence. Primers used in the present study are also shown; Supplementary Figure S5: DNA sequencing of *Pel*B enhancers from threespine sticklebacks and a reference sequence. Primers used in the present study are shown; Supplementary Figure S6: DNA sequence alignments upstream and downstream of TG-repeats I–III of *Pel*A, and *Pel*B sequences of specimens from Lake Storvatnet, Lake Gjerdhaugvatnet, and the marine site.

Author Contributions: Conceptualization, D.A., S.D.J., T.B.M., and J.T.N.; methodology, D.A., S.D.J., T.B.M., and J.T.N.; acquisition, all authors; lab work, D.A. and I.K.H.; validation, all authors; formal analysis, all authors; investigation, all authors; resources, S.D.J., T.B.M., and J.T.N.; data curation, D.A.; writing—original draft preparation, all authors; visualization, D.A. and S.D.J.; supervision, S.D.J., T.B.M., and J.T.N.; project administration, J.T.N.; funding acquisition, J.T.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the study being carried out according to ethical guidelines stated by the Norwegian Ministry of Agriculture and Food through the Animal Welfare Act. According to these guidelines, we were not required to—and therefore do not—have a specific approval or approval number.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

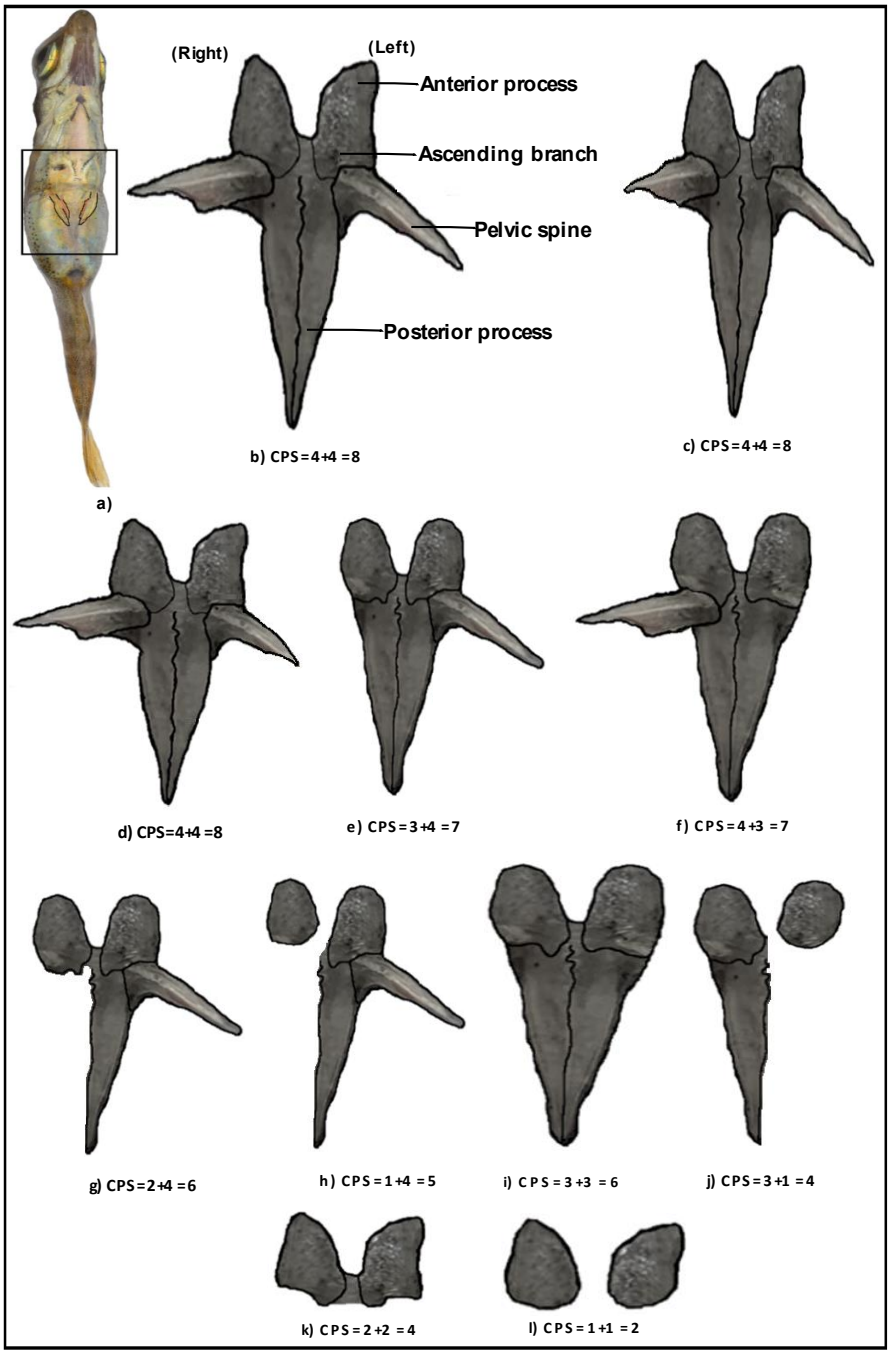
References

1. Futuyma, D.J. *Evolutionary Biology*, 2nd ed.; Sinauer: Sunderland, MA, USA, 1986.
2. Schluter, D.; Conte, G.L. Genetics and ecological speciation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9955–9962. [CrossRef]
3. Barrett, D.H.; Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **2008**, *23*, 38–44. [CrossRef]
4. Innan, H.; Kim, Y. Pattern of polymorphism after strong selection in a domestication event. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10667–10672. [CrossRef]
5. Schluter, D.; Clifford, E.A.; Nemethy, M.; McKinnon, J.S. Parallel evolution and inheritance of quantitative traits. *Am. Nat.* **2004**, *163*, 809–822. [CrossRef]
6. Arendt, J.; Reznick, D. Convergence and parallelism reconsidered: What have we learned about genetics of adaptations? *Trend Ecol. Evol.* **2007**, *23*, 38–44. [CrossRef] [PubMed]
7. Little, C. *The Terrestrial Invasion: An Ecophysiological Approach to the Origins of Land Animals*; Cambridge University Press: Cambridge, UK, 1990.
8. Lee, C.E.; Bell, M.A. Causes and consequences of recent freshwater invasions by saltwater animals. *Trends Ecol. Evol* **1999**, *14*, 284–288. [CrossRef] [PubMed]
9. Reimchen, T.E. Spine deficiency and polymorphism in a population of *Gasterosteus aculeatus*: An adaptation to predators? *Can. J. Zool.* **1980**, *58*, 1232–1244. [CrossRef]
10. Giles, N. The possible role of environmental calcum levels during the evolution of phenotypic diversity in Outer Hebridean populations of the three-spined stickleback, *Gasterosteus aculeatus*. *J. Zool.* **1983**, *199*, 535–544. [CrossRef]
11. Baumgartner, J.V.; Bell, M.A. Lateral plate morph variation in California populations of the threespine stickleback, *Gasterosteus aculeatus*. *Evolution* **1984**, *38*, 665–674. [CrossRef] [PubMed]
12. Bell, M.A.; Orti, G.; Walker, J.A.; Koenings, J.P. Evolution of pelvic reduction in threespine stickleback fish: A test of competing hypotheses. *Evolution* **1993**, *47*, 906–914. [CrossRef] [PubMed]
13. Barrett, R.D.H.; Paccard, A.; Healy, T.M.; Bergek, S.; Schulte, P.M.; Schluter, D.; Rogers, S.M. Rapid evolution of cold tolerance in stickleback. *Proc. R. Soc. Lond. B Biol. Sci.* **2011**, *278*, 233–238. [CrossRef] [PubMed]
14. Aguirre, W.E.; Bell, M.A. Twenty years of body shape evolution in a threespine stickleback population adapting to a lake environment. *Biol. J. Linn. Soc. Lond.* **2012**, *105*, 817–831. [CrossRef]
15. Shapiro, M.D.; Marks, M.E.; Peichel, C.L.; Blackman, B.K.; Nereng, K.S.; Jonsson, B.; Schluter, D.; Kingsley, D.M. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **2004**, *428*, 717–723. [CrossRef]
16. Klepaker, T.O.; Østbye, K. Pelvic anti-predator armour reduction in Norwegian populations of the threespine stickleback: A rare phenomenon with adaptive implications? *J. Zool.* **2008**, *26*, 81–88. [CrossRef]
17. Fang, B.; Kemppainen, P.; Momigliano, P.; Feng, X.; Merilä, J. On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nat. Ecol. Evol.* **2020**, *4*, 1105–1115. [CrossRef] [PubMed]
18. Bell, M.A.; Orti, G. Pelvic reduction in threespine stickleback from Cook Inlet Lakes: Geographical distribution and intrapopulation variation. *Copeia* **1994**, *2*, 314–325. [CrossRef]
19. Wootton, R.J. *The Biology of the Sticklebacks*; Academic Press: London, UK, 1976.
20. Klepaker, T. Morphological changes in a marine population of threespine stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can. J. Zool.* **1993**, *71*, 1251–1258. [CrossRef]
21. Roberts Kingman, G.A.; Vyas, D.N.; Jones, F.C.; Brady, S.D.; Chen, H.I.; Reid, K.; Milhaven, M.; Bertino, T.S.; Aguirre, W.E.; Heins, D.C.; et al. Predicting future from past: The genomic basis of recurrent and rapid stickleback evolution. *Sci. Adv.* **2021**, *7*, eabg5285. [CrossRef]
22. Reimchen, T.E. Predators and morphological evolution in threespine stickleback. In *The Evolutionary Biology of the Threespine Stickleback*; Bell, M.A., Foster, S.A., Eds.; Oxford University Press: Oxford, UK, 1994; pp. 240–276.
23. Hagen, D.W.; Gilbertson, L.G. Geographic variation and environmental selection in *Gasterosteus aculeatus* L. in the Pacific Northwest, America. *Evolution* **1972**, *26*, 32–51. [CrossRef]
24. Moodie, G.E.E. Predation, natural selection and adaptation in an unusual threespine stickleback. *Heredity* **1972**, *28*, 115–167. [CrossRef]
25. Bell, M.A.; Francis, G.W.; Havens, A.C. Pelvic reduction and its directional asymmetry in threespine sticklebacks from the Cook Inlet Region, Alaska. *Copeia* **1985**, *2*, 437–444. [CrossRef]
26. Klepaker, T.; Østbye, K.; Bell, M.A. Regressive evolution of the pelvic complex in stickleback fishes: A study of convergent evolution. *Evol. Ecol. Res.* **2013**, *15*, 413–435.
27. McPhail, J.D. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): Evidence for a species-pair in Paxton Lake, Texada Island, British Columbia. *Can. J. Zool.* **1992**, *70*, 361–369. [CrossRef]
28. Chan, Y.F.; Marks, M.E.; Jones, F.C.; Villarreal, G., Jr.; Shapiro, M.D.; Brady, S.D.; Southwick, A.M.; Absher, D.M.; Grimwood, J.; Schmutz, J.; et al. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **2010**, *327*, 302–305. [CrossRef]
29. Peichel, C.L.; Nereng, K.S.; Ohgi, K.A.; Cole, B.L.E.; Colosimo, P.F.; Buerkle, C.A.; Schluter, D.; Kingsley, D.M. The genetic architecture of divergence between threespine stickleback species. *Nature* **2001**, *414*, 901–905. [CrossRef]
30. Reimchen, T.E.; Bergstrom, C.; Nosil, P. Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evol. Ecol. Res.* **2013**, *15*, 241–269.

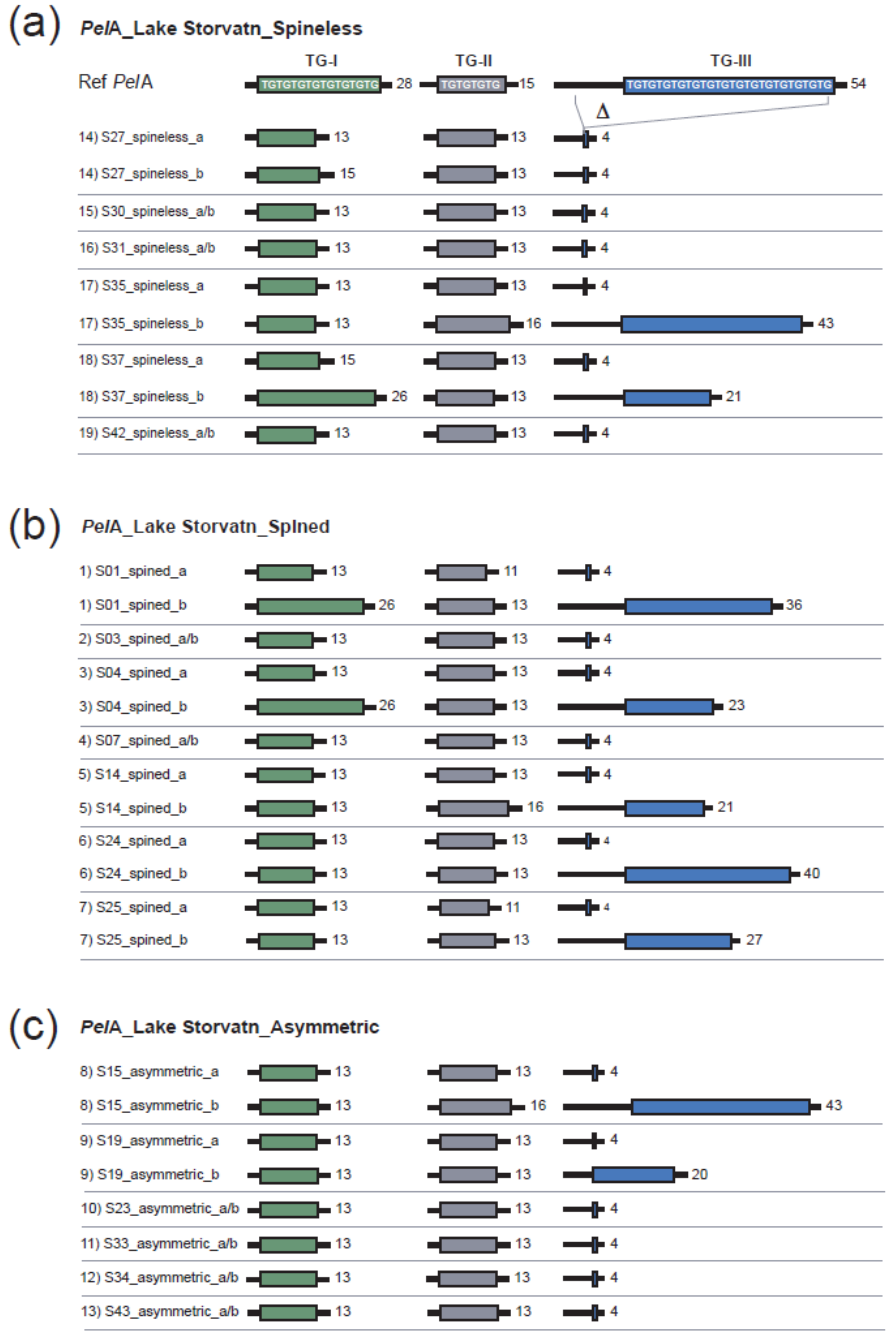
31. Coyle, S.M.; Huntingford, F.A.; Peichel, C.L. Parallel evolution of *Pitx*1 underlies pelvic reduction in Scottish threespine stickleback (*Gasterosteus aculeatus*). *J. Hered.* **2007**, *98*, 581–586. [CrossRef]
32. Marchinko, K.B. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* **2009**, *63*, 127–138. [CrossRef] [PubMed]
33. Zeller, M.; Lucek, K.; Haesler, M.; Seehausen, O.; Sivasundar, A. Little evidence for a selective advantage of armour-reduced threespined stickleback individuals in an invertebrate predation experiment. *Evol. Ecol.* **2012**, *26*, 1293–1309. [CrossRef]
34. Cresko, W.A.; Amores, A.; Wilson, C.; Murphy, J.; Currey, M.; Phillips, P.; Bell, M.A.; Kimmel, C.B.; Postlethwait, J.H. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6050–6055. [CrossRef]
35. Kingman, G.A.R.; Lee, D.; Jones, F.C.; Desmet, D.; Bell, M.A.; Kingsley, D.M. Longer or shorter spines: Reciprocal trait evolution in stickleback via triallelic regulatory changes in *Stanniocalcin2a*. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2100694118. [CrossRef] [PubMed]
36. Petit, F.; Sears, K.E.; Ahituv, N. Limb development: A paradigm of gene regulation. *Nat. Rev. Genet.* **2017**, *18*, 245–258. [CrossRef]
37. Thompson, A.C.; Capellini, T.D.; Guenther, C.A.; Chan, Y.F.; Infante, C.R.; Menke, D.B.; Kingsley, D.M. A novel enhancer near the Pitx1 gene influences development and evolution of pelvic appendages in vertebrates. *eLife* **2018**, *7*, e38555. [CrossRef]
38. Liu, R.; Liu, H.; Chen, X.; Kirby, M.; Brown, P.O.; Zhao, K. Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* **2001**, *106*, 309–318. [CrossRef] [PubMed]
39. Xie, K.T.; Wang, G.; Thompson, A.C.; Wucherpfennig, J.I.; Reimchen, T.E.; MacColl, A.D.C.; Schluter, D.; Bell, M.A.; Vasquez, K.M.; Kingsley, D.M. DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* **2019**, *363*, 81–84. [CrossRef]
40. Kratochwil, C.F.; Meyer, A. Fragile DNA contributes to repeated evolution. *Genome Biol.* **2019**, *20*, 39. [CrossRef] [PubMed]
41. Cole, N.J.; Tanaka, M.; Prescott, A.; Tickle, C. Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. *Curr. Biol.* **2003**, *13*, R951–R952. [CrossRef] [PubMed]
42. Tickle, C.; Cole, N.J. Morphological diversity: Taking the spine out of three-spine stickleback. *Curr. Biol.* **2004**, *14*, R422–R424. [CrossRef]
43. Klepaker, T.; Ostbye, K.; Bernatchez, L.; Vollestad, L.A. Spatio-temporal patterns in pelvic reduction in threespine stickleback (*Gasterosteus aculeatus* L.) in Lake Storvatnet. *Evol. Ecol. Res.* **2012**, *14*, 169–191.
44. Bell, M.A.; Khalef, V.; Travis, M.P. Directional asymmetry of pelvic vestiges in threespine stickleback. *J. Exp. Zool. B Mol. Dev. Evol.* **2007**, *308*, 189–199. [CrossRef]
45. Hindar, A.; Garmo, Ø.; Austnes, K.; Sample, J.E. *Nasjonal innsjøundersøkelse 2019*; Rapport L.Nr. 7530-2020; Norwegian Institute for Water Research: Oslo, Norway, 2020. (In Norwegian)
46. Fang, B.; Merilä, J.; Ribeiro, F.; Alexandre, C.M.; Momigliano, P. Worldwide phylogeny of three-spined sticklebacks. *Mol. Phylogenet. Evol.* **2018**, *127*, 613–625. [CrossRef] [PubMed]
47. Fang, B.; Kemppainen, P.; Momigliano, P.; Merilä, J. Population Structure Limits Parallel Evolution in Sticklebacks. *Mol. Biol. Evol.* **2021**, *38*, 4205–4221. [CrossRef] [PubMed]
48. Kemppainen, P.; Li, Z.; Rastas, P.; Löytynoja, A.; Fang, B.; Yang, J.; Guo, B.; Shikano, T.; Merilä, J. Genetic population structure constrains local adaptation in sticklebacks. *Mol. Ecol.* **2021**, *30*, 1946–1961. [CrossRef]
49. Hohenlohe, P.A.; Bassham, S.; Etter, P.D.; Stiffler, N.; Johnson, E.A.; Cresko, W.A. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **2010**, *6*, e1000862. [CrossRef]
50. Karlsen, B.O.; Klingan, K.; Emblem, Å.; Jørgensen, T.E.; Jueterbock, A.; Furmanek, T.; Hoarau, G.; Johansen, S.D.; Nordeide, J.T.; Moum, T. Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Mol. Ecol.* **2013**, *22*, 5098–5111. [CrossRef] [PubMed]

Supplementary materials

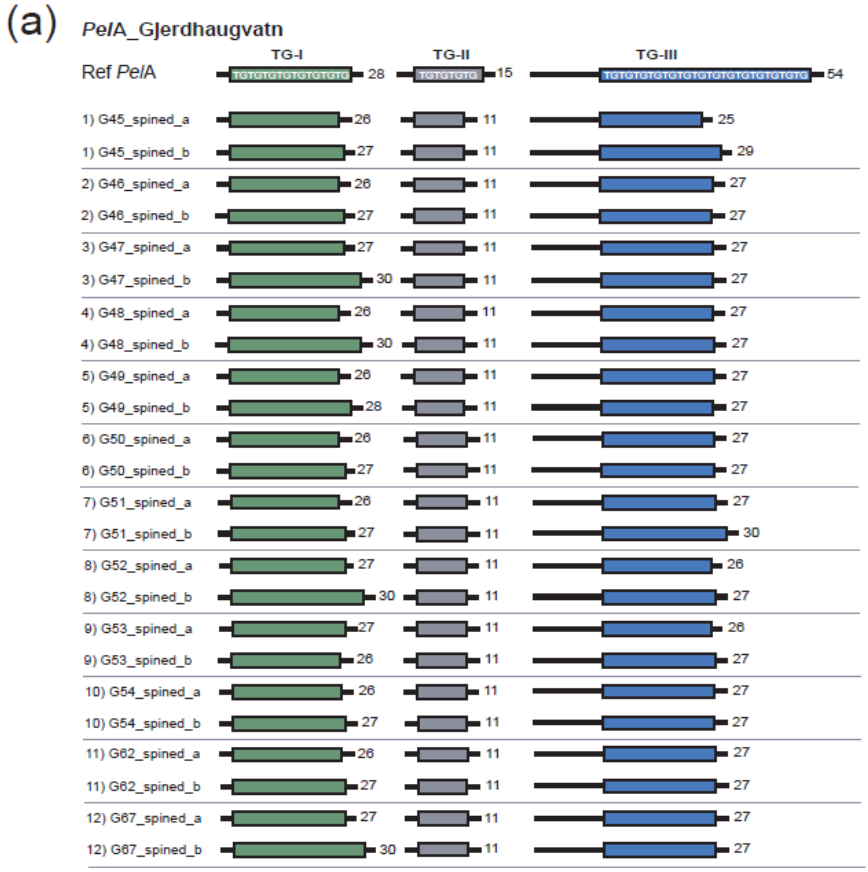
S1. (a) Ventral view of a three-spined stickleback and (b) the pelvis structure (also termed "pelvic girdle") consisting of pairs of "anterior process", "ascending branch", "pelvic spine" and "posterior process" [18, 19]. The following varieties were found in Lake Storvatnet: (b) a symmetric specimen pelvis with complete pairs of pelvic spines, posterior processes, ascending branches, and anterior processes and a combined pelvic score (CPS) of 8; (c) a specimen with left-biased asymmetric pelvis and CPS 8; (d) right-biased asymmetric pelvis with CPS 8); (e) left-biased asymmetric pelvis with CPS 7; (f) right-biased asymmetric pelvis structure with CPS 7; (g) left-biased asymmetric pelvis with CPS 6; (h) left-biased asymmetric pelvis with CPS 5; (i) spineless pelvis structure with CPS 6; (j) spineless pelvis structure with CPS 4; (k) spineless pelvis structure with CPS 4; (l) spineless pelvis structure with CPS 2. Specimens with more than 0.2 mm difference in length between the right and left pelvic spine were categorized as "asymmetric spined" specimens (Fig. 2c – h) and varied from CPS 8 (Fig. 2c) to CPS 5 (Fig. 2h).


(Right)
(Left)
Anterior process
Ascending branch
Pelvic spine
Posterior process
a)
b) CPS = 4+4 = 8
c) CPS = 4+4 = 8
d) CPS=4+4 =8
e) CPS = 3+4 = 7
f) CPS = 4+3 = 7
g) CPS = 2+4 = 6
h) CPS = 1+4 = 5
i) CPS = 3+3 = 6
j) CPS = 3+1 = 4
k) CPS = 2+2 = 4
l) CPS = 1+1 = 2

S2. Number of $(TG)_n$ at TG-repeat I (green), II (grey), and III (blue) from (a) six spineless, (b) seven symmetric spined, and (c) six asymmetric spined specimens from Lake Storvatnet. Numbers to the right of TG-repeat I, II and III show the number of repeats. The reference sequence (GU130435) from Salmon River, British Columbia is shown in (a). Note that all the examined specimens from Lake Storvatnet (but not from Lake Gjerdhaugvatnet in Suppl. S3a or the marine site in Suppl. S3b) have an additional 58 bp deletion (which does not contain a TG-repeat) upstream of and flanking TG-repeat III. These deleted nucleotides are not shown here (but see Fig. 5, and Suppl. S6). Fragment analyses were carried out to identify the allelic variation of TG-repeats I-III. The two alleles in the diploid sticklebacks are indicated by a and b. Note that the alleles of TG-repeats I-III are unphased.


(a)
PelA_Lake Storvatn_Spineless
TG-I
TG-II
TG-III
Ref PelA
TGTGTGTGTGTGTGTG 28
TGTGTGTG 15
TGTGTGTGTGTGTGTGTGTGTGTGTG 54
Δ
14) S27_spineless_a 13 13 4
14) S27_spineless_b 15 13 4
15) S30_spineless_a/b 13 13 4
16) S31_spineless_a/b 13 13 4
17) S35_spineless_a 13 13 4
17) S35_spineless_b 13 16 43
18) S37_spineless_a 15 13 4
18) S37_spineless_b 26 13 21
19) S42_spineless_a/b 13 13 4
(b)
PelA_Lake Storvatn_Spined
1) S01_spined_a 13 11 4
1) S01_spined_b 26 13 36
2) S03_spined_a/b 13 13 4
3) S04_spined_a 13 13 4
3) S04_spined_b 26 13 23
4) S07_spined_a/b 13 13 4
5) S14_spined_a 13 13 4
5) S14_spined_b 13 16 21
6) S24_spined_a 13 13 4
6) S24_spined_b 13 13 40
7) S25_spined_a 13 11 4
7) S25_spined_b 13 13 27
(c)
PelA_Lake Storvatn_Asymmetric
8) S15_asymmetric_a 13 13 4
8) S15_asymmetric_b 13 16 43
9) S19_asymmetric_a 13 13 4
9) S19_asymmetric_b 13 13 20
10) S23_asymmetric_a/b 13 13 4
11) S33_asymmetric_a/b 13 13 4
12) S34_asymmetric_a/b 13 13 4
13) S43_asymmetric_a/b 13 13 4

S3. Number of TG-repeats at TG-repeat I (green), II (grey), and III (blue) from (a) 12 specimens from Lake Gjerdhaugvatnet and (b) seven specimens from the marine site. All examined specimens from these two sites were spined (CPS is 8). Numbers to the right of TG-repeat I, II and III show the number of repeats. The reference sequence (GU130435.1) from Salmon River, British Columbia is shown in the upper line in (a). Fragment analyses were carried out to identify the allelic variation of TG-repeats I-III. The two alleles in the diploid sticklebacks are indicated by a and b. Note that the alleles of TG-repeats I-III are unphased.


(a)
PelA_GJerdhaugvatn
TG-I
TG-II
TG-III
Ref PelA 28 15 54
1) G45_spined_a 26 11 25
1) G45_spined_b 27 11 29
2) G46_spined_a 26 11 27
2) G46_spined_b 27 11 27
3) G47_spined_a 27 11 27
3) G47_spined_b 30 11 27
4) G48_spined_a 26 11 27
4) G48_spined_b 30 11 27
5) G49_spined_a 26 11 27
5) G49_spined_b 28 11 27
6) G50_spined_a 26 11 27
6) G50_spined_b 27 11 27
7) G51_spined_a 26 11 27
7) G51_spined_b 27 11 30
8) G52_spined_a 27 11 26
8) G52_spined_b 30 11 27
9) G53_spined_a 27 11 26
9) G53_spined_b 26 11 27
10) G54_spined_a 26 11 27
10) G54_spined_b 27 11 27
11) G62_spined_a 26 11 27
11) G62_spined_b 27 11 27
12) G67_spined_a 27 11 27
12) G67_spined_b 30 11 27
(b)
PelA_Marine
1) M260_spined_a 13 13 31
1) M260_spined_b 29 13 40
2) M262_spined_a 13 12 31
2) M262_spined_b 30 12 39
3) M263_spined_a 25 11 36
3) M263_spined_b 25 11 37
4) M265_spined_a 25 12 24
4) M265_spined_b 27 12 37
5) M266_spined_a 17 12 31
5) M266_spined_b 27 12 32
6) M268_spined_a 26 9 24
6) M268_spined_b 18 9 39
7) M269_spined_a 27 16 47
7) M269_spined_b 32 12 46

S4. Partial *Pitx*1 reference sequence encompassing the *Pel*A enhancer elements

Partial *Pitx*1 reference sequence from a marine pelvic-complete stickleback from Salmon River, British Columbia, GenBank accession GU130435; 377 852 nt [128 457-131 624] with the *Pel*A enhancer region extending from 128 457-131 382 (indicated by **[ ]**). Bold-Underlined-Red sequences indicate the forward and reverse primer sites for the PCR amplicon (2813 bp). TG-repeat I [$(TG)_{25}$], TG-repeat II [$(TG)_{15}$], and TG-repeat III [$(TG)_{50}$] are underlined. Nucleotide sequences in bold within brackets, (1) and (2), (3) and **(4)**, (5) and (6), indicate sites for the forward and reverse sequencing primers for TG-repeat I, TG-repeat II, and TG-repeat III, respectively.

[CACCGATGTTCCAACAGATGTTTGTCAACACCATCCGGCTGTGGTGGCCCTTGTTGTTTTATTTGGT
TGCTGCCATGGCGACGGGGGCCGGGGGGGGATTCTCTCTTGCCGATGATGACGCTACGTCGGGGCCTC
GCCGCTGCGTGAGAATATGGGCATTAGGCGCTAAGCTGGGCGCCGTGCGGCGCTTTGAAGGGATTTGG
GGGGGGGATTATCAAGGGGAGGGGGCGGTGCTAGTAACACACAAATGGCCTTTAATGAAGTGAGGGGA
GGAGCCTAAAGCTGCAAATGGACCGGTTTGATTATTAGCACATATAAGATGTGTAGTTTGGGGTCTTT
CTGCTGCTGGTCGACA**GCCCAAAACTGACAAAGCA**CATTGCTCTTTTCATGGGGGGGTCGGGGGGGGG
GTTCAGAGATGGAAATGAAGCAGCTCTCTAAACACTGCTGATCTCAGCGGCTCCTCTGACATCACGTT
GGAGCATCAAACCTCGTCCTCTCAGACTGAACCCTG(1**:AGGTCCACAGTACAGTGCAG)**GTTCAAAG
GTCACAAACCCCCCCATCGACCCCCCCCCTCCCTCGCCCGCGCCGTGACGTGTGCGCGCAGGTCAGAC
GTGACGCGGCGCTCCATCACCGAGCCGCTTTGATGTGGGCCTAATATGGCTTGTAAAGAAGGGGAACC
CTGCACTTTCAAGTGTTGCCTGACGCTGTCCAGACACAAGTGTGTGTGTGTGTGTGTGTGTGCGTGTG
TGTGCGTGTGTGTGTGTGTGTGTGTGTGAGTCTGCAGCTGTCATCCTGGAGTTATAAAACGTGATGTA
AACATTGAGAGGGTCTGGAGGAGCAGGTCTGGCTGGAGAATATGTCTCTGTCATGGAGCACCACCTCG
TTC(2**:CTGAAGGCATCTTCTCGTCCCA)**GCAGCAGTAACAACTATGTGAGTTTTCCCCTTTATTCAC
CCAATAAAACCCATTTACTAAAAATGCTCAACTCGATTTCCATCACGTTGTTTATTAATGTTCTCTTC
TGCACATGAAGGATCATTAACAGTAATTCAGGTGCACAAACCTTTATTTTCACCTTTTCACCCCTCAA
ATCTGCAGCGTTTTCTCTGTTTTACAGCCTGATGTGCAGCACACCTGGAGGATCCTTTCACTGTGCAA
TCAGCTCTGAGATAAAAAGCAGTTACGACACACATAAACAGCCGAGATAAGCTGCTCTGCCCCCACCC
CCGAAATCCCCCCCC**(**3**:GTCGAAGCAAAGAGGCGAGACATC)**ACGGCGTGAAGCGACTGTTATTCAT
AAACATTCAGATATCAAACCAACAAGAAACTGTTCAAATGTGTTCAAACAAACAGCGGGACGGGGGAG
TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGTGGTCGGGCGGGAGTGTGTAGATCATGGTGTGCAG
GGGGAACGGGGGGGATGGGGGGGGGGGGGTCAGGGCCGGACCATCTAACTCTTCCACTGATT
{5**:*GTTATGAAGG*(**4**:*GCCGAGCGAC*}CACTTTAGAA)**CTCTGCACAAG*GTCCTAATGCATTCTGTTT*
*GACCTCGCCGGAGTAAATCAAATACTGGCCTCAGATCTCC*TGTGTGTGCGTGTGTGTGTGTGTGTGTG
TGTGTGTGTGTGTGTGCGTGTGCGTGTGTGTGTGTGTGCGTGTGTGTGTGTGTGTGTGTGTGTG
TGTGTGTGTGTGACAGCACGCTAGACACAAGGAAGGAACGGAGGGAGAGAATTAGGAAAGTGGGATAA
ATGAAGAAAGGAGGCTTATAAAACATTTGAGCTTCCA(6**:CGGATTGTTGTGGTCACGC**)CACACGTT
GTCATAGCAACCAATAGTCATTTTGTACATCAGCATCATTTGTTCTGGAGACTGGGTCACCAGCTAGC
CGCTAACAGGTAGCATTAGCGGCTAACAGGTAGCATTAGCGGCTAACAAGTAGCGTTCAACTCTTTCT
AGGCGAGACAGAACCAGAACCCGGCTGAGGAGCTGAGACGTGCGCGATGAAAGGCAGCAGGTTTATTG
TCTCTATTTAAATAACCACCGTGTGATGAAAGAGACGTCTGATAGACCAGCATGTCCCCCCCCCCCCC
CCCCCCCCCCCCATGTGTTCTTCATAATACAGAATCAGCATCTTCTGTATGTGTGTGTCTCCTGTTGA
TCCGTCCTGAAAGACGTCCTCCGTGCGAGTTGTCCCCCATGGGTACTGCGATAGATCTGAGGTCCACC
TGCACCGGGGACACCGGGGACAATAATGAGAGAGAGAGACAGGCCGTCCCCAGAGGGAGGTCTGAGGG
ACTAAGGTCCTCCAGCAGGAGGAAACCTCTCTGCTGACAAACACCACATCGTTTTGGTTGGAGGGGGG
GGGTCAGCCAGCAGCTGCTGTGGTGGGAACGCCAAACCAAACAGAGGGAGGGGGGCTCACCTCTGCCC
AAACACACGCCCCGTCCTCCCGACCGCAGGAGACATATCAAAATAAAAAGGGTCAAATCTACAGAGGG
ATTTAAGGCCTTATCGTTATCTGTCTCCACGGAGGACGGAGGCCTGCTCAGGACCAGCGGCTGTCCTC
TGTCTCCACGGAGGACGGAGGCCTCCTCAGGACCAGCGGCTGTCCTCGTCTCCACGGAGGACGGAGGC
CTCCTCAGGACCAGCGGTTGTCCTCTGTCTCCACGGAGGACGGAGGCCTCCTCGGGACCAGCGGCTGT
CCTCGTCTCCACGGAGGACGGAGGCCTCCTCAGGACCAGCGGTTGTCCTCTGTCTCCACGGAGGACGG
AGGCCTCCTCAGGACCAGCGGCTGTCCTCTGTCTCCACGGAGGACGGAGGCCTCTTGTGACTGGAGGA
ATCAGAGAGGAAAGACGTCCGTTTTCTTCCCGTCTGAATCGTTTCTAAAAGCCTCCGTTGAGGATTTG
CTCGTTATTCCGTAAAGCAGCTCGCTGTGAAAAAC**]**ATTAATAACCGTCTTAATAACTGTTAGTTTGG

GTTAAAGGACAGAAAAGTAGATGTGAAATGTTCTCATTTCGTTTTTACATGTCATTTGTCTAATTGTC
TGAAATATGTTATAGGTGCAAAATGCTGATAAAGAATGTTTGAGGTTTTAGTTTTGTGATTCCCTACA
AACTGTTTTCTTTATTTATAATTAAAACCACGTGCTTCATTATAACAGAAATG**TCTCATTTTGCTTTT**
GCTGCT

S5. DNA sequencing of *Pel*B enhancers from 9, 6 and 6 threespine sticklebacks from Lake Storvatnet, Lake Gjerdhaugvatnet and the marine site, respectively. All specimens were homozygous identical to the reference sequence for the 671 bp interrogated. Red sequences show primer sites used to produce an amplicon of 1009 bp. Black-bold-underlined sequences indicate the sites for sequencing primers.

*Pel*B enhancer region sequence from *Pitx*1 reference GU130435; 377 852 [176 680-177 688].

CACGGATTACTGAGCAGCAATGGAAGGAGGACGCGTGGAGGCTCGCAGGGTTCCTGAAGCTC
TTCTGCCCCCTGGAGGCCTCCAGAGGGTCTTCACCCTCATCTCAAACTCTCGGGGTCTGATC
TCCTCGAGCCGATCGCTCGTCAGAGCGCTGAG**ACAGACAGACAGACAGACAG**ACAGACGGAG
AAGGCTGCCAACATGTAGGGTTGGAACAACATGAGCTCATTCCCCACTAACATGAGGGAGGC
CTGGGGGGGGGCGAGGGGGCGAGGGGGGGGGCAGCGGTCTGCAAACCGGAGGAGACGTTTCTA
CCAATTTGATTGTTTGTGAAAGAAATCATTCTCATCAAAACCCCCAATTGCTGCCAACTGGG
CGCCTGGCGTGTCTCAGGGTAATGCGGCCTGACGTGTCGGGGTAAGAGGAGAGGCCGCTCCA
GCAGATGCCCCCCCGACGCCGCCCCCCCATCCCCCAACCACGACTTCAAACGCGGGGAAGTG
TAAAGACAGACTAGTTTAGGGCCTAAAACAACTAAAGACGGGATTAGAGGAAGTCCGCTAAG
TGGCCTTGAGATGCTGCAGCGAGTCCATGGCGGGACTCAGTGTCTCTGTTACCCCTGGGGGC
GGGGGGGGGGGAGGGGGGGGTGATACTGTGATACTTGTGTTAAGTTTCCAGTGAGCTGGTAA
AGTTAAACTACACACATCAAGTTGTTTGCACTAAATTATTACACTAAAGAGGCAAAACCCAA
AGGACGTCCTCAAGAAAACCCTTTCACAATAAAAGAGGCTGCATTCTCTGTAGTGACCAGCA
GGGGGCGACTCCTCCGACCCATAGACGTTTATGAGAAGATC**CCTCTTCCTCTCTCGATTGAT**
ATATTGTTGGTTGTCAGTCCTCAACACACGCACGCACGCACGCACACACACACACACACACC
TGGTTGAACGTTTTGAAGTGGAGCTCTTTGAAGATGGCGTCGCGCTCCTCCACCTCCACCCA
TCCAGAGGTCTTGAGCT

S6. DNA sequence alignments

Sequences upstream and downstream of TG-repeats I-III (1 – 3), and *Pel*B sequences (4a-b), aligned to the reference sequence GU130435. S, G and M represent the specimens from Lake Storvatnet, Lake Gjerdhaugvatnet and the marine site, respectively. '+', '-', and '*' indicate spined, spineless, and asymmetric spined specimens, respectively, from Lake Storvatnet. Ambiguous nucleotide positions are indicated by 'N'. The number above each alignment represents the position of bases according to the reference. 4c shows the agarose gel bands of *Pel*B amplicons (approximately 1009 bp).

1. TG repeat I

Upstream to the TG-I repeat

```
    (129,082)                                                          (129,153)
Ref  CACCGAGCCGCTTTGATGTGGGCCTAATATGGCTTGTAAAGAAGGGGAACCCTGCACTTTCAAGTGTTGCC
S03+ ...........................................................................
S07+ ...........................................................................
S14+ ..........................................................................T..
S24+ ..........................................................................T..
S15* ...........................................................................
S19* ...........................................................................
S23* .T.........................................................................
S33* .T.........................................................................
S43* ...........................................................................
S30- ...........................................................................
S31- .T.........................................................................
S35- ...........................................................................
```

2. TG repeat II

Upstream of TG-II repeat

```
      (129,805)                                          (129,862)
Ref    TATCAAACCAACAAGAAACTGTTCAAATGTGTTCAAACAAACAGCGGGACGGGGGAG(TG-II)
S01+   ..................................................
S03+   ..................................................
S04+   ..........................................T.......
S07+   ..........................................T.......
S14+   ..........................................T.......
S24+   ..........................................T.......
S25+   ..........................................T.......
S15*   ..........................................T.......
S19*   ..........................................T.......
S23*   ..........................................T.......
S33*   ..........................................T.......
S34*   ..........................................T.......
S43*   ..........................................T.......
S27-   ..........................................T.......
S30-   ..........................................T.......
S31-   ..........................................T.......
S35-   ..........................................T.......
S37-   ..........................................T.......
S42-   ..........................................T.......
G45    ..................................................
G46    ..................................................
G47    ..................................................
G48    ..................................................
G49    ..................................................
G50    ..................................................
G51    ...........T......................................
G52    ..................................................
G53    ..................................................
G54    ..................................................
G62    ...........T......................................
G67    ..................................................
```

```
M260   ..........................................T.............
M262   ..................................................... 
M263   ..................................................... 
M265   ..........................................T.............
M266   ...........T.................................... 
M268   ..........................................T.............
M269   ..........................................T.............
```

3. TG-III repeats
   a. Deletion of unique 58-bp upstream of TG-III repeats found in Lake Storvatnet only

```
        (129,974)                                                 (130,033)
Ref: GGTCCTAATGCATTCTGTTTGACCTCGCCGGAGTAAATCAAATACTGGCCTCAGATCTCC(TG-III)
S04+:GG-----------------------------------------------------------
S07+:GG-----------------------------------------------------------
S23*:GG-----------------------------------------------------------
S33*:GG-----------------------------------------------------------
S34*:GG-----------------------------------------------------------
S43*:GG-----------------------------------------------------------
S27-:GG-----------------------------------------------------------
S30-:GG-----------------------------------------------------------
S42-:GG-----------------------------------------------------------
G53: ..................................................... 
G54: ..................................................... 
M268: .....................................................
```

   b. Downstream of TG-III repeats

```
          (130,134)                                               (130,194)
Ref       TGTGTGTGACAGCACGCTAGACACAAGGAAGGAACGGAGGGAGAGAATTAGGAAAGTGGG (60)
S01+      ...............A.......................................... 
S14+      ...............A.......................................... 
S43*      ...............A.......................................... 
S27-      ...............A.......................................... 
G53       ...............A.......................................... 
G54       ...............A.......................................... 
M269      ...............A.......................................... 

         (130,195)                                                 (130,254)
Ref       ATAAATGAAGAAAGGAGGCTTATAAAACATTTGAGCTTCC-CGGATTGTTGTGGTCACGC (120)
S01+      ........................................A..................... 
S14+      ........................................A..................... 
S43*      ........................................A..................... 
S27-      ........................................A..................... 
G53       ........................................A..................... 
G54       ........................................A..................... 
M269      ........................................A..................... 

         (130,255)                                                 (130,314)
Ref       CACACGTTGTCATAGCAACCAATAGTCATTTTGTACATCAGCATCATTTGTTCTGGAGAC (180)
S01+      ............................................................ 
S14+      .........................................................G.... 
S43*      ............................................................ 
S27-      ............................................................ 
G53       ............................................................ 
G54       ............................................................ 
M269      ............................................................
```

4. *Pel*B

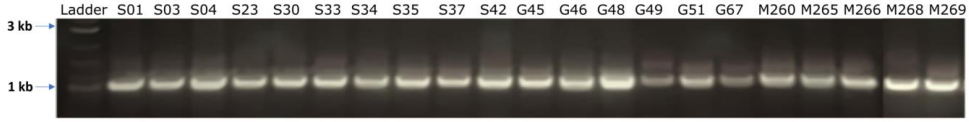
a. Upstream of first poly-G tract

```
          (176,895)                                                (176,948)
Ref        CATGAGCTCATTCCCCACTAACATGAGGGAGGCCTGGGGGGGGGCGAGGGGGCG
SO1+      ..................................................N.........
SO3+      ..................................................N.........
SO4+      ..................................................N.........
S23*      ..................................................N.........
S33*      ..................................................N.........
S34*      ..................................................N.........
S30-      ..................................................N.........
S35-      ..................................................N.........
S37-      ..................................................N.........
S42-      ..................................................N.........
G45       ..................................................N.........
G46       ..................................................N.........
G48       ..................................................N.........
G49       ..................................................N.........
G51       ..................................................N.........
G67       ..................................................N.........
M260      ..................................................A...N.....
M265      ..................................................N.........
M266      ..................................................N.........
M268      ..................................................N.........
M269      ..................................................N.........
```

b. Downstream of last poly-G tract

```
          (177,313)                                                (177,378)
Ref        GGGGGGGGTGATACTGTGATACTTGTGTTAAGTTTCCAGTGAGCTGGTAAAGTTAAACTACACAC
S03+      ............................G.....................................
S23*      ............................G.....................................
S37-      ............................G.....................................
G45       ............................G.....................................
M269      ............................G.....................................
```

c. Agarose gel of *Pel*B amplicons (1009 bp)



List of previously published theses for PhD in Aquaculture / PhD in Aquatic Biosciences / PhD in Biosciences, Nord University

No. 1 (2011)
PhD in Aquaculture
Chris André Johnsen
Flesh quality and growth of farmed Atlantic salmon (*Salmo salar* L.) in relation to feed, feeding, smolt type and season
ISBN: 978-82-93165-00-2

No. 2 (2012)
PhD in Aquaculture
Jareeporn Ruangsri
Characterization of antimicrobial peptides in Atlantic cod
ISBN: 978-82-93165-01-9

No. 3 (2012)
PhD in Aquaculture
Muhammad Naveed Yousaf
Characterization of the cardiac pacemaker and pathological responses to cardiac diseases in Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-02-6

No. 4 (2012)
PhD in Aquaculture
Carlos Frederico Ceccon Lanes
Comparative Studies on the quality of eggs and larvae from broodstocks of farmed and wild Atlantic cod
ISBN: 978-82-93165-03-3

No. 5 (2012)
PhD in Aquaculture
Arvind Sundaram
Understanding the specificity of the innate immune response in teleosts: Characterisation and differential expression of teleost-specific Toll-like receptors and microRNAs
ISBN: 978-82-93165-04-0

No. 6 (2012)
PhD in Aquaculture
Teshome Tilahun Bizuayehu
Characterization of microRNA during early ontogeny and sexual development of Atlantic halibut (*Hippoglossus hippoglossus* L.)
ISBN: 978-82-93165-05-7

No. 7 (2013)
PhD in Aquaculture
Binoy Rajan
Proteomic characterization of Atlantic cod skin mucosa – Emphasis on innate immunity and lectins
ISBN: 978-82-93165-06-04

No. 8 (2013)
PhD in Aquaculture
Anusha Krishanthi Shyamali Dhanasiri
Transport related stress in zebrafish: physiological responses and bioremediation
ISBN: 978-82-93165-07-1

No. 9 (2013)
PhD in Aquaculture
Martin Haugmo Iversen
Stress and its impact on animal welfare during commercial production of Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-08-8

No. 10 (2013)
PhD in Aquatic Biosciences
Alexander Jüterbock
Climate change impact on the seaweed *Fucus serratus*, a key foundational species on North Atlantic rocky shores
ISBN: 978-82-93165-09-5

No. 11 (2014)
PhD in Aquatic Biosciences
Amod Kulkarni
Responses in the gut of black tiger shrimp *Penaeus monodon* to oral vaccine candidates against white spot disease
ISBN: 978-82-93165-10-1

No. 12 (2014)
PhD in Aquatic Biosciences
Carlo C. Lazado
Molecular basis of daily rhythmicity in fast skeletal muscle of Atlantic cod *(Gadus morhua)*
ISBN: 978-82-93165-11-8

No. 13 (2014)
PhD in Aquaculture
Joanna Babiak
Induced masculinization of Atlantic halibut (*Hippoglossus hippoglossus* L.): towards the goal of all-female production
ISBN: 978-82-93165-12-5

No. 14 (2015)
PhD in Aquaculture
Cecilia Campos Vargas
Production of triploid Atlantic cod: A comparative study of muscle growth dynamics and gut morphology
ISBN: 978-82-93165-13-2

No. 15 (2015)
PhD in Aquatic Biosciences
Irina Smolina
Calanus in the North Atlantic: species identification, stress response, and population genetic structure
ISBN: 978-82-93165-14-9

No. 16 (2016)
PhD in Aquatic Biosciences
Lokesh Jeppinamogeru
Microbiota of Atlantic salmon (*Salmo salar L.*), during their early and adult life
ISBN: 978-82-93165-15-6

No. 17 (2017)
PhD in Aquatic Biosciences
Christopher Edward Presslauer
Comparative and functional analysis of microRNAs during zebrafish gonadal development
ISBN: 978-82-93165-16-3

No. 18 (2017)
PhD in Aquatic Biosciences
Marc Jürgen Silberberger
Spatial scales of benthic ecosystems in the sub-Arctic Lofoten-Vesterålen region
ISBN: 978-82-93165-17-0

No. 19 (2017)
PhD in Aquatic Biosciences
Marvin Choquet
Combining ecological and molecular approaches to redefine the baseline knowledge of the genus Calanus in the North Atlantic and the Arctic Oceans
ISBN: 978-82-93165-18-7

No. 20 (2017)
PhD in Aquatic Biosciences
Torvald B. Egeland
Reproduction in Arctic charr – timing and the need for speed
ISBN: 978-82-93165-19-4

No. 21 (2017)
PhD in Aquatic Biosciences
Marina Espinasse
Interannual variability in key zooplankton species in the North-East Atlantic: an analysis based on abundance and phenology
ISBN: 978-82-93165-20-0

No. 22 (2018)
PhD in Aquatic Biosciences
Kanchana Bandara
Diel and seasonal vertical migrations of high-latitude zooplankton: knowledge gaps and a high-resolution bridge
ISBN: 978-82-93165-21-7

No. 23 (2018)
PhD in Aquatic Biosciences
Deepti Manjari Patel
Characterization of skin immune and stress factors of lumpfish, *Cyclopterus lumpus*
ISBN: 978-82-93165-21-7

No. 24 (2018)
PhD in Aquatic Biosciences
Prabhugouda Siriyappagouder
The intestinal mycobiota of zebrafish – community profiling and exploration of the impact of yeast exposure early in life
ISBN: 978-82-93165-23-1

No. 25 (2018)
PhD in Aquatic Biosciences
Tor Erik Jørgensen
Molecular and evolutionary characterization of the Atlantic cod mitochondrial genome
ISBN: 978-82-93165-24-8

No. 26 (2018)
PhD in Aquatic Biosciences
Yangyang Gong
Microalgae as feed ingredients for Atlantic salmon
ISBN: 978-82-93165-25-5

No. 27 (2018)
PhD in Aquatic Biosciences
Ove Nicolaisen
Approaches to optimize marine larvae production
ISBN: 978-82-93165-26-2

No. 28 (2019)
PhD in Aquatic Biosciences
Qirui Zhang
The effect of embryonic incubation temperature on the immune response of larval and adult zebrafish (*Danio rerio*)
ISBN: 978-82-93165-27-9

No. 29 (2019)
PhD in Aquatic Biosciences
Andrea Bozman
The structuring effects of light on the deep-water scyphozoan *Periphylla periphylla*
ISBN: 978-82-93165-28-6

No. 30 (2019)
PhD in Aquatic Biosciences
Helene Rønquist Knutsen
Growth and development of juvenile spotted wolffish (*Anarhichas minor*) fed microalgae incorporated diets
ISBN: 978-82-93165-29-3

No. 31 (2019)
PhD in Aquatic Biosciences
Shruti Gupta
Feed additives elicit changes in the structure of the intestinal bacterial community of Atlantic salmon
ISBN: 978-82-93165-30-9

No. 32 (2019)
PhD in Aquatic Biosciences
Peter Simon Claus Schulze
Phototrophic microalgal cultivation in cold and light-limited environments
ISBN: 978-82-93165-31-6

No. 33 (2019)
PhD in Aquatic Biosciences
Maja Karoline Viddal Hatlebakk
New insights into Calanus glacialis and C. finmarchicus distribution, life histories and physiology in high-latitude seas
ISBN: 978-82-93165-32-3

No. 34 (2019)
PhD in Aquatic Biosciences
Arseny Dubin
Exploration of an anglerfish genome
ISBN: 978-82-93165-33-0

No. 35 (2020)
PhD in Aquatic Biosciences
Florence Chandima Perera Willora Arachchilage
The potential of plant ingredients in diets of juvenile lumpfish (*Cyclopterus lumpus*)
ISBN: 978-82-93165-35-4

No. 36 (2020)
PhD in Aquatic Biosciences
Ioannis Konstantinidis
DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication
ISBN: 978-82-93165-36-1

No. 37 (2021)
PhD in Aquatic Biosciences
Youngjin Park
Transcriptomic and cellular studies on the intestine of Atlantic salmon
Discovering intestinal macrophages using omic tools
ISBN: 978-82-93165-34-7

No. 38 (2021)
PhD in Aquatic Biosciences
Purushothaman Kathiresan
Proteomics of early embryonic development of zebrafish (Danio rerio)
ISBN: 978-82-93165-37-8

No. 39 (2021)
PhD in Aquatic Biosciences
Valentin Kokarev
Macrobenthic communities of sub-Arctic deep fjords: composition, spatial patterns and community assembly
ISBN: 978-82-93165-38-5

No. 40 (2021)
PhD in Aquatic Biosciences
Aurélien Delaval
Population genomics of a critically endangered data-deficient elasmobranch, the blue skate *Dipturus batis*
ISBN: 978-82-93165-39-2

No. 41 (2021)
PhD in Aquatic Biosciences
Isabel Sofía Abihssira García
Environmental impact of microplastics in relation to Atlantic salmon farming
ISBN: 978-82-93165-40-8

No. 42 (2022)
PhD in Aquatic Biosciences
Yousri Abdelmutalab Ahmed Abdelhafiz
Insights into the bacterial communities of Nile tilapia – core members and intergenerational transfer
ISBN: 978-82-93165-41-5

No. 43 (2022)
PhD in Aquatic Biosciences
Fredrik Ribsskog Staven
Interaction studies on lumpfish exposed to Atlantic salmon: behavioural observations and the underlying physiological and neurobiological mechanisms
ISBN: 978-82-93165-42-2

No. 44 (2022)
PhD in Aquatic Biosciences
Solveig Lysfjord Sørensen
Influence of feed ingredients and additives on mucosal health with focus on the intestine of Atlantic salmon (*Salmo salar*)
ISBN: 978-82-93165-43-9

No. 45 (2022)
PhD in Aquatic Biosciences
Apollo Marco Dalonos Lizano
Examining challenges in species-level taxonomy among Calanus copepods in the Northern seas using genome and transcriptome data
ISBN: 978-82-93165-44-6

No. 46 (2022)
PhD in Aquatic Biosciences
Sowmya Ramachandran
Ribosomal RNA, ribose methylation, and box C/D snoRNAs during embryonic development of teleosts zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*)
ISBN: 978-82-93165-45-3

No. 47 (2023)
PhD in Aquatic Biosciences
William John Hatchett
The brown algal genus Fucus: A unique insight into reproduction and the evolution of sex-biased genes
ISBN: 978-82-93165-46-0

No. 48 (2023)
PhD in Aquatic Biosciences
Ying Yen
Application of ensiled *Saccharina latissima* and *Alaria esculenta* as feed: ensilbility, digestibility and bioactivity
ISBN: 978-82-93165-47-7

No. 49 (2023)
PhD in Aquatic Biosciences
Deepak Pandey
Marine macroalgae as an alternative, environment-friendly, and bioactive feeding resource for animals
ISBN: 978-82-93165-48-4

No. 50 (2023)
PhD in Aquatic Biosciences
Nimalan Nadanasabesan
Dietary approaches to improve mucosal health of Atlantic salmon (*Salmo salar*)
ISBN: 978-82-93165-49-1

No. 51 (2023)
PhD in Aquatic Biosciences
Likith Reddy Pinninti
Biosystematics and evolutionary genomics of deep-sea fish (lumpsuckers, snailfishes, and sculpins) (Perciformes: Cottoidei)
ISBN: 978-82-93165-50-7

No. 52 (2023)
PhD in Biosciences
Mie Prik Arnberg
Directed endozoochory: a hitchhiker's guide to successful sexual reproduction in clonal ericaceous plants
ISBN: 978-82-93165-51-4

No. 53 (2023)
PhD in Aquatic Biosciences
Adnan Hussain Gora
Insights from a zebrafish model to combat dyslipidemia using microbe-derived bioactive compounds
ISBN: 978-82-93165-52-1

No. 54 (2023)
PhD in Aquatic Biosciences
Saima Rehman
Diet-induced inflammation in zebrafish and its alleviation by functional oligo- and polysaccharides
ISBN: 978-82-93165-53-8

No. 55 (2023)
PhD in Biosciences
Clara Isabel Wagner
Exploring the nuclear genome of the spiny dogfish (*Squalus acanthias*)
ISBN: 978-82-93165-54-5

No. 56 (2023)
PhD in Aquatic Biosciences
Miiro Ilmari Virtanen
Internal tagging in Atlantic salmon (*Salmo salar L.*): A study of welfare, wounds, and stress
ISBN: 978-82-93165-55-2

No. 57 (2023)
Md Golam Rbbani
PhD in Biosciences
Characterization of Circular RNAs in Nile Tilapia Muscle and Thermal Modulation of their Expression in Relation to Growth
ISBN: 978-82-93165-56-9

No. 58 (2023)
Èric Jordà Molina
PhD in Aquatic Biosciences
Spatio-temporal dynamics of soft-bottom macrobenthic communities in a rapidly changing Arctic: a case study of the Northwestern Barents Sea
ISBN: 978-82-93165-57-6

No. 59 (2024)
Shubham Varshney
PhD in Aquatic Biosciences
From nanoplastics to chemical pollutants: Exploring mixture toxicity in zebrafish
ISBN: 978-82-93165-58-3

No. 60 (2024)
PhD in Aquatic Biosciences
Charles Patrick Lavin
The cumulative impacts of climate change and fishing on marine communities
ISBN: 978-82-93165-59-0