The Effects of Neutral and Adaptive Evolutionary Processes on Genetic and Phenotypic Variation in Small Populations of Icelandic Arctic charr (*Salvelinus alpinus*)

by

Braden Judson

A Thesis

presented to

The University of Guelph

In partial fulfilment of requirements
for the degree of

Master of Science

in

Integrative Biology

Guelph, Ontario, Canada

ABSTRACT

THE EFFECTS OF NEUTRAL AND ADAPTIVE EVOLUTIONARY PROCESSES ON GENETIC AND PHENOTYPIC VARIATION IN SMALL POPULATIONS OF ICELANDIC ARCTIC CHARR (*SALVELINUS ALPINUS*)

Braden J. Judson
University of Guelph, 2022

Advisor:
Moira M. Ferguson

I examined Arctic charr (*Salvelinus alpinus*) sampled from 19 lava caves near Lake Mývatn, Iceland to assess the hypotheses that patterns of genetic and phenotypic variation are affected by: (i) colonization history; (ii) genetic drift and gene flow coupled with contemporary landscape features; and (iii) divergent natural selection caused by ecological variation as an agent of selection. I detected significant signatures of colonization history based on the genetic clustering of the cave populations and their relationship to the lake fish. Signatures of genetic bottlenecks and the association of genetic variation with geography and population size suggest that genetic drift has contributed to population divergence. Estimates of gene flow and fish movement are positively related to the geographic proximity of populations. In contrast, I found limited evidence that patterns of genetic and phenotypic variation are related to ecological variation.

ACKNOWLEDGEMENTS

First and foremost, I would like to give my thanks to my advisor, Dr. Moira Ferguson, for her guidance, patience, and encouragement throughout this degree. Under her advisement, I have grown both personally and as a scientist. I would also like to extend thanks to Anne Easton for her invaluable help in the laboratory. My committee members, Dr. Elizabeth Mandeville, and Dr. Camille Leblanc provided counsel and were instrumental in completing this thesis. I also owe thanks to Dr. Bjarni Kristjánsson for providing helpful feedback and conducting the field work for this project (with Camille and many volunteers!).

It is to my family and friends that I owe the greatest thanks. This was, by far, the most challenging experience of my life and without their constant support this endeavor simply would not have been possible. The constant challenges and restrictions imposed by the Covid-19 pandemic made this degree infinitely more difficult than I had ever anticipated. Primarily, I owe thanks to my incredible partner Megan, who's love and unwavering support kept me going. Many phone calls to my parents Dave and Sherri kept me thinking of better times and their lifelong support has helped me complete this program and so much more. I owe thanks to my grandparents, Jack, Lorna, Karen and Jim for their inspiration, guidance, and care. I am deeply indebted to my friends and colleagues for their friendship and patience when I was not at my best. Without the contributions of the above people, I would have not even started this program, let alone finish it. I hope to one day give you all the support that you have shown me. Lastly, thanks to my calico cat for keeping my lap warm and my spirits high.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUPPLEMENTARY TABLES

SUPPLEMENTARY FIGURES

INTRODUCTION

Understanding the factors that promote or constrain adaptive divergence is a fundamental problem in evolutionary biology. This is particularly relevant in small populations where adaptability is often considered limited, which has implications for the conservation of biodiversity in response to environmental change (Willi et al. 2006; Willi and Hoffmann 2009; Hoffmann et al. 2017). One reason for limited adaptability is the expectation that genetic drift is higher in small populations, which can lead to a loss of potentially adaptive alleles as the raw material for selection (Allendorf et al. 2013; Fraser et al. 2014; Wood et al. 2015; Ferchaud et al. 2020). Gene flow, however, can counteract the effects of drift by introducing allelic variation but can also constrain local adaptation if the alleles introduced by migrants are maladaptive and/or gene flow is high (Tallmon et al. 2004; Garant et al. 2007; Räsänen and Hendry 2008). Nevertheless, if selection is sufficiently strong and local ecological conditions favor established locals over maladapted migrants, gene flow may be reduced as offspring of these migrants suffer reduced fitness (Nosil et al. 2008; Wang and Bradburd 2014).

The relative effects of drift, migration, and selection on the distribution of genetic variation differ among species and landscapes because of several factors (Orsini et al. 2013a; Fenderson et al., 2020). These factors include: (i) historical events that impact the colonization of novel habitats (Koblmãller et al. 2011; Kolbe et al. 2012; Ventura et al. 2014), (ii) contemporary landscape features that affect the likelihood for drift and migration (Seymour et al. 2013; Salisbury et al. 2016; Chan and Brown 2020), and (iii) the amount of variation in ecological conditions that can lead to selection (Schluter 2009; Wang and Bradburd 2014; Forester et al. 2016; Ye et al. 2021). Disentangling how these factors interact to partition genetic variation is

required to provide a better understanding of adaptive potential, with an emphasis on the evolution of small populations given their vulnerability to environmental change.

The colonization of organisms into novel habitats can leave genetic signatures across varying spatial scales, including entire continents (Spellman and Klicka 2007; Aldenhoven et al. 2010; Nigenda-Morales et al. 2019), and within a single geological feature, such as an archipelago (Deagle et al. 2013; Spurgin et al. 2014). The pattern of interpopulation differentiation arising from colonization history is referred to as *Isolation by Colonization* (IBC) (Orsini et al. 2013b; Spurgin et al. 2014). Historical events such as the movement and contraction of quaternary ice sheets, have led to many large-scale patterns of genetic variation such as those arising from populations expanding out from isolated refugia and subsequently colonizing newly accessible habitats (Hewitt 1996, 2000, 2004; Excoffier et al. 2009; Griffiths 2010; April and Hanner 2013). For example, small numbers of founding individuals in many species of northern freshwater fishes have colonized areas after glacial recession, and despite considerable post-colonization divergence, the genetic footprints of these founding events are still apparent (Bernatchez and Wilson 1998; Brunner et al. 2001; Wilson et al. 2004; Moore et al. 2015). However, despite the considerable attention that large-scale colonization events have received, dispersal limitations on contemporary timescales can reinforce founder effects and yield similar outcomes over small spatial scales (Stelkens et al. 2012; Richardson et al. 2014). Small scale IBC has been observed in many organisms with limited dispersal options, such as plants (Trapnell et al. 2013; Helsen et al. 2015; Ngeve et al. 2017). However, IBC could also occur in animal populations, such as those occupying isolated lakes with little or no ability to disperse between populations (Ventura et al. 2014; Barrera-moreno et al. 2015; Kautt et al. 2018)

or islands (Kolbe et al. 2012; Spurgin et al. 2014). Thus, understanding the factors that govern patterns of genetic variation requires investigation of the effects of colonization and dispersal on historical and contemporary timescales.

The evolution of populations after the colonization of novel habitats is also contingent upon contemporary landscape features that can both facilitate or restrict migration depending on the system (Seymour et al. 2013; Lanier et al. 2015; Salisbury et al. 2016). For instance, species in barrier free marine environments often show relatively high genetic diversity and admixture, whereas species in fragmented freshwater systems typically have lower within population genetic variation that is more patchily distributed across the landscape (DeWoody and Avise 2000; Grummer et al. 2019). Populations in insular freshwater systems typically exhibit stronger genetic structuring as physical barriers inhibit migration, which would otherwise facilitate admixture (Salisbury et al. 2018; Grummer et al. 2019). In the absence of physical barriers where migration is comparatively unrestricted, the predicted pattern is that of *Isolation by Distance* (IBD), wherein genetic differentiation increases as a function of geographic distance due to dispersal limitations (Wright 1942; Jenkins et al. 2010; Cameron et al. 2019). However, IBD rarely occurs in the absence of potentially adaptive processes given the frequent association of geography and the environment (Shafer and Wolf 2013). Thus, it is often difficult to disentangle the effects of colonization and migration from the adaptive effects of divergent natural selection (Nadeau et al. 2016).

Ecological conditions may impose selective pressures that act in conjunction with drift and migration to shape contemporary patterns of biological variation (Dewoody et al. 2015; Seymour et al. 2019). If local adaptation is sufficiently strong, migrants from dissimilar environments may

incur negative fitness consequences, thereby resulting in the gradual reduction of gene flow and the development of a pattern of *Isolation by Adaptation* (IBA) (Nosil et al. 2008; Wang and Bradburd 2014). IBA is expected to result in a positive correlation between genetic distance and environmental dissimilarity as observed in many species (Shafer and Wolf 2013; Bond et al. 2014; Wang and Bradburd 2014; Haileselasie et al. 2016; Dowle et al. 2017; Ye et al. 2021). Studies of IBA have typically focused on populations that span large spatial scales that traverse strong ecological gradients. However, there is growing recognition that subtle environment variation can shape population structure (Richardson et al. 2014). A possible mechanism for fine-scale IBA is monopolization, wherein first colonizers rapidly adapt to local conditions and incur a slight selective advantage over subsequent waves of migrants (de Meester et al. 2002; Orsini et al. 2013a). By limiting the establishment of migrants through competition, monopolization can yield longstanding concordance between genetic structuring and subtle patterns of ecological variation (de Meester et al. 2002; de Meester et al. 2016). However, early colonizers may not adapt to local conditions given that they have a subset of the original population's genetic variation due to founder effects and subsequent genetic drift can override selection. Thus, the circumstances that facilitate fine scale adaptive differentiation are not fully understood and therefore warrant additional investigation across a variety of taxa and settings (Richardson et al. 2014).

Studies of adaptation often evaluate the relationships between environmental factors and fitness related traits, such as those relating to foraging performance or movement (Rüber and Adams 2001; Cooper et al. 2010; Torres-Dowdall et al. 2012; Astudillo-Clavijo et al. 2015). Phenotype-environment associations may reflect local adaptation and/or adaptive plasticity as the

result of selection due to the biomechanical trade-offs associated with contrasting habitats or niches (Kingsolver et al. 2012), thereby facilitating directional selection towards some phenotypic optimum and favoring specialization (Skúlason and Smith 1995; Smith and Skúlason 1996). Resource availability is often associated with patterns of adaptive phenotypic variation and this relationship is thought to underlie the diversification of many disparate taxa (Smith and Skúlason 1996; Monteiro and Nogueira 2010; Rosamond et al. 2020). One well-studied example is that of lacustrine fishes that show differential resource use along a benthic-pelagic axis of prey types (Robinson and Wilson 1994; Cooper et al. 2010; Burress et al. 2017; Brachmann et al. 2021). Among many systems occupied by disparate taxa, pelagic specialists display slender, fusiform bodies with terminal mouths presumably adapted for long-distance cruising and planktivory, whereas benthic specialists are deeper bodied with subterminal mouths to encourage maneuverability and foraging on the benthos (Malmquist 1992; Motta et al. 1995; Ferry-Graham et al. 2001; Higham 2007). The adaptive nature of these phenotype-environment associations has been demonstrated by translocations and laboratory experiments which suggest that fitness is higher when morphology and environment are matching (e.g., benthic phenotypes and benthic prey), whereas phenotype-environment mismatches (e.g., benthic phenotypes and pelagic prey) suffer reduced fitness as a result (Schluter 1993, 1995; Hatfield and Schluter 1999). Accordingly, resource availability is a strong pervasive agent of selection, which has resulted in the phenotypic diversification of many disparate taxa (Smith and Skúlason 1996; Monteiro and Nogueira 2010; Rosamond et al. 2020). However, colonization history and the subsequent effects of genetic drift and gene flow may inhibit adaptation, and therefore understanding the factors

that sort and distribute adaptive variation requires both historical and contemporary perspectives (Kolbe et al. 2012; Weese et al. 2012).

Disentangling the effects of neutral and adaptive processes on phenotypic and genetic variation is often difficult given the interrelated nature of the factors governing these processes (Shafer and Wolf 2013). One powerful approach is to assess the contrasting predictions of IBC, IBD and IBA on the relationships among genotypic, environmental, and phenotypic variation of populations (Orsini et al., 2013b). In brief, a positive relationship between genetic distance and geographic distance is consistent with IBD, a positive relationship between ecological distance and either genetic or phenotypic distance is consistent with IBA, and no relationship between genetic distance and either geographic or ecological distance is consistent with IBC. The power to disentangle the effects of selection, drift and gene flow is highest among simple and tractable systems where ecological and geographic variables are not highly correlated (Shafer and Wolf 2013; Nadeau et al. 2016). Such a system is Icelandic Arctic charr (*Salvelinus alpinus*) given their simple colonization of Iceland (Brunner et al. 2001) and many of the inland water bodies throughout Iceland. Furthermore, Icelandic Arctic charr have diversified into phenotypically discrete morphs that utilize different resources (Malmquist et al. 1992; Knudsen et al. 2016; Guðbrandsson et al. 2019; Brachmann et al. 2021), in addition to considerable within-morph variation associated with varying environmental characteristics (Kristjánsson et al. 2011, 2012). For example, ecological variation is associated with discrete patterns of phenotypic variation that is associated with lentic and lotic environments in benthic Arctic charr (Kristjánsson et al. 2011, 2012). In many taxa, determining the scale and extent of local adaptation may be difficult given that their colonization histories are often complex and may affect contemporary patterns of

variation (Shikano et al. 2010; Willing et al. 2010; Nadeau et al. 2016; Bagley et al. 2017). The ability to evaluate the roles of IBD and IBA is potentially stronger in populations with relatively simple colonization histories from single glacial refugia (Ramstad et al. 2004), such as Icelandic Arctic charr (Brunner et al. 2001).

My goal was to determine the relative roles of neutral and adaptive microevolutionary processes in shaping patterns of genetic and phenotypic variation in cave-dwelling populations of Arctic charr located near Lake Mývatn, Iceland. My project had three specific hypotheses. First, I tested the hypothesis that colonization history has contributed to patterns of genetic differentiation among the cave populations (IBC). Genetic clustering of geographically proximate populations, lower levels of genetic variation in the cave populations relative to the lake fish and a positive relationship between genetic differentiation and geographic distance of the cave and the lake populations are expected genetic signatures of colonization history. Second, I tested the hypothesis that genetic drift and gene flow coupled with contemporary landscape features contribute to patterns of genetic diversity and differentiation among populations. Signatures of historical bottlenecks and a positive association between genetic diversity and population size are both evidence for the effects of genetic drift. A positive relationship between geographic and genetic distances (IBD) and significant estimates of gene flow and fish movement among contemporary populations would indicate that gene flow has contributed to patterns of genetic differentiation. Third, I tested the hypothesis that divergent natural selection caused by ecological differences among populations (IBA) has contributed to genetic and phenotypic variation among populations. The finding that patterns of ecological variation predict patterns of genetic and phenotypic variation after controlling for the effects of

colonization is support for this hypothesis. This study will contribute to our understanding of the evolutionary processes that influence patterns of intraspecific variation with an emphasis on those affecting small, fragmented populations.

METHODS

STUDY SYSTEM

Arctic charr were captured in 2014 and 2019 from ponds in 19 lava caves located in the Haganes and Vindbelgur regions around the perimeter of the southern basin of Lake Mývatn, Iceland (Figure 1; Table 1). The caves are located between 16 and 4270m from each other. Some of the caves are connected through subterranean groundwater passages that allow the movement of Arctic charr between them based on mark recapture data (Leblanc unpubl. data). Despite the movement of fish between some caves, the fish living in a lava cave are referred to as a population in this study. The lava caves were likely formed approximately 2300 years ago following a series of volcanic eruptions and the subsequent cooling of the lava flow around Lake Mývatn (Thorarinsson 1979). Roughly 1500 years after the caves were formed, several earthquakes caused the water level of Lake Mývatn to drop by more than 3.5 meters (Thorarinsson 1979), likely rendering the caves inaccessible from the lake. Analysis of microsatellite genetic variation in fish from 2012 suggests that the lava caves were initially colonized by fish from the lake (Leblanc. unpubl. data). These data also indicate that population sizes, fish movement over three years, and morphological traits vary among the cave populations.

The fish were collected as part of a long-term project where each population has been sampled twice annually (June and August) from 2012 to 2021. The fish were captured by electrofishing and with unbaited traps (fyke and minnow). Small numbers of three-spined stickleback (*Gasterosteus aculeatus*) were also captured in some of the caves (Leblanc unpubl. data). Upon capture, fish were anaesthetized in a buffered solution of 2-phenoxyethanol

(300ppm) and scanned with an electronic tag reader to determine if fish had been previously captured. Untagged fish were marked with a PIT-tag (HDX; Oregon RFID; 8mm for fish between 45-65mm fork length (FL) or 12mm for fish ≥ 65mm FL). Some larger fish were captured without PIT-tags and the presence of scars/clipped fins suggested they had lost their tags. These individuals were retagged, and an additional tissue sample was collected. Anesthetized fish were weighed (nearest 0.1g), measured (FL, nearest 1mm) and photographed on the left lateral side with a scale bar. For newly captured individuals, a small (< 30mg) portion of the upper lobe of the caudal fin was removed with scissors. Tissue samples were preserved in 95% ethanol and stored at -20^{o}C, although some samples (~ 50%) were decanted and stored dry at -80^{o}C for 1-8 years prior to processing. Additional tissue samples were also available from a subset of 50 individuals sampled from Lake Mývatn in June of 2012 (Table 1). Twenty-eight of these were a benthic-like morph designated as Krús and were captured using electrofishing along the shore. The remaining fish had a more pelagic morphology (designated as generalists) and were supplied by the Marine and Freshwater Institute, Iceland after they were captured with gill nets as part of an annual monitoring project (Guðbergsson, pers. comm.). The Krus and generalist fish collectively are referred to as lake fish in this study.

There was very little overlap in the identity of fish captured in 2014 and 2019 in that only 33 fish (< 0.1% of the fish tagged in 2014) were recaptured in 2019. This suggests the passage of approximately one generation between 2014 and 2019. The FLs of fish sampled in 2019 were on average 10mm shorter ($p < 0.01$) than those sampled in 2014 (Figure S1, Table S1). After removing individuals captured in both sampling years, we obtained individual data from all unique individuals captured in either 2014 or 2019 (Table S1). Phenotypic data were therefore

available for 1783 individuals. Between five and 30 individuals from each cave per year (Tables 1 and 2) were genotyped. Some populations had fewer than 30 fish captured within a year.

GENETIC VARIATION

Genetic variation was quantified with genotypes at 815 to 1401 single nucleotide polymorphisms (SNPs) per population after (Table 2). Tissue samples were rinsed with chilled dH_2O and digested for approximately 12 hours in a proteolytic solution at 37^0C. Following digestion, a modified version of a phenol-chloroform DNA extraction protocol (Bardakci and Skibinski 1994) was performed using phase lock gel tubes (5PRIME, Quantabio). The digestion mixture was transferred to a phase-lock gel tube and washed with a chloroform/isoamyl mixture and the aqueous layer was retained following several rounds of centrifugation. Once centrifuged, the aqueous layer was washed with ethanol and spun down until a pellet of DNA was visible. This DNA pellet was washed with ethanol and left to air-dry prior to hydration. Once DNA was suspended and dissolved in dH_2O, subsamples were subjected to a series of quality control processes. DNA was quantified using an Invitrogen Qubit Fluorometer (ThermoFisher Scientific) and quality (i.e., presence of high molecular weight DNA) was visualized via agarose gel electrophoresis before concentrations were standardized to 10ng/μL. DNA samples were genotyped at the Clinical Genomics Centre at Mount Sinai Hospital (Toronto, Canada) using 87k Affymetrix Axiom Array (Nugent et al. 2019). This genotyping array was developed from markers sourced from North American aquaculture populations and wild Icelandic fish (Nugent et al. 2019).

Using the Axiom Analysis Suite v5.1.1, genotypic output files (*.CEL and *.ARR) were processed following the best practices workflow for a diploid genome with the exception that

average call rates for a sample to pass was adjusted from 98.5% to 96.5%. Fourteen individuals were discarded prior to subsequent analyses: one individual failed quality control checks, and thirteen had distinctly different allelic compositions from all other samples and lacked complementary capture information. After removing these samples, 1055 individuals from the cave and lake populations were retained for further genetic analyses (Tables 1 and 2). I then implemented conservative SNP filtering steps to minimize the potentially erroneous effects of uninformative loci in inferring population structure (Roesti et al. 2012). SNPs that were discarded were either monomorphic, had fewer than three copies of the minor allele (Linck and Battey 2017), scored inconsistently among twice genotyped samples (N = 12), or have demonstrated non-Mendelian inheritance in North American populations (Danzmann pers. comm.) (Table S2). Biallelic SNPs passing quality control checks were positioned to the *Salvelinus sp.* genome (Christensen et al. 2021) or a pseudochromosome contig hereafter referred to as AC38 (Table S3). This filtering process was first undertaken with all individuals to elucidate the genetic relationships between the lake and cave fish. Genotyping and filtering steps were conducted a second time including only cave fish (N = 1005) to evaluate the relationships between genetics, ecology, and distance among cave populations alone (Table S2). After filtering, a total of 1962 and 1752 SNPs remained for the complete and cave-specific datasets, respectively (Table S2). Among these SNPs, genotype missingness was generally low (< 1%) and over 60% of loci were positioned to the genome, although coverage was fewer than one SNP per megabase (Table S3).

ECOLOGICAL VARIATION

Ecological variation was quantified using both abiotic and biotic variables available from another study (Kristjánsson and Leblanc, unpubl. data). Four abiotic variables (temperature, pH, O_2 saturation, and conductivity) were used for the current study (Table S4). Temperature was measured four times daily from 2013 onwards in each cave with HOBO temperature loggers (UA-001-64 Pendant; Onset Corporation). Water pH, conductivity and oxygen saturation were also recorded during sampling in June and August of each year using a calibrated multiparameter probe (Hydrolab DS5 (46711) Water Quality Probes). Due to the temporal stability of the abiotic parameters in this system (Kristjánsson and Leblanc unpubl. data), I used the mean values of each variable over 2013 to 2019 for each population. These data also included the minimum linear distance between each cave and the perimeter of Lake Mývatn.

Two types of biotic variables were available based on the type of invertebrate community composition (termed Benthic and Aerial). These data were obtained in 2014 and were used as a proxy for prey item availability (Table S5). For the benthic biotic variables, benthic invertebrate community composition was characterized in 15 populations (Table 1). In most cases, three stones were removed from each cave, photographed, scrubbed to remove the benthic organisms, and sieved (≤ 125 µm). The invertebrates collected from stone scrubs were identified to the lowest taxonomic level possible and enumerated. The benthic invertebrates were then categorized into taxonomic groups using a similar approach to Kreiling et al. (2021). The groups were Chironimidae, Cladocera, Copepoda, Nematoda, Oligochaeta, Ostracoda, Arachnida and an eighth miscellaneous group composed of rare miscellaneous taxa such as Chaetogaster, Collembola, Coleoptera, Trichoptera, Plecoptera, Tardigrada and assorted insect pupae.

Invertebrates were categorized into these higher-level groups to prevent zero-inflating the dataset. Arachnid abundances were removed from the dataset given they were collinear (R^2 = 0.84) with Cladocerans, which are a known prey item for Icelandic Arctic charr (Gudbergsson 2004; Kristjánsson et al. 2012; Kristjánsson and Leblanc 2018). Thus, seven benthic biotic variables were available for analysis. For the single aerial biotic variable, the input of aerial invertebrates was estimated for each cave in 2014 by placing fall-in traps built from clear buckets (2300 ml, 196 cm^2 surface area) at the water surface of the entrance of each cave, and adding a mixture of 30% propylene glycol, aroma-free soap, and water. Trap contents were filtered through a 125 μm sieve and plant materials were removed prior to estimating the total gross volume of invertebrates (ml). For each cave, the number and area (m^2) of the openings to the sky were previously measured. Invertebrate volumes were thus divided by the total area of the openings of each cave to obtain density estimates ($ml \cdot m^{-2}$) (Kristjánsson and Leblanc unpubl. data). Aerial invertebrates were primarily the carcasses of blackflies (Simuliidae) and midges (Chironimidae).

PHENOTYPIC VARIATION

Body shape variation was characterized from photos of 1783 individuals across two sampling years (Table 1) using 24 homologous landmarks (Figure 2) with the software tpsDIG v2.31 (Rohlf 2015). Landmark locations were similar to those used in previous studies (Kristjánsson et al. 2011; Kristjánsson et al. 2012; Kristjánsson and Leblanc 2018; Leblanc unpubl. data) and I placed all landmarks. Eighteen landmarks were fixed in position and the remaining 6 were sliding. Sliding landmarks were positioned along a curve that allowed them to move to minimize the average shape difference among specimens. Shape variation associated

with dorso-ventral arching was minimized using the "unbend" utility in tpsUTIL v1.47 (Rohlf 2015). This process fits landmarks 23, 24, 8 and 9 along a quadratic curve wherein all other landmarks are adjusted along this curved axis to statistically account for shape variation attributed to the up and down bending of each specimen. Following unbending, landmarks 23 and 24 were removed. All subsequent morphometric analyses were performed using the R package Geomorph (Adams and Otárola-Castillo 2013) unless stated otherwise. For individuals photographed with their mouth agape, the placement of the landmarks at the anterior tip of the dentary and the posterior end of the maxilla (landmarks 18 and 20, respectively) were imputed on a population-by-population basis with the function estimate.missing. This procedure uses shape data from within each population to estimate the average position of these landmarks and ensures that morphometric variation is attributed to shape rather than position.

Shape data were then superimposed and transformed to a common scale, position, and rotation by performing a generalized Procrustes analysis (GPA). GPA removes isometric effects of body size on shape but retains allometric effects. I investigated the extent of allometric effects across populations and sampling dates using a series of multivariate Procrustes ANOVAs (function procD.lm). Using both FL and centroid size as proxies of body size, I tested for spatial and temporal variation among the existing allometric relationships using Procrustes ANOVAs with 10,000 randomized residual permutations. I found that allometric effects are significant and vary among populations, months, and years (Table S6). Although many studies minimize allometric variation, I retained it as it may be biologically meaningful developmental variation. Morphometric analyses were performed twice; once with whole body shape (22 landmarks), and once with craniofacial shape only (11 landmarks; Figure 2). Fish sampled in June will have

smaller condition factors (a measure of body girth) than those in August. Thus, performing the analysis on craniofacial features alone will facilitate the detection of potentially adaptive phenotypic variation rather than variation due to short-term changes in feeding.

STATISTICAL ANALYSES

All subsequent analyses were conducted in R v4.1.1 (R Core Team 2021) unless indicated otherwise.

HYPOTHESIS 1: THE EFFECTS OF COLONIZATION HISTORY ON GENETIC DIFFERENTIATION

I tested for the effects of colonization history on genetic differentiation among population by characterizing genetic population structure. However, as SNPs affected by selection can bias neutral population structure, I identified and excluded outlier loci through a pair of complementary genome scans. First, pcadapt (Luu et al. 2017) was used, which estimates the Mahalanobis distances between SNPs and finds loci statistically associated with population structure. The significance of the pcadapt outputs was assessed at a *q*-value threshold of 0.05. Secondly, I used BayeScan (Foll and Gaggiotti 2008), which detects outlier loci using F_{ST} values and the posterior probability that SNPs under selection exceed background levels of genetic differentiation. BayeScan analyses were performed with 50,000 iterations, 200,000 burn-in steps, and prior odds of 100 and 1000, representing both relaxed and conservative parameters, respectively. SNPs identified as outliers by either approach were removed (Table S2). Lastly, to minimize the number of non-informative SNPs in each dataset, SNPs in linkage disequilibrium were identified using PLINK (Purcell et al. 2007). Correlations between each pair of SNPs in a window of 50 loci were estimated, shifting five loci each iteration of the analysis. For correlated

pairs of SNPs ($R^2 > 0.80$), one SNP was discarded at random (Table S3). All remaining SNPs were presumed to be selectively neutral and are retained for all subsequent analyses.

I tested for significant genetic differences between individuals from different populations and in different sampling years with two analyses of molecular variance (AMOVAs). The AMOVAs were performed using the R package poppr (Kamvar et al. 2014) and significance was assessed using ade4 (Dray and Dufour 2007) following 10,000 repetitions at a significance threshold of $\alpha = 0.05$. I found no genetic differences between fish from the same population but sampled in different years (Table S7). Therefore, all subsequent genetic analyses were performed on the 2014 and 2019 samples grouped together. Levels of genetic diversity were quantified using estimates of expected heterozygosity (H_e), observed heterozygosity (H_o) (from the R package adegenet (Jombart 2008)), and allelic richness (A_r) (using the R package hierfstat (Goudet 2005)). Genetic differentiation among pairs of populations was estimated with Weir and Cockerham's (1984) F_{ST} through the R package HIERFSTAT (Goudet 2005) (Table S8). This metric is appropriate given that it is well suited for biallelic markers and is not biased by varying sampling sizes (Meirmans and Hedrick 2011). Using the function boot.ppfst and 1,000 bootstrap replicates, 95% confidence intervals were calculated around each estimate of F_{ST}. Pairwise estimates of F_{ST} were not considered significant if the lower limit was less than or equal to zero. As some populations are not genetically differentiated, F_{ST} values were estimated with these populations considered separately and then again with data from these populations grouped together.

I searched for the genetic signatures of colonization history through the characterization of fine and broad-scale patterns of genetic differentiation between populations using three

analytical approaches. Among systems with hierarchically arranged genetic structuring, fine-scale differentiation is often the product of post-colonization divergence, whereas higher-level patterns (higher-level genetic clusters of populations, HLGCs) are more likely to reflect common ancestry as a result of colonization history (Brunner et al. 2001; Willing et al. 2010). I first characterized genetic population structure with sparse non-negative matrix factorization (sNMF) using the R package LEA (Frichot and François 2015). The sNMF algorithm estimates ancestry coefficients for an unknown number of ancestral gene pools and uses cross-entropy coefficients to determine the optimal number of genetic clusters. The sNMF algorithm is well suited for analyzing small populations given that the algorithm makes no assumptions about populations being in Hardy-Weinberg equilibrium (Frichot et al. 2014). Individual ancestry coefficients were estimated for 1 to 19 ancestral gene pools (K) using 50 replicates for each value. Cross-entropy coefficients were computed for each value of K, where lower coefficients indicate better model support. Visualization of the resulting admixture data was obtained using the R package pophelper (Francis 2016). I then conducted principal components analyses (PCAs) and constructed neighbor-joining trees. Unlike the sNMF analysis, these second and third approaches illustrate genetic relationships among individuals with *a priori* population assignment. Using neutral SNPs only, PCAs were performed using the function dudi.pca from the R package ade4 (Dray and Dufour 2007). Neighbor-joining trees were constructed with SplitsTree v4.17.1 (Huson 1998; Bryant and Moulton 2004) using pairwise Euclidean distances among populations. Together, these three approaches i) identify patterns of cluster membership and admixture (sNMF), ii) visualize the major axes of variation among individuals (PCA), and iii) estimate genetic differences between populations (Neighbor-joining tree). Using analytical approaches

that use both *a priori* and *a posteriori* population assignment increases the probability of obtaining an accurate picture of genetic population structure and inferred colonization history.

To further assess the effects of colonization history, I determined whether the observed patterns of interpopulation genetic and geographic variation are consistent with the colonization of the caves by fish from the lake as suggested previously (Leblanc unpubl. data). The previous analysis showed that the microsatellite alleles detected in the cave populations were a subset of those found in the lake. This scenario of colonization is based on the assumption that there was a lower probability of fish colonizing the caves further from the lake than the more adjacent ones. For this analysis, I grouped samples from the two morphs of lake fish together (F_{ST} = 0.058) so that the sample size of the lake fish was comparable to that of the cave populations. If the lake fish indeed colonized the lava caves, I expected that the cave populations would share a large number of SNPs with the lake fish as shown with the microsatellite data and have lower levels of genetic diversity. I then determined if the geographic distance of each population to the lake (minimum linear distance, Table S4) is negatively associated with levels of genetic diversity and the magnitude of genetic differentiation (F_{ST}) using a series of simple linear regression models.

HYPOTHESIS 2: THE EFFECTS OF GENETIC DRIFT, GENE FLOW AND LANDSCAPE FEATURES ON GENETIC DIFFERENTIATION

I tested for the effects of genetic drift caused by founder effects at colonization and bottlenecks on genetic diversity and differentiation by searching for signatures of historic bottlenecks. This analysis was first performed across the populations within each of the three HLGCs to detect bottlenecks during the initial phase of colonization, and then for each individual population to detect bottlenecks during the later phase of post-colonization isolation. I

used the software Bottleneck v1.2.02 (Cornuet and Luikart 1996; Piry et al. 1999), which tests for excess heterozygosity levels given that rare alleles are typically lost rapidly during a population bottleneck. Using SNPs with a MAF above 5%, heterozygosity expected under the infinite allele model was compared to observed levels using a one-tailed Wilcoxon signed rank test. The minor allele frequency (MAF) spectra were then visualized to infer the intensity of the population bottlenecks detected above.

To further assess the effects of genetic drift, I determined if population size is associated with estimates of genetic diversity. I estimated census (N_c) and effective (N_e) population sizes for each population. Excluding young-of-the-year individuals (< 65mm FL), individual recapture data between June and August of 2014 and 2019 was used to estimate N_c using the Lincoln-Petersen method through the R package FSA (Ogle et al. 2021). The chapman modifier was applied to account for small population sizes (i.e., < 30) as applicable. The harmonic mean of N_c estimates from 2014 and 2019 was calculated for each population to minimize the effects of outlier variation and facilitate comparison with estimates of N_e. However, estimates of N_c were generally consistent between sampling years (adjusted $R^2 = 0.94$, $p < 0.001$). N_e was estimated for each population using the temporal change in allele frequencies with NeEstimator v2.1 (Do et al. 2014). This software utilizes a moment-based estimate of temporal allelic variation (F_s, Jorde and Ryman 2007) and infers N_e given the assumption that N_e is the primary factor driving temporal genetic variation when migration is low (Waples 1989). This analysis assumes that five years between sampling years represents one generation, which is consistent with analysis of otolith data (Leblanc unpubl. data) and the observation that less than 0.1% of the fish tagged in 2014 were recaptured in 2019. SNPs with a minor allele frequency less than 5% were excluded.

The Plan I approach outlined by Waples (2005) was used as it is based on estimating the harmonic mean of N_e between temporally separated, non-lethally obtained samples and does not assume that individuals in the later generation are strictly the progeny of the earlier one (Waples 1989, 2005). Jackknife analysis (Waples and Do 2008) were performed over all loci to estimate confidence intervals for all population specific estimates of N_e. Lastly, the relationships between N_c, N_e, and all metrics of genetic diversity (H_o, H_e, and A_r) were quantified using simple linear regression models. As C25 has the largest N_c by far, models with N_c as a response variable were performed with and without this population in the dataset.

I searched for a positive relationship between geographic and genetic distances (IBD, Wright 1942) with canonical redundancy analyses (RDAs) and spatial eigenvectors. This approach is more likely to detect spatial patterns present among the response data and yield more informative outputs than Mantel tests (Legendre and Fortin 2010; Legendre et al. 2015). In brief, RDAs extend on multiple linear regression by detecting linear (i.e., "redundant") relationships among multiple response and explanatory variables (Rao 1964; Borcard et al. 2011; Legendre and Legendre 2012). RDAs, therefore, have the potential to elucidate the relative importance (or lack thereof) of single variables in explaining a multivariate dataset. RDAs can account for a third dataset (i.e., partial RDAs) by conditioning predictor variables upon an additional set of data prior to analysis. Here, the response data is a matrix comprised of PCoA values derived from pairwise F_{ST} estimates (Weir and Cockerham 1984, see above), and the response data were a matrix of distance-based Moran's eigenvector maps (dbMEMs) calculated from pairwise, linear distances between all populations. Here, PCoA loadings represent variation in genetic distances between populations. Only PCoA loadings explaining more than 5% of the total

genetic variation were used as response variables, and a forward selection procedure with 10,000 permutations was applied to the model to retain only the most informative dbMEMs. dbMEMs are orthogonal eigenvectors that describe spatial relationships among the data in a rectangular matrix suitable for RDAs (Borcard and Legendre 2002; Dray et al. 2006). dbMEMs are well suited to describe spatial relationships among hierarchically arranged systems given that they serve as multi-scale spatial eigenvectors, meaning that large and small dbMEMs account for broad and fine-scale spatial structures, respectively. To statistically account for the effects of colonization history, the simple RDAs were followed up with partial RDAs wherein a third matrix comprised of ancestry coefficients representing membership probability to each HLGC (K = 3, see results). The R package vegan v2.5-7 (Oksanen et al. 2020) was used to derive dbMEMs and perform the RDAs and model selection processes.

I tested the effects of gene flow on population differentiation by estimating rates of contemporary gene flow in both directions between pairs of caves. Rates of contemporary gene flow were estimated using BA3-SNPs (Wilson and Rannala 2003; Mussmann et al. 2019). The BA3-SNPs algorithm extends on the original BayesAss (Wilson and Rannala 2003) program to incorporate large-scale SNP data to estimate the extent and direction of rates of effective migration. Using default Mixed chain Monte Carlo parameters (10,000,000 iterations, burn-in of 1,000,000 iterations, and sampling every 100 iterations), the analyses were performed separately for the Haganes and Vindbelgur populations as preliminary runs suggest demographic isolation between populations in these regions. BA3-SNPs-autotune (Mussmann et al. 2019) was utilized to determine optimal mixing parameters for allele frequencies, inbreeding coefficients, and migration rates for each dataset. Following Wilson and Rannala (2003), each analysis was run

five times and mean values were retained for both migration rates and confidence levels. Estimates with confidence intervals containing negative values were considered not significant. Average rates of contemporary gene flow were compared to migration rates, which were estimated from the longitudinal mark-recapture data. For this analysis, individuals PIT-tagged in one cave and recaptured in another were considered migrants. A positive association between estimates of gene flow and migration would suggest that migrants are successful breeders.

HYPOTHESIS 3a: THE EFFECTS OF SELECTION ON GENETIC DIFFERENTIATION

To test for the effects of selection on genetic differentiation, I first searched for relationships among patterns of genetic and ecological variation taking colonization history or geographic distances into account when relevant. However, as C25 has by far the largest N_c and receives a disproportionate number of aerial invertebrates each summer, the following analyses were repeated with and without data from this population. Furthermore, data from pairs of populations that were not significantly differentiated (based on F_{ST} values) were considered both independently and grouped together as a single population. All tests were conducted using data from the subset of 15 populations where ecological data were available (Table 1). I used both distance-based and canonical approaches. Relationships among pairwise estimates of genetic, ecological, and geographical distances were assessed using multiple matrix regression with randomization (MMRR (Wang 2013)) as implemented by the R package ecodist v2.0.7 (Goslee and Urban 2007). The MMRR analysis extends the principals of multiple regression to incorporate distance (or dissimilarity) matrices, thus simultaneously quantifying the relationship between one response matrix and several predictor matrices. By simultaneously accounting for

multiple response matrices, this process accounts for some of the variation resulting from autocorrelation and thus offers greater statistical power than other approaches tests (Legendre and Fortin 2010, Legendre et al. 2015). These analyses return information on the strength (regression coefficients) and significance (p-values) of each predictor matrix and the fit of the overall model (R^2). For these distance-based analyses, response variables were pairwise estimates of genetic differentiation (Weir and Cockerham's F_{ST}) and predictor variables were ecological distance matrices. Distance matrices derived from abiotic and aerial invertebrate data were estimated by using the R function dist to compute pairwise Euclidean distances between populations. Likewise, Hellinger-transformed benthic invertebrate abundances were used to estimate pairwise Bray-Curtis (1957) community dissimilarity indices between populations using the R package vegan. To ensure that variables and regression coefficients are comparable, all distance measures were Z-transformed before analysis.

I followed up the distance-based analyses by evaluating relationships among individual genetic, ecological, and geographic variables through a series of canonical RDAs and partial RDAs (pRDAs). The justification for conducting a second analytical approach is that while comparing distance matrices has the potential to reveal coarse relationships between datasets, more subtle associations may evade detection. For example, patterns of genetic variation may covary with environmental distances and this signal may be overwhelmed or counteracted by other variables (Magalhaes et al. 2016).

For the canonical RDAs, genetic PCoA scores were used as the response matrix and the predictor matrices were comprised of the transformed values from the abiotic, benthic biotic and aerial biotic variables. To account for colonization history, the simple RDA models were

complemented by a series of pRDAs wherein the response data were conditioned upon ancestry coefficients corresponding to each HLGC (K = 3, see Results). The performance of each RDA and pRDA was assessed using ANOVA significance testing and model fit estimates (adjusted R^2).

HYPOTHESIS 3b: THE EFFECTS OF SELECTION ON PHENOTYPIC VARIATION

Analysis of the morphological data detected body and craniofacial shape differences between sampling years (Table S9). Therefore, all subsequent analyses were performed with phenotypic data from 2014 and 2019 considered together and then separately. However, as the biotic data (benthic and aerial invertebrates) were only available for 2014, these data were not compared to phenotypic variables from 2019. These analyses were also performed using whole body and then craniofacial shape data considered separately given that body shape can vary over the summer. As before, I used both distance-based and canonical approaches to assess relationships. For the distance-based analyses, response data were morphological distance matrices and response data were the same ecological distance matrices described previously. Pairwise morphological distances were estimated between populations using the function morphol.disparity from the R package Geomorph (Adams and Otárola-Castillo 2013).This function estimates morphological distances as the average Procrustes variance between groups and significance is assessed via randomized residual permutations (Zelditch 2012). For the canonical RDAs, the same ecological predictor variables were used as described previously and the response matrices were comprised of scores along the first three PC axes following a PCA on GPA superimposed landmark coordinates. These axes were selected as they each explained

approximately 10% of the total phenotypic variation or more and together explain an average of 47 and 62% of the total body and craniofacial shape, respectively. To account for the effects of colonization history, the canonical RDAs were repeated with predictor variables conditioned upon ancestry coefficients representing each HLGC (K = 3, see Results). Model significance was assessed using 10,000 permutations.

RESULTS

HYPOTHESIS 1: EFFECTS OF COLONIZATION HISTOR ON GENETIC DIFFERENTIATION

Most cave population differ genetically from each other based on patterns of neutral SNP variation. A significant proportion of the total genetic variation (29.8%, $p < 0.01$) was distributed between populations based on an AMOVA (Table S7). In addition, F_{ST} values between most pairs of populations are highly significant (mean $F_{ST} = 0.29$, standard deviation = 0.11) (Table S8). The most genetically differentiated pair of populations is situated within the Vindbelgur region (C18 and C23, $F_{ST} = 0.455$). However, pairwise estimates of F_{ST} were not significant between two pairs of caves known to exchange fish (C1 and C2; C17 and C18). All cave populations were highly differentiated from the lake fish (mean $F_{ST} = 0.33$, standard deviation = 0.11, all $p < 0.05$). Metrics of genetic diversity ranged considerably between cave populations, with C18 and C19 having the lowest levels of genetic variation, and the highest estimates of diversity observed in C7 (Table 2). Estimates of genetic diversity were on average 15% (A_r) to 45% (H_e) higher among the fish from Lake Mývatn.

Genetic clustering among populations was detected at both fine and broad geographic scales. Fine scale population structure was indicated by the curved distribution of cross-entropy values estimated in the sNMF analysis, with many large values of K performing similarly well (i.e., K = 10 through to K = 19) (Figure S2). However, the elbow of the distribution suggests higher-level patterns of genetic structuring (Figure S2). For example, the Haganes and Vindbelgur populations were separated at K=2 with clear differentiation while K=3 detected

differences between the eastern and western populations in Vindbelgur (Figure 3). In contrast, K = 4 detected a fourth genetic cluster with relatively small differences between the northern and southern populations in Haganes. Thus, these results suggest that there are three higher level genetic clusters (HLGCs) that are highly differentiated from each other. The existence of three distinct HLGCs was also supported by the genetic PCAs (Figure 4). The Haganes and Vindbelgur populations were differentiated along the first PC axis (7.3% of the total genetic variation), whereas the second PC axis (6.7%) depicted strong separation between the eastern and western Vindbelgur populations. Similar to the sNMF analysis, the third PC axis (5.3%) showed only modest separation between the northern and southern Haganes populations. The grouping of the Haganes populations into one cluster and the eastern and western populations in Vindbelgur into two additional clusters was also suggested by the Neighbor-joining trees (Figure 5). Given that genetic differences between the northern and southern populations in Haganes are far more subtle than the differences between the other areas, the most parsimonious interpretation is that there are thee HLGCs (Haganes, Vindbelgur East, and Vindbelgur West) among the cave populations. Individual membership to these HLGCs was therefore used as a conservative indicator of possible colonization history in subsequent analyses.

Patterns of genetic variation in the cave populations relative to the lake fish are consistent with the scenario of colonization of the cave populations by the lake fish. First, 3099 of the 3386 alleles (91.5%) found in the cave fish were also detectable in the lake fish and the lake fish showed greater genetic diversity than the cave populations (Table 2). Second, the degree of differentiation (F_{ST}) between the cave populations is significantly and positively associated with geographic distance from the lake (Figure 6d). Finally, there is a significant negative relationship

between each metric of within population genetic diversity in the cave populations and the degree of differentiation from the lake fish (Figure 6a, b and c).

HYPOTHESIS 2: EFFECTS OF GENETIC DRIFT, GENE FLOW AND LANDSCAPE FEATURES ON GENETIC DIFFERENTIATION

I detected signatures of genetic bottlenecks at each of the three HLGCs and the individual populations based on the detection of excess heterozygosity (all $p < 0.001$). Visualization of the MAF spectra also identified the genetic signatures of population bottlenecks within some of the individual cave populations (Figure 7).

Estimates of N_c varied markedly among populations, with mean values ranging from 19 to 365 individuals in C17 and C25, respectively (Table 2). The census size of C25 was much larger than that of any other population. Estimates of N_e were considerably lower than those of N_c (average of 3.3-fold difference) although these estimates are significantly correlated (adjusted $R^2 = 0.38$, $p = 0.003$). However, there are notable differences in N_c and N_e within some caves. For example, C25 had by far the largest N_c (mean $N_c = 365$), and yet C26 had a higher N_e (mean $N_e = 59$). Furthermore, very small values of N_e were estimated for C21 (mean $N_e = 4$) and C24 (mean $N_e = 7$), despite these populations not being the smallest (see N_c estimates, Table 2).

The relationships between either N_e or N_c and estimates of genetic diversity were in the expected positive direction (Figure S3). Both estimates of population size were significant predictors of allelic richness (both $p < 0.05$) and showed a suggestive relationship with H_e and H_o ($p < 0.1$). However, N_c was no longer correlated with any estimate of genetic diversity when C25

was removed from the dataset. These data indicate that N_c is a poor predictor of genetic diversity, whereas modest correlations between N_e and each metric of genetic diversity are observed. Some of the variance in genetic diversity that is not explained by N_e may be the result of varying colonization histories or population bottlenecks.

I found strong evidence for IBD as geographic distance (dbMEMs) was a significant predictor of genetic differentiation (PcoA axes) among populations (adjusted $R^2 = 0.39$, $F_{(2,16)} = 6.65$, $p < 0.001$) based on the RDA. However, evidence of IBD is stronger when high level genetic structuring (HLGCs) is taken into account (adjusted $R^2 = 0.45$, $F_{(2,13)} = 12.06$, $p < 0.001$). These data indicate that genetic differentiation increases with geographic distance at multiple spatial scales. For this analysis, genetic differentiation is represented by four PCoA axes that explained 33.2, 24.3, 13.4 and 8.4% of the total genetic variation, respectively. Similarly, geographic relationships between populations were described using two dbMEMs (dbMEM3 and dbMEM6), which were identified as optimally describing spatial relationships between populations. The identification mid and small-scale dbMEMs reflects the hierarchical arrangement of populations, with populations within each HLGC being relatively close to each other and populations between HLGCs being comparatively distant.

Estimates of contemporary gene flow between 11 pairs of neighboring populations (mean geographic distance = 56m, standard deviation = 29m) were statistically significant (Table 3). The estimated rates of effective migration varied from 3% (from C12 to C7) to 25% (from C7 to C25) per generation (Table S10). The mark-recapture data indicated that 96 individuals moved between caves at least once between 2014 and 2019. The highest number of migrants were detected between C7 and C25 (N = 51), whereas C19 and C20 exchanged the fewest migrants (N

= 2) (Table 3). The movement data are consistent with estimates of gene flow for five populations. However, the estimates of gene flow were significant for five pairs of populations (C5 and C10, C7 and C12, C21 and C22, C23 and C24, and C22 and C24) even though the movement of PIT-tagged individuals has not been detected (Table 3). Additionally, four individuals were observed to move between C17 and C17b, despite non-significant estimates of gene flow. For the eleven pairs of populations with significant estimates of contemporary gene flow, between zero and 51 (mean = 9.7, standard deviation = 15.4) migrants are estimated to have moved between caves. Estimates of gene flow (average in both directions) was significantly related to the proportion of detected migrants (adjusted $R^2 = 0.38$, $p = 0.025$). Relative to contemporary population sizes, migration rates were the highest between C1 and C2. Estimates of gene flow between populations were asymmetric and nearly twice as high (14.4 vs. 6.4%) in the direction facing away from the lake (i.e., North in Vindbelgur and West in Haganes). Lastly, estimates of gene flow (average of both directions) is inversely related to the degree of genetic differentiation (F_{ST}) (adjusted $R^2 = 0.60$, $p = 0.003$) across the eleven pairs of populations with significant estimates of gene flow.

HYPOTHESIS 3a: EFFECTS OF SELECTION ON GENETIC DIFFERENTIATION

Abiotic distances were predictors of genetic distance based on the MMRR analyses ($p < 0.001$, Table 4). In contrast, between population differences in the benthic and aerial invertebrate ecological distances were not significant predictors of genetic differentiation. These findings were not affected by the inclusion or exclusion of C25. Together, the full MMRR model explained a considerable proportion of the observed genetic differentiation (adjusted $R^2 = 0.336$).

In contrast, the canonical RDAs failed to detect relationships between patterns of genetic differentiation and the abiotic and biotic ecological variables, and this outcome was also unaffected by inclusion or exclusion of data from C25 (Table 5). Accounting for colonization history (differences between HLGCs), did not qualitatively affect these results. Furthermore, these RDAs generally explained a low proportion of the total genetic differences observed (all adjusted $R^2 < 0.13$). The results of both the MMRR analyses and the RDAs were not affected by the grouping or splitting of non-differentiated population pairs (i.e., C1 and C2, and C17 and C18), and therefore these data are not presented.

HYPOTHESIS 3b: EFFECTS OF SELECTION ON PHENOTYPIC VARIATION

The major axes of body shape variation reflected subtle differences in condition factor, body depth and head size (Figure S4 and S6). Regardless of the sampling year, the two first PC axes of body shape variation were correlated with body size (using FL as a proxy, all adjusted $R^2 > 0.15$, $p < 0.001$). The retained craniofacial shape variables were also captured fine morphological differences, and illustrated differences in snout shape, operculum morphology and mouth position (Figure S5). The retained shape variables (PCA scores) encompassed more than 45% and 60% of the total body and craniofacial shape variation, respectively (Figures S4 and S5).

Between population differences in the input of aerial invertebrates was associated with variation in body and craniofacial shape from 2014, but not for the separate analyses of shape data from 2019 or both years combined (Table 6). However, when C25 was removed from the dataset, phenotypic distances from 2014 were no longer significantly related to patterns of

abiotic variation (Table 6). This observation suggests that the significance of these relationships is biased by the disproportionate amount of aerial invertebrates that C25 receives during the summer. Otherwise, estimates of body and craniofacial shape differentiation from 2014 were unrelated to both abiotic and benthic invertebrate differences between populations, regardless of whether C25 was included in the dataset. Similarly, the MMRR analyses indicate that abiotic differences between populations were unrelated to both body and craniofacial shape differences in 2019 and when data from 2014 and 2019 were grouped together (Table 6).

The canonical RDAs generally corroborated the results of the MMRR analyses (Table 7). Using phenotypic data from 2014, the only relationship detected was between body shape and the input of aerial invertebrates while accounting for colonization history. This relationship, however, was not detectable when C25 was removed from the dataset, again suggesting that C25 was driving this relationship. Patterns of phenotypic variation from 2019 were not related to patterns of abiotic variation, and this finding was unaffected by the inclusion of C25 or whether colonization history was accounted for. Similarly, abiotic factors were not a significant predictor of phenotypic variation when shape data collected in 2014 and 2019 were grouped together. Notably, however, when phenotypic data are grouped across sampling years, a proportion of the variation is explained by differences between the HLGCs (all $p < 0.06$). This result is consistently detected with both body and craniofacial shape is unaffected by data from C25. Together, the distance-based and canonical approaches both yield little evidence that patterns of phenotypic and ecological variation are related, and that the few statistically significant relationships detected are biased by the inclusion of C25. However, there is evidence that phenotypic variation differs somewhat among populations belonging to different HLGCs. The

phenotypic differences between the HLGCs were modest and were mostly related to body depth (which was, on average, slightly higher in Haganes). The results produced by both analytical approaches were also unaffected by the grouping or splitting of non-genetically differentiated population pairs (data not shown).

DISCUSSION

In this study, I tested the hypotheses that patterns of genetic and phenotypic variation are affected by: (i) colonization history; (ii) genetic drift and gene flow coupled with contemporary landscape features; and (iii) divergent natural selection caused by ecological variation as an agent of selection. First, the clustering of populations into three high level genetic clusters (HLGCs) suggests that colonization history has contributed to system-wide patterns of genetic differentiation. Second, genetic drift and gene flow combined with the contemporary landscape appear to influence patterns of post-colonization differentiation. Small population sizes, the signatures of historical population bottlenecks, and the observation that the populations with the least genetic diversity are the most differentiated from the putative source population are all consistent with founders' events and the subsequent effects of genetic drift. Moreover, gene flow appears restricted to pairs of neighboring populations, is asymmetric and proportional to the extent of genetic admixture. Third, I found little support for the effects of natural selection or adaptive plasticity given that ecological factors are generally unrelated to patterns of genetic and phenotypic variation. However, fish from different HLGCs vary in body shape, suggesting that colonization may have contributed to patterns of phenotypic variation. Together, these data suggest that patterns of differentiation are largely the product of neutral evolutionary processes acting in concert with contemporary landscape features over a small spatial scale. These findings are therefore consistent with the longstanding notion that the genetic drift plays a major role in the evolution of small populations (Lande 1988, Willi et al. 2006).

COLONIZATION HISTORY

The detection of three high-level genetic clusters (HLGCs) by three complementary analytical approaches suggests that the lava caves in the Haganes, Vindbelgur East and Vindbelgur West geographic regions were colonized in three independent events. However, the geological history of this system makes it difficult to determine the relative timing of colonization. As the charr must have colonized the lava caves between their formation (approximately 2300 years ago) and the decline in Lake Mývatn's water level (roughly 1500 years ago) (Thorarinsson 1979), the window for colonization is recent and relatively brief (≤ 800 years). It is therefore difficult to estimate the relative timing of colonization given that colonization events of a few generations and hundreds of generations often yield similar genetic signatures depending on the time since colonization and the initial levels of genetic diversity (Launey et al. 2010; Habel et al. 2014). While demogenetic modelling has been used to estimate colonization times, these analyses are typically most successful for systems spanning larger timescales (e.g., Lanier et al. 2015; Boria et al. 2021) or taxa with shorter generation times (e.g., Laurent et al. 2011; Rey et al. 2015) than the one here. For instance, estimates of demographic parameters such as bottleneck timing show relatively large confidence intervals in taxa evolving in a similar number of generations (McCoy et al. 2014; Nunziata et al. 2017). Therefore, it will be challenging to determine the relative timing of the colonization of the caves around Lake Mývatn.

The observed patterns of genetic and geographic variation indicate that fish from the lake colonized the lava caves. First, the cave populations had lower levels of genetic diversity than the lake fish, which is consistent with many island biogeographical studies where the putative

source (ancestral) population is more genetically diverse than the smaller descendant populations (Jordan and Snell 2008; Wang et al. 2014; Funk et al. 2016). Second, the degree of genetic differentiation of the cave populations from the lake fish is associated with estimates of genetic diversity and the distance from the lake. Third, over 90% of the SNPs found in the cave populations are shared with the lake fish. This pattern of shared SNPs between the lake and cave populations is consistent with the microsatellite data of Leblanc (unpubl. data), where alleles found in the cave populations were a subset of those found in the lake. However, the current study also identified private alleles in the cave populations unlike the Leblanc (unpubl. data) study. One explanation for the detection of private alleles is that other source populations contributed to the cave populations, which could be investigated through additional sampling of larger water bodies in the Mývatn area. Another explanation is that the genetic variation estimated from sampling the lake fish is not representative of the historical source population given the dynamic nature of the landscape in the Mývatn region. This, combined with the limited number of lake fish analyzed relative to the cave populations may have led to incomplete sampling of the genetic variants in the presumed source population. Regardless, these findings further emphasize the importance of colonization history in understanding contemporary patterns of genetic variation (Caldera and Bolnick 2008; Croucher et al. 2012; Spurgin et al. 2014; Machado et al. 2022), even at a finer spatial scale than what is commonly appreciated in animals.

GENETIC DRIFT, GENE FLOW AND LANDSCAPE FEATURES

My data strongly supports the hypothesis that genetic drift and gene flow interacting with the landscape has influenced patterns of genetic variation in the cave populations. Detection of genetic signatures of historical bottlenecks, the small estimates of N_e and the lower levels of

genetic diversity in the cave populations suggest that genetic drift would have been prevalent in the post-colonization evolution of the cave populations. Indeed, no populations had an estimated average N_e much greater than 50, suggesting that the effects of genetic drift are expected to be strong (Allendorf et al. 2013; Frankham et al. 2014). Moreover, the significant positive relationship between genetic diversity and N_e is consistent with the premise that the lower genetic diversity of the cave populations is primarily governed by N_e given that migration is low (Frankham 1996; McCusker and Bentzen 2010; Devillard et al. 2022). Populations within the size range estimated here are considered to have an elevated risk of extirpation (or extinction) given their susceptibility to stochastic demographic changes (Shafer 1981) and/or presumed loss of long-term adaptive potential (Franklin and Frankham 1998; Allendorf et al. 2013; Frankham et al. 2014).

The observation that gene flow is only detectable between populations that are in close proximity to each other suggests that the matrix of lava and sediments act as a barrier to movement. Adjacent populations show signatures of admixture whereas more physically distant populations are more strongly differentiated from each other. These findings support previous studies in a variety of taxa and settings where landscape heterogeneity and barriers to gene flow affect the propensity for drift to occur (Seymour et al. 2013; Cameron et al. 2019; Gallego-García et al. 2019). The current study detected evidence of gene flow between several pairs of populations that was not apparent from the microsatellite analysis of population structure (Leblanc unpubl. data) and movement data based on mark recapture analysis of PIT-tagged fish. The increased resolution of the current genetic analysis likely reflects the increased sample sizes and number of genetic markers used combined with the longer sampling period (genotypes from

one year vs two here, and mark-recapture from three years vs. six years here). The lower ability of the movement data to infer gene flow can be partially attributed to our inability to detect the movement of small juveniles (< 45mm FL) who are not PIT-tagged. Thus, the current study has improved our understanding of the extent and effects of gene flow and its relationship to fish movement in this system.

Between population estimates of gene flow were asymmetric, indicating that the landscape may be affecting the direction and intensity of gene flow between populations. For nine pairs of populations where the direction of gene flow relative to the lake could be inferred, significant estimates of gene flow were more than twice as high in the direction away from Lake Mývatn (paired t-test, $t = 2.8$, $df = 8$, $p = 0.024$) (i.e., approximately North and West in Vindbelgur and Haganes, respectively). On a coarse scale, rates of gene flow appear to coincide with the direction of groundwater flow (Einarsson et al. 2004) (Figure 2). The direction of water flow is commonly used to explain patterns of asymmetric gene flow among fluvial fishes (Morrissey and Kerckhove 2009; Whiteley et al. 2010; Salisbury et al. 2016). However, gene flow can also be directionally biased by physical barriers (Torterotot et al. 2014) and varying selection against migrant establishment (Räsänen and Hendry 2014). Therefore, additional research on the physical characteristics of the landscape near Lake Mývatn and how it influences groundwater flow will be required to identify the potential drivers of asymmetric gene flow.

The detection of a strong relationship between estimates of genetic and geographic distances between populations provides support for IBD, thus suggesting that reduced gene flow via dispersal limitations has affected patterns of genetic variation. Accordingly, gene flow is expected to be highest among the most proximal populations, where the probability of gene flow

decreases with increasing geographic distance (Wright 1943). However, a pattern of increasing genetic distance with increasing geographic distance can also be initiated or reinforced by the effects of sequential founder events along a geographic gradient (Pruett and Winker 2005; Orsini et al. 2013b). This scenario has been reported in several non-migratory bird species, where the effects of IBD are better explained by colonization history than contemporary patterns of gene flow and dispersal (Pruett and Winker 2005; Clegg and Phillimore 2010; Sendell-Price et al. 2021). Likewise, the finding that the cave charr populations most distant from the lake are the most differentiated and have the lowest genetic diversity supports a model of serial colonization within each HLGC. Additional study is needed to better understand how colonization history has influenced patterns of genetic variation in this system.

ISOLATION BY ADAPTATION

I found limited support for the hypothesis that ecological variation influences patterns of genetic variation among populations. Neither benthic nor aerial invertebrate availabilities were significant predictors of genetic divergence among populations based on the findings from both distance-based (MMRR) and canonical (RDA) analyses. However, the distance-based analyses identified abiotic distances as a significant predictor of pairwise genetic differentiation among populations. In contrast, the canonical analyses failed to provide support for a relationship between abiotic ecological variation and genetic variation even when controlling for presumed colonization. Collectively, this suggests that evidence for selection arising from variation in abiotic ecological factors on patterns of genetic variation in this system is weak at best. These findings are therefore consistent with other studies that detect limited support for the effects of adaptive differentiation among finite, wild populations spanning modest ecological gradients

(e.g., Rogell et al. 2010, Spurgin et al. 2014, Chan and Brown 2020). Observations like these may suggest that subtle ecological variations are less likely to incite adaptive differentiation among populations when the effects of IBC (and genetic drift) are strong.

The detected association between genetic and abiotic distances may be the result of autocorrelation with geographic distance and not adaptation to local abiotic factors. The subdivision of system wide genetic variation into clusters of populations, which are in turn located in different geographical regions that vary in abiotic characteristics (Kristjánsson and Leblanc unpubl. data), may result in the statistical association between genetic and abiotic distances. Therefore, the observed autocorrelation in this system makes it difficult to disentangle the relative effects of the abiotic ecological variables on genetic variation (Shafer and Wolf 2013; Orsini et al. 2013a). Furthermore, the observation that the canonical analyses did not detect a relationship between patterns of genetic and abiotic variation suggests that the abiotic factors considered in this study do not have a strong effect on patterns of genetic variation. The study of populations where abiotic and geographic distances are not autocorrelated may help parse out the relative effects of geographic distance and abiotic variation in shaping population structure especially when ecological gradients are subtle.

PHENOTYPE-ENVIRONMENT ASSOCIATIONS

The observation that patterns of phenotypic variation are largely unrelated to variation in contemporary ecological conditions provides limited support for the hypothesis that ecological variation influences phenotypic variation. The only significant association detected was that both body and craniofacial shape variation in 2014 were associated with the input of aerial invertebrates with the distance-based analyses. However, this relationship appears to be

dependent on the inclusion of C25 as its removal rendered the relationship not significant. All patterns of phenotypic variation were otherwise unrelated to patterns of ecological variation, regardless of the analytical approach used or whether colonization history was accounted for. These results are consistent with previous studies on small benthic charr in Iceland, where relationships between body shape and ecological factors (i.e., water chemistry, substrate composition and diet) were mostly observed between different habitat types (pond vs streams) (Kristjánsson et al. 2012), rather than among populations occupying the same type of habitat like those in the present study (but see Kristjánsson & Leblanc, 2018).

The identity and features of the ecological factors measured as well as data availability may have limited our ability to detect support for our hypotheses. Examination of the breeder's equation indicates that the rate of evolutionary change (R) is proportional to the heritability (h^2) of the target of selection and the identity and strength of contemporary selection differentials (S) (Falconer and Mackay 1996). The subtle ecological differences between the populations in our system may result in low S, and indeed the studies that have detected evolutionary change in response to selection have evaluated populations over stronger environmental gradients (Shafer and Wolf 2013; Bond et al. 2014; Tobler et al. 2015). Furthermore, the ecological factors measured here may not accurately characterize evolutionarily relevant ecological variation given that other unmeasured factors, such as parasites (Karvonen and Seeuhausen 2012; Hayward et al. 2017) or varying resource availability over the winter (MacMeans et al. 2015; Kreiling et al. 2021), are also likely to be important. It is also unlikely that the ecological variables used in this study have captured all the ways that the environment could influence the evolution of these populations. For example, the input of aerial invertebrates is highly variable from year to year

(Einarsson et al. 2004; Ives et al. 2008) and thus our estimates from 2014 only may not reflect the long-term significance of this potential resource. Likewise, sampling benthic invertebrates was limited to 2014. Additional monitoring of the invertebrate communities in these lava caves may be required to increase our understanding of the temporal ecological variation in this system, and thus improve our ability to detect a signature of selection patterns of phenotypic variation if it exists.

Another explanation for the limited evidence for the effects natural selection is that the phenotypic variation measured in this study may not be the target of selection. Variation in body and craniofacial shape are considered as a possible target of selection given that they are often coupled with varying resource utilization (Malmquist 1992; Gíslason et al. 1999; Garduño-Paz and Adams 2010; Skoglund et al. 2015; Brachmann et al. 2021). However, other studies have demonstrated that the relationship between individual resource use and body shape is indirect (Franklin et al. 2018), or that body shape variation is better explained by fish community composition and the type of water source (Seymour et al. 2019; Kristjánsson and Leblanc 2018; Kristjánsson 2008). Furthermore, other traits not measured in the current study, such as fin morphology, and physiological factors have been identified as possible targets of selection in other fish taxa (e.g., Evans et al. 2012; Zastavniouk et al. 2017).

The detection of significant phenotypic differences among populations from the three HLGCs in the combined analysis across years with the canonical analyses suggests that colonization history may have influenced patterns of contemporary phenotypic variation. A relatively large proportion of the total phenotypic variation in body and craniofacial shape (all adjusted $R^2 > 0.19$) is explained by the HLGCs. Although there is a paucity of studies wherein

the distribution of complex morphological traits can be attributed to natural founder events, there are a few examples that have detected this relationship at a roughly similar timescale to the system studied here (Kolbe et al. 2012; Spurgin et al. 2014). These studies contrast the view that founder events and populations bottlenecks are unlikely to leave a detectable imprint on complex phenotypic patterns given that the effects of drift, mutation, and selection will rapidly erode these signatures (Coyne et al. 1997; Zhang 2018). Thus, this study is a potentially rare example where the colonization of the landscape has helped shape patterns of contemporary phenotypic variation in natural populations. It may therefore be important for future studies to consider recent colonization history when searching for phenotype-environment associations.

CONCLUSIONS

Using comprehensive genetic, phenotypic, ecological, and geographic datasets, my study of the cave populations of Arctic charr has revealed that contemporary patterns of genetic variation are largely the result of neutral evolutionary processes despite variation in ecological conditions. Colonization history appears to have resulted in the genetic divergence between populations in different geographical regions, and the subsequent effects of genetic drift and gene flow interacting with the landscape are associated with patterns of post-colonization divergence. I was able to disentangle the relative effects of colonization (IBC), gene flow (IBD), and adaptation (IBA) at a smaller spatial scale than is typical because of the relatively recent and colonization history of this system. This study may also represent an uncommon example where historical colonization events can have a lasting imprint on the distribution of phenotypic variation. Historical contingencies may therefore influence contemporary patterns of variation, which may in turn have implications for the long-term adaptability of small and/or isolated

populations. Understanding how the effects of historical events interact with contemporary ecological conditions is an important step in conserving intraspecific biodiversity in fragmented habitats.

LITERATURE CITED

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

Aldenhoven, J. T., M. A. Miller, P. S. Corneli, and M. D. Shapiro. 2010. Phylogeography of ninespine sticklebacks (*Pungitius pungitius*) in North America: Glacial refugia and the origins of adaptive traits. Molecular Ecology 19:4061–4076.

Allendorf, F. W., G. Luikart, and S. N. Aitken. 2013. Conservation and the genetics of populations. 2nd ed. John Wiley & Sons, Ltd, West Sussex.

April, J., and R. H. Hanner. 2013. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. Molecular Ecology 22:409–422.

Arbour, J. H., D. C. Hardie, and J. A. Hutchings. 2011. Morphometric and genetic analyses of two sympatric morphs of arctic char (*Salvelinus alpinus*) in the Canadian high arctic. Canadian Journal of Zoology 89:19–30.

Astudillo-Clavijo, V., J. H. Arbour, and H. López-Fernández. 2015. Selection towards different adaptive optima drove the early diversification of locomotor phenotypes in the radiation of Neotropical geophagine cichlids Evolutionary developmental biology and morphology. BMC Evolutionary Biology 15:1–13.

Bagley, R. K., V. C. Sousa, M. L. Niemiller, and C. R. Linnen. 2017. History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (*Neodiprion lecontei*). Molecular Ecology 26:1022–1044.

Bardakci, F., and D. O. F. Skibinski. 1994. Application of the RAPD technique in tilapia fish: Species and subspecies identification. Heredity 73:117–123.

Barrera-moreno, O. A., J. Ciros-pérez, and E. Ortega-Mayagoitia. 2015. From local adaptation to ecological speciation in copepod populations from neighboring lakes. PLoS ONE 10:1–25.

Bernatchez, L., and C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Molecular Ecology 7:431–452.

Bernatchez, S., M. Laporte, C. Perrier, P. Sirois, and L. Bernatchez. 2016. Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. Molecular ecology 25:4773–4792.

Bond, M. H., P. A. Crane, W. A. Larson, and T. P. Quinn. 2014. Is isolation by adaptation driving genetic divergence among proximate Dolly Varden char populations? Ecology and Evolution 4:2515–2532.

Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical Ecology with R. Springer, New York, USA.

Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51–68.

Boria, R. A., S. K. Brown, M. D. Matocq, and J. L. Blois. 2021. Genome-wide genetic variation coupled with demographic and ecological niche modeling of the dusky-footed woodrat (*Neotoma fuscipes*) reveal patterns of deep divergence and widespread Holocene expansion across northern California. Heredity 126:521–536.

Brachmann, M. K., K. Parsons, S. Skúlason, and M. M. Ferguson. 2021. The interaction of resource use and gene flow on the phenotypic divergence of benthic and pelagic morphs of Icelandic Arctic charr (*Salvelinus alpinus*). Ecology and Evolution 0:1–20.

Bray, J., and J. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

Brunner, P. C., M. R. Douglas, A. Osinov, C. C. Wilson, and L. Bernatchez. 2001. Holarctic Phylogeography of Arctic charr (*Salvelinus alpinus* L.) inferred from Mitochondrial DNA Sequences. Evolution 55:573–586.

Bryant, D., and V. Moulton. 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. Molecular Biology and Evolution 21:255–265.

Burress, E. D., J. M. Holcomb, M. Tan, and J. W. Armbruster. 2017. Ecological diversification associated with the benthic-to-pelagic transition by North American minnows. Journal of Evolutionary Biology 30:549–560.

Caldera, E. J., and D. I. Bolnick. 2008. Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (*Gasterosteus aculeatus*) populations in a single watershed. Evolutionary Ecology Research 10:575–598.

Cameron, A. C., R. B. Page, J. I. Watling, C. A. M. Hickerson, and C. D. Anthony. 2019. Using a comparative approach to investigate the relationship between landscape and genetic connectivity among woodland salamander populations. Conservation Genetics 20:1265–1280.

Chan, K. O. N. N., and R. M. Brown. 2020. Elucidating the drivers of genetic differentiation in Malaysian torrent frogs (Anura: Ranidae: *Amolops*): A landscape genomics approach. Zoological Journal of the Linnean Society 190:65–78.

Charlesworth, B. 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics 10:195–205.

Christensen, K. A., E. B. Rondeau, D. R. Minkley, J. S. Leong, C. M. Nugent, R. G. Danzmann, M. M. Ferguson, A. Stadnik, R. H. Devlin, R. Muzzerall, M. Edwards, W. S. Davidson, and B. F. Koop. 2021. Retraction: The Arctic charr (*Salvelinus alpinus*) genome and transcriptome assembly. PLoS ONE 16:247083.

Clegg, S. M., and A. B. Phillimore. 2010. The influence of gene flow and drift on genetic and phenotypic divergence in two species of *Zosterops* in Vanuatu. Philosophical Transactions of the Royal Society B: Biological Sciences 365:1077–1092.

Cooper, W. J., K. Parsons, A. McIntyre, B. Kern, A. McGee-Moore, and R. Craig. 2010. Bentho-pelagic divergence of cichlid feeding architecture was prodigious and consistent during multiple adaptive radiations within African rift-lakes. PLoS ONE 5:e9551.

Cornuet, J., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014.

Courtwright, J., and C. L. May. 2013. Importance of terrestrial subsidies for native brook trout in Appalachian intermittent streams. Freshwater Biology 58:2423–2438.

Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: A critique of Sewall Wright's shifting balance theory of evolution. Evolution 51:643–671.

Croucher, P. J. P., G. S. Oxford, A. Lam, N. Mody, and R. G. Gillespie. 2012. Colonization History and Population Genetics of the Color-polymorphic Hawaiian Happy-face Spider *Theridion grallator* (Araneae, Theridiidae). Evolution 66:1–19.de Meester, L., A. Gómez, B. Okamura, and K. Schwenk. 2002. The Monopolization Hypothesis and the dispersal-gene flow paradox in aquatic organisms. Acta Oecologica 23:121–135.

Deagle, B. E., F. C. Jones, D. M. Absher, D. M. Kingsley, and T. E. Reimchen. 2013. Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. Molecular Ecology 22:1917–1932.

Devillard, S., M. Jacquier, J. M. Vandel, F. Léger, J. Duhayer, S. Pardonnet, L. Say, and S. Ruette. 2022. Genetic variability and population size covary positively across nine badgers (*Meles meles*) populations in France. Mammal Research.

DeWoody, J., and J. C. Avise. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fish Biology 56:461–473.

Dewoody, J., H. Trewin, and G. Taylor. 2015. Genetic and morphological differentiation in *Populus nigra* L.: isolation by colonization or isolation by adaptation? Molecular Ecology 24:2641–2655.

Dowle, E. J., R. R. Bracewell, M. E. Pfrender, K. E. Mock, B. J. Bentz, and G. J. Ragland. 2017. Reproductive isolation and environmental adaptation shape the phylogeography of mountain pine beetle (*Dendroctonus ponderosae*). Molecular Ecology 26:6071–6084.

Dray, S., and A. Dufour. 2007. The ade4 package: Implementing the duality diagram for ecologists. Journal of Statistical Software 22:1–20.

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196:483–493.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. Már Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Evans, M. L., K. Præbel, S. Peruzzi, and L. Bernatchez. 2012. Parallelism in the oxygen transport system of the lake whitefish: the role of physiological divergence in ecological speciation. Molecular Ecology 21:4038–4050.

Excoffier, L., M. Foll, and R. J. Petit. 2009. Genetic consequences of range expansions. Annual Review of Ecology, Evolution, and Systematics 40:481–501.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longman, Essex, UK.

Fenderson, L. E., A. I. Kovach, and B. Llamas. 2020. Spatiotemporal landscape genetics: Investigating ecology and evolution through space and time. Molecular Ecology 29:218–246.

Ferchaud, A. L., M. Leitwein, M. Laporte, D. Boivin-Delisle, B. Bougas, C. Hernandez, É. Normandeau, I. Thibault, and L. Bernatchez. 2020. Adaptive and maladaptive genetic diversity in small populations: Insights from the Brook Charr (*Salvelinus fontinalis*) case study. Molecular Ecology 29:3429–3445.

Ferry-Graham, L. A., G. v. Lauder, and C. D. Hulsey. 2001. Aquatic prey capture in ray-finned fishes: A century of progress and new directions. Journal of Morphology 248:99–119.

Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180:977–993.

Forester, B. R., M. R. Jones, E. St, and J. R. Lasky. 2016. Detecting spatial genetic signatures of local adaptation in heterogenous landscapes. Molecular Ecology 25:104–120.

Francis, R. M. 2016. POPHELPER: An R package and web app to analyse and visualize population structure. Molecular Ecology Resources2 17:27–32.

Frankham, R., C. J. A. Bradshaw, and B. W. Brook. 2014. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biological Conservation 170:56–63.

Franklin, O. D., S. Skúlason, M. B. Morrissey, and M. M. Ferguson. 2018. Natural selection for body shape in resource polymorphic Icelandic Arctic charr. Journal of Evolutionary Biology 31:1498–1512.

Fraser, D. J., P. v. Debes, L. Bernatchez, and J. A. Hutchings. 2014. Population size, habitat fragmentation, and the nature of adaptive variation in a stream fish. Proceedings of the Royal Society B: Biological Sciences 281.

Fraser, D. J., L. K. Weir, L. Bernatchez, M. M. Hansen, and E. B. Taylor. 2011. Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity 106:404–420.

Frichot, E., and O. François. 2015. LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution 6:925–929.

Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973–983.

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

Gallego-García, N., G. Forero-Medina, M. Vargas-Ramírez, S. Caballero, and H. B. Shaffer. 2019. Landscape genomic signatures indicate reduced gene flow and forest-associated adaptive divergence in an endangered neotropical turtle. Molecular Ecology 28:2757–2771.

Garant, D., S. E. Forde, and A. P. Hendry. 2007. The multifarious effects of dispersal and gene flow on. Functional Ecology 21:434–443.

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

Gíslason, D., M. M. Ferguson, S. Skúlason, and S. S. Snorrason. 1999. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). Canadian Journal of Fisheries and Aquatic Sciences 56:2229–2234.

Gomulkiewicz, R., and D. Houle. 2009. Demographic and genetic constraints on evolution.

Goslee, S., and D. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22:1–19.

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

Griffiths, D. 2010. Pattern and process in the distribution of North American freshwater fish. Biological Journal of the Linnean Society 100:46–61.

Grummer, J. A., L. B. Beheregaray, L. Bernatchez, B. K. Hand, G. Luikart, S. R. Narum, and E. B. Taylor. 2019. Aquatic landscape genomics and environmental effects on genetic variation. Trends in Ecology and Evolution 34:641–654.

Gudbergsson, G. 2004. Arctic charr in Lake Myvatn: The centennial catch record in the light of recent stock estimates. Aquatic Ecology 38:271–285.

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

Habel, J. C., M. Husemann, A. Finger, P. D. Danley, and F. E. Zachos. 2014. The relevance of time series in molecular ecology and conservation biology. Biological Reviews 89:484–492.

Haileselasie, T. H., J. Mergeay, L. J. Weider, R. Sommaruga, T. A. Davidson, M. Meerhoff, H. Arndt, K. Jürgens, E. Jeppesen, and L. de Meester. 2016. Environment not dispersal limitation drives clonal composition of Arctic Daphnia in a recently deglaciated area. Molecular Ecology 25:5830–5842.

Hatfield, T., and D. Schluter. 1999. Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. Evolution 53:866–873.

Hayward, A., M. Tsuboi, C. Owusu, A. Kotrschal, and S. D. Buechel. 2017. Evolutionary associations between host traits and parasite load: insights from Lake Tanganyika cichlids. Journal of Evolutionary Biology 46:1056–1067.

Helsen, K., H. Jacquemyn, and O. Honnay. 2015. Hidden founder effects: small-scale spatial genetic structure in recently established populations of the grassland specialist plant *Anthyllis vulneraria*. Molecular Ecology 24:2715–2728.

Hewitt, G. 2000. The genetic legacy of the quaternary ice ages. Nature 405:907–913.

Hewitt, G. M. 2004. A climate for colonization. Heredity 92:1–2.

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

Higham, T. E. 2007. The integration of locomotion and prey capture in vertebrates: Morphology, behavior, and performance. Integrative and Comparative Biology 47:82–95.

Hoffmann, A. A., C. M. Sgrò, and T. N. Kristensen. 2017. Revisiting adaptive potential, population size, and conservation. Trends in Ecology and Evolution 32:506–517.

Huson, D. H. 1998. SPLITSTREE: Analyzing and visualizing evolutionary data. Bioinformatics 14:68–73.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Mývatn. Nature 452:84–87.

Jacobs, A., M. Carruthers, A. Yurchenko, N. v. Gordeeva, S. S. Alekseyev, O. Hooker, J. S. Leong, D. R. Minkley, E. B. Rondeau, B. F. Koop, C. E. Adams, and K. R. Elmer. 2020. Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. PLoS Genetics 4:1–34.

Jenkins, D. G., M. Carey, J. Czerniewska, J. Fletcher, T. Hether, A. Jones, S. Knight, J. Knox, T. Long, M. Mannino, M. McGuire, A. Riffle, S. Segelsky, L. Shappell, A. Sterner, T. Strickler, and R. Tursi. 2010. A meta-analysis of isolation by distance: Relic or reference standard for landscape genetics? Ecography 33:315–320.

Jordan, M. A., and H. L. Snell. 2008. Historical fragmentation of islands and genetic drift in populations of Galápagos lava lizards (*Microlophus albemarlensis* complex). Molecular Ecology 17:1224–1237.

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ e281.

Karvonen, A., and O. Seehausen. 2012. The role of parasitism in adaptive radiations — When might parasites promote and when might they constrain ecological speciation? International Journal of Ecology Volume 2012: 1–20.

Kautt, A. F., G. Machado-Schiaffino, and A. Meyer. 2018. Lessons from a natural experiment: Allopatric morphological divergence and sympatric diversification in the Midas cichlid species complex are largely influenced by ecology in a deterministic way. Evolution Letters 2:323–340.

Kawaguchi, Y., and S. Nakano. 2001. Contribution of terrestrial invertebrates to the annual resource budget for salmonids in forest and grassland reaches of a headwater stream. Freshwater Biology 46:303–316.

Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters 7:1225–1241.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kingsolver, J. G., S. E. Diamond, A. M. Siepielski, and S. M. Carlson. 2012. Synthetic analyses of phenotypic selection in natural populations: Lessons, limitations and future directions. Evolutionary Ecology 26:1101–1118.

Knudsen, R., A. Klemetsen, S. Alekseyev, C. E. Adams, and M. Power. 2016. The role of Salvelinus in contemporary studies of evolution, trophic ecology and anthropogenic change. Hydrobiologia 783:1–9.

Koblmãller, S., W. Salzburger, B. Obermãller, E. Eigner, C. Sturmbauer, and K. M. Sefc. 2011. Separated by sand, fused by dropping water: Habitat barriers and fluctuating water levels steer the evolution of rock-dwelling cichlid populations in Lake Tanganyika. Molecular Ecology 20:2272–2290.

Kolbe, J. J., T. W. Schoener, D. A. Spiller, and J. B. Losos. 2012. Founder effects persist despite adaptive differentiation: A field experiment with lizards. Science 335:1086–1090.

Kreiling, A. K., E. J. O'Gorman, S. Pálsson, D. Benhaïm, C. A. Leblanc, J. S. Ólafsson, and B. K. Kristjánsson. 2021. Seasonal variation in the invertebrate community and diet of a top fish predator in a thermally stable spring. Hydrobiologia 848:531–545.

Kristjánsson, B. K. 2008. Fine scale phenotypic diversity of Arctic charr (*Salvelinus alpinus*) in relation to ecological characters. PhD. thesis. Zoology Department, University of Guelph.

Kristjánsson, B. K., and C. A. Leblanc. 2018. Variation in the magnitude of morphological and dietary differences between individuals among populations of small benthic Arctic charr in relation to ecological factors. Ecology and Evolution 8:1573–1581.

Kristjánsson, B. K., H. J. Malmquist, F. Ingimarsson, T. Antonsson, S. U. S. Snorrason, and S. Skúlason. 2011. Relationships between lake ecology and morphological characters in Icelandic Arctic charr, *Salvelinus alpinus*. Biological Journal of the Linnean Society 103:761–771.

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

Lande, R. 1988. Genetics and demography in biological conservation. Science 241:1455-1460.

Lanier, H. C., R. Massatti, Q. He, L. E. Olson, and L. Lacey Knowles. 2015. Colonization from divergent ancestors: Glaciation signatures on contemporary patterns of genomic variation in Collared Pikas (*Ochotona collaris*). Molecular Ecology 24:3688–3705.

Launey, S., G. Brunet, R. Guyomard, and P. Davaine. 2010. Role of introduction history and landscape in the range expansion of brown trout (*Salmo trutta* L.) in the Kerguelen Islands. Journal of Heredity 101:270–283.

Laurent, S. J. Y., A. Werzner, L. Excoffier, and W. Stephan. 2011. Approximate Bayesian Analysis of *Drosophila melanogaster* polymorphism data reveals a recent colonization of Southeast Asia. Molecular Biology and Evolution 28:2041–2051.

Legendre, P., and M. J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10:831–844.

Legendre, P., M. J. Fortin, and D. Borcard. 2015. Should the Mantel test be used in spatial analysis? Methods in Ecology and Evolution 6:1239–1247.

Legendre, P., and L. Legendre. 2012. Canonical analysis. Developments in Environmental Modelling 24:625–710.

Linck, E., and C. J. Battey. 2017. Minor allele frequency thresholds dramatically affect population structure inference with genomic datasets. Molecular Ecology Resources 19:639–647.

Luu, K., E. Bazin, and M. G. B. Blum. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources 17:67–77.

Machado, A. P., A. Topaloudis, T. Cumer, E. Lavanchy, V. Bontzorlos, R. Ceccherelli, M. Charter, N. Kassinis, P. Lymberakis, F. Manzia, A. L. Ducrest, M. Dupasquier, N. Guex, A. Roulin, and J. Goudet. 2022. Genomic consequences of colonisation, migration and genetic drift in barn owl insular populations of the eastern Mediterranean. Molecular Ecology 31:1375–1388.

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

Malmquist, A. H. J., S. S. Snorrason, S. Skúlason, B. Jonsson, and O. T. Sandlund. 1992. Diet differentiation in polymorphic Arctic charr in Thingvallavatn, Iceland. Journal of Animal Ecology 61:21–35.

Malmquist, H. J. 1992. Phenotype-specific feeding behaviour of two Arctic charr *Salvelinus alpinus* morphs. Oecologia 92:354–361.

McCoy, R. C., N. R. Garud, J. L. Kelley, C. L. Boggs, and D. A. Petrov. 2014. Genomic inference accurately predicts the timing and severity of a recent bottleneck in a nonmodel insect population. Molecular Ecology 23:136–150.

McCusker, M. R., and P. Bentzen. 2010. Positive relationships between genetic diversity and abundance in fishes. Molecular Ecology 19:4852–4862.

McMeans, B. C., K. S. McCann, M. Humphries, N. Rooney, and A. T. Fisk. 2015. Food web structure in temporally-forced ecosystems. Trends in Ecology and Evolution 30:662–672.

Meester, L. de, J. Vanoverbeke, L. J. Kilsdonk, and M. C. Urban. 2016. Evolving perspectives on monopolization and priority effects. Trends in Ecology and Evolution 31:136–146.

Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: FST and related measures. Molecular Ecology Resources 11:5–18.

Monteiro, L. R., and M. R. Nogueira. 2010. Adaptive radiations, ecological specialization, and the evolutionary integration of complex morphological structures. Evolution 64:724–744.

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

Morrissey, M. B., D. T. de Kerckhove. 2009. The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. The American Naturalist 174:875–889.

Motta, P. J., S. F. Norton, and J. J. Luczkovich. 1995. Perspectives on the ecomorphology of bony fishes. Environmental Biology of Fishes 44:11–20.

Mussmann, S. M., M. R. Douglas, T. K. Chafin, and M. E. Douglas. 2019. BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. Methods in Ecology and Evolution 10:1808–1813.

Nadeau, S., P. G. Meirmans, S. N. Aitken, K. Ritland, and N. Isabel. 2016. The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. Ecology and Evolution 6:8649–8664.

Ngeve, M. N., T. van der Stocken, D. Menemenlis, N. Koedam, and L. Triest. 2017. Hidden founders? Strong bottlenecks and fine-scale genetic structure in mangrove populations of the Cameroon Estuary complex. Hydrobiologia 803:189–207.

Nigenda-Morales, S. F., M. E. Gompper, D. Valenzuela-Galván, A. R. Lay, K. M. Kapheim, C. Hass, S. D. Booth-Binczik, G. A. Binczik, B. T. Hirsch, M. McColgin, J. L. Koprowski, K. McFadden, R. K. Wayne, and K. P. Koepfli. 2019. Phylogeographic and diversification

patterns of the white-nosed coati (*Nasua narica*): Evidence for south-to-north colonization of North America. Molecular Phylogenetics and Evolution 131:149–163.

Nosil, P., S. P. Egan, and D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. Evolution 62:316–336.Nugent, C. M., J. S. Leong, K. A. Christensen, E. B. Rondeau, M. K. Brachmann, A. A. Easton, C. L. Ouellet-Fagg, M. T. T. Crown, W. S. Davidson, B. F. Koop, R. G. Danzmann, and M. M. Ferguson. 2019. Design and characterization of an 87k SNP genotyping array for Arctic charr (*Salvelinus alpinus*). PLoS ONE 14:1–17.

Nunziata, S. O., S. L. Lance, D. E. Scott, E. M. Lemmon, and D. W. Weisrock. 2017. Genomic data detect corresponding signatures of population size change on an ecological time scale in two salamander species. Molecular Ecology 26:1060–1074.

Oksanen, J., G. F. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2020. vegan: Community Ecology Package. R package version 2.5-7.

Orsini, L., J. Mergeay, J. Vanoverbeke, and L. de Meester. 2013a. The role of selection in driving landscape genomic structure of the waterflea *Daphnia magna*. Molecular Ecology 22:583–601.

Orsini, L., J. Vanoverbeke, I. N. E. Swillen, and J. Mergeay. 2013b. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. Molecular Ecology 22:5983–5999.

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

Præbel, K., R. Knudsen, A. Siwertsson, M. Karhunen, K. K. Kahilainen, O. Ovaskainen, K. Østbye, S. Peruzzi, S. E. Fevolden, and P. A. Amundsen. 2013. Ecological speciation in postglacial European whitefish: Rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. Ecology and Evolution 3:4970–4986.

Pruett, C. L., and K. Winker. 2005. Northwestern song sparrow populations show genetic effects of sequential colonization. Molecular Ecology 14:1421–1434.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81:559–575.

Rao, C. 1964. The use and interpretation of principal components analysis in applied research. The Indian Journal of Statistics, Series A 26:329–358.

Räsänen, K., and A. P. Hendry. 2008. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. Ecology Letters 11:624–636.

Räsänen, K., and A. P. Hendry. 2014. Asymmetric reproductive barriers and mosaic reproductive isolation: insights from Misty lake-stream stickleback. Ecology and Evolution 4:1166–1175.

Rey, O., L. Fourtune, I. Paz-Vinas, G. Loot, C. Veyssière, B. Roche, and S. Blanchet. 2015. Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. Molecular Ecology 24:5348–5363.

Richardson, J. L., M. C. Urban, D. I. Bolnick, and D. K. Skelly. 2014. Microgeographic adaptation and the spatial scale of evolution. Trends in Ecology and Evolution 29:165–176.

Robinson, B. W., and D. S. Wilson. 1994. Character release and displacement in fishes: A neglected literature. The American Naturalist 144:596–627.

Roesti, M., W. Salzburger, and D. Berner. 2012. Uninformative polymorphisms bias genome scans for signatures of selection. BMC Evolutionary Biology 12.

Rohlf, J. F. 2015. The tps series of software. Hystrix 26:9–12.

Rosamond, K. M., C. M. Kent, S. C. Hunter, and T. W. Sherry. 2020. Morphological traits influence prey selection by coexisting species of New World warblers (Parulidae). Journal of Field Ornithology 91:393–408.

Rüber, L., and D. C. Adams. 2001. Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika. Journal of Evolutionary Biology 14:325–332.

Salisbury, S. J., C. Booker, G. R. Mccracken, T. Knight, D. Keefe, R. Perry, and D. E. Ruzzante. 2018. Genetic divergence among and within Arctic char (*Salvelinus alpinus*) populations inhabiting landlocked and sea-accessible sites in Labrador, Canada. Canadian Journal of Fisheries and Aquatic Sciences 1269:1256–1269.

Salisbury, S. J., G. R. McCracken, D. Keefe, R. Perry, and D. E. Ruzzante. 2016. A portrait of a sucker using landscape genetics: how colonization and life history undermine the idealized dendritic metapopulation. Molecular ecology 25:4126–4145.

Salisbury, S. J., G. R. McCracken, R. Perry, D. Keefe, K. K. S. Layton, T. Kess, C. M. Nugent, J. S. Leong, I. R. Bradbury, B. F. Koop, M. M. Ferguson, and D. E. Ruzzante. 2020. Limited genetic parallelism underlies recent, repeated incipient speciation in geographically proximate populations of an Arctic fish (*Salvelinus alpinus*). Molecular Ecology 29:4280–4294.

Schluter, D. 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. Ecology 74:699–709.

Schluter, D. 1995. Adaptive radiation in sticklebacks: Trade-offs in feeding performance and growth. Ecology 76:82–90.

Schluter, D. 2009. Evidence for ecological speciation and its alternative. Science 323:737–741.

Sendell-Price, A. T., K. C. Ruegg, B. C. Robertson, and S. M. Clegg. 2021. An island-hopping bird reveals how founder events shape genome-wide divergence. Molecular Ecology 30:2495–2510.

Seymour, M., K. Räsänen, R. Holderegger, and B. K. Kristjánsson. 2013. Connectivity in a pond system influences migration and genetic structure in threespine stickleback. Ecology and Evolution 3:492–502.

Seymour, M., K. Räsänen, and B. K. Kristjánsson. 2019. Drift versus selection as drivers of phenotypic divergence at small spatial scales: The case of Belgjarskógur threespine stickleback. Ecology and Evolution 9:8133–8145.

Shafer, A. B. A., and J. B. W. Wolf. 2013. Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. Ecology Letters 16:940–950.

Shaffer, M. K. 1981. Minimum viable populations size for species conservation. Bioscience 31:131–134.

Shikano, T., Y. Shimada, G. Herczeg, and J. Merila. 2010. History vs. habitat type: explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. Molecular Ecology 19:1147–1161.

Skoglund, S., A. Siwertsson, P. A. Amundsen, and R. Knudsen. 2015. Morphological divergence between three Arctic charr morphs - the significance of the deep-water environment. Ecology and Evolution 5:3114–3129.

Skúlason, S., and T. B. Smith. 1995. Resource polymorphisms in vertebrates. Trends in Ecology and Evolution 4:238–239.

Smith, T. B., and S. Skúlason. 1996. Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. Annual Review of Ecology and Systematics 27:111–133.

Spellman, G. M., and J. Klicka. 2007. Phylogeography of the white-breasted nuthatch (*Sitta carolinensis*): Diversification in North American pine and oak woodlands. Molecular Ecology 16:1729–1740.

Spurgin, L. G., J. C. Illera, T. H. Jorgensen, D. A. Dawson, and D. S. Richardson. 2014. Genetic and phenotypic divergence in an island bird: Isolation by distance, by colonization or by adaptation? Molecular Ecology 23:1028–1039.

Stelkens, R. B., G. Jaffuel, M. Escher, and C. Wedekind. 2012. Genetic and phenotypic population divergence on a microgeographic scale in brown trout. Molecular Ecology 21:2896–2915.

Sweka, J. A., and K. J. Hartman. 2008. Contribution of terrestrial invertebrates to yearly brook trout prey consumption and growth. Transactions of the American Fisheries Society 137:224–235.

Tallmon, D. A., G. Luikart, and R. S. Waples. 2004. The alluring simplicity and complex reality of genetic rescue. Trends in Ecology & Evolution 19:489–496.

Team, R. C. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Tobler, M., K. Scharnweber, R. Greenway, C. N. Passow, L. Arias-Rodriguez, and F. K. García-de-León. 2015. Convergent changes in the trophic ecology of extremophile fish along replicated environmental gradients. Freshwater Biology 60:768–780.

Torterotot J., C. Perrier, N. E. Bergeron, and L. Bernatchez. 2014. Influence of forest road culverts and waterfalls on the fine-scale distribution of brook trout genetic diversity in a boreal watershed. Transactions of the American Fisheries Society 143:1577–1591.

Torres-Dowdall, J., C. A. Handelsman, D. N. Reznick, and C. K. Ghalambor. 2012. Local adaptation and the evolution of phenotypic plasticity in Trinidadian guppies (*Poecilia Reticulata*). Evolution 66:3432–3443.

Trapnell, D. W., J. L. Hamrick, C. D. Ishibashi, and T. R. Kartzinel. 2013. Genetic inference of epiphytic orchid colonization; It may only take one. Molecular Ecology 22:3680–3692.

Ventura, M., A. Petrusek, A. Miró, E. Hamrová, D. Buñay, L. de Meester, and J. Mergeay. 2014. Local and regional founder effects in lake zooplankton persist after thousands of years despite high dispersal potential. Molecular Ecology 23:1014–1027.

Wang, I. J. 2013. Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution 67:3403–3411.

Wang, I. J., and G. S. Bradburd. 2014. Isolation by environment. Molecular Ecology 23:5649–5662.

Wang, S., W. Zhu, X. Gao, X. Li, S. Yan, X. Liu, J. Yang, Z. Gao, and Y. Li. 2014. Population size and time since island isolation determine genetic diversity loss in insular frog populations. Molecular Ecology 23:637–648.

Weese, D. J., M. M. Ferguson, and B. W. Robinson. 2012. Contemporary and historical evolutionary processes interact to shape patterns of within-lake phenotypic divergences in polyphenic pumpkinseed sunfish, *Lepomis gibbosus*. Ecology and Evolution 2:574–592.

Weir, B., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 1358–1370.

Willi, Y., and A. A. Hoffmann. 2009. Demographic factors and genetic variation influence population persistence under environmental change. Journal of Evolutionary Biology 22:124–133.

Willi, Y., J. van Buskirk, and A. A. Hoffmann. 2006. Limits to the adaptive potential of small populations. Annual Review of Ecology, Evolution, and Systematics 37:433–458.

Willing, E. M., P. Bentzen, C. van Oosterhout, M. Hoffmann, J. Cable, F. Breden, D. Weigel, and C. Dreyer. 2010. Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. Molecular Ecology 19:968–984.

Wilson, A. J., D. Gíslason, S. Skúlason, S. S. Snorrason, C. E. Adams, G. Alexander, R. G. Danzmann, and M. M. Ferguson. 2004. Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. Molecular Ecology 13:1129–1142.

Wilson, G. A., and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191.

Wood, J. L. A., D. Tezel, D. Joyal, and D. J. Fraser. 2015. Population size is weakly related to quantitative genetic variation and trait differentiation in a stream fish. Evolution 69:2303–2318.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

Ye, H., Z. Wang, H. Hou, J. Wu, Y. Gao, W. Han, W. Ru, G. Sun, and Y. Wang. 2021. Localized environmental heterogeneity drives the population differentiation of two endangered and endemic *Opisthopappus* Shih species. BMC Ecology and Evolution 21:1–20. BioMed Central.

Zastavniouk, C., L. K. Weir, and D. J. Fraser. 2017. The evolutionary consequences of habitat fragmentation: Body morphology and coloration differentiation among brook trout populations of varying size. Ecology and Evolution 7:6850–6862.

Zelditch, M. L. 2012. Geometric morphometrics for biologists. 2nd ed. Elsevier Inc.

Zhang, J. 2018. Neutral theory and phenotypic evolution. Molecular Biology and Evolution 35:1327–1331.

TABLES

Table 1. Geographic information and sample sizes for 1782 Arctic charr (*Salvelinus alpinus*) collected in June and August of 2014 and 2019 from 19 lava caves located in the Haganes (H) and Vindbelgur (V) regions near Lake Mývatn, Iceland. Ecological data refers to biotic (benthic and aerial) and abiotic variables. The forward slash separates the number of individuals genotyped and the number of individuals with phenotypic data, respectively. Additionally, two morphs (Krús and Generalist, see text for more information) were collected in 2012 from the lake, but these samples have genetic data only.

Population	Region	Geographic cluster	Ecological data?	2014		2019		Total
				June	August	June	August	
Cave 1	H	South (H-S)	Y	13 / 13	7 / 7	7 / 7	14 / 14	41 / 41
Cave 2	H	South (H-S)	Y	18 / 18	10 / 10	13 / 13	14 / 14	55 / 55
Cave 5	H	North (H-N)	Y	17 / 25	13 / 21	8 / 31	22 / 22	60 / 99
Cave 7	H	Central (H-C)	Y	13 / 27	17 / 29	12 / 22	17 / 33	59 / 111
Cave 10	H	North (H-N)	Y	17 / 22	13 / 13	13 / 18	17 / 21	60 / 74
Cave 11	H	North (H-N)	Y	16 / 22	15 / 15	13 / 13	17 / 17	60 / 67
Cave 12	H	Central (H-C)	Y	21 / 22	8 / 8	8 / 8	12 / 12	49 / 50
Cave 17	V	West (V-W)	N	5 / 5	9 / 9	6 / 6	10 / 10	30 / 30
Cave 17b	V	West (V-W)	Y	7 / 7	7 / 7	9 / 9	16 / 16	39 / 39
Cave 18	V	West (V-W)	Y	17 / 31	13 / 19	8 / 19	22 / 33	60 / 102
Cave 19	V	West right (V-W r)	Y	11 / 11	2 / 2	16 / 16	12 / 12	41 / 41
Cave 20	V	West (V-W)	Y	21 / 23	9 / 12	12 / 22	18 / 25	60 / 82
Cave 21	V	East (V-E)	N	21 / 22	5 / 5	10 / 21	20 / 26	56 / 74
Cave 22	V	East (V-E)	Y	17 / 17	8 / 8	11 / 18	18 / 37	54 / 80
Cave 23	V	East (V-E)	Y	22 / 27	8 / 10	10 / 14	20 / 28	60 / 79
Cave 24	V	East (V-E)	N	17 / 17	6 / 6	5 / 5	13 / 13	41 / 41
Cave 25	H	Central (H-C)	Y	22 / 134	8 / 128	20 / 134	10 / 87	60 / 483
Cave 26	H	Central-South (H-CS)	N	15 / 44	15 / 33	14 / 31	16 / 34	60 / 142
Cave 27	H	Central-South (H-CS)	Y	16 / 20	14 / 29	17 / 24	13 / 19	60 / 92
			Total:	304 / 507	188 / 371	213 / 424	301 / 480	1005 / 1782

Table 2. Estimates of genetic diversity and number of private alleles based on neutral SNPs and census (N_c) and genetic effective (N_e) population sizes of 19 cave (C) populations of Icelandic Arctic charr (*Salvelinus alpinus*). Individuals from Lake Mývatn (L, both Krús and Generalist morphs) were sampled in 2012, whereas cave-dwelling individuals from the Haganes (H) and Vindbelgur (V) regions near the lake were sampled in June and August of 2014 and 2019. Three high-level genetic clusters (HLGCs) were identified, and population membership to Haganes, Vindbelgur East (VE) and Vindbelgur West (VW) is also indicated (see Results for justification). Genetic diversity was quantified as observed heterozygosity (H_o), expected heterozygosity (H_e) and allelic richness (A_r). Estimates were repeated by grouping individuals from C1 and C2, and C17 and C18 together (termed "combined") as these populations were not significantly genetically differentiated (see text for more information).

Population	Region	HLGC	Samples genotyped	Biallelic SNPs	Private Alleles	A_r	H_o	H_e	N_c Mean (95%	N_e Mean (95% CI)
C1	H	H	41	937	8	1.470	0.159	0.153	33 (26 – 49)	13 (11 – 15)
C2	H	H	57	1030	4	1.500	0.164	0.160	30 (23 – 49)	18 (16 – 22)
C5	H	H	61	1054	6	1.543	0.184	0.178	77 (57 – 116)	43 (34 – 58)
C7	H	H	59	1213	2	1.621	0.203	0.199	103 (77 – 147)	48 (36 – 71)
C10	H	H	60	1056	3	1.547	0.184	0.179	51 (38 – 83)	22 (19 – 27)
C11	H	H	60	1093	2	1.566	0.186	0.184	37 (29 – 55)	17 (15 – 19)
C12	H	H	49	1149	1	1.591	0.200	0.192	41 (23 – 81)	11 (10 – 12)
C17	V	VW	30	874	6	1.448	0.141	0.132	19 (13 – 34)	10 (9 – 11)
C17b	V	VW	39	893	3	1.451	0.146	0.140	32 (18 – 74)	15 (13 – 18)
C18	V	VW	59	908	11	1.427	0.131	0.127	80 (55 – 129)	24 (20 – 31)
C19	V	VW	41	815	2	1.423	0.150	0.141	29 (22 – 48)	12 (11 – 14)
C20	V	VW	60	925	3	1.438	0.145	0.143	68 (47 – 114)	21 (18 – 25)
C21	V	VE	57	975	0	1.505	0.184	0.174	40 (34 – 59)	4 (3 – 4)
C22	V	VE	58	1032	8	1.534	0.183	0.181	65 (40 – 134)	18 (15 – 21)
C23	V	VE	59	976	7	1.470	0.157	0.151	53 (38 – 81)	12 (11 – 14)
C24	V	VE	41	946	0	1.511	0.178	0.172	39 (21 – 88)	7 (6 – 8)
C25	H	H	60	1217	2	1.616	0.199	0.196	365 (298 -	51 (36 – 90)
C26	H	H	60	940	0	1.510	0.175	0.171	99 (78 – 133)	59 (43 – 97)
C27	H	H	60	1013	5	1.517	0.183	0.174	56 (44 – 82)	32 (27 – 39)
Mývatn	L	NA	50	1435	60	1.797	0.239	0.252	NA	NA
Combined										
C1 and C2	H	H	98	1114	15	1.546	0.162	0.158	63 (52 – 83)	19 (17 – 21)
C17 and C18	V	VW	89	988	22	1.495	0.134	0.130	96 (71 – 140)	32 (26 – 40)

Table 3. Migration rates and population sizes among nine pairs of populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland. Average census and effective population sizes (N_c and N_e, respectively) are presented for each pair of populations. Also presented is the number of PIT-tagged migrant individuals detected in all years between 2014 and 2019. Average rates of effective migration (percentage per generation (standard deviation)) are presented for both directions. Estimates of gene flow and genetic differentiation (F_{ST}) are statistically significant ($p < 0.05$) unless indicated otherwise (*NS*).

Population pair	Distance (m)	Average N_c	Average N_e	Tagged migrants	Gene flow		F_{ST}
C1 and C2	16	32	16	15	C1 → C2: 0.161 (0.025)	C2 → C1: 0.198 (0.027)	*NS*
C5 and C10	106	64	32	0	C5 → C10: 0.019 (0.009)	C10 → C5: *NS*	0.061
C5 and C11	57	57	30	11	C5 → C11: 0.176 (0.026)	C11 → C5: 0.159 (0.023)	0.002
C7 and C25	46	234	50	51	C7 → C25: 0.251 (0.030)	C25 → C7: 0.100 (0.018)	0.002
C12 and C7	99	30	11	0	C12 → C7: 0.026 (0.012)	C7 → C12: *NS*	0.096
C17 and C18	42	99	17	14	C17 → C18: 0.231 (0.024)	C18 → C17: 0.011 (0.008)	*NS*
C17 and C17b	17	25	12	4	C17 → C17b: *NS*	C17b → C17: *NS*	0.036
C17b and C18	59	56	19	1	C17b → C18: 0.126 (0.044)	C18 → C17b: 0.014 (0.009)	0.053
C19 and C20	72	44	17	2	C20 → C19: 0.143 (0.025)	C19 → C20: 0.104 (0.023)	0.018
C21 and C22	50	59	11	0	C21 → C22: 0.174 (0.041)	C22 → C21: *NS*	0.040
C22 and C24	54	52	12	0	C22 → C24: 0.067 (0.016)	C24 → C22: 0.045 (0.020)	0.063
C23 and C24	14	46	9	0	C23 → C24: 0.110 (0.040)	C24 → C23: *NS*	0.092

Table 4. Relationships between genetic (F_{ST}), geographic, and ecological distances among 15 populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland. The relationship between the distance matrices was assessed via multiple matrix regression with randomization (MMRR). Response and predictor variables are separated by a tilde (~). Regression coefficients (β) and estimates of significance (P(>F)) are presented for each predictor variable. Genetic distances were represented using Weir and Cockerham's F_{ST} and neutral SNPs. Benthic invertebrate dissimilarities were estimated using Bray-Curtis indices, and Euclidean distances represent distances between abiotic factors and aerial invertebrates. Significant values ($p < 0.05$) are indicated in bold.

Model	***All* populations**			**Without Cave 25**		
	R^2	**β**	***P* (>F)**	**R^2**	**β**	***P* (>F)**
Genetic differentiation ~						
Abiotic distance	0.336	0.550	**< 0.001**	0.296	0.523	**< 0.001**
Benthic invertebrates	"	0.035	0.682	"	0.039	0.670
Aerial invertebrates	"	-0.046	0.493	"	-0.022	0.796

Table 5. Relationships among geographic distance, abiotic and biotic ecological variation, and genetic differentiation (F_{ST}) among 15 populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland based on redundancy analyses. All ecological variables were collected in the summer of 2014. Response and each predictor variable are separated by a tilde (~). High-level Genetic Clusters (HLGCs) are comprised of ancestry coefficients reflecting the three main genetic clusters (see text for justification). Analyses were conducted with and without Cave 25 (see text for justification). Partial redundancy analyses (pRDAs) were used to control for the effects of the matrix after the '|' symbol. In this case, the effects of colonization history are accounted for by using ancestry coefficients corresponding to each of three HLGCs. No Significant values ($p < 0.05$) were detected.

Model	All populations			Without Cave 25		
	Adjusted R^2	**F**	***P* (>F)**	**Adjusted R^2**	**F**	***P* (>F)**
Genetic differentiation						
~ Abiotic factors	0.129	1.517	0.139	0.097	1.351	0.222
~ Benthic invertebrates	-0.417	0.412	0.989	-0.490	0.389	0.993
~ Aerial invertebrates	-0.042	0.436	0.745	0.015	1.192	0.340
Genetic differentiation						
~ Abiotic factors | HLGCs	0.061	0.122	0.323	0.087	1.513	0.198
~ Benthic invertebrates | HLGCs	-0.594	0.312	0.987	-0.253	0.654	0.772
~ Aerial invertebrates | HLGCs	-0.042	0.461	0.669	0.046	1.601	0.223

Table 6. Phenotype-environment associations for populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland based on distance analyses. Relationships between phenotypic, ecological, and genetic distance matrices were assessed via multiple matrix regression with randomization (MMRR). Response and predictor variables are separated by a tilde (~). Regression coefficients (β) and estimates of significance (P(>F)) are presented for each predictor variable. The degree of significance is indicated by one or two asterisks, indicating p-values less than 0.10 and 0.05, respectively. Benthic invertebrate dissimilarities were estimated using Bray-Curtis indices, and Euclidean distances represent dissimilarities among aerial invertebrates and abiotic variables. Biotic variables were only available from 2014. Genetic distances were calculated using Weir and Cockerham's F_{ST} and neutral SNPs. Analyses were conducted with and without Cave 25 (see text for justification).

Model	All populations			Without Cave 25		
	R^2	β	P(>F)	R^2	β	P(>F)
2014						
Body shape ~						
Abiotic distance	0.031	-0.045	0.630	0.025	-0.105	0.290
Benthic invertebrates	"	0.024	0.812	"	0.068	0.527
Aerial invertebrates	"	0.143	0.055 *	"	0.093	0.286
Craniofacial shape ~						
Abiotic distance	0.049	0.071	0.440	0.007	0.065	0.529
Benthic invertebrates	"	-0.042	0.677	"	0.023	0.832
Aerial invertebrates	"	0.167	0.033 **	"	-0.039	0.670
2019						
Body shape ~						
Abiotic distance	0.011	-0.099	0.291	0.011	-0.098	0.323
Craniofacial shape ~						
Abiotic distance	0.015	0.117	0.192	0.016	0.125	0.202
2014 and 2019 combined						
Body shape ~						
Abiotic distance	0.013	-0.098	0.241	0.014	-0.105	0.236

Craniofacial shape ~						
Abiotic distance	0.075	0.078	0.364	0.002	0.042	0.652

Table 7. Phenotype-environment associations for populations of Arctic charr (Salvelinus alpinus) sampled from lava caves around Lake Mývatn, Iceland based on redundancy analyses (RDAs). Relationships were detected between multivariate phenotypic data (first three PC axes) and groups of ecological variables. Response and each predictor variable are separated by a tilde (~). The degree of significance is indicated by one or two asterisks, indicating p values less than 0.10 and 0.05, respectively. High-level Genetic Clusters (HLGCs) are comprised of ancestry coefficients reflecting membership to the three main genetic clusters (see text for justification). Partial redundancy analyses were used to control for the effects of the matrix after the '|' symbol. Models were assessed with phenotypic data from both sampling years together and then separate by year. Analyses were conducted with and without Cave 25 (see text for justification).

Model	All populations			Cave 25 excluded		
	Adjusted R^2	F	*p* value	Adjusted R^2	F	*p* value
2014						
Whole body shape						
~ Abiotic factors	0.040	1.144	0.348	-0.026	0.917	0.522
~ Benthic invertebrates	-0.086	0.842	0.643	-0.136	0.778	0.735
~ Aerial invertebrates	0.010	1.136	0.374	0.040	1.535	0.188
~ HLGCs	0.093	1.715	0.153	0.010	1.720	0.159
Whole body shape						
~ Abiotic factors | HLGCs	-0.015	0.949	0.534	-0.089	0.75	0.648
~ Benthic invertebrates | HLGCs	0.104	1.223	0.359	0.193	1.429	0.290
~ Aerial invertebrates | HLGCs	0.115	2.742	0.060 *	-0.080	0.102	0.948
Craniofacial shape						
~ Abiotic factors	-0.045	0.850	0.553	-0.054	0.832	0.564
~ Benthic invertebrates	0.248	1.661	0.214	0.202	1.468	0.314
~ Aerial invertebrates	0.022	1.313	0.244	0.022	1.290	0.254
~ HLGCs	0.089	1.683	0.189	0.100	1.772	0.171
Craniofacial shape						
~ Abiotic factors | HLGCs	-0.114	0.69	0.681	-0.025	0.920	0.515
~ Benthic invertebrates | HLGCs	0.262	1.646	0.250	0.307	1.870	0.156

~ Aerial invertebrates | HLGCs	0.013	1.164	0.308	0.072	2.036	0.160
2019						
Whole body shape						
~ Abiotic factors	-0.050	0.834	0.612	0.023	1.075	0.424
~ HLGCs	-0.056	0.628	0.690	-0.020	0.870	0.542
~ Abiotic factors | HLGCs	0.011	1.031	0.455	0.110	1.333	0.283
Craniofacial shape						
~ Abiotic factors	-0.083	0.733	0.685	-0.022	0.929	0.521
~ HLGCs	-0.110	0.320	0.899	-0.084	0.499	0.829
~ Abiotic factors | HLGCs	0.215	1.681	0.173	0.212	1.709	0.139
2014 and 2019 combined						
Whole body shape						
~ Abiotic factors	0.172	1.725	0.146	0.200	1.813	0.162
~ HLGCs	0.194	2.682	0.052 *	0.244	3.10	0.052 *
~ Abiotic factors | HLGCs	0.052	1.206	0.378	0.081	1.330	0.312
Craniofacial shape						
~ Abiotic factors	0.097	1.377	0.262	0.088	1.313	0.310
~ HLGCs	0.225	3.037	0.054 *	0.236	3.011	0.059 *
~ Abiotic factors | HLGCs	-0.031	0.883	0.537	-0.093	0.701	0.685

FIGURES

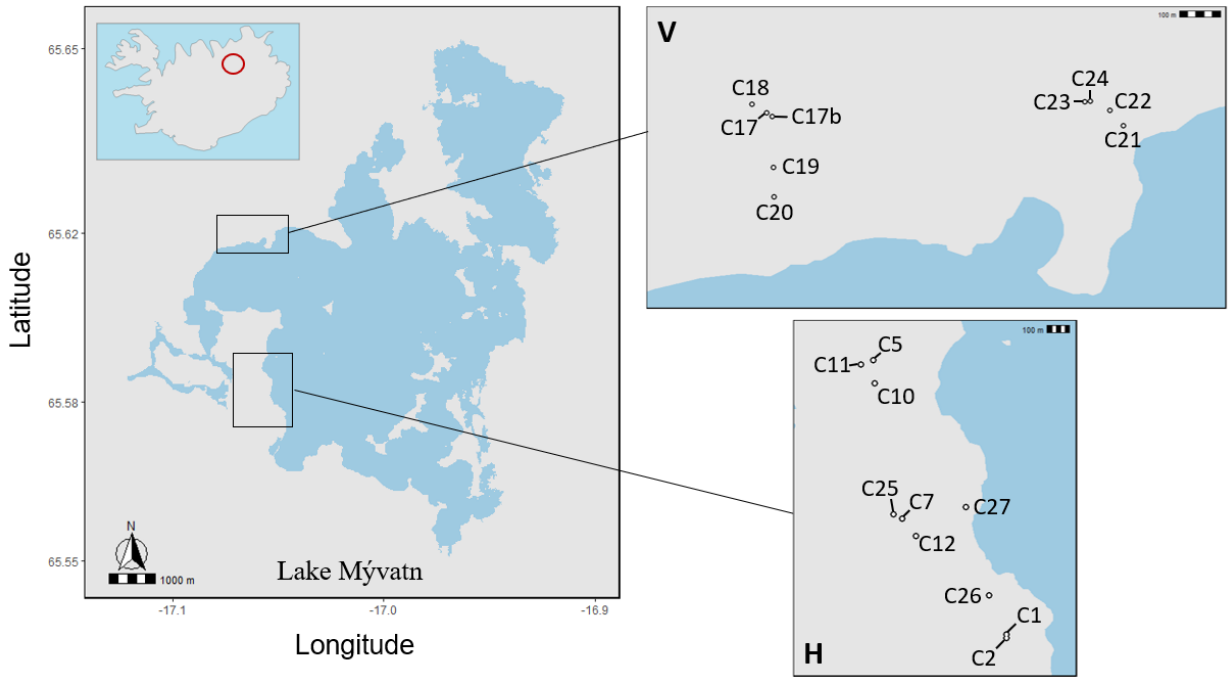

Figure 1. Location of nineteen lava caves around Lake Mývatn, Iceland (top left inset, red circle) where populations of Arctic charr (*Salvelinus alpinus*) have been monitored since 2012. Nine caves are situated within the Vindbelgur region (V, top right) and ten caves are situated within the Haganes region (H, bottom right).

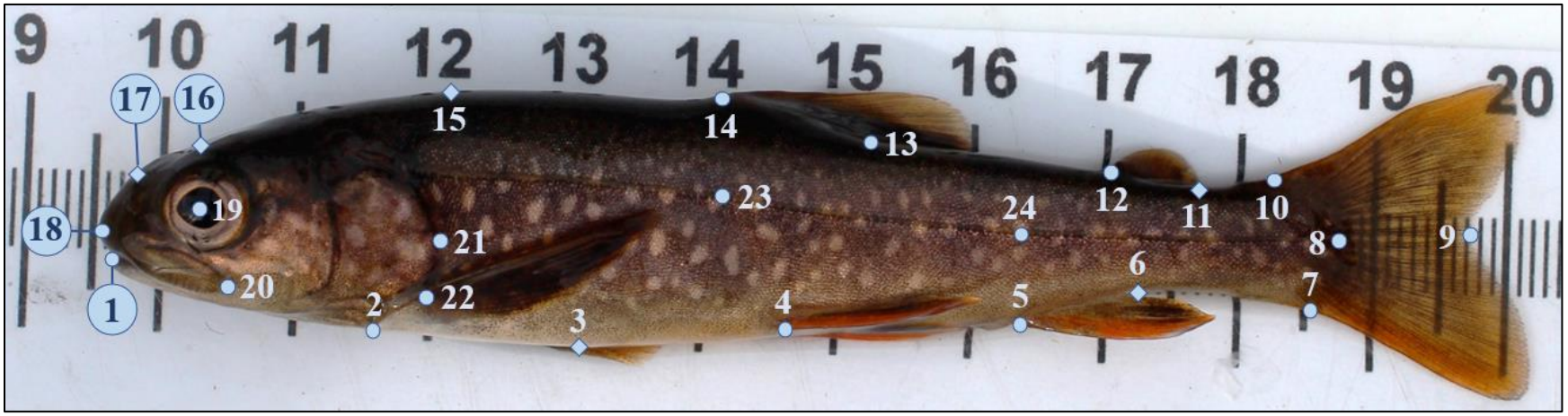

Figure 2. Positions of the landmarks used to characterize shape variation in Arctic charr (*Salvelinus alpinus*), sampled from lava caves around Lake Mývatn, Iceland. Landmarks 6, 11, 15, 16 and 17 are sliding landmarks (diamonds), whereas the remaining 19 landmarks are fixed (circles). Landmarks 23 and 24 were used only for unbending and were removed prior to subsequent analyses. Landmarks 1 to 22 were used to characterize body shape variation, whereas landmarks 1, 2 and 14 through to 22 were used to characterize craniofacial shape variation (modified from Leblanc et al. unpublished manuscript). Landmark positions are as follows: (1) Anterior tip of the lower mandible, (2) Ventral point of the opercula opening, (3) Half distance between landmarks 2 and 4, (4) Anterior insertion of the pelvic fin, (5) Anterior insertion of the anal fin, (6) Posterior insertion of the anal fin, (7) Ventral insertion of the caudal fin, (8) Posterior point of the hypural bone at the lateral line, (9) Fork of the caudal fin, (10) Dorsal insertion of the caudal fin, (11) Posterior insertion of the adipose fin, (12) Anterior insertion of the adipose fin, (13) Posterior insertion of the dorsal fin, (14) Anterior insertion of the dorsal fin, (15) Posterior edge of the cranium, (16) Top of the cranium at the midpoint of the eye, (17) Middle of the snout, (18) Upper tip of the snout, (19) Center of the bony orbit of the eye, (20) Posterior point of the maxilla, (21) The most posterior point on the curve of the operculum, (22) Anterior insertion of the pectoral fin, (23) Lateral line below the anterior insertion of the dorsal fin, (24) Lateral line above the anterior insertion of the anal fin.

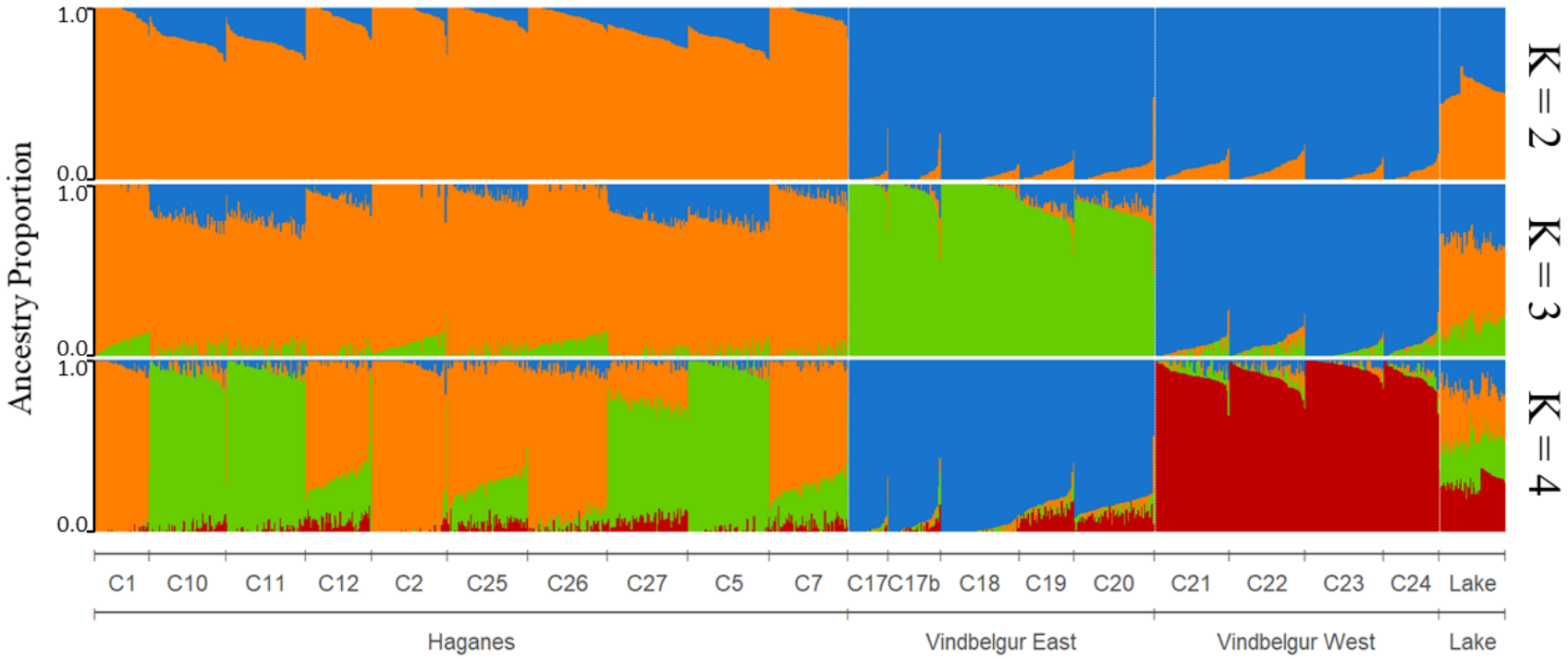

Figure 3. Patterns of genetic admixture among populations of Arctic charr (*Salvelinus alpinus*) sampled from 19 lava caves around Lake Mývatn, Iceland and two morphs (Krús and Generalist) from within the lake. Ancestry coefficients were estimated using the sNMF algorithm and neutral SNP frequencies. Populations are distributed across three geographic regions: Haganes, Vindbelgur and the Lake. Vertical bars represent individuals, and the y-axis depicts the proportion of each genome assigned to each genetic cluster. Patterns of admixture are presented for two, three and four genetic clusters (K) (see text for justification).

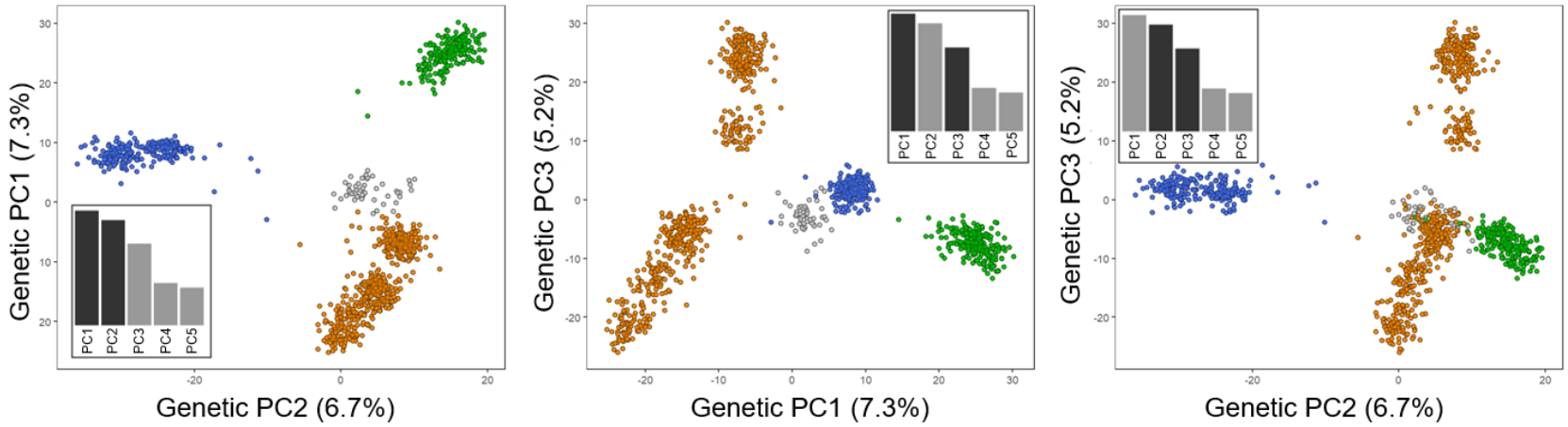

Figure 4. Genetic differentiation between individual Arctic charr (*Salvelinus alpinus*) sampled from and near Lake Mývatn, Iceland. Principal components analyses were conducted using neutral SNP allele frequencies. Colors indicate High-level Genetic Cluster (HLGC; see text for justification) membership to Haganes (orange), Vindbelgur East (green), Vindbelgur West (blue), and the lake (gray). Patterns of genetic differentiation are illustrated using PCA scores along the first three axes of variation. Scree plots denoting the proportion of variation explained by each axis are inset within each plot (dark bars indicate which axes are represented).

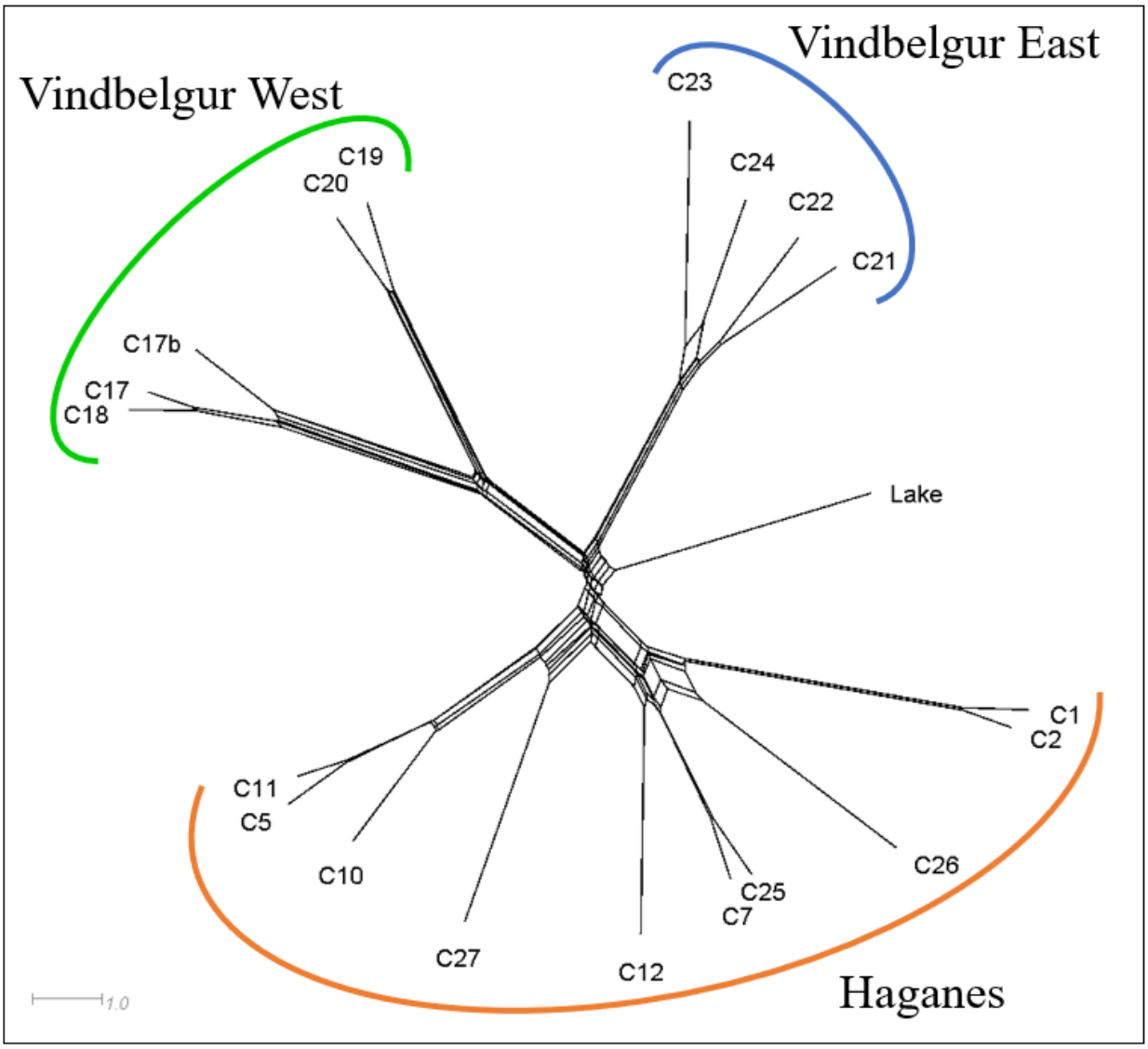

Figure 5. Patterns of genetic differentiation between populations of Arctic charr (*Salvelinus alpinus*), as indicated by Neighbor-joining trees. Individuals were sampled from 19 lava caves around Lake Mývatn in addition to samples being collected from within the lake. The tree was constructed from Euclidean distances between all populations. The three high-level genetic clusters are depicted as Haganes (orange), Vindbelgur West (green), and Vindbelgur East (blue) (see text for justification).

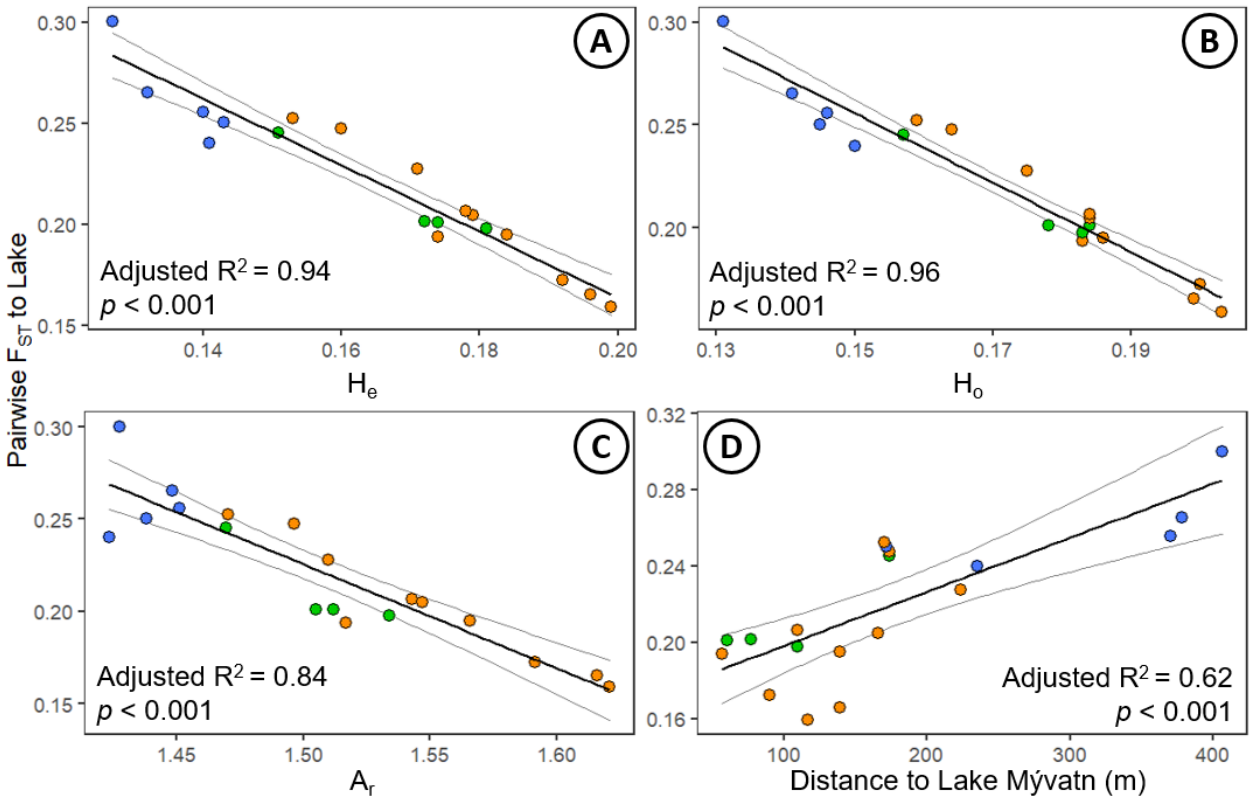

Figure 6. Relationships between estimates of within population genetic diversity and the degree of differentiation (Weir and Cockerham's F_{ST}) for 19 populations of Arctic charr (*Salvelinus alpinus*) from lava caves and charr from Lake Mývatn. Genetic diversity was quantified using expected heterozygosity (H_e) (A), observed heterozygosity (H_o) (B), and allelic richness (A_r) (C). Pairwise estimates of genetic differentiation were also compared to linear distances to the edge of Lake Mývatn (D). Simple linear regression models were used to assess relationship strength (adjusted R^2) and significance (p-values). The colour of each point reflects membership to each of three high-level genetic clusters: Haganes (orange), Vindbelgur West (green) and Vindbelgur East (blue) (see text for justification).

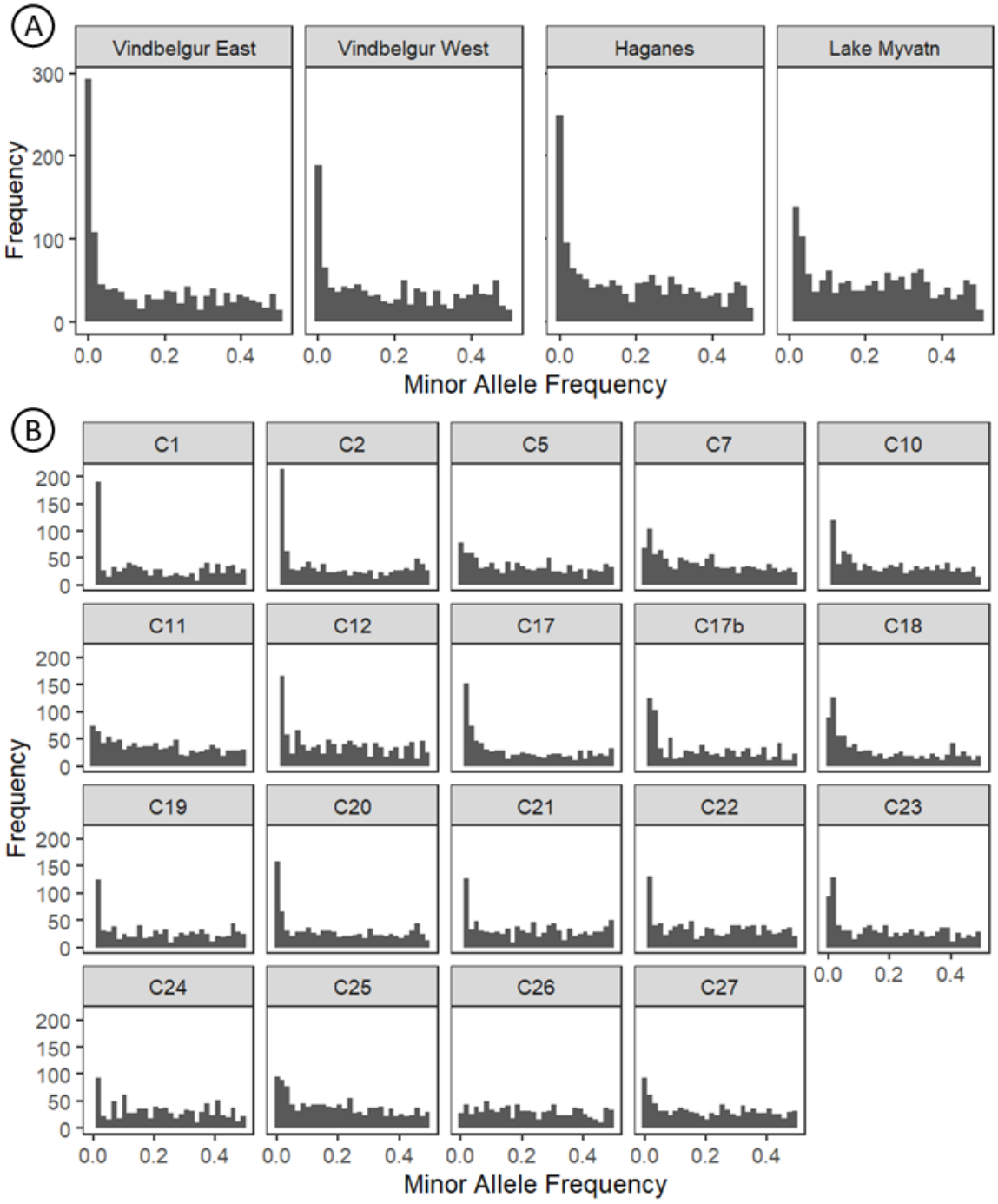

Figure 7. Minor allele frequency spectra among four high-level genetic clusters of Icelandic Arctic charr (*Salvelinus alpinus*) (top; A) and 19 populations considered independently (bottom; B). The intensity of historical population bottlenecks is proportional to how L-shaped allele frequency distributions are.

SUPPLEMENTARY TABLES

Table S1. Minimum, maximum, and mean fork length and centroid size among Arctic charr (*Salvelinus alpinus*) sampled in June and August of 2014 and 2019 from 19 lava caves around Lake Mývatn, Iceland. Shape data were superimposed prior to estimating centroid size. Note that both centroid size and fork length differ between sampling years, but not months (see Results).

.Sampling event		Fork Length (mm)			Centroid Size		
		Minimum	**Maximum**	**Mean (± SD)**	**Minimum**	**Maximum**	**Mean (± SD)**
2014	June	44	210	92.9 ± 25.7	4.34	4.74	4.46 ± 0.046
	August	50	171	100.9 ± 19.7	4.34	4.60	4.45 ± 0.041
2019	June	45	191	80.6 ± 26.3	4.33	4.60	4.47 ± 0.045
	August	44	207	87.8 ± 27.0	4.35	4.60	4.48 ± 0.042

Table S2. SNP genotyping and filtering among Arctic charr (*Salvelinus alpinus*) sampled from the Mývatn basin, Iceland. All individuals were genotyped using the 87k SNP Axiom Genotyping Array for Arctic charr (Nugent et al., 2019). Filtering steps were first applied to data from all individuals (Cave and Lake Populations, N = 1055) and cave-dwelling individuals only (Cave Populations Only, N = 1005) (refer to text for more information and sample sizes). Refer to text for information regarding the software and parameters used to filter SNPs.

Filtering step	Cave and Lake Populations		Cave Populations Only	
	Loci removed	Loci remaining	Loci removed	Loci remaining
Genotyping array	0	86504	0	86504
Retain recommended loci	19949	66555	20413	66091
Retain biallelic loci	60701	5854	62659	3432
Minor allele count ≥ 3	2954	2900	1382	2025
Remove non-segregating loci	65	2835	65	1960
Remove inconsistent genotypes	208	1962	208	1752
Remove Pcadapt outliers	39	1945	28	1724
Remove BayeScan outliers	0	1945	0	1724
Subset SNPs in LD	223	1722	171	1553

Table S3. Distribution of 1581 biallelic SNPs genotyped among 1005 Arctic charr (*Salvelinus alpinus*) from 19 caves around Lake Mývatn, Iceland. The software BayeScan and pcadapt were employed to identify loci that significantly deviate from neutrality, whereas selectively neutral loci were those not detected by either approach. Where possible, SNPs were positioned to the Salvelinus sp. genome (Christensen et al., 2018). SNPs not positioned to the genome were designated to a pseudochromosomal contig, referred here to as AC38.

Linkage Group	Length (Mbp)	Outlier SNPs	Neutral SNPs	SNP density (Mb^{-1})
AC01	58.02	0	19	0.33
AC02	43.54	0	26	0.60
AC03	36.00	0	30	0.83
AC04p	28.29	0	22	0.78
AC04q.1	50.52	0	31	0.61
AC04q.2	29.60	0	7	0.24
AC05	37.08	1	38	1.02
AC06.1	30.25	0	20	0.66
AC06.2	26.03	0	35	1.34
AC07	34.30	0	19	0.55
AC08	54.84	0	49	0.89
AC09	32.65	1	26	0.80
AC10	22.46	0	12	0.53
AC11	51.12	2	52	1.02
AC12	13.98	0	11	0.79
AC13	50.98	2	49	0.96
AC14	54.10	0	32	0.59
AC15	67.33	0	38	0.56
AC16	42.87	1	25	0.58
AC17	41.84	0	22	0.52
AC18	72.74	1	52	0.72
AC19	38.23	0	44	0.15
AC20	80.00	4	69	0.86
AC21	6.91	0	5	0.72
AC22	37.60	0	31	0.82
AC23	49.63	2	24	0.48
AC24	11.43	0	14	1.22
AC25	26.20	0	17	0.65
AC26	49.93	0	41	0.82
AC27	38.73	2	41	1.06
AC28	32.73	2	39	1.19
AC29	40.00	0	28	0.70
AC30	26.19	0	27	1.03
AC31	32.01	0	45	1.41
AC32	38.48	0	42	1.09
AC33	38.08	0	32	0.84
AC34	8.96	1	8	0.89

AC35	21.60	0	6	0.28
AC36	41.23	0	26	0.63
AC37	19.55	1	27	1.38
AC38	NA	6	565	NA

Table S4. Physical characteristics of the lava cave habitats near Lake Mývatn, Iceland. Water temperature, pH, oxygen saturation and conductivity were estimated using a calibrated multi-probe sonde. Water chemistry values are averages from data collected in June and August annually between 2013 and 2019 (Kristjánsson et al., unpublished manuscript). Note that water chemistry data are unavailable for Cave 24 and Cave 26.

Cave	Region	HLGC	Distance to lake (m)	Temperature (ºC)	pH	O_2 saturation (%)	Conductivity (µS)
C1	H	H	170	6.7	8.6	63.3	149.2
C2	H	H	174	5.7	8.7	57.6	149.9
C5	H	H	110	6.6	7.9	56.4	138.5
C7	H	H	117	6.4	8.6	43.1	148.3
C10	H	H	166	6.4	8.3	65.4	154.0
C11	H	H	139	6.3	8.1	53.7	137.2
C12	H	H	90	7.0	8.3	44.4	133.2
C17	V	VW	378	5.9	8.6	73.3	103.8
C17b	V	VW	370	5.8	8.6	72.9	102.9
C18	V	VW	406	6.1	8.5	73.7	105.5
C19	V	VW	236	6.2	8.6	73.4	101.7
C20	V	VW	172	6.7	8.4	76.8	102.4
C21	V	VE	61	7.6	8.5	58.2	112.0
C22	V	VE	110	6.6	8.4	57.6	107.8
C23	V	VE	174	6.9	7.7	68.5	108.3
C24	V	VE	162				
C25	H	H	139	6.6	8.5	50.7	149.5
C26	H	H	102				
C27	H	H	57	6.4	8.3	54.9	151.6

Table S5. Average invertebrate densities among 15 lava caves around Lake Lake Mývatn, Iceland in the summer of 2014. The input of aerial invertebrates ($ml \cdot m^{-2}$) was estimated using fall-in traps. All other categories are benthic invertebrate densities (individuals per 100 cm^2) estimated from stone scrubs. Benthic invertebrates were identified to the lowest taxonomic level possible (Kristjánsson et al., unpublished data) and were grouped into taxonomic groups similar to Kreiling et al., (2021) (see text for justification).

Cave	Aerial invertebrates	Arachnida	Chironimidae	Cladocera	Copepoda	Nematoda	Oligochaeta	Ostracoda	Miscellaneous
C1	27.0	0.9	12.8	0.7	99.1	14.0	5.1	19.0	3.5
C2	412.5	0.8	3.9	11.7	31.2	0.8	4.2	5.2	0.0
C5	120.4	0.7	2.8	19.3	37.9	2.8	2.0	43.6	3.6
C7	4681.3	0.2	8.9	13.4	23.1	0.0	5.6	33.7	8.4
C10	4.2	2.3	9.4	47.3	36.0	1.8	0.0	11.2	0.2
C11	7.2	2.0	5.4	11.0	44.0	30.8	11.2	19.6	1.1
C12	3720.2	1.2	3.9	32.9	45.8	5.5	1.4	63.7	2.0
C17b	2960.2	0.6	6.8	4.1	71.3	17.2	23.4	87.6	0.0
C18	300.1	1.0	6.5	4.6	32.4	7.1	5.4	49.9	2.1
C19	1028.7	0.9	14.1	9.0	43.1	9.5	6.1	38.9	1.7
C20	206.3	1.4	3.2	23.8	35.2	15.1	0.2	9.2	2.5
C22	4337.1	1.7	9.0	16.8	16.8	14.3	0.0	56.2	0.2
C23	787.3	1.0	3.1	17.9	15.0	82.1	9.5	20.1	1.5
C25	22192.6	2.3	3.4	68.5	68.2	30.7	6.6	73.8	3.7
C27	2385.5	3.5	6.9	80.2	53.9	0.0	2.6	11.2	0.3

Table S6. Spatial and temporal partitioning of allometric variation of body shape using Procrustes ANOVAs. Shape data were obtained from 1782 Arctic charr (*Salvelinus alpinus*) sampled from 19 lava caves around Lake Mývatn, Iceland. Body shape was characterized using 22 homologous landmarks. Fork length (FL) and centroid size (Csize) are used as proxies for body size. Significance was assessed using 10,000 randomized residual permutations. The degree of significance is indicated by one, two or three asterisks, indicating p values less than 0.05, 0.01 and 0.001, respectively.

Model	R^2	F	Z	Pr(>F)
Body shape ~ FL	0.117	274.51	14.95	< 0.001 ***
Body shape ~ FL * Cave	0.015	1.90	6.28	< 0.001 ***
Body shape ~ FL * Year	0.006	13.53	7.09	< 0.001 ***
Body shape ~ FL * Month	< 0.001	2.45	2.26	0.0116 *
Body shape ~ FL * Year * Cave	0.014	1.89	6.49	< 0.001 ***
Body shape ~ FL * Month * Cave	0.010	1.28	2.61	0.0054 **
Body shape ~ Csize	0.104	242.50	14.05	< 0.001 ***
Body shape ~ Csize * Cave	0.012	1.57	4.60	< 0.001 ***
Body shape ~ Csize * Year	0.009	21.57	8.66	< 0.001 ***
Body shape ~ Csize * Month	0.001	2.51	2.54	0.0055 **
Body shape ~ Csize * Year * Cave	0.012	1.53	4.44	< 0.001 ***
Body shape ~ Csize * Month * Cave	0.009	1.21	1.98	0.0246 *

Table S7. Spatial and temporal distribution of genetic variation among 1005 Arctic charr (*Salvelinus alpinus*) sampled from 19 lava caves around Lake Mývatn, Iceland. Patterns of genetic variation were inferred from neutral SNPs using the function poppr.amova from the R package poppr (Kamvar et al., 2014). Significance values were assessed using 10,000 repetitions and a significance threshold of 0.05. Phi (ϕ) is proportional to the degree of differentiation.

Group	Variance component	Variance	% of Total	*p*-value	ϕ
Population	Within samples	306.56	71.43	0.01	0.286
	Between samples within populations	-5.46	-1.27	0.98	-0.018
	Between populations	128.08	29.84	0.01	0.298
Year	Within samples	306.56	72.57	0.01	0.274
	Between samples within years	115.70	27.39	0.01	0.274
	Between years	0.19	0.04	0.13	< 0.001
HLGC	Within samples	306.56	65.87	0.01	0.341
	Between samples within HLGCs	53.99	11.58	0.01	0.150
	Between HLGCs	104.90	22.54	0.01	0.225

Table S8. Genetic differentiation (F_{ST}) (Weir and Cockerman 1984) between 19 populations of Arctic charr (*Salvelinus alpinus*) from the Mývatn basin of Northern Iceland. Genetic distances and significance levels were estimated using neutral SNPs. The R package HIERFSTAT (Goudet 2005) was used to estimate F_{ST} and 95% confidence intervals were derived from 1000 bootstrap replicates. Cells shaded gray indicate that differentiation is not significant.

	C2	C5	C7	C10	C11	C12	C17	C17b	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	Krús	Generalist
C1	0.001	0.346	0.212	0.339	0.331	0.244	0.420	0.411	0.440	0.412	0.412	0.382	0.372	0.437	0.392	0.217	0.227	0.328	0.266	0.306
C2		0.340	0.203	0.331	0.324	0.236	0.406	0.398	0.427	0.396	0.396	0.372	0.363	0.423	0.381	0.207	0.216	0.319	0.261	0.297
C5			0.223	0.061	0.002	0.244	0.360	0.350	0.385	0.349	0.353	0.290	0.288	0.359	0.296	0.227	0.301	0.184	0.214	0.252
C7				0.212	0.208	0.099	0.329	0.317	0.357	0.305	0.307	0.273	0.267	0.318	0.275	0.002	0.124	0.183	0.156	0.207
C10					0.055	0.232	0.362	0.352	0.387	0.345	0.349	0.288	0.287	0.353	0.292	0.216	0.289	0.172	0.208	0.255
C11						0.232	0.346	0.336	0.372	0.336	0.341	0.289	0.275	0.348	0.283	0.212	0.286	0.175	0.200	0.240
C12							0.353	0.343	0.382	0.335	0.341	0.289	0.282	0.339	0.295	0.111	0.157	0.200	0.173	0.220
C17								0.036	0.000	0.276	0.267	0.373	0.359	0.432	0.373	0.332	0.373	0.377	0.288	0.319
C17b									0.053	0.241	0.232	0.353	0.340	0.408	0.355	0.319	0.369	0.361	0.279	0.305
C18										0.301	0.293	0.397	0.384	0.450	0.398	0.360	0.396	0.401	0.328	0.356
C19											0.019	0.338	0.332	0.390	0.339	0.307	0.361	0.348	0.259	0.295
C20												0.343	0.337	0.393	0.344	0.309	0.362	0.353	0.269	0.303
C21													0.040	0.132	0.068	0.275	0.340	0.284	0.208	0.250
C22														0.128	0.061	0.269	0.328	0.286	0.202	0.246
C23															0.092	0.320	0.387	0.339	0.256	0.303
C24																0.277	0.324	0.297	0.207	0.253
C25																	0.132	0.186	0.165	0.212
C26																		0.280	0.233	0.279
C27																			0.203	0.239
Krús																				0.058

Table S9. Spatial and temporal partitioning of phenotypic variation using a pair of Procrustes ANOVAs. Shape data were obtained from 1782 Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland in June and August of 2014 and 2019. Morphology for body and craniofacial shape are characterized by 22 and 11 landmarks, respectively (Figure 2). The HLGC factor refers to three high-level genetic clusters identified (see Results). Significance was assessed using 10,000 randomized residual permutations and the degree of significance is indicated by one, two or three asterisks, indicating p values less than 0.05, 0.01 and 0.001, respectively.

Factor	Df	R^2	F	Z	Pr(>F)
Body shape					
HLGC	2	0.0024	2.28	2.92	0.0012 **
Month	1	0.0032	6.01	4.64	< 0.0001 ***
Year	1	0.0459	86.14	11.27	< 0.0001 ***
HLGC * Month	2	0.0013	1.23	0.83	0.2067
HLGC * Year	2	0.0018	1.73	1.96	0.0242 *
Month * Year	1	0.0013	2.47	2.38	0.0099 **
HLGC * Month * Year	2	0.0015	1.39	1.25	0.1029
Craniofacial shape					
HLGC	2	0.0035	3.48	3.61	0.0004 ***
Month	1	0.0024	4.81	3.46	0.0003 ***
Year	1	0.1006	200.72	14.88	< 0.0001 ***
HLGC * Month	2	0.0014	1.38	1.03	0.1550
HLGC * Year	2	0.0020	1.95	1.96	0.0242 *
Month * Year	1	0.0014	2.80	2.28	0.0115 *
HLGC * Month * Year	2	0.0012	1.21	0.69	0.2468

Table S10. Effective migration rates (proportion of individuals per generation) between lava cave populations of Arctic charr (*Salvelinus alpinus*) sampled from the Mývatn basin, Iceland. Effective migration rates were estimates using BA3-SNPs (Wilson and Rannala 2003; Mussmann et al., 2019). Estimates (standard deviation) here are average values obtained from 5 runs of each dataset. Significant values are indicated in shaded cells with bold text.

Haganes

Sending \ Receiving	C1	C2	C5	C7	C10	C11	C12	C25	C26	C27
C1		**0.161 (0.025)**	0.007 (0.007)	0.007 (0.006)	0.013 (0.009)	0.006 (0.006)	0.007 (0.007)	0.007 (0.006)	0.007 (0.006)	0.007 (0.007)
C2	**0.198 (0.027)**		0.005 (0.005)	0.005 (0.005)	0.009 (0.007)	0.006 (0.006)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.009 (0.007)
C5	0.005 (0.005)	0.005 (0.005)		0.005 (0.005)	**0.019 (0.009)**	**0.176 (0.026)**	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)
C7	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)		0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	**0.251 (0.030)**	0.006 (0.006)	0.010 (0.007)
C10	0.005 (0.005)	0.005 (0.005)	0.013 (0.008)	0.008 (0.006)		0.008 (0.007)	0.005 (0.005)	0.006 (0.005)	0.005 (0.005)	0.005 (0.005)
C11	0.005 (0.005)	0.005 (0.005)	**0.159 (0.023)**	0.006 (0.005)	0.026 (0.030)		0.005 (0.005)	0.010 (0.007)	0.005 (0.005)	0.005 (0.005)
C12	0.006 (0.005)	0.006 (0.005)	0.006 (0.005)	**0.026 (0.012)**	0.016 (0.009)	0.006 (0.006)		0.006 (0.006)	0.008 (0.007)	0.006 (0.006)
C25	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	**0.100 (0.018)**	0.007 (0.006)	0.007 (0.006)	0.005 (0.005)		0.009 (0.007)	0.005 (0.005)
C26	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)		0.005 (0.005)
C27	0.005 (0.005)	0.005 (0.005)	0.008 (0.006)	0.005 (0.005)	0.012 (0.007)	0.009 (0.007)	0.005 (0.005)	0.005 (0.005)	0.006 (0.005)	

Vindbelgur

Sending \ Receiving	C17	C17b	C18	C19	C20	C21	C22	C23	C24
C17		0.042 (0.023)	**0.231 (0.024)**	0.009 (0.008)	0.009 (0.008)	0.009 (0.009)	0.009 (0.009)	0.009 (0.008)	0.009 (0.008)
C17b	0.014 (0.009)		**0.126 (0.044)**	0.011 (0.009)	0.010 (0.008)	0.007 (0.007)	0.007 (0.010)	0.007 (0.007)	0.007 (0.007)
C18	0.011 (0.008)	**0.063 (0.023)**		0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)
C19	0.007 (0.007)	0.006 (0.006)	0.007 (0.007)		**0.104 (0.023)**	0.008 (0.007)	0.008 (0.009)	0.007 (0.006)	0.008 (0.007)
C20	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	**0.143 (0.025)**		0.005 (0.005)	0.005 (0.007)	0.005 (0.005)	0.005 (0.005)
C21	0.006 (0.005)	0.005 (0.005)	0.005 (0.005)	0.006 (0.005)	0.009 (0.007)		**0.174 (0.041)**	0.005 (0.005)	0.010 (0.007)
C22	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.010 (0.007)		0.005 (0.005)	**0.067 (0.016)**
C23	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)		**0.110 (0.040)**
C24	0.007 (0.007)	0.006 (0.007)	0.007 (0.006)	0.007 (0.007)	0.007 (0.007)	0.013 (0.012)	**0.045 (0.020)**	0.013 (0.009)	

Table S11. Relationships among geographic distance, abiotic and biotic ecological variation among 15 lava caves sampled around Lake Mývatn, Iceland based on simple Mantel tests. Benthic invertebrate dissimilarities were calculated as Bray-Curtis dissimilarity indices. Euclidean distances were computed among aerial invertebrate inputs and abiotic factors. Mantel statistics (r) and *p* values are presented below and above the diagonal, respectively. Significant values ($p < 0.05$) are indicated in bold. Analyses were conducted with and without Cave 25 (see text for justification).

Dataset	**All populations**				**Without Cave 25**			
	Benthic invertebrates	Aerial invertebrates	Abiotic factors	Geographic distance	Benthic invertebrates	Aerial invertebrates	Abiotic factors	Geographic distance
Benthic invertebrates		0.757	**0.036**	0.081		0.318	**0.030**	0.065
Aerial invertebrates	-0.163		0.726	0.758	0.052		0.492	0.579
Abiotic factors	0.236	-0.140		**< 0.001**	0.246	-0.009		**< 0.001**
Geographic distance	0.138	-0.080	0.616		0.148	-0.040	0.575	

SUPPLEMENTARY FIGURES

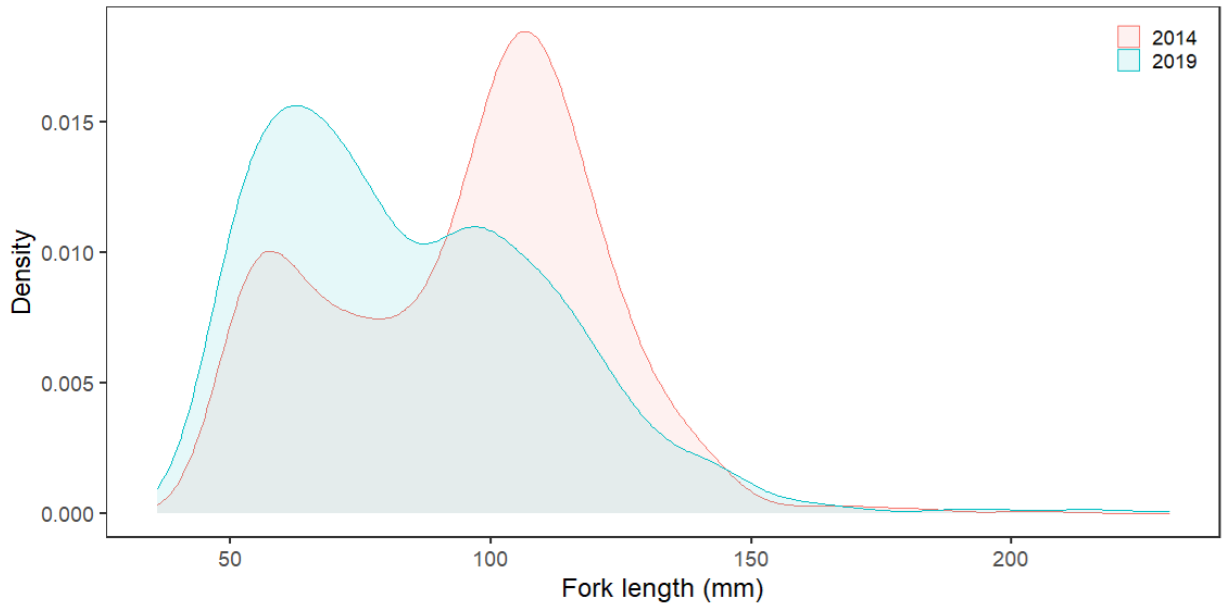

Figure S1. Distribution of fork lengths for 1782 Arctic (*Salvelinus alpinus*) sampled in June and August of 2014 and 2019 from 19 lava caves around Lake Mývatn, Iceland.

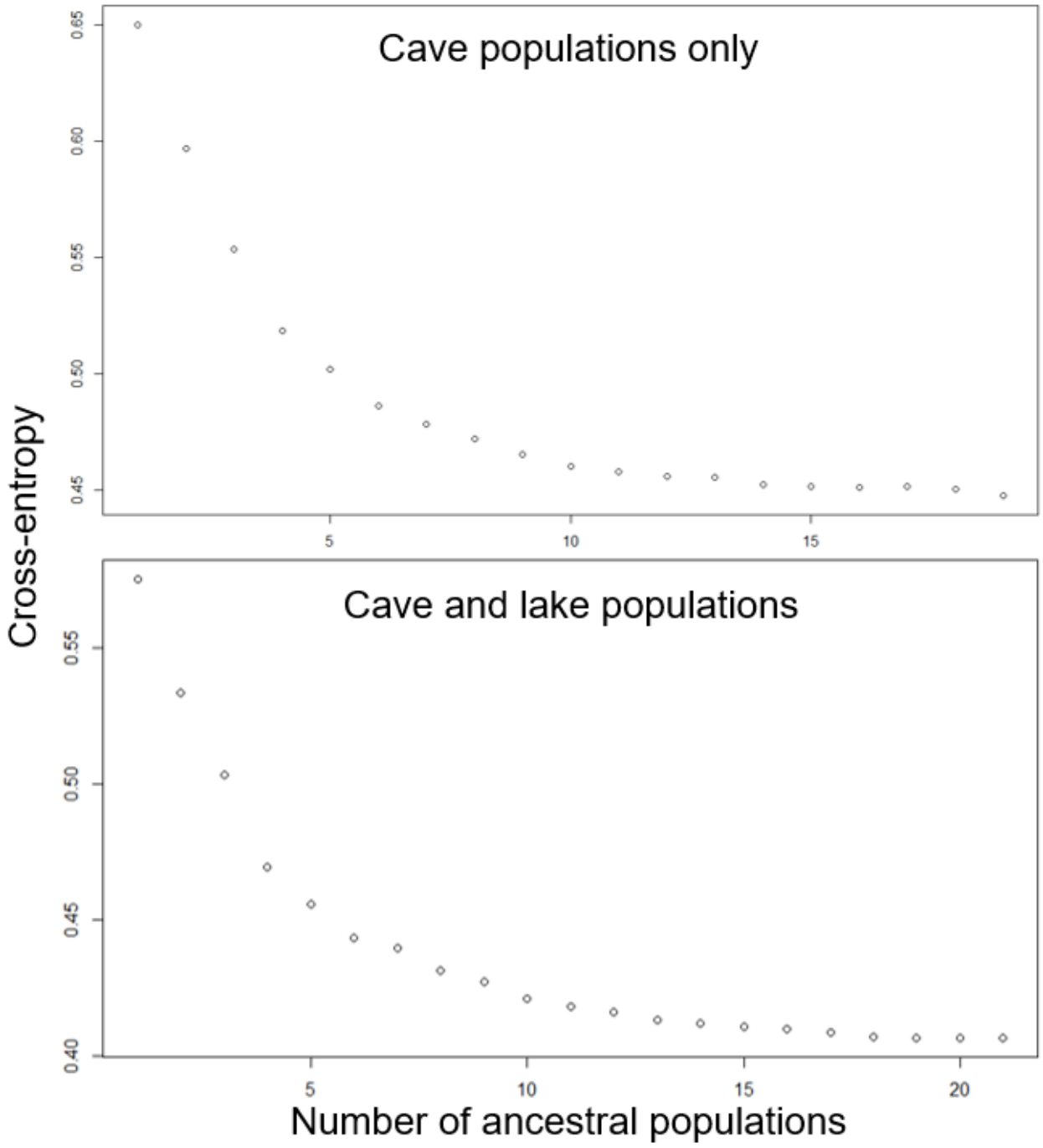

Figure S2. Detection of high-level genetic clusters among Arctic charr (*Salvelinus alpinus*) sampled the Mývatn region of northeastern Iceland. Cross-entropy values were estimated from samples from 19 lava caves around Lake Mývatn (top). This analysis was conducted a second time using data from all cave-dwelling individuals and samples from within the lake (bottom). As indicated by the elbow in the cross-entropy values, there are high-level genetic clusters present at K = 3 (see text for justification).

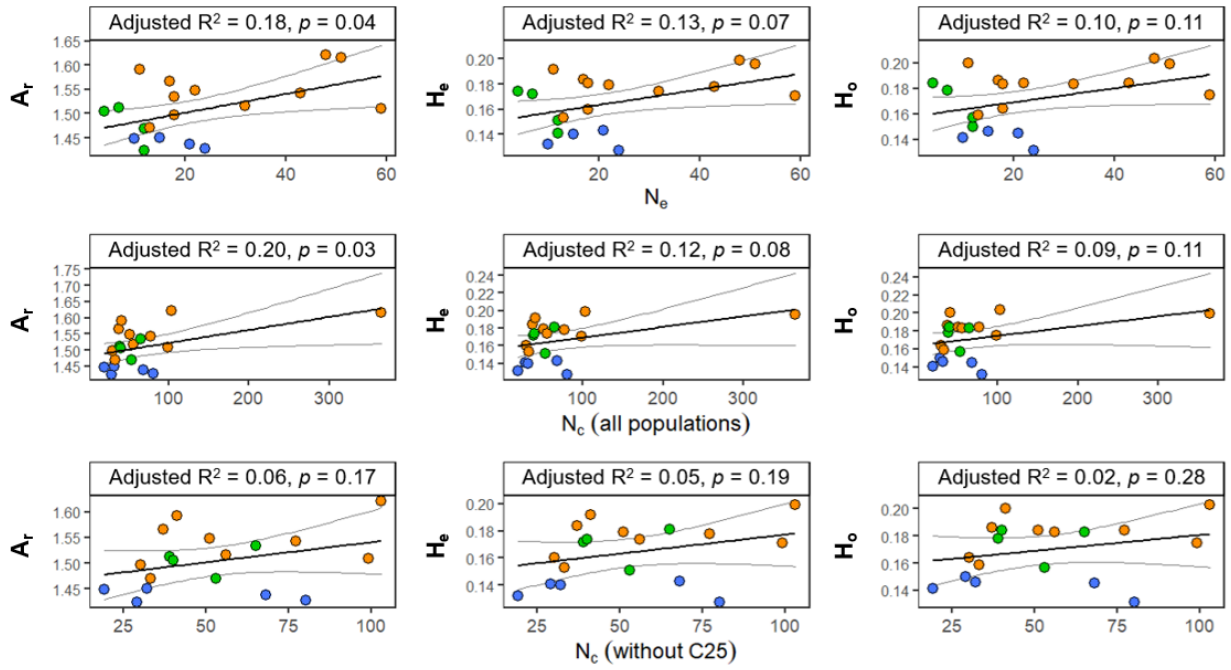

Figure S3. Relationships between census (N_c) and effective (N_e) population sizes and estimates of genetic diversity among 19 populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland. As C25 has a disproportionately large N_c, analyses were repeated with and without this population. Using neutral SNPs, allelic richness (A_r), expected heterozygosity (H_e), and observed heterozygosity (H_o) are used as estimates of genetic diversity. Gray lines indicate 95% confidence intervals. Colours represent membership to the high-level genetic clusters, identified as Haganes (orange), Vindbelgur East (green) and Vindbelgur West (blue).

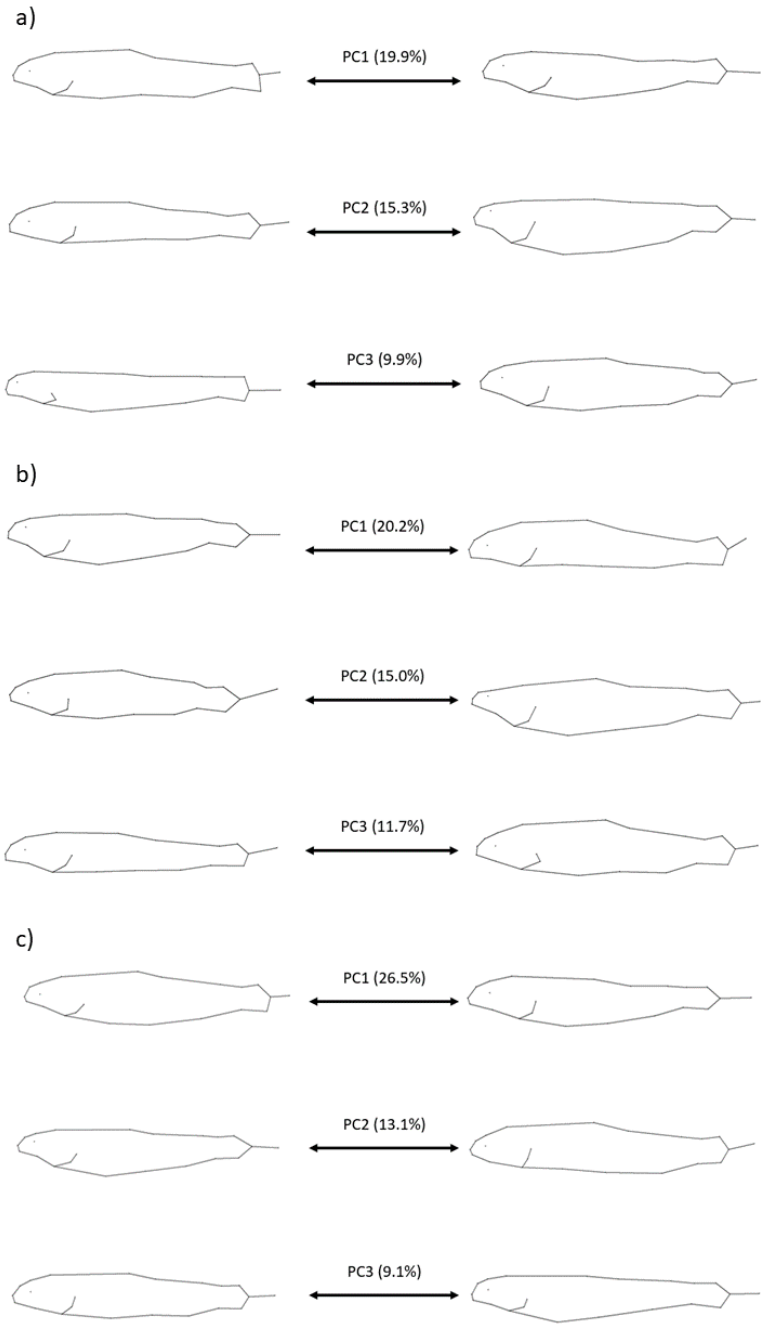

Figure S4. Body shape variation among Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn in 2014 and 2019. Data are presented separately for both sampling years grouped together (a), as well as 2014 and 2019 separately (b and c, respectively). Variation in body shape was assessed using 22 homologous landmarks. Superimposed landmark coordinates were subjected to a principal components analysis, where first three axes were retained for each grouping. Smaller values are on the left, whereas larger values are on the right of each axis. The proportion of the total phenotypic variation explained by each axis is depicted above each arrow.

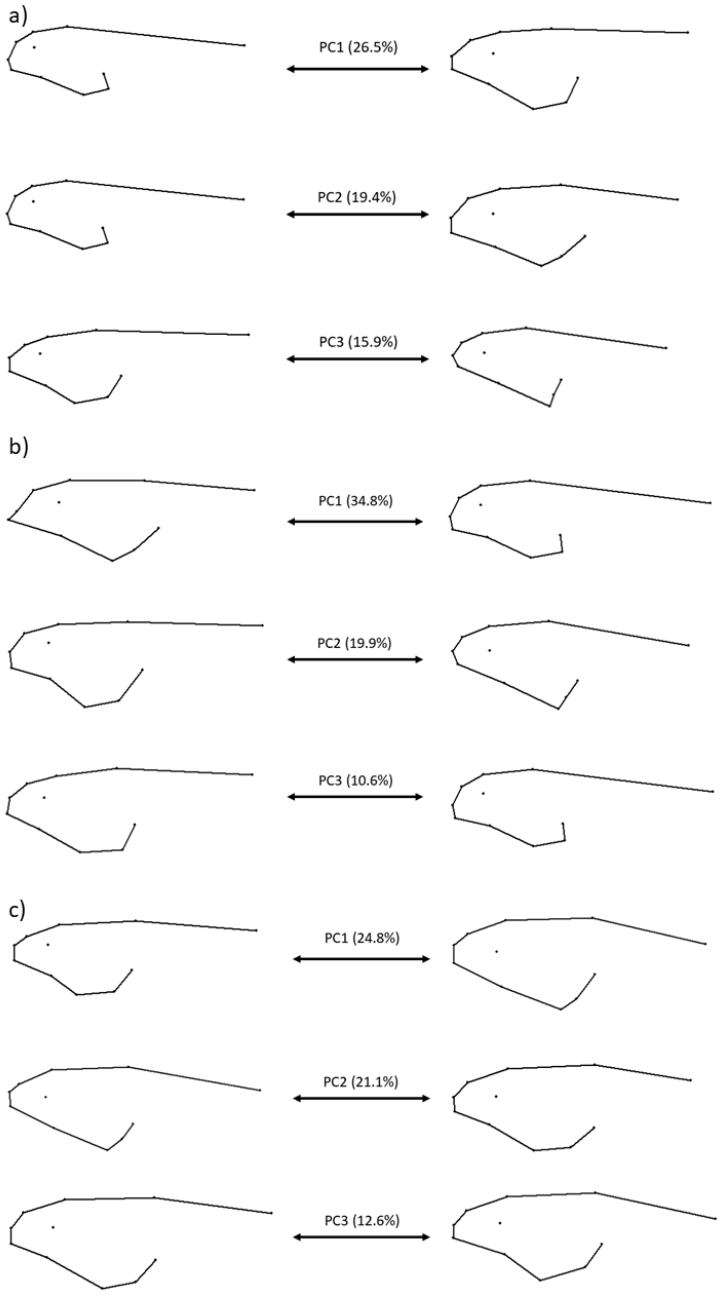

Figure S5. Craniofacial shape variation among Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn in 2014 and 2019. Data are presented separately for both sampling years grouped together (a), as well as 2014 and 2019 separately (b and c, respectively). Variation in craniofacial shape was assessed using 11 homologous landmarks. Superimposed landmark coordinates were subjected to a principal components analysis, where the three most informative axes were retained for each grouping. Smaller values are on the left, whereas larger values are on the right of each axis. The proportion of the total phenotypic variation explained by each axis is depicted above each arrow.

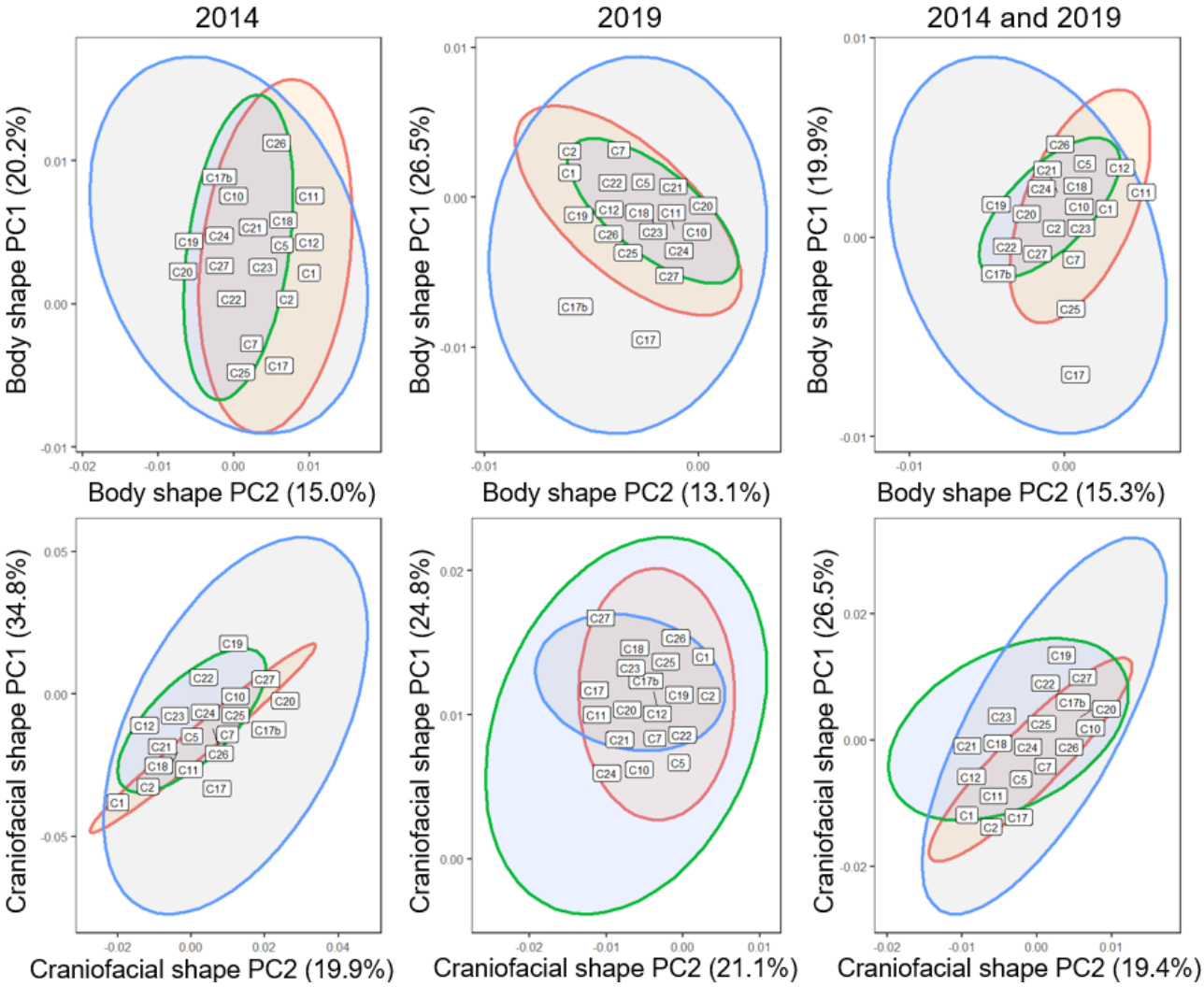

Figure S6. Patterns of phenotypic variation among 19 populations of Arctic charr (*Salvelinus alpinus*) sampled from lava caves around Lake Mývatn, Iceland in the summer of 2014 and 2019. Patterns of phenotypic differentiation are presented along the two most informative axes of variation. Labels indicate average values and ellipses indicate 95% confidence intervals for the high-level genetic clusters, identified as Haganes (orange), Vindbelgur East (green), and Vindbelgur West (Blue).